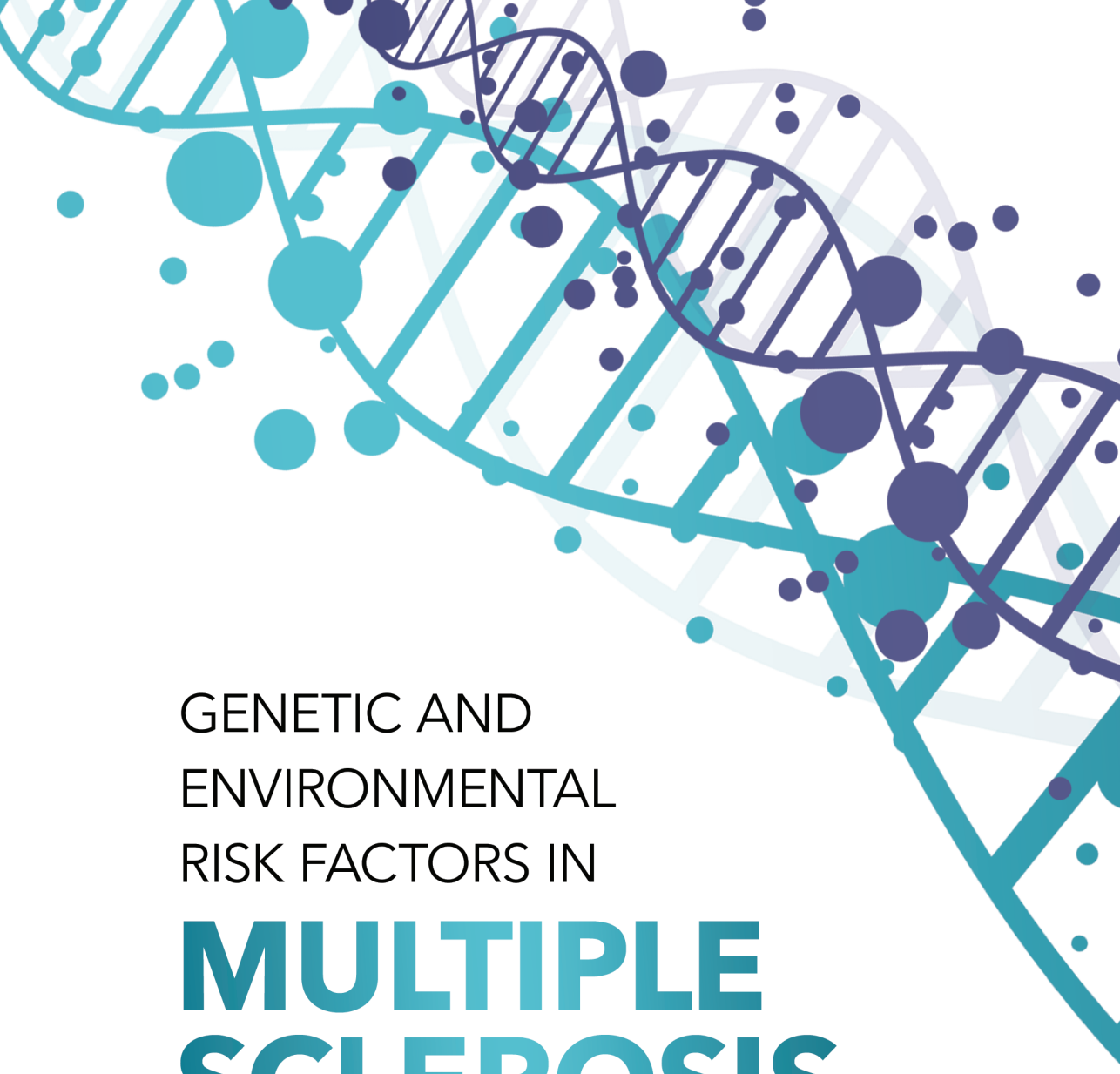


GENETIC AND ENVIRONMENTAL RISK FACTORS IN

MULTIPLE SCLEROSIS MULTIPLEX FAMILIES

Julia Y. Mescheriakova

Genetic and Environmental Risk Factors in Multiple Sclerosis Multiplex Families

Julia Y. Mescheriakova

The research in this thesis was supported by the Dutch MS Research Foundation. Printing of this thesis was financially supported by Erasmus University Medical Center Rotterdam.

ISBN 978-94-6423-550-0
Author Julia Y. Mescheriakova
Layout and cover design Bregje Jaspers | www.proefschriftontwerp.nl
Print ProefschriftMaken | www.proefschriftmaken.nl

© Julia Y. Mescheriakova, 2021

All rights are reserved. No part of this book may be reproduced, distributed, stored or transmitted in any form or by any means, without prior written permission of the author.

Genetic and Environmental Risk Factors in Multiple Sclerosis Multiplex Families

Genetische en omgevingsrisicofactoren in multiplex families met multiple sclerose

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof. dr. A.L. Bredenoord

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 1 december 2021 om 10.30 uur

door

Julia Y. Mescheriakova
geboren te Ashkhabad, Turkmenistan

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotoren:	prof. dr. P.A. van Doorn
	prof. dr. ir. C.M. van Duijn
Overige leden:	prof. dr. C.C.W. Klaver
	prof. dr. A.G. Uitterlinden
	prof. dr. J.D. Laman

TABLE OF CONTENTS

PART I

INTRODUCTION

Chapter 1

GENERAL INTRODUCTION AND SCOPE OF THIS THESIS

Adapted from:

Mescheriakova, J.Y, Kreft, K.L. and Hintzen, R.Q. (2013). *Genetics of Multiple Sclerosis* in T. Yamamura & B. Gran (Ed.).

Multiple Sclerosis Immunology - A Foundation for Current and Future Treatments

EPIDEMIOLOGY AND CLINICAL FEATURES OF MS

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system (CNS), characterised by inflammation, demyelination and neurodegeneration, leading to progressive neurological dysfunction. It is the most common cause of disability among young adults. Many patients are unemployed within 15 years of diagnosis, more than half experience depression,[1] and divorce, and suicide rates are substantially higher compared to healthy population. About 40-65% of MS patients experience cognitive problems which can even be present in early phases of the disease,[2,3] and about the same number of patients need assistance with mobility within 20 years of MS diagnosis. [4]

Worldwide, about 2.8 million people are affected by MS.[5] The prevalence of MS in North America and Northern Europe is around 0.1%.[6] The incidence is low in childhood, peaks between 20 and 40 years of age, and becomes rare at age older than 50.[6,7] As in other autoimmune diseases, MS affects more often women than men. In the last decades the prevalence of MS has been rising, especially in women, causing an increased gender bias with a current female to male ratio of 3:1.[8,9]

A majority of patients (about 80-90%) present with an (sub)acute episode of inflammatory demyelination affecting one or more CNS sites. Clinical manifestations in motor, sensory, visuals and/or autonomic systems can be present, but more diffuse complains such as fatigue[10] or cognitive impairment can also be the first sign.[11] This first (sub)acute episode, which is called a clinically isolated syndrome (CIS) when affecting only one CNS site (with no other clinical or MRI demonstration of dissemination in space or time, or the presence of CSF-specific oligoclonal bands), can remain a single event, but can also be followed by the relapsing disease MS. Diagnosis of MS after CIS can be made by clinical and magnetic resonance imaging (MRI) evidence of dissemination of the disease in space and time. Through years different diagnostic criteria and their revisions have been developed and implemented in the clinical practice.[12–16] The chances on the second clinical episode after CIS vary between 50 % in 2-years to 82% in 20 years of follow-up.[17] MS is a complex and very heterogeneous disease in terms of progression with some patients being mildly affected for many years, while others have a very rapid disease progression with irreversible neurological decline. Approximately 85% of MS patients have a relapsing-remitting disease course with recurrent and reversible neurological symptoms (Relapsing-Remitting MS, RRMS).[18] Eventually, within 20-25 years around 60-70% of patients enter the secondary progressive phase (Secondary Progressive MS,

SPMS). In 20%, the illness is progressive from disease onset (Primary Progressive MS, PPMS). In SPMS and PPMS progression starts on average at around 40 years of age.[19] The female predominance is observed in all phenotypes of MS except for PPMS in which males are slightly more often affected than females.[20,21] Furthermore, females tend to have an earlier disease onset[21,22], more clinical relapses[23], later progression from RRMS to SPMS[21,24] and less disability progression than men[25–27].

FAMILIAL AGGREGATION OF MS

Family studies assessing MS risks in relatives of MS patients have revealed a marked familial aggregation of this disease. About 20% of MS patients have an affected family member.[28,29] It appears that MS recurrence risk in a family correlates with degree of relatedness (**Figure 1**). The highest risk for developing MS is for a monozygotic twin (~18% lifetime risk) and for an offspring of 2 affected parents (~18% lifetime risk). High concordance for MS in monozygotic twins is almost entirely dependent on female gender.[30] Siblings of MS patients have a slightly but significantly higher lifetime risk for MS (2.7%) than their parents (1.5%) or offspring (2%). Although first-degree relatives are generally at 10–20 times greater risk of developing MS compared to general population, [31] one should keep in mind that absolute risks are low: 90-98% of the first degree relatives do not develop MS. These numbers are relevant when counselling people with MS and their families. This especially applies to young women with MS diagnosis and their reproductive decision making, as many of them are concerned about MS being inherited by their child.[32] In addition, MS risks to relatives rise proportionately with increasing population prevalence as seen at higher latitudes.[31]

Many studies attempted to examine the presence of parental origin-of-effect in MS. There is a tendency towards the maternal transmission of MS,[33–35] with the exception of one study that showed a paternal origin of transmission .[36]

The familial aggregation of MS supports the role of genetic influences and has prompted studies aiming to identify these.

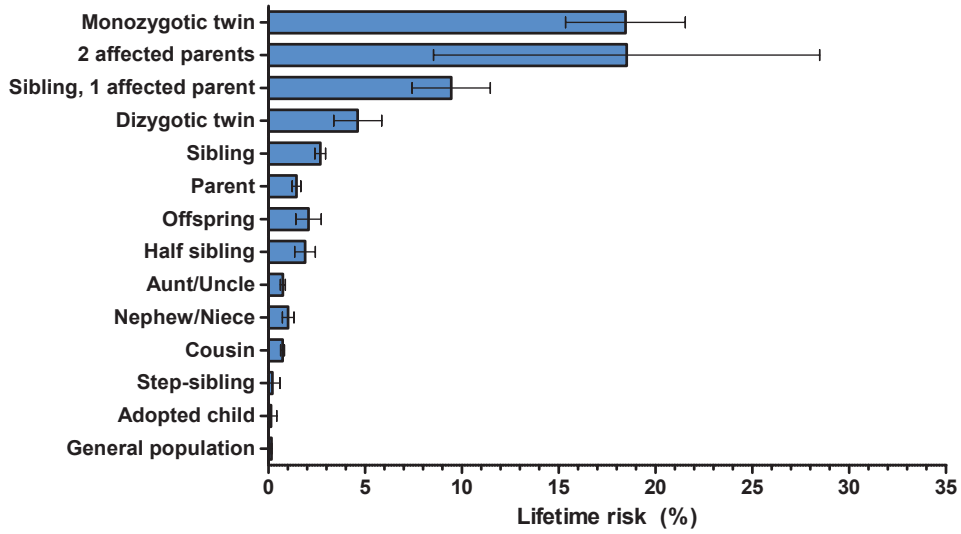


Figure 1. Lifetime risks for various types of relative of MS patients compared with a population lifetime risk. Based on systematic meta-analysis data[31].

GENETIC STUDIES IN MS

The HLA region

The major histocompatibility complex (MHC) on chromosome 6 makes the single largest contribution to MS susceptibility. The classical *HLA-DRB1*1501* allele has the strongest association to MS risk, and its role has been studied and replicated extensively.[37] The complex structure of the MHC region has made it challenging to pinpoint variants that play a causal role in MS because this region is in a very strong linkage disequilibrium.

Earlier studies described the *HLA-DRB1*1501-DQA1*0102-DQB1*0602* extended haplotype in the North European population with estimated MS risk ratios of approximately 3; homozygosity for this haplotype increases the risk by six-fold.[38] The association with this haplotype was shown to be driven entirely by the *HLA-DRB1*1501* allele.[39] The *HLA-DRB1*1501* allele is carried by approximately 30% of MS patients of Northern European descent compared to 12% in general population.

There is accumulating evidence that HLA class I loci may also influence MS susceptibility. The long-suspected HLA class-I protective effect has been confirmed in many studies[40,41] and shown to be driven mainly by the *HLA-A*0201* allele[39].

Epistasis describes a certain relationship between genes, where an allele of one gene attenuates or masks the visible output, or phenotype, of another gene. Dominant negative epistasis is seen in the HLA region as *HLA-DRB1*14* completely abrogates any risk associated with *HLA-DRB1*1501* when these are inherited together (**Figure 2**). [42,43] This interaction brings the relative risk from 3 down to 1. *HLA-DRB1*10* and *HLA-DRB1*01* provide even a greater protection against *HLA-DBR1*1501* than *HLA-DRB1*14*.[37] No evidence was found for an epistatic interaction between protective *HLA-A*0201* and MS associated other classical HLA class I and II alleles (*DRB1*1501, DRB1*1303, DRB1*0301, HLA-B*44:02, HLA-B*38:01*). [44]

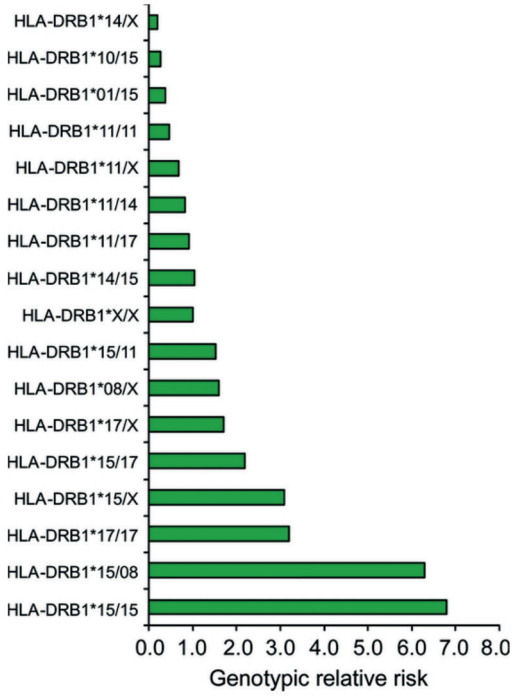


Figure 2. Genotypic relative risks for MS for combination of alleles at the *HLA-DRB1* locus.

With persmission from Ramagopalan, S.V. and Ebers, G.C. Epistasis: multiple sclerosis and the major histocompatibility complex. Neurology. 2009;72:566-567.

In other ethnical populations, the MS risk allele can be different, such as in Sardinians where MS is associated with the *HLA-DRB1*0301, DRB1*0405* and *DRB1*1303* alleles[45] or in African–Brazilian MS patients where the strongest association was observed with *DQB1*0602* rather than *DRB1*1501*[46]. In those of Northern European descent, there is evidence for a role of other variants than *HLA-DRB1*1501*. The *HLA-DRB1*17* allele has been associated with increased risk of MS, but to a lesser extent than *HLA-DRB1*1501.*[42,47] Also *DRB1*0301, *1303, *0404, *0401, *1401* and *HLA-DPB1*0301* alleles were shown to be independently associated with MS risk.[39,48,49] On contrary**,** *HLA-DRB1*14* was identified as strongly protective allele with a risk ratio of 0.3.[42] The prevalence of MS is low in Asia. In this population the prevalence of the *HLA-DRB1*14* allele is very high.[43]

Genome-wide association studies and common variants

Identifying susceptibility genes for MS has been a major challenge in the past four decades. In early years genome-wide linkage screens were applied to investigate several hundred microsatellite markers in order to identify alleles that were linked to MS. For many years the only region that was consistently linked to MS with genome-wide significance was the HLA- class II region. Later on, candidate gene studies dominated the genetic research area. Candidate gene studies are thought to have a greater ability to detect common alleles with a modest effect compared to linkage studies.[50] However, they rely on the selection of potential candidate genes, that are assumed to be involved in MS aetiology. A candidate gene study is often a case-control comparison where allele and genotype frequencies of single nucleotide polymorphisms (SNP; a variation of a single nucleotide at a specific position in the genome) or other markers in the past are compared between 2 groups. Many such studies for MS have been published to date, but the results are hardly consistent. Mainly due to smaller sample size that these studies have, and the fact that often no adjustments were made for multiple testing or population stratification. A major breakthrough in the genetic research of complex diseases such as MS, was made through the application of the genome-wide association studies (GWAS), introduced in the mid-2000s. Array-based technology allowed a very efficient genotyping of a hundred thousand SNPs throughout the genome simultaneously allowing a hypothesis-free approach to identify the location of MS genes. Using this approach, many new susceptibility loci have been discovered in MS. More than 10 GWAS have been completed for MS.[48,51–61] Unsurprisingly, the success and reliability of these studies is directly related to the number of subjects included. The latest GWAS[53] included 47,351 MS subjects and 68,284 control subjects and uncovered 200 autosomal susceptibility variants outside the MHC, one chromosome X variant, and 32 independent associations

within the extended MHC, bringing us to 233 genome-wide variants associated with MS (**Figure 3**). Most of the susceptibility variants lie in the intronic (a region inside a gene that is removed by RNA splicing before translation) and intergenic

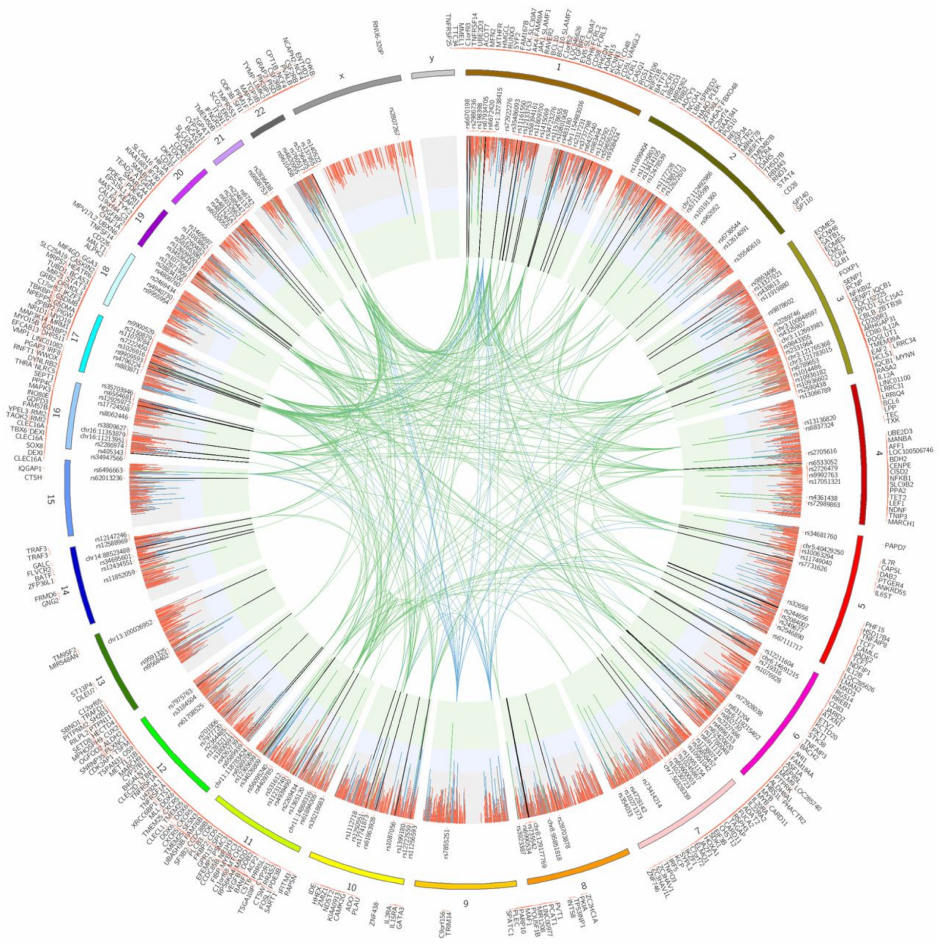


Figure 3. Current genetic map of multiple sclerosis.

From [International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019;365:eaav7188]. Reprinted with permission from AAAS.

(a stretch of DNA sequence located between genes) regions of nearby genes, and are believed to be involved in regulatory mechanisms. A few functional studies following GWAS have suggested that the identified genetics variants could alter splicing of exons, and hereby inhibit downstream signaling[62,63], or could change the expression of nearby genes [64–66]. Cell-specific protein network analyses revealed that the genetic susceptibility variants are primarily engaged in the cellular pathways in myeloid, B and T cell arms of the human adaptive immune response. Also, a smaller but not negligible contribution of CNS pathways to MS risk was found with the monocyte-macrophage-microglia axis as a key player in directing the autoimmune process to the CNS.[67]

Rare variants

Available recurrence risk data for MS are consistent with a polygenic model of inheritance involving one locus of moderate effect and many loci of small effect. To date all known genetic loci explain about 39 % of MS heritability.[53] Much of the speculation about missing heritability in complex diseases such as MS has focused on the possible contribution of variants with minor allele frequency (MAF) below < 5%, and rare variants defined as MAF below < 1%. Low frequency variants could increase disease risk by two- to threefold without demonstrating clear Mendelian segregation, and contribute substantially to missing heritability.[68] However, very few examples of such variants contributing to complex traits exist, possibly due to the fact that it takes large samples sizes to identify those. Moreover, the arrays used did not contain rare variants nor did the variants on the arrays flag the rare variants by linkage disequilibrium.[69]

Current GWAS are not well equipped to capture these variants, nor do these variants capture the effects of variants with sufficiently large effects to be easily detected by classical linkage analysis in family studies.[69] To identify those, a different approach is needed. Whole exome sequencing (WES) is a genomic technique for sequencing of the protein-coding regions in the human genome (known as exomes). Severe disease-causing variants are much more likely (but by no means exclusively) to be in the protein coding sequence, which makes this technique attractive for the search of the rare variants. WES has been effective in finding causative genes in rare Mendelian diseases.[70,71]

A family-based approach is one of ways to investigate rare variants. Despite familial segregation, extended pedigrees with many disease cases are extremely rare in MS and no example of a Mendelian variant of MS has yet been found.[72] Indeed, there are few reports of WES successfully identifying rare variants in several MS families.[73–80]

However, most of these variants could not be replicated in other populations. Recently, the International Multiple Sclerosis Genetics Consortium (IMSGC) identified 4 new genes containing rare-variants in a large case-control sample of 68379 participants using an exome chip array, and showed that up to 5% of heritability of MS is explained by low-frequency variation in gene coding sequence. None of these variants were the previously identified variants in the familial studies segregating MS.[81]

PREDICTION OF DISEASE SUSCEPTIBILITY

One of the goals of understanding genetic variants is to investigate their use in prediction of MS so that environmental changes or therapeutic interventions can be applied before the disease process starts. None of these variants on their own can predict susceptibility for MS because of the polygenic architecture of MS. A concept of combined genetic risk score has been introduced to deal with the complex genetic predisposition and to consider the additive effect of many variants.[82] The genetic variants are aggregated into a genetic risk score (GRS), in an unweighted (simple summation of number of risk alleles in one's genome) or weighted fashion (risks alleles are multiplied with the effect size of each variant and summed up with each other). First weighted GRS (wGRS) that was used to predict MS in a case-control dataset, was based on 19 risk loci and had an area-under-the -curve (AUC) of 0.70.[83] The wGRS could not distinguish MS from CIS. [83]Implementation of wGRS in a family study[84] showed a higher wGRS in familial MS patients compared to sporadic MS, and for the first time suggesting that within multiplex MS families a higher genetic load of common risk variants is present. Hilven et al.[85] have shown for the Belgian patients that wGRS was associated with increased relapse rate and shorter relapse-free intervals after disease onset. In other studies wGRS was associated with MS subphenotypes and clinical course.[86,87] These discoveries suggest that GRS might have a more extensive utility in the clinical setting.

CHILDHOOD-ONSET MS

Of all adult MS patients, about 10% had their first attack during childhood.[88] A first demyelinating event in children can present with a broad spectrum of neurological symptoms caused by inflammation and damage of the central nervous system. In contrast to adults, not all children with relapsing CNS demyelination will develop MS.

This has implications for treatment decisions. In children, the symptoms are therefore categorised into syndromes, the so-called acquired demyelinating syndromes (ADS). The spectrum of ADS includes monophasic demyelinating events such as: isolated (recurrent) optic neuritis (ON), transverse myelitis (TM), other clinically isolated syndromes (CIS), monophasic neuromyelitis optica spectrum disorders (NMOSD), myelin oligodendrocyte glycoprotein (MOG) antibody disease and acute disseminated encephalomyelitis (ADEM). Based on previous cohort studies, it is estimated that 21-32% of children presented with a first event of ADS will have a future diagnosis of MS.[89,90] Accurate diagnosis is therefore essential in terms of prognosis and initiation of treatment.

*HLA-DRB1*15* is strongly associated with childhood-onset MS.[91–93] There is limited evidence that non-HLA risk loci that confer risk for adult-onset MS, are also associated with paediatric MS.[93] As childhood-onset MS is a rare occurrence, GWAS with large enough case/control samples are lacking.

ROLE OF ENVIRONMENTAL FACTORS IN MS

The increase in the incidence of MS (especially in women[94]) over the past century cannot be explained by genetic factors, because the distribution and frequency of genetic risk factors is unlikely to change over such a short period of time. As the gene pool stays constant over time, this rapid increase likely reflects changes in the lifestyle, improved access to MS diagnosis[95] and environment. Environmental factors implicated in MS aetiology that have been consistently confirmed by many studies, include low levels of vitamin D[96–100], smoking[101–107], Epstein-Barr virus (EBV) infection[107–110], and obesity[111–113]. Contradicting results were found for factors such as salt intake[89–93], alcohol[119–122] and stress[123–127]. The association of MS and EBV is well validated in various epidemiological and serological studies.[107–110] It appears that EBV antibodies are elevated already before the initial clinical manifestations of MS.[128,129] In young adult patients a primary EBV infection manifests clinically as infectious mononucleosis (IM) in 25-70%.[130] A history of IM increases the risk of MS two-fold.[109] Next to IM, there is a consistent finding that all patients with MS (>99%) are infected with EBV in comparison to 90-95% of healthy controls.[131] MS risk appears to be extremely low in EBV negative individuals.[131]

The mechanism that relates MS to EBV is still unclear. It has been shown that IgG response to Epstein-Barr virus nuclear antigen 1 (EBNA-1) has a heritability of 22-43%, suggesting that host-genetic factors are important in the immune response to EBV.[132–134] Indeed, *HLA-DRB1*1501* allele was shown to act independently and synergistically with high levels of EBV antibodies to increase risk of MS.[135] More recently, genetic loci influencing both EBNA-1 IgG titres and MS risk have been identified. [133,136,137] A meta-analysis in twins has shown that shared environmental influences during early stages of life partly affect MS susceptibility in families.[138] In relation to EBV antibody titres only few small-sized studies were conducted in MS patients, their siblings and twins, which showed somewhat variable results.[139–142]

MS AND CO-OCCURRENCE OF OTHER AUTOIMMUNE DISEASES

Clinical and epidemiological studies have shown that immune-mediated inflammatory disease can co-occur in the same individual or in closely related family members. Many autoimmune diseases share common risk genes and pathways which may explain the observed co-segregation of these diseases.[143] However, it has been shown that several genetic loci that predispose for one autoimmune disease may increase or decrease the risk of another autoimmune disease.[144]

About one third of the MS-associated genetic loci is associated with one other autoimmune disease at a genome-wide level.[58] A meta-analysis found that thyroid disease is the most frequently observed autoimmune disease in MS patients and their first-degree relatives.[145] A similar association was found between MS and inflammatory bowel disease and psoriasis, although not in relatives of MS patients.[145] These findings should be interpreted with caution because most studies in autoimmune diseases are based on self-reported diagnoses and are therefore very much prone for misclassification and reporting bias. Nevertheless, there is overwhelming evidence for shared genetic background between autoimmune diseases.

SCOPE OF THE THESIS

MS is a heterogeneous and complex disease that clusters in families. Despite extensive research including the discovery of several biomarkers and genetic risk variants, MS heritability is not completely explained and prediction of MS in individuals is still limited. One of the ways to study heritability of a disease, is to assess genetic associations within families. In this thesis, Dutch multiplex MS families (families in which a person with MS has a first- or second-degree relative with the same diagnosis) were studied in which two or more family members have MS. The aim of the studies described in this thesis is to explore genetic (known and unknown) and environmental factors in MS patients and their family members, and how these factors relate to MS risk.

In the past 12 years several genome-wide studies comprising several thousands of patients and controls from different countries including the Netherlands, discovered several genetic risk factors that increase susceptibility to MS. One of the goals of understanding genetic variants is to investigate their use in prediction of MS. Because MS has a polygenic architecture and none of the variants acts on its own, a concept of combined genetic risk score (GRS) has been introduced to deal with the complex genetic predisposition. In **chapter 2** we investigate whether GRS differs between familial and sporadic MS, and whether it can distinguish between these two, and in relation to a healthy population. In **chapter 3** we discuss the utility of the genetic risk score for the prediction of childhood-onset MS in children who first had an acquired demyelinating syndrome. Common genetic variants identified by genome-wide association studies explain only a small part of MS heritability. A part of missing heritability is believed to be hidden in rare genetic variants. In **chapter 4.1** we explore a large MS Dutch family and seek for rare coding variants in the affected individuals by means of linkage analysis and whole exome sequencing. **Chapter 4.2** explains why the used modern technique has its pitfalls. As siblings of MS patients have higher chances of developing MS compared to normal population and elevated EBNA-1 antibodies have shown to increase MS risk, we examined the immune response to EBV in MS patients and their healthy full-siblings in relation to genetic factors in **chapter 5.** MS has increased comorbidity with other autoimmune diseases and is believed to share about one third of its genes with other autoimmune diseases. Some studies suggested that first-degree relatives of MS patients could be at higher risk of autoimmune diseases other than MS. In **chapter 6** we explore to what extent MS and other autoimmune diseases co-occur in MS patients and their

first-degree relatives from Dutch MS multiplex families. In **chapter 7**, the main findings of this thesis are discussed and suggestions are made for further research.

REFERENCES

1. Feinstein, A. The neuropsychiatry of multiple sclerosis. Can J Psychiatry. 2004;49:157-63.
2. Achiron, A. & Barak, Y. Cognitive changes in early MS: A call for a common framework. J Neurol Sci. 2006;245:47-51.
3. Bobholz, J. A. & Rao, S. M. Cognitive dysfunction in multiple sclerosis: A review of recent developments. Curr Opin Neurol. 2003;16:283-8.
4. Compston, A. et al. McAlpine's Multiple Sclerosis. 2006.
5. MSIF. Atlas of MS, 3rd Edition 2020 .
6. Compston, A. & Coles, A. Multiple sclerosis. Lancet.2008;372:1502-1517.
7. Koch-Henriksen, N. The Danish Multiple Sclerosis Registry: a 50-year follow-up. Mult Scler J. 1999;5:293-296.
8. Koch-Henriksen, N. & Sørensen, P. S. The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol. 2010;9: 520-532.
9. Orton, S. M. *et al.* Sex ratio of multiple sclerosis in Canada: a longitudinal study. Lancet Neurol. 2006;5:932-6.
10. Runia, T. F. *et al*. Fatigue at time of CIS is an independent predictor of a subsequent diagnosis of multiple sclerosis. J Neurol Neurosurg Psychiatry. 2015;86:543–546
11. Feuillet, L. *et al.* Early cognitive impairment in patients with clinically isolated syndrome suggestive of multiple sclerosis. Mult Scler. 2007;13:124-7.
12. Poser, C. M. *et al.* New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. Ann Neurol. 1983;13:227-31.
13. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Ann Neurol. 2001;50:121–127.
14. Polman, C. H. *et al.* Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol. 2005;58:840-846.
15. Polman, C.H. *et al.* Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. Ann Neurol. 2011;69:292-302.
16. Thompson, P. A. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17:162-173.
17. Fisniku, L. K. *et al.* Disability and T2 MRI lesions: A 20-year follow-up of patients with relapse onset of multiple sclerosis. Brain. 2008; 131:808-817.
18. Hauser, S. L. & Oksenberg, J. R. The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. Neuron. 2006;52:61-76.
19. Confavreux, C. & Vukusic, S. Age at disability milestones in multiple sclerosis. Brain. 2006;129:595-605.
20. Thompson, A. J. *et al.* Primary progressive multiple sclerosis. Brain. 1997;120:1085-96.
21. Miclea, A. *et al.* Age-dependent variation of female preponderance across different phenotypes of multiple sclerosis: A retrospective cross-sectional study. CNS Neurosci Ther. 2019; 25:527-531.
22. Cossburn, M. *et al.* Age at onset as a determinant of presenting phenotype and initial relapse recovery in multiple sclerosis. Mult Scler. 2012;18:45-54.
23. Ribbons, K. A. *et al.* Male sex is independently associated with faster disability accumulation in relapse-onset MS but not in primary progressive MS.PLoSOne.2015;10:e0122686.
24. Confavreux, C. *et al*. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. Brain. 2003;126:770-82.

25. Kalincik, T. *et al.* Sex as a determinant of relapse incidence and progressive course of multiple sclerosis. Brain. 2013;136:3609-3617.
26. Golden, L. C. & Voskuhl, R. The importance of studying sex differences in disease: The example of multiple sclerosis. J Neurosci Res. 2017;95:633-643.
27. Koch-Henriksen, N. et al. Worsening of disability caused by relapses in multiple sclerosis: A different approach. Mult Scler Relat Disord. 2019;32:1-8.
28. Sadovnick, A. D. *et al*. Multiple sclerosis: Updated risks for relatives. Am J Med Genet. 1988;29(3):533-41.
29. Robertson, N. P. *et al.* Age-adjusted recurrence risks for relatives of patients with multiple sclerosis. Brain. 1996;119:449-55.
30. Willer, C. J. *et al*. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci. 2003;100:12877-82.
31. O'Gorman, C. *et al*. Modelling genetic susceptibility to multiple sclerosis with family data. Neuroepidemiology. 2012;40:1-12.
32. Alwan, S. *et al.* Reproductive decision making after the diagnosis of multiple sclerosis (MS). Mult Scler. 2013;19:351-8.
33. Hoppenbrouwers, I. A. *et al.* Maternal transmission of multiple sclerosis in a Dutch population. Arch Neurol. 2008;65:345-8.
34. Ebers, G. C. *et al.* Parent-of-origin effect in multiple sclerosis: Observations in half-siblings. Lancet. 2004;363:1773-4.
35. Herrera, B. M. *et al.* Parent-of-origin effects in MS: Observations from avuncular pairs. Neurology. 2008;71:799-803.
36. Kantarci, O. H. *et al.* Men transmit MS more often to their children vs women: The Carter effect. Neurology. 2006;67:305-10.
37. Ramagopalan, S. V. & Ebers, G. C. Epistasis: Multiple sclerosis and the major histocompatibility complex. Neurology. 2009; 72:566-7.
38. Dyment, D.*et al*. Review Genetics of multiple sclerosis. Lancet Neurol. 2004;3:104-10.
39. Patsopoulos, N. A. *et al.* Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and Non-HLA Effects. PLoS Genet. 2013;9:e1003926.
40. Harbo, H. F. *et al.* Genes in the HLA class I region may contribute to the HLA class II-associated genetic susceptibility to multiple sclerosis. Tissue Antigens. 2004;63:237-47.
41. Yeo, T. W. *et al.* A second major histocompatibility complex susceptibility locus for multiple sclerosis. Ann. Neurol. 2007; 61:228-36.
42. Dyment, D. A. *et al.* Complex interactions among MHC haplotypes in multiple sclerosis: Susceptibility and resistance. Hum Mol Genet. 2005;14:2019-26.
43. Ramagopalan, S. V. *et al.* The inheritance of resistance alleles in multiple sclerosis. PLoS Genet. 2007;3:1607-13.
44. Moutsianas, L. *et al.* Class II HLA interactions modulate genetic risk for multiple sclerosis. Nat Genet. 2015;47:1107-1113.
45. Marrosu, M. G. *et al.* Dissection of the HLA association with multiple sclerosis in the founder isolated population of Sardinia. Hum Mol Genet. 2001;10:2907-16.
46. Caballero, A. *et al.* DQB1*0602 confers genetic susceptibility to multiple sclerosis in Afro-Brazilians. Tissue Antigens. 1999;54:524-6.
47. Campbell, A. Modes of action of HLA-DR susceptibility specificities in sultiple sclerosis. Am J Hum Genet. 2004;74:1321-2.
48. Sawcer, S. *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476:214-9.
49. Field, J. *et al.* A polymorphism in the HLA-DPB1 gene is associated with susceptibility to multiple sclerosis. PLoS One. 2010;5:e13454.

50. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. Science. 1996;273:1516-7.

51. Sanna, S. *et al.* Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. Nat Genet. 2010;42:495-497.

52. Nischwitz, S. *et al.* Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. J Neuroimmunol. 2010;227:162-6.

53. Patsopoulos, N. A. *et al.* Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019;365:eaav7188.

54. International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide atudy. New Engl J Med. 2009;357:851-62.

55. Beecham, A. H. *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353-60.

56. De Jager, P. L. *et al.* Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet. 2009;41:776-82.

57. Comabella, M. *et al.* Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. PLoS One. 2008; 3:e3490.

58. Patsopoulos, N. A. *et al.* Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol. 2011;70:897-912.

59. Jakkula, E. *et al.* Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. Am J Hum Genet. 2010;86:285-91.

60. Bahlo, M. *et al.* Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet. 2009;41:824-8.

61. Baranzini, S. E. *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet. 2009;18:767-78.

62. Gregory, A. P. *et al.* TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. Nature. 2012; 488:508-511.

63. Maier, L. M. *et al.* Effect of soluble IL-2RA levels in multiple sclerosis subjects and the effect of soluble IL-2RA on immune responses. J. Immunol. 2011;182:1541-7.

64. De Jager, P. L. *et al.* The role of the CD58 locus in multiple sclerosis. Proc Natl Acad Sci. 2009;106:5264-9.

65. Raj, T. *et al.* Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science. 2014;344:519-23.

66. Kreft, K. L. *et al.* Decreased systemic IL-7 and soluble IL-7Rα in multiple sclerosis patients. Genes Immun. 2012;13:587-92.

67. International Multiple Sclerosis Genetics Consortium (IMSGC). A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. Nat Commun. 2019;10:2236.

68. McCarthy, M. I. *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet. 2008;9:356-69.

69. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. Nature. 2009; 461:747-53.

70. Ng, S. B. *et al.* Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 2010;42:30-5.

71. Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011;2:745-55.

72. Willer, C. J. *et al.* A genome-wide scan in forty large pedigrees with multiple sclerosis. J Hum Genet. 2007;52:955-962.

73. Dyment, D. A. *et al.* Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. Neurology. 2012;79:406-11.

74. Ramagopalan, S. V. *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. Ann Neurol. 2011;70:881-6.

75. Ross, J. P. *et al.* Analysis of CYP27B1 in multiple sclerosis. J Neuroimmunol. 2014; 266:64-6.

76. Wang, Z. *et al.* Nuclear Receptor NR1H3 in Familial Multiple Sclerosis. Neuron. 2016; 90:948-54.

77. Sadovnick, A. D. *et al.* Purinergic receptors P2RX4 and P2RX7 in familial multiple sclerosis. Hum Mutat. 2017;38:736-744.

78. Sadovnick, A. D. *et al.* Analysis of Plasminogen Genetic Variants in Multiple Sclerosis Patients. G3. 2016;6:2073-9.

79. Zrzavy, T. *et al.* A rare P2RX7 variant in a Hungarian family with multiple sclerosis. Mult Scler Relat Disord. 2019;27:340-341.

80. Maver, A. *et al.* Identification of rare genetic variation of NLRP1 gene in familial multiple sclerosis. Sci Rep. 2017;7:3715.

81. Mitrovič, M. *et al.* Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. Cell. 2018;75:1679-1687.

82. Wray, N. R. *et al.* Prediction of individual genetic risk of complex disease. Curr Opin Genet Dev. 2008;18:257-63.

83. De Jager, P. L. *et al.* Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol. 2009;8:1111-9.

84. Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann Neurol. 2011;69:65-74.

85. Hilven, K. *et al.* Burden of risk variants correlates with phenotype of multiple sclerosis. Mult Scler J. 2015;21:1670-80.

86. Harbo, H. F. *et al.* Oligoclonal bands and age at onset correlate with genetic risk score in multiple sclerosis. Mult Scler J. 2014;20:660-8.

87. Pan, G. *et al.* Role of genetic susceptibility variants in predicting clinical course in multiple sclerosis: A cohort study. J Neurol Neurosurg Psychiatry. 2016;87:1204-1211.

88. Banwell, B. L. Multiple sclerosis in children. Handb Clin Neurol. 2014;122:427-41.

89. Banwell, B. *et al.* Clinical, environmental, and genetic determinants of multiple sclerosis in children with acute demyelination: A prospective national cohort study. Lancet Neurol. 2011;10:436-45.

90. Neuteboom, R. F. *et al.* Prognostic factors after a first attack of inflammatory CNS demyelination in children. Neurology. 2008;71:967-73.

91. Boiko, A. N. *et al.* Association and linkage of juvenile MS with HLA-DR2(15) in Russians. Neurology. 2002;58:658-60.

92. Disanto, G. *et al.* HLA-DRB1 confers increased risk of pediatric-onset MS in children with acquired demyelination. Neurology. 2011; 76:781-6.

93. Gianfrancesco, M. A. *et al.* Genetic risk factors for pediatric-onset multiple sclerosis. Mult Scler J. 2018;24:1825-1834.

94. Boström, I. & Landtblom, A. M. Does the changing sex ratio of multiple sclerosis give opportunities for intervention? Acta Neurol Scand. 2015;132:42-5.

95. Nakken, O. *et al.* Sex ratio in multiple sclerosis mortality over 65 years; an age-period-cohort analysis in Norway. J Neurol. 2018;265:1295-1302.

96. Salzer, J. *et al.* Vitamin D as a protective factor in multiple sclerosis. Neurology. 2012;79;21:2140-5.

97. Runia, T. F. *et al.* Lower serum vitamin D levels are associated with a higher relapse risk in multiple sclerosis. Neurology. 2012;79:261-6.

98. Ascherio, A. *et al.* Vitamin D as an early predictor of multiple sclerosis activity and progression. JAMA Neurol. 2014;71:306-14.

99. Martinelli, V. *et al.* Vitamin D levels and risk of multiple sclerosis in patients with clinically isolated syndromes. Mult Scler. 2014;20:147-55.

100. Munger, K. L. *et al.* 25-Hydroxyvitamin D deficiency and risk of MS among women in

the Finnish Maternity Cohort. Neurology. 2017;89:1578-1583.

101. Di Pauli, F. *et al.* Smoking is a risk factor for early conversion to clinically definite multiple sclerosis. Mult Scler. 2008;14:1026-30.

102. Manouchehrinia, A. *et al.* Tobacco smoking and disability progression in multiple sclerosis: United Kingdom cohort study. Brain. 2013;136:2298-304.

103. Ramanujam, R. *et al.* Effect of smoking cessation on multiple sclerosis prognosis. JAMA Neurol. 2015;72:1117-23.

104. Hedström, A. K. *et al.* Exposure to environmental tobacco smoke is associated with increased risk for multiple sclerosis. Mult Scler J. 2011;17:788-93.

105. Hedström, A. K. *et al.* Smoking and multiple sclerosis susceptibility. Eur J Epidemiol. 2013;28:867-74.

106. Handel, A. E. *et al.* Smoking and multiple sclerosis: An updated meta- analysis. PLoS One 2011;6:e16149.

107. Belbasis, L. *et al.* Environmental risk factors and multiple sclerosis: An umbrella review of systematic reviews and meta-analyses. Lancet Neurol. 2015;14:263-73.

108. Ascherio, A. *et al.* Epstein-Barr virus antibodies and risk of multiple sclerosis: A prospective study. J Am Med Assoc. 2001;286:3083-8.

109. Thacker, E. L *et al*. Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis. Ann Neurol. 2006;59:499-503.

110. Lünemann, J. D. & Münz, C. EBV in MS: guilty by association? Trends Immunol. 2009;30:243-8.

111. Munger, K. L. *et al.* Body size and risk of MS in two cohorts of US women. Neurology. 2009;73:1543-50.

112. Munger, K. L. *et al.* Childhood body mass index and multiple sclerosis risk: A long-term cohort study. Mult Scler. 2013;19:1323-9.

113. Hedström, A. K. *et al.* Body mass index during adolescence, rather than childhood, is critical in determining MS risk. Mult Scler. 2016; 22:878-83.

114. Nourbakhsh, B. *et al.* Dietary salt intake and time to relapse in paediatric multiple sclerosis. J Neurol Neurosurg Psychiatry. 2016; 87:1350-1353.

115. Kleinewietfeld, M. *et al.* Sodium chloride drives autoimmune disease by the induction of pathogenic TH 17 cells. Nature. 2013;496:518-22.

116. Cortese, M. *et al.* No association between dietary sodium intake and the risk of multiple sclerosis. Neurology. 2017;89:1322-1329.

117. Farez, M. F. *et al.* Sodium intake is associated with increased disease activity in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2015;86:26-31.

118. Fitzgerald, K. C. *et al.* Sodium intake and multiple sclerosis activity and progression in BENEFIT. Ann Neurol. 2017;82:20-29.

119. Zhu, T. *et al.* Association between alcohol consumption and multiple sclerosis: a meta-analysis of observational studies. Neurol Sci. 2015;36:1543-50.

120. Foster, M. *et al.* Associations of moderate alcohol consumption with clinical and MRI measures in multiple sclerosis. J Neuroimmunol. 2012;243:61-8.

121. Hedström, A. K. *et al.* Alcohol as a modifiable lifestyle factor affecting multiple sclerosis risk. JAMA Neurol. 2014;71:300-5.

122. Paz-Ballesteros, W. C. *et al.* Cigarette Smoking, Alcohol Consumption and Overweight in Multiple Sclerosis: Disability Progression. Arch Med Res. 2017;48:113-120.

123. Li, J. *et al.* The risk of multiple sclerosis in bereaved parents: A nationwide cohort study in Denmark. Neurology. 2004;62:726-9.

124. Buljevac, D. *et al.* Self reported stressful life events and exacerbations in multiple sclerosis: prospective study. BMJ. 2003;327:646.

125. Nielsen, N. M. *et al.* Major stressful life events in adulthood and risk of multiple

sclerosis. J Neurol Neurosurg Psychiatry. 2014;85:1103-8.

126. Artemiadis, A. K. *et al.* Stress as a risk factor for multiple sclerosis onset or relapse: A systematic review. Neuroepidemiology. 2011;36:109-20.

127. Lovera, J. & Reza, T. Stress in multiple sclerosis: Review of new developments and future directions. Curr Neurol Neurosci Rep. 2013;13:398.

128. Décard, B. F. *et al.* Low vitamin D and elevated immunoreactivity against Epstein-Barr virus before first clinical manifestation of multiple sclerosis. J Neurol Neurosurg Psychiatry. 2012;83:1170-3.

129. Levin, L. I. *et al.* Temporal relationship between elevation of Epstein-Barr virus antibody titers and initial onset of neurological symptoms in multiple sclerosis. J Am Med Assoc. 2005;293:2496-500.

130. Crawford, D. H. *et al.* A cohort study among university students: identification of risk factors for Epstein-Barr virus seroconversion and infectious mononucleosis. Clin Infect Dis. 2006;43:276-82.

131. Ascherio, A. & Munger, K. L. Environmental risk factors for multiple sclerosis. Part I: The role of infection. Ann Neurol. 2007; 61:288-99.

132. Rubicz, R. *et al.* Genetic factors influence serological measures of common infections. Hum Hered. 2011;72:133-41.

133. Rubicz, R. *et al.* A Genome-wide integrative genomic study localizes genetic factors influencing antibodies against Epstein-Barr virus nuclear antigen 1 (EBNA-1). PLoS Genet. 2013;9:e1003147.

134. Sallah, N. *et al.* Whole-genome association study of antibody response to Epstein-Barr virus in an African population: a pilot. Glob Heal Epidemiol Genom. 2017;2:e18.

135. De Jager, P. L. *et al.* Integrating risk factors: HLA-DRB1*1501 and Epstein-Barr virus in multiple sclerosis. Neurology. 2008;70:1113-8. 136. Kreft, K. L. *et al.* Elevated EBNA-1 IgG in MS is associated with genetic MS risk variants. Neurol Neuroimmunol NeuroInflamm. 2017;4:e406.

137. Zhou, Y. *et al.* Genetic loci for Epstein-Barr virus nuclear antigen-1 are associated with risk of multiple sclerosis. Mult Scler. 2016;22:1655-1664.

138. Fagnani, C. *et al.* Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. Mult Scler. 2015;21:1404-13.

139. Bergkvist, M. & Sandberg-Wollheim, M. Serological differences in monozygotic twin pairs discordant for multiple sclerosis. Acta Neurol Scand. 2001;104:262-5.

140. Comabella, M. *et al.* Antiviral immune response in patients with multiple sclerosis and healthy siblings. Mult Scler. 2010;16:355-8.

141. Kinnunen, E. *et al.* Viral antibodies in Multiple Sclerosis: A Nationwide co-twin study. Arch Neurol. 1990;47:743-6.

142. Sumaya, C. V. *et al.* Increased prevalence and titer of Epstein-Barr virus antibodies in patients with multiple sclerosis. Ann Neurol. 1985;17:371-7.

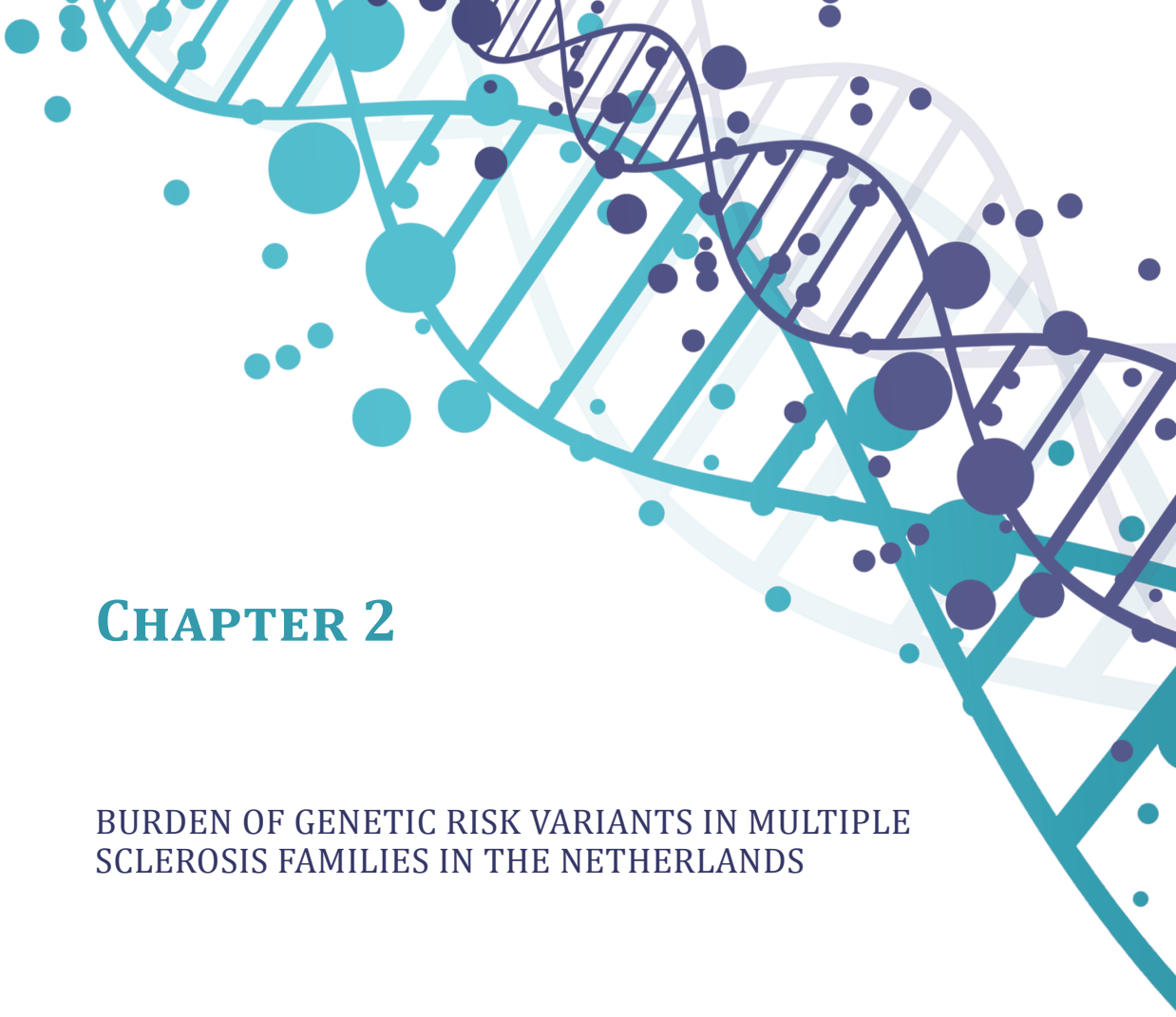
143. Cotsapas, C. *et al.* Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 2011;7:e1002254.

144. Sirota, M. *et al.* Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet. 2009;5:e1000792.

145. Dobson, R. & Giovannoni, G. Autoimmune disease in people with multiple sclerosis and their relatives: A systematic review and meta-analysis. J. Neurol. 2013;260:1272-85.

PART II

COMMON GENETIC RISK FACTORS IN MS



Chapter 2

BURDEN OF GENETIC RISK VARIANTS IN MULTIPLE SCLEROSIS FAMILIES IN THE NETHERLANDS

J.Y. Mescheriakova, L. Broer, S. Wahedi, A.G. Uitterlinden, C.M. van Duijn and R.Q. Hintzen.

Multiple Sclerosis Journal Experimental, Translational and Clinical, 2016

ABSTRACT

Background

Approximately 20% of Multiple Sclerosis patients have a family history of MS. Studies of MS aggregation in families are inconclusive.

Objective

To investigate the genetic burden based on currently discovered genetic variants for MS risk in patients from Dutch MS multiplex families versus sporadic MS cases, and to study its influence on clinical phenotype and disease prediction.

Methods

Our study population consisted of 283 sporadic multiple sclerosis cases, 169 probands from multiplex families and 2028 controls. A weighted genetic risk score based on 102 non-HLA loci and *HLA-DRB1*1501* was calculated.

Results

The weighted genetic risk score based on all loci was significantly higher in familial than in sporadic cases. The *HLA-DRB1*1501* contributed significantly to the difference in genetic burden between the groups. A high weighted genetic risk score was significantly associated with a low age of disease onset in all multiple sclerosis patients, but not in the familial cases separately. The genetic risk score was significantly but modestly better in discriminating familial versus sporadic multiple sclerosis cases from controls.

Conclusion

Familial multiple sclerosis patients are more loaded with the common genetic variants than sporadic cases. The difference is mainly driven by the *HLA-DRB1*1501*. The predictive capacity of genetic loci is poor and unlikely to be useful in clinical settings.

INTRODUCTION

Multiple sclerosis (MS) is an inflammatory and neurodegenerative disease of the central nervous system. The etiology of MS is complex with both genetic and environmental factors contributing to its etiology. [1] Evidence of genetic contribution is found in family studies where approximately 20% of MS patients have an affected family member. [2] In addition, the recurrence risk in siblings of MS patients is higher than in the general population. [3] The main genetic locus of MS risk is the human leukocyte antigen (HLA) class II region (the classical *HLA-DRB1*1501* allele). This major genetic locus has been studied and replicated extensively. [4, 5] In the most recent collaborative genome-wide association study (GWAS) 110 non-HLA MS risk loci were established [6] and the list is still expanding.

Despite the success of the GWAS, a large part of MS heritability remains unexplained. Understanding the familial aggregation of MS will help to unravel the mechanisms of the MS development. Because environmental factors probably exert their effect on population level, although this has not been well determined for MS families, it seems more likely that these families are more loaded with the genetic variants. This could result in a higher disease susceptibility in families. Several studies investigated the genetic burden in families, but the results are inconclusive. [7-12]

A robust way to measure the genetic loading is by a combination of the genetic variants into a cumulative genetic risk score. [13, 14] In this study we investigated the accumulated genetic risk in MS multiplex families and compared it to the sporadic cases in the Netherlands. We also explored the influence of genetic burden on clinical characteristics and assessed the prediction of the disease status in MS.

MATERIALS AND METHODS

Subjects

All patients suspected of MS who visited the neurological outpatient clinic of the Erasmus Medical Centre (EMC) between 2004 and 2009 were asked to participate in the study. During the visit to the hospital every patient was interviewed about the family history of MS. MS patients with positive family history of MS were defined as multiplex MS cases, whereas patients with no family history of MS were defined as sporadic MS

cases. A large proportion of multiplex cases is derived from the still ongoing study on gene-environment interaction in MS (GEMS) in the Netherlands. In this study multiplex MS families are included in which at least one first-degree or second-degree relative of an affected proband also has clinically definite MS. The participants of the GEMS study fill in the questionnaire about their family structure and their family history of MS. The diagnosis of MS in all patients was evaluated according to the standard diagnostic criteria. [15, 16] The following clinical information was collected: gender, age at disease onset, disease duration, clinical course, Expanded Disability Status Scale score (EDSS) and MS Severity Score (MSSS). A control group consisted of unrelated healthy adults from the general population enrolled in the longitudinal Rotterdam Study III (RSIII). [17] Written informed consent was obtained from all participants with approval from the medical ethics committee of the EMC.

Quality control and genetic risk score

Whole blood samples were collected from all participants and DNA was isolated by a standardised method. [18] Samples were genotyped on the Illumina 610-Quad Bead array (n=2466) and the ImmunoChip (n=302). Both arrays were subjected to the standard quality control (QC). [19] Only subjects of European descent were included in the analysis. Study participants with unexpected relatedness (PI_HAT > 0.35) were excluded. When an individual's genotype for a SNP used for the calculation of the genetic risk score (see further) was missing, the sample was excluded from further analysis (see Appendix 1 for QC-steps). After QC, 569 MS patients and 2028 controls were eligible for the analysis. Of all cases, 286 patients reported a positive family history of MS and 283 MS patients were identified as sporadic cases. For all comparative analyses we used 169 probands from the multiplex families.

After QC, 102 out of 110 MS risk single nucleotide polymorphisms (SNPs) [6] were available for extraction from the arrays. When a SNP was not present on the Illumina 610-Quad Bead array, we used tagging SNPs with $r^2 > 0.6$ (see Appendix 2). The tagging SNP rs9271366 was used for the *HLA-DRB1*1501* locus ($r^2 = 0.957$). [20, 21] All tested SNPs were in Hardy-Weinberg Equilibrium.

The weighted genetic risk score (wGRS) was calculated as previously described [13] implementing 102 available SNPs. The analyses were also conducted using only the original SNPs. Because the results (data not shown) were in the same direction, we used all available SNPs in the calculation of the weighted genetic risk score. To assess the additional effect of the *HLA-DRB1*1501* risk allele, we optionally included it into the

model. To determine how well the $wGRS_{HLA}$, $wGRS_{102}$ or $wGRS_{102+HLA}$ models discriminate between the studied groups receiver operating characteristic (ROC) curves were constructed and the area under the curve (AUC) was calculated.

Statistical analysis

Quality control was completed using PLINK. SPSS statistical software (IBM Company, version 21) and GraphPad software v5 were used to analyse clinical variables (gender, age of onset, MS course, EDSS and MSSS) using chi-square test, non-parametric Mann-Whitney test and Spearman correlation. The wGRS scores and ROC curves were computed in SPSS and in R by usage of PredictABEL package.[22] Differences in predicting performance of ROC curves were tested with De Long test using the R package "pROC".[23] We applied correction for multiple testing using Benjamini-Hochberg procedure for controlling the false-positive rate.

RESULTS

The case and control characteristics are presented in Table 1. There were no significant differences between probands from MS multiplex families (fMS) and sporadic MS cases (sMS) in terms of age of disease onset, gender ratio, disease course, EDSS or MSSS (all $p > 0.05$).

Table 1. Demographic and clinical characteristics of patients with sporadic MS, multiplex MS and healthy controls.

	Controls (n=2028)	Sporadic MS (n=283)	Multiplex MS (all) (n=286)	Multiplex MS (probands) (n=169)	P-value[c]
Female:male ratio (n:n)	1.27:1 (1136:892)	2.88:1 (210:73)	2.18:1 (196:90)	2.60:1 (122:47)	NS
Disease course (*n*)	N/A	RR 181 SP 53 PP 49	RR 193 SP 41 PP 52	RR 117 SP 25 PP27	NS
Relapsing-Remitting onset (%)		82.7	81.8	84	NS
Disease duration (yrs)[a]	N/A	12.94 ±8.80	15.50 ±10.59	14.59 ± 9.77	NS
Age at onset (yrs)[a]	N/A	34.59 ±10.43	32.82 ± 9.76	33.15 ± 9.21	NS
EDSS[b]	N/A	3.0	3.5	3.5	NS
MSSS[a]	N/A	4.75 ± 3.01	4.93 ± 2.86	4.81 ± 2.65	NS

EDSS: Expanded Disability Status Scale; MSSS: Multiple Sclerosis Severity Score; NS: not significant; N/A: not applicable; RR/SP/PP: relapsing-remitting, secondary progressive, primary progressive MS; SD: standard deviation.

[a] *For disease duration, age of onset and MSSS the mean and ±SD are indicated.*

[b] *For EDSS the median is reported.*

[c] *For comparisons of sporadic MS with multiplex MS probands.*

The $wGRS_{102}$ and $wGRS_{102+HLA}$ were significantly higher in sporadic and multiplex MS patients than in controls ($p<0.0001$; Figure 1 and Table 2). We observed a trend towards a higher $wGRS_{102}$ in fMS than in sMS ($p=0.08$). The addition of HLA into the $wGRS_{102}$ score resulted in a significant higher score in fMS than in sMS ($p< 0.0001$). The risk allele frequency of the *HLA-DRB1*1501* was significantly higher in fMS (0.35) than in sMS (0.25, $p=0.001$) and healthy controls (0.13, $p< 0.0001$). There was no compensatory aggregation of non-HLA SNPs in affected individuals not carrying the *HLA-DRB1*1501* risk allele ($p=0.31$). We found no differences in genetic risk scores between males and females in multiplex MS and sporadic MS patients.

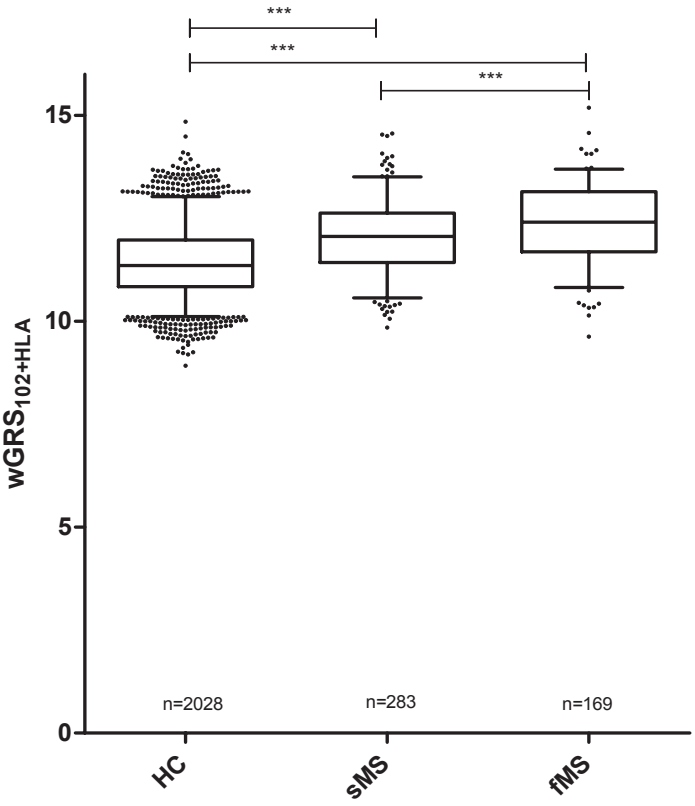


Figure 1. Distribution of genetic risk score in MS patients and healthy controls.

*The figure shows a distribution of the weighted genetic risk score based on 102 non-HLA loci and the HLA-DRB1*1501 allele ($wGRS_{102+HLA}$) in healthy controls (HC), sporadic MS (sMS) and multiplex MS (fMS). The significance is indicated in the figures as *** p<0.0001.*

Table 2. Mean weighted genetic risk score in healthy controls, sporadic and multiplex MS patients.

	Controls n=2028	Sporadic MS n=283	Multiplex MS n=169	p-value Controls vs sporadic MS	p-value Controls vs multiplex MS	p-value sporadic MS vs multiplex MS
$wGRS_{102}$	11.14 ± 0.67	11.48 ± 0.70	11.60 ± 0.65	< 0.0001	< 0.0001	0.08
$wGRS_{102+HLA}$	11.44 ± 0.86	12.04 ± 0.88	12.39 ± 0.96	< 0.0001	< 0.0001	< 0.0001

wGRS: weighted genetic risk score. Means and standard deviations are indicated for wGRS.

There was a trend towards a correlation between a high $wGRS_{102}$ and a low age at disease onset (AAO) in all patients with MS (p=0.09, Figure 2(a)). Inclusion of HLA into the genetic score showed a significant correlation between a high genetic score and a low AAO in this group. After stratification, we found that a low age of onset correlated significantly with a high $wGRS_{102}$ only in sporadic MS (p=0.02) but not in familial cases (p=0.61; Figure 2(b) and 2(c)). The same significant correlation was found for $wGRS_{102+HLA}$ and age of onset (p=0.02) in sporadic MS but not in multiplex MS cases (p=0.72).

There was no correlation between genetic risk scores and MSSS (p> 0.05). No differences were found in $wGRS_{102}$ and $wGRS_{102+HLA}$ between primary progressive and remitting-relapsing MS patients (p> 0.05).

Both the $wGRS_{HLA}$ (AUC=0.63 [0.61-0.66]), and the $wGRS_{102}$ (AUC = 0.66 [0.63-0.68]) had a poor predictive capacity in discriminating between all patients and controls (Figure 3(a) and Table 3). The AUC increased significantly to 0.72 (p< 0.001) when both non-HLA SNPs and the *HLA-DRB1*1501* were included into the model. The predictive capacity of the genetic models was also tested in sporadic and multiplex MS patients separately (Figure 3(b) and 3(c)). The $wGRS_{102+HLA}$ had a statistically better predictive performance than the model considering only *HLA-DRB1*1501* or only 102 non-HLA SNPs in both sporadic and familial MS versus controls (p< 0.001, Table 3). In addition, the $wGRS_{102+HLA}$ was better in discriminating familial MS from controls (AUC=0.77 [0.73-0.81]) than the same model discriminating sporadic MS from controls (AUC=0.69 [0.66-0.72], p=0.0033). These results parallel the Nagelkerke's R^2 results where the proportion of genetic variability explained by different wGRS scores is represented. The maximum R^2 was 0.167 in fMS and 0.0892 in sMS for $wGRS_{102+HLA}$(see Appendix 3).

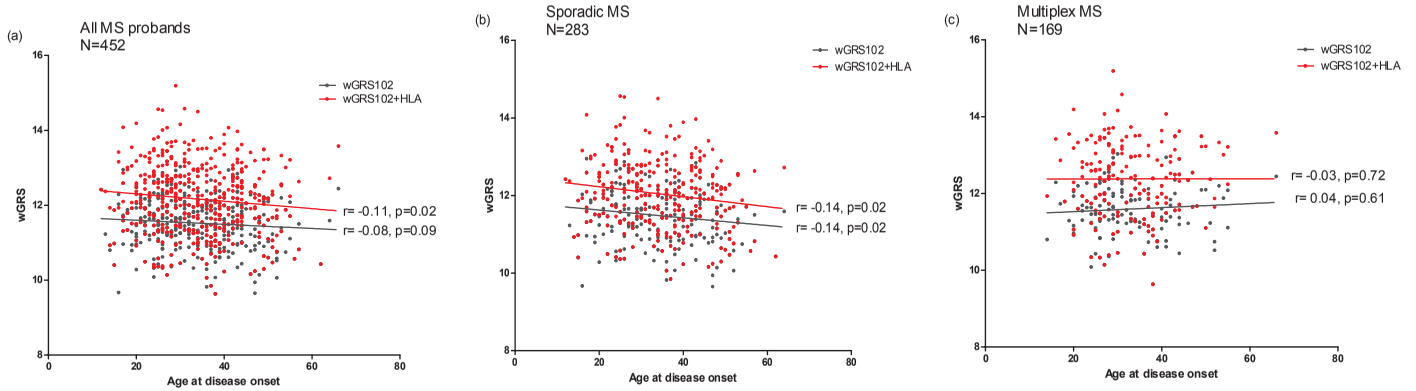


Figure 2. Correlation between age of disease onset and wGRS.

The figures show the correlation between weighted genetic risk scores (wGRS) and age at disease onset in (a) all MS probands, (b) sporadic MS and (c) multiplex MS. The linear regression lines are drawn in black for $wGRS_{102}$ and red for $wGRS_{102+HLA}$. The r estimates from the Spearman correlation test and p-values are shown next to the regression lines.

We also calculated ROC for different wGRS– models to be able to discriminate between familial and sporadic MS (Figure 3(d)), but the AUC values were disappointingly low and not suitable to be used in clinical practice (Table 3).

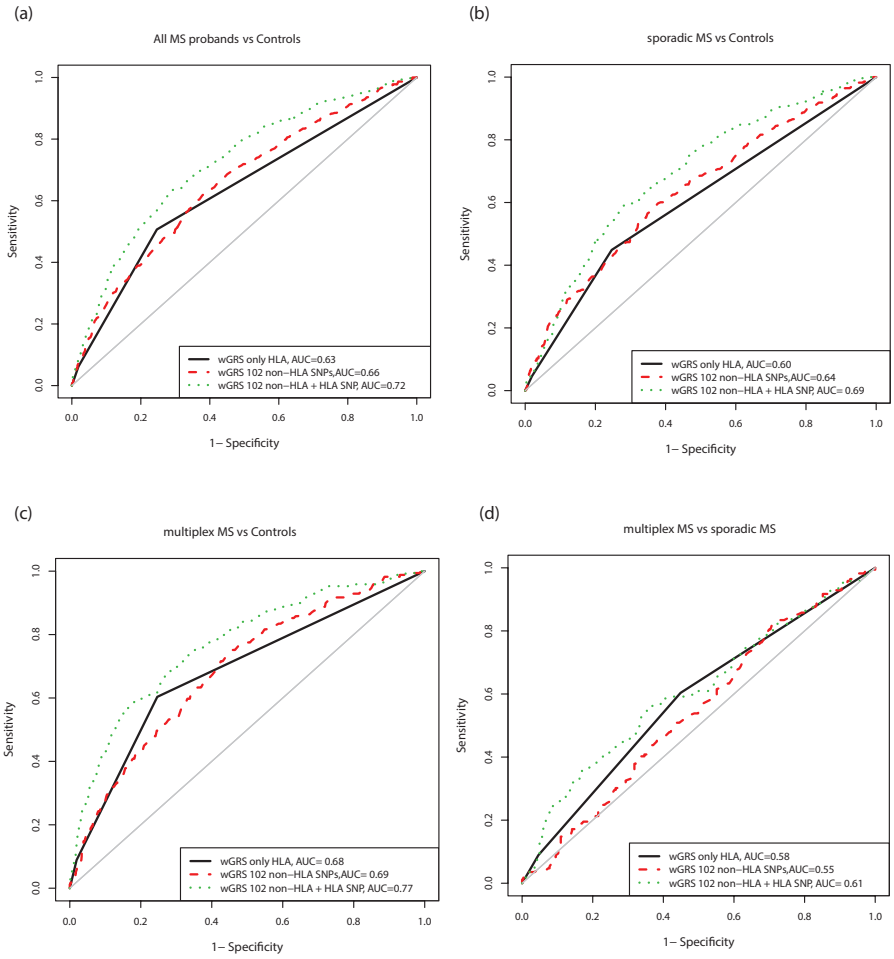


Figure 3. Predictive performance of wGRS.

*The receiver operating characteristics curves are plotted considering 3 models: wGRS based on only HLA-DRB1*1501 (solid black line), 102 non-HLA loci (red dashed line) and 102 non-HLA loci + HLA-DRB1*1501 allele (green dotted line). The area under the curve (AUC) was calculated to assess the predictive capacity in (a) all MS probands, (b) sporadic MS and (c) multiplex MS compared to controls. Figure (d) shows the prediction of multiplex MS patients from sporadic cases.*

Table 3. AUC values for wGRS risk models in sporadic and multiplex MS patients.

wGRS based on	All MS probands vs Controls		Sporadic MS vs Controls		Multiplex MS vs Controls		Sporadic MS vs multiplex MS	
	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI
HLA-DRB1*1501	0.63	0.61-0.66	0.60	0.57-0.63	0.68	0.65-0.72	0.58	0.55-0.63
102 non-HLA SNPs	0.66	0.63-0.68	0.64	0.60-0.67	0.69	0.65-0.73	0.55	0.49-0.60
102 non-HLA SNPs *+HLA-DRB1*1501*	0.72	0.69-0.75	0.69	0.66-0.72	0.77	0.73-0.81	0.61	0.55-0.66

wGRS: weighted genetic risk score; AUC: area under the receiver operating curve; CI: confidence interval; HLA: human leukocyte antigen; SNP: single nucleotide polymorphisms.

DISCUSSION

In this study we showed that the cumulative genetic risk score based on all currently known genetic risk factors including the *HLA-DRB1*1501*, was significantly higher in MS multiplex patients than in sporadic cases and healthy controls in Dutch population. These results are in agreement with previous reports. [7-9, 12] Two studies found no difference in genetic load between multiplex and sporadic MS. [10, 11] These studies were less powered than other cohorts because of a lower number of multiplex cases. Moreover, population related differences in genetic background could result in different allele frequencies and thus a distinct genetic load in these studies.

The difference in genetic burden between Dutch multiplex and sporadic MS patients was primarily driven by a higher *HLA-DRB1*1501* allele frequency in familial cases. A strong association between MS risk and *HLA-DRB1*1501* has been shown previously in some MS families. [24, 25] Gourraud et.al [8] also found a higher genetic risk score in multiplex families based solely on non-HLA variants. Our study showed a similar trend. We have conducted a post-hoc power analysis which showed that our study was underpowered (power of 0.43) to detect a significant difference between sporadic and multiplex MS for wGRS based on only non-HLA SNPs. The study numbers for both groups would need to increase to the study numbers used by Gourraud et al. [8] to have the power of at least 0.8 at the α-level of 0.05.

Because in our cohort there was a large part of patients without the *HLA-DRB1*1501* risk allele, it was interesting to look at the aggregation of non-HLA SNPs in these patients.

We did not observe a compensatory increase in non-HLA SNPs in the HLA non-carriers, as earlier reported.[8]

The factors contributing to the difference between men and women in MS susceptibility are unknown. Our study showed that the genetic risk score is not different between men and women with MS, and is thus probably not responsible for the gender differences in the disease susceptibility. Our results are in line with previous studies. [10, 14]

The accumulation of a higher number of susceptibility alleles, such as the case in multiplex MS, might predispose to an earlier disease onset. However, we did not find a younger age at onset in familial MS than in sporadic MS, supported by several other studies. [8, 26, 27] Furthermore, the genetic risk score was not associated with the age at onset in Dutch multiplex cases. According to a post-hoc power analysis the participant number in the multiplex MS group was too low to detect a statistically significant result. When all sporadic and multiplex cases were taken together a significant inverse correlation was found between genetic risk score and age at onset. This finding is consistent with some [10, 11] but not all studies. [8, 14] The result of our study probably reflects the known association of the *HLA-DRB1*1501* with a younger age at onset. [28] We did not find any association between genetic risk score and other clinical parameters such as MSSS and the disease course, as previously reported. [8, 11, 14]

Finally, the genetic risk model based on all risk SNPs was modest in discriminating MS patients from controls and this result was comparable with previous findings. [13, 29] Our study showed that the model based on all risk loci was better in discriminating multiplex MS than sporadic MS from healthy controls. In the treatable and preventable disease like coronary heart disease an AUC of about 0.77 is used (Framingham Risk Score). [30] Although, the AUC for MS found in our study is approaching this value, the value with a higher specificity is required for a relatively rare condition such as MS. In order to obtain higher AUCs a considerable number of additional common variants or stronger associated variants with higher odds ratios are needed. [31] Addition of environmental factors into the model could also be beneficial in the disease prediction. [13]

We have conducted correction for multiple testing using Benjamini-Hochberg procedure for controlling false-positive rate, and can confirm that all results with significant p-values remained significant after application of this correction.

There were a few limitations to our study. First, our current findings are based on a simplistic modeling of only one tagging SNP of the major HLA–class II locus. As well as the *HLA-DRB1*1501*, other alleles are also associated with MS risk.[5] Although these alleles are independently associated with MS, their contribution to MS risk is smaller than that of the *HLA-DRB1*1501*. Second, the sample size of our study was moderate and lacked power for some comparisons, but despite this the results were in the same direction as in the largest study. [9] Third, the effect sizes used for the calculation of the genetic risk score originate from a GWAS done in general MS population. Effect sizes of the SNPs might be different for specifically multiplex cases. However, it cannot be ruled out that some of the GWAS participants have a family history of MS and already account for the effect size of the found associations in GWAS. In our own data (data not shown) we observed slightly higher odds ratios (ORs) for the majority of the original SNPs in fMS vs controls than in sMS vs controls. Because of the low numbers in both case-groups we could not prove whether the differences in ORs were significant. D'Netto et al.[7] also suggest that the effect sizes of genetic variants in multiplex families are probably increased compared to sporadic cases. Unfortunately, also this study was underpowered to be able to make reliable conclusions about the mentioned differences. Fourth, the accuracy of the wGRS calculation might be influenced by the use of the proxy SNPs which might blur the differences between the groups. In our study the impact of the use of proxy SNPs is limited because of two reasons: 1) even with the use of the proxy SNPs we were able to observe differences between cases and controls; 2) we also conducted calculations of wGRS based on only original SNPs and observed that the effect was in the same direction as with the wGRS based on original and proxy SNPs.

In conclusion, our results indicate that patients from multiplex MS families are more genetically loaded with common variants, especially with the *HLA-DRB1*1501*, than the sporadic MS patients. A greater genetic burden leads to an increased risk of MS in families. Because of the elevated frequencies of the genetic variants, patients from multiplex families are valuable and may increase the power of case-control studies. For further understanding of familial MS aggregation the research should also focus on rare variants and environmental influences on MS risk in families. Combining this information with the genetic risk score might help us to explain familial MS aggregation and also to achieve a better disease prediction.

REFERENCES

1. Compston, A. & Coles, A. Multiple sclerosis. Lancet. 2008;372:1502-17.
2. Sadovnick, A. D. *et al.* Multiple sclerosis: updated risks for relatives. Am J Med Genet. 1988;29:533-41.
3. O'Gorman, C. *et al*. Modelling genetic susceptibility to multiple sclerosis with family data. Neuroepidem. 2013;40:1-12.
4. Jersild, C. *et al.* Histocompatibility determinants in multiple sclerosis, with special reference to clinical course. Lancet. 1973;2:1221-5.
5. Patsopoulos, N. A. *et al*. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013;9:e1003926.
6. International Multiple Sclerosis Genetics Consortium. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353-60.
7. D'Netto, M. J. *et al.* Risk alleles for multiple sclerosis in multiplex families. Neurology. 2009;72:1984-8.
8. Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann Neurol. 2011;69:65-74.
9. Isobe, N. *et al.* Genetic burden in multiple sclerosis families. Genes Immun. 2013;14:434-40.
10. Harbo, H. F. *et al.* Oligoclonal bands and age at onset correlate with genetic risk score in multiple sclerosis. Mult Scler. 2014;20:660-8.
11. Esposito, F. *et al.* Impact of MS genetic loci on familial aggregation, clinical phenotype, and disease prediction. Neurol Neuroimmunol Neuroinflamm. 2015;2:e129.
12. Xia, Z. *et al.* Genes and environment in Multiple Sclerosis project: A platform to investigate multiple sclerosis risk. Ann Neurol. 2016;79:178-89.
13. De Jager, P. L. *et al.* Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol. 2009;8:1111-9.
14. Hilven, K. *et al.* Burden of risk variants correlates with phenotype of multiple sclerosis. Mult Scler. 2015;21:1670-80.
15. Poser, C. M. *et al.* New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol. 1983;13:227-31.
16. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50:121-7.
17. Hofman, A. *et al.* The Rotterdam Study: 2012 objectives and design update. Eur J Epidemiol. 2011;26:657-86.
18. Miller, S. A. *et al.* A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.
19. Anderson, C. A. *et al.* Data quality control in genetic case-control association studies. Nat Protoc. 2010;5:1564-73.
20. de Bakker, P. I. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet. 2006;38:1166-72.
21. Field, J. *et al.* A polymorphism in the HLA-DPB1 gene is associated with susceptibility to multiple sclerosis. PLoS One. 2010;5:e13454.
22. Kundu, S. *et al.* PredictABEL: an R package for the assessment of risk prediction models. Eur J Epidemiol. 2011;26:261-4.
23. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
24. Sawcer, S. *et al.* A high-density screen for linkage in multiple sclerosis. Am J Hum Genet. 2005;77:454-67.

25. Barcellos, L. F. *et al.* Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet. 2006;15:2813-24.

26. Ebers, G. C. *et al.* The natural history of multiple sclerosis: a geographically based study: 8: familial multiple sclerosis. Brain. 2000;3:641-9.

27. Weinshenker, B. G. *et al*. A comparison of sporadic and familial multiple sclerosis. Neurology. 1990;40:1354-8.

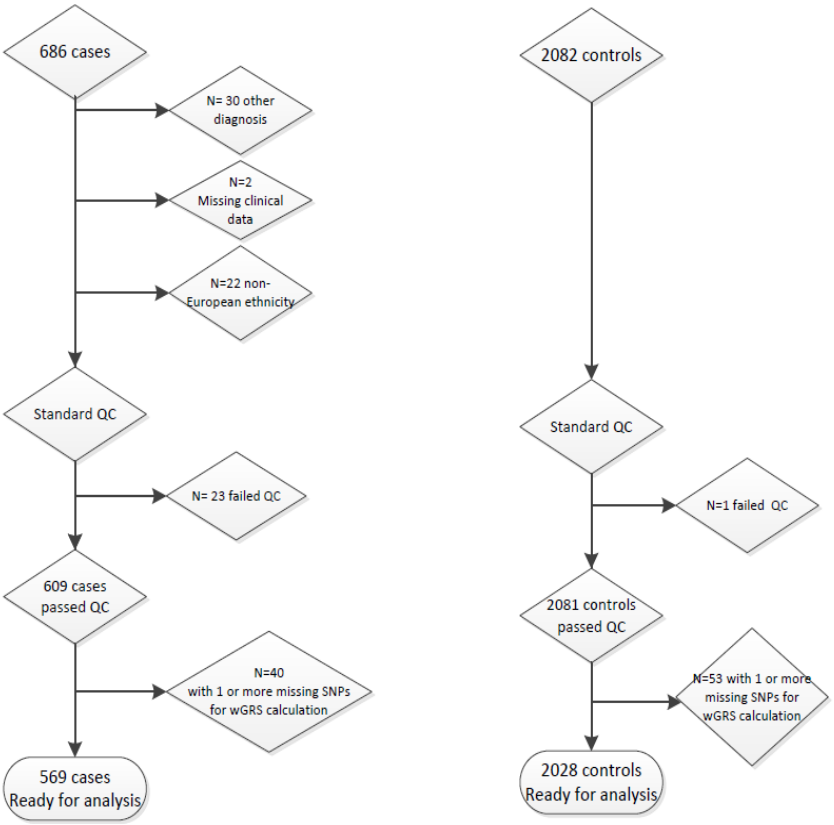
28. International Multiple Sclerosis Genetics Consortium. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476:214-9.

29. Disanto, G. *et al.* The refinement of genetic predictors of multiple sclerosis. PLoS One. 2014;9:e96578.

30. Wilson, P. W. *et al.* Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97:1837-47.

31. Jafari, N. *et al*. Perspectives on the use of multiple sclerosis risk genes for prediction. PLoS One. 2011;6:e26493.

SUPPLEMENTARY MATERIAL



Appendix 1. QC-steps during the analysis of the genetic data.

Appendix 2. List of single nucleotide polymorphisms and their weights used for the calculation of the genetic risk score.

Chr	Gene	Original SNP	RA	OR	Best Proxy[b]	RA	r^2
1	*MMEL1*	rs3748817	A	1.14	rs10752747	C	0.962
1	*PLEKHG5*	rs3007421	A	1.12		A	
1	*BCL10 (dist=4406, DDAH1 (dist=37175)*	rs12087340	A	1.22	rs4074156	G	0.736
1	*DDAH1*	rs11587876	A	1.12		A	
1	*EVI5*	rs41286801	A	1.20	rs6662618	A	0.752
1	*VCAM1 (dist=36292), EXTL2 (dist=97035)*	rs7552544	A	1.08		A	
1	*VCAM1*	rs11581062	G	1.05		G	
1	*CD58*	rs6677309	A	1.34	rs758518	A	1
1	*PHGDH*	rs666930	G	1.09		G	
1	*FCRL1*	rs2050568	G	1.08		G	
1	*SLAMF7*	rs35967351	A	1.09	rs518721	G	0.861
1	*RGS21 (dist=205058), RGS1 (dist=3385)*	rs1359062	C	1.18	rs1323292	A	0.945
1	*C1orf106*	rs55838263	A	1.12	rs7522462	G	1
2	*CENPO*	rs4665719	G	1.09	rs2891409	C	1
2	*HAAO (dist=341505), ZFP36L2 (dist=88285)*	rs2163226	A	1.10	rs12466022	C	0.771
2	*FLJ16341*	rs842639	A	1.11		A	
2	*CNRIP1 (dist=40294), PLEK (dist=4845)*	rs7595717	A	1.10		A	
2	*MERTK*	rs17174870	G	1.03		G	
2	*STAT4*	rs9967792	G	1.11	rs6738544	C	0.964
2	*SP140*	rs9989735	C	1.17	rs9989899	A	0.948
3	*SATB1 (dist=305320), KCNH8 (dist=404432)*	rs11719975	C	1.09	rs9284846	A	1
3	*EOMES*	rs2371108	A	1.08	rs11129295	A	0.801
3	*EOMES (dist=314786, CMC1 (dist=204553)*	rs1813375	A	1.15	rs427221	G	1
3	*CCR4 (dist=17080),GLB1 (dist=24617)*	rs4679081	G	1.08	rs6550177	A	0.652
3	*FOXP1*	rs9828629	G	1.08		G	
3	*CBLB*	rs2028597	G	1.04		G	
3	*TIMMDC1*	rs1131265	C	1.19	rs2293370	G	1
3	*IQCB1*	rs1920296	C	1.14		C	
3	*ILDR1 (dist=29412), CD86 (dist=3670)*	rs2255214	C	1.11		C	
3	*CD86*	rs9282641	G	1.12		G	
3	*IQCJ-SCHIP1 (dist=75957), IL12A (dist=15511)*	rs1014486	G	1.11		G	
4	*NFKB1 (dist=13144), MANBA (dist=1040)*	rs7665090	G	1.08		G	
4	*TET2*	rs2726518	C	1.09	rs7678440	G	0.967
5	*IL7R (dist=2233), CAPSL (dist=25242)*	rs6881706	C	1.12	rs11742240	C	1

5	*DAB2 (dist=973761), PTGER4 (dist=280936)*	rs6880778	G	1.10	rs1373692	C	0.835
5	*ANKRD55*	rs71624119	G	1.12	rs10065637	G	0.904
5	*VDAC1 (dist=105751), TCF7 (dist=3827)*	rs756699	A	1.12		A	
5	*LOC285626*	rs2546890	A	1.06		A	
5	*RGS14*	rs4976646	G	1.13	rs4075958	A	0.778
6	*CD83 (dist=582348), JARID2 (dist=526710)*	rs17119	A	1.11	rs1267499	G	0.917
6	*PXT1*	rs941816	G	1.13		G	
6	*BACH2*	rs72928038	A	1.11	rs10944479	A	0.932
6	*THEMIS (dist=39022), PTPRK (dist=11126)*	rs802734	A	1.03		A	
6	*AHI1*	rs11154801	A	1.11		A	
6	*IL20RA (dist=86610), IL22RA2 (dist=12049)*	rs17066096	G	1.14		G	
6	*OLIG3 (dist=147124), TNFAIP3 (dist=225670)*	rs7769192d	G	1.08	rs600469	A	0.902
6	*TNFAIP3 (dist=40367), PERP (dist=164826)*	rs67297943	A	1.12	rs7746779	A	0.609
6	*TAGAP (dist=4375), FNDC1 (dist=119870)*	rs212405	T	1.15	rs212402	A	0.78
7	*CARD11 (dist=29525), SDK1 (dist=228046)*	rs1843938	A	1.08		A	
7	*SKAP2 (dist=110647), HOXA1 (dist=117626)*	rs706015	C	1.14	rs774250	A	0.929
7	*JAZF1*	rs917116	C	1.12	rs917117	A	0.838
7	*ELMO1*	rs60600003	C	1.16	rs11984075	G	0.867
7	*ZNF767*	rs354033	G	1.03		G	
8	*PKIA (dist=58302), ZC2HC1A (dist=2478)*	rs1021156	A	1.12	rs1384804	C	1
8	*PCAT1 (dist=159722), POU5F1B (dist=234876)*	rs2456449	G	1.10		G	
8	*MIR1204(dist=6755), PVT1(dist=87845)*	rs4410871	G	1.12		G	
8	*PVT1 (dist=45446), MIR1208 (dist=3417)*	rs759648	C	1.09		C	
10	*IL2RA*	rs2104286	A	1.21		A	
10	*ZNF438 (dist=94240), ZEB1-AS1 (dist=190351)*	rs793108	A	1.09		A	
10	*CAMK2G (dist=24000), C10orf55 (dist=11378)*	rs2688608	A	1.07		A	
10	*ZMIZ1*	rs1782645	A	1.09	rs1250552	G	0.687
10	*HHEX (dist=26509), EXOC6 (dist=112553)*	rs7923837	G	1.11		G	
11	*AGBL2*	rs7120737	G	1.13		G	
11	*CD6 (dist=5482), CD5 (dist=76600)*	rs34383631	A	1.11	rs2905517	G	0.639
11	*PRDX5 (dist=7938), CCDC88B (dist=10457)*	rs694739	A	1.08		A	
11	*TREH (dist=16365), DDX6 (dist=51727)*	rs533646	G	1.10	rs519982	G	1
11	*CXCR5*	rs523604	A	1.09		A	
12	*TNFRSF1A*	rs1800693	G	1.14		G	
12	*LTBR (dist=2768), CD27-AS1 (dist=44667)*	rs12296430	C	1.14	rs2364482	C	1
12	*CD69*	rs11052877	G	1.10		G	
12	*PITPNM2*	rs7132277	A	1.10	rs655293	G	0.947
13	*MIR548AN (dist=27705), TM9SF2 (dist=67469)*	rs4772201	A	1.12		A	

14	*ZFP36L1*	rs2236262	A	1.08	rs4899263	G	0.748
14	*JDP2 (dist=22107), BATF (dist=27273)*	rs4903324	A	1.10		A	
14	*GALC*	rs74796499	C	1.31	rs7157457	G	0.744
14	*TRAF3*	rs12148050	A	1.08		A	
15	*MORF4L1 (dist=17385), CTSH (dist=6626)*	rs59772922	A	1.11	rs9806693	G	1
15	*IQGAP1*	rs8042861	A	1.08		A	
16	*SOX8 (dist=36573), SSTR5-AS1 (dist=40530)*	rs2744148	G	1.09		G	
16	*CLEC16A*	rs12927355	G	1.21	rs2041670	G	1
16	*CLEC16A (dist=12760), SOCS1 (dist=59468)*	rs4780346	A	1.09	rs7203535	A	1
16	*PRM1 (dist=60798), RMI2 (dist=3321)*	rs6498184	G	1.15	rs12927773	C	1
16	*MAPK3 (dist=22333), CORO1A (dist=37768)*	rs7204270	G	1.09		G	
16	*CDH3*	rs1886700	A	1.11		A	
16	*WWOX*	rs12149527	A	1.08	rs7199945	A	0.901
17	*GRB7 (dist=8839), IKZF3 (dist=1591)*	rs12946510	A	1.08	rs907092	A	0.905
17	*STAT3*	rs4796791	A	1.10	rs9891119	C	0.965
17	*MRPL45P2 (dist=27112), NPEPPS (dist=11346)*	rs4794058	A	1.07	rs11079784	A	1
17	*VMP1*	rs8070345	A	1.14	rs2777899	A	0.966
18	*MALT1*	rs7238078	A	1.05		A	
19	*TNFSF14*	rs1077667	G	1.16		G	
19	*SLC44A2*	rs2288904	G	1.14		G	
19	*EPS15L1*	rs1870071	G	1.12		G	
19	*IFI30*	rs11554159	G	1.15	rs874628	A	0.806
19	*DKKL1*	rs8107548	G	1.09	rs2303759	C	0.918
20	*CD40*	rs4810485	A	1.08	rs1569723	C	0.956
20	*SLC9A8*	rs17785991	A	1.09	rs4809760	G	0.687
20	*CYP24A1 (dist=1002), PFDN (dist=32984)*	rs2248359	G	1.07		G	
20	*SLC2A4RG*	rs2256814	A	1.11	rs2315654	G	0.941
20	*ZBTB46*	rs6062314	A	1.10		A	
22	*MAPK1*	rs2283792	C	1.08		C	
22	*TYMP*	rs470119	A	1.07		A	
6	*HLA-DRB1*1501*	rs3135388	T	3.1	rs9271366	G	0.957

Chr: chromosome; RA: risk allele; SNP: single nucleotide polymorphism; OR: odds ratio.

*[a] Total of 102 out of 110 non-HLA MS susceptibility loci and a tagging SNP for the HLA-DRB1*1501 (rs9271366) were available for the calculation of wGRS.*

[b] When an original associated SNP was not available on the Illumina 610-Quad Bead array, we used the best proxy (with $r^2 > 0.6$) according to www.BroadInstitute.org/SNAP.

Appendix 3. Nagelkerke's R^2 values for various wGRS.

	$wGRS_{HLA}$	$wGRS_{102}$	$wGRS_{102+HLA}$
fMS	0.0946	0.0765	0.1674
sMS	0.0372	0.051	0.0892

*The proportion of genetic variability is explained by different wGRS as measured by Nagelkerke'sR^2. The wGRS is calculated for HLA only, 102 non-HLA loci and 102 non-HLA loci + HLA-DRB1*1501 risk allele for sporadic MS patients (sMS) and multiplex MS cases (fMS).*

Chapter 3

RISK GENES ASSOCIATED WITH PEDIATRIC-ONSET MS BUT NOT WITH MONOPHASIC ACQUIRED CNS DEMYELINATION

E.D. van Pelt*, J.Y. Mescheriakova*, N. Makhani, I.A. Ketelslegers, R.F. Neuteboom, S. Kundu, L. Broer, A.C.J.W. Janssens, C.E. Catsman-Berrevoets, C.M. van Duijn, B. Banwell, A. Bar-Or and R.Q. Hintzen.

* These authors contributed equally to this work

Neurology, 2013

ABSTRACT

Objective

To investigate whether 57 genetic risk loci recently identified in a large-scale genome-wide association study in adult patients with multiple sclerosis (MS) are also associated with a risk for pediatric-onset MS and whether they can predict MS diagnosis in children presenting with acquired demyelinating syndromes (ADS).

Methods

We included 188 children with ADS, of which 53 were diagnosed with MS, 466 patients with adult-onset MS and 2046 adult controls in our cohort study. Weighted genetic risk scores (wGRS) were calculated to evaluate genetic effects.

Results

Mean wGRS was significantly higher for patients with pediatric-onset MS (7.32 ± 0.53) as compared with patients with monophasic ADS (7.10 ± 0.47, p=0.01) and controls (7.11 ± 0.53, p<0.01). We found no difference in mean wGRS of participants with monophasic ADS (7.10 ± 0.47) and controls (7.11 ± 0.53). The ability of the wGRS for the 57 single nucleotide polymorphisms (SNPs) to discriminate between children with MS and those with monophasic ADS was moderate (AUC =0.64), but improved with the addition of sex and *HLA-DRB1*15* (AUC =0.70). The combined effect of 57 SNPs exceeded the effect of *HLA-DRB1*15* alone in our risk models for pediatric- and adult-onset MS.

Conclusion

The previously reported 57 SNPs for adult-onset MS also confer increased susceptibility to pediatric-onset MS, but not to monophasic ADS.

INTRODUCTION

Multiple sclerosis (MS) is being increasingly diagnosed in childhood.[1] In children, an initial attack of central nervous system (CNS) demyelination (acquired demyelinating syndrome or ADS) frequently remains monophasic. Approximately 21-32% of the children with ADS will display further MRI or clinical evidence of inflammatory CNS demyelination meeting diagnostic criteria for MS.[2,3] This is in contrast with adults, where the majority of patients are diagnosed with MS after an initial event of acute CNS demyelination.[4] Pediatric-onset MS has been proposed as a unique time window for the study of early MS disease mechanisms. However, it is not known whether pediatric- and adult-onset MS share the same genetic risk factors. A recent Canadian study of a large ADS cohort reported that children harboring one or more *HLA-DRB1*15* alleles were more likely to be confirmed to have MS, compared to ADS children lacking *HLA-DRB1*15* alleles.[5] While HLA alleles thus contribute to the risk of both pediatric- and adult-onset MS[6], large-scale genome-wide association studies (GWAS) recently identified 57 non-HLA genetic risk loci in adult MS patients compared to controls.[7] Whether these 57 single nucleotide polymorphisms (SNPs) also contribute to the risk of either MS in children, or childhood-onset ADS more generally is not known. The objective of this study was to investigate whether the 57 SNPs identified in adult-onset MS are associated with increased risk of pediatric-onset MS, and whether such SNPs distinguish children with MS from children with monophasic ADS. We utilized a previously published approach to generate compound-weighted genetic risk scores[8], and compared the predictive value of the 57 SNPs with the predictive value of *HLA-DRB1*15* alone, in distinguishing children with MS from children with monophasic ADS and controls from the general population.

METHODS

Patients and definitions

Children younger than 16 years 0 days (Canadian cohort) and 17 years 0 days (Dutch cohort) of age who presented with ADS between 2001 and 2009 were enrolled in either the prospective Canadian Pediatric Demyelinating Disease Study or the Dutch Study group for Pediatric MS study. The Canadian National Pediatric Demyelinating Disease Study Group consists of 23 participating pediatric health care centers across Canada. The Dutch Study Group for Pediatric MS consists of 15 participating pediatric healthcare centers in the Netherlands. Initial phenotypes were characterized by clinical history and

physical examination as clinically monofocal optic neuritis (ON), clinically monofocal transverse myelitis (TM), other clinically monofocal disease, clinically polyfocal disease, or acute disseminated encephalomyelitis (ADEM[9]). Children were diagnosed with MS based either on evidence of a second clinical attack after at least 30 days or MRI evidence of dissemination in time.[9,10] The recent 2010 iteration of the McDonald criteria for MS diagnosis was not available at the time of clinical classification, and was not used for this work. All participants had a minimum follow-up of 2 years from initial ADS. A total of 466 adult-onset (> 18 years) MS patients with either whole blood or saliva available to extract DNA were identified through the Rotterdam MS center. DNA from a control group of 2046 unrelated European adults from the general population was obtained from individuals enrolled in the longitudinal Rotterdam Study.[11] In order to control the effect of genetic variation due to ancestry, only participants with self-reported European ancestry were included in the analyses presented in the paper. Potential stratification was corrected for by genomic control and principal component analysis.

Standard protocol approvals, Registration and Patient Consents

Institutional ethical approval by an ethical standard committee on human experimentation was obtained at all 23 sites participating in the Canadian National Pediatric Demyelinating Disease Study and all 15 sites participating in the Dutch national study and for the Rotterdam study. Written informed consent for genetic analysis was obtained from all participants and/ or their families.

SNP Selection and Genotyping

DNA isolation and purification from saliva (Oragene DNA Purification kit, DNA Genotek ®) or whole blood samples[12] was performed. Genotyping was performed using the Illumina Human610-Quad Bead array and the 57 risk SNPs of interest were extracted.[7] An overview of these 57 risk SNPs and their odds ratios (ORs) obtained from a recent GWAS[7] are presented in supplementary table 1. A tagging SNP (rs9271366) for the *HLA-DRB1*15* locus was used. This SNP is in linkage disequilibrium (LD) (r^2=0.957)[13] with the most often described *HLA-DRB1*15* tagging SNP (rs3135388)[14] in MS and is strongly correlated with the presence of *HLA-DRB1*15* itself.[15] We confirmed this linkage with *HLA-DRB1*15* in our pediatric patients with an overlap of 97.4% for *HLA-DRB1*15* allele typed using PCR amplification[5] and the tagging SNP rs9271366 (r^2 =0.95, p<0.01). All genotyping was carried out blinded to clinical data.

Genetic risk score computation

Unweighted genetic risk scores (uwGRS) were calculated by adding the total number of risk alleles for the 57 non-HLA SNPs carried by each individual. Weighted genetic risk scores (wGRS) were calculated by multiplying the number of risk alleles for each SNP with the effect size (log odd ratios) obtained from the literature[7,15] (supplementary table 1) and then taking the sum across all 57 risk SNPs.[8,16] To assess the additional effect of *HLA-DRB1*15* status (as determined by presence of the rs9271366 tagging risk SNP), we also calculated the wGRS for the 57 risk SNPs with and without including *HLA-DRB1*15* status. To determine how well our genetic risk scores discriminated between children with MS and monophasic ADS we constructed receiver operating characteristic (ROC) curves by plotting the sensitivity of the continuous wGRS scores against '1 -specificity' and calculated the area under the ROC curve (AUC). AUC is a measure of how well the model is able to distinguish between patients and non-patients and varies between 0.5 (no discrimination) and 1 (perfect discrimination). Sex, a known risk factor in MS[1,17,18], was also included in the final adjusted models. For comparative purposes, we tested the performance of the wGRS using a risk model for our adult-onset MS patients and controls. This study was reported based on the guideline for the Reporting of Genetic Risk Prediction Studies (GRIPS).[19]

Statistical analysis

Statistical analyses were performed using R software.[20] The Welch two-sample t-test was used to compare the means of the GRS. PredictABEL package[21] was used to compute univariate ORs. For the construction of the ROC curves and computation and comparison of the AUC values, we used PredictABEL[21] and ROCR[22] packages for the R software. SPSS Statistical Software (IBM Company, version 20) was used to analyze categorical and continuous variables (e.g., sex and mean age) using Chi-Square and one-way-ANOVA tests and to calculate Spearman's Rank Correlation Coefficient for *HLA-DRB1*15* allele and tagging SNP rs9271366.

RESULTS

We identified 209 children with European background and with ADS. We excluded 8 patients with relapsing diseases that were not MS (multiphasic / recurrent ADEM n = 2, optic neuritis after ADEM n = 1, and recurrent ON n = 5). Two patients with neuromyelitis optica (NMO) and 5 patients with other alternative diagnoses (vasculitis n = 1, cerebellitis

n = 1, CNS infection n = 2 and one patient with progressive unexplained visual loss) were excluded. Three patients withdrew from the Canadian prospective study within 6 months of ADS and were excluded. Two patients were excluded because poor DNA quality did not permit accurate genotyping and one patient was excluded because data for one risk SNP was missing. A total of 188 children with ADS were therefore included in this study. Of these 188 pediatric ADS patients, 53 children were diagnosed with MS during a mean overall follow-up period of 8.1 months (range 1.1 – 37.0 months, median 5.4 months, 90th percentile 17.6 months). The mean age at ADS (13.1 years ± 3.08) of children subsequently diagnosed with MS was higher than the mean age (9.0 years ± 4.57) of those who remained monophasic ($F = 21.82$, $p<0.01$). There was a greater proportion of females in the MS group as compared to the monophasic ADS group ($\chi^2=7.880$, $p<0.01$). Of the 135 children who had monophasic ADS, 52 had ADEM as their ADS phenotype. Patient and control characteristics are presented in detail in Table 1. We did not find any clinical differences between the children from the Canadian and the Dutch cohort.

Table 1: Characteristics of patients and controls.

	Pediatric-onset MS	**Pediatric-onset Monophasic ADS**	**Controls**	**Adult-onset MS**
Total (n)	53	135	2046	466
Sex (%female)	68	45	56	72
Age at onset (mean, SD)	13.1 (± 3.1)	9.0 (±4.6)	NA	NA
Type of onset				
Optic Neuritis	10 (18.9%)	26 (19.3%)	NA	NA
Transverse Myelitis	4 (7.5%)	28 (20.7%)		
Clinically monofocal disease	17 (32.1%)	11 (8.1%)		
Clinically polyfocal disease	19 (35.8%)	18 (13.3%)		
ADEM	3 (5.7%)	52 (38.5%)		
Follow-up time months (range)	63.5 (25 - 115)	66.0 (24 - 160)	NA	NA

** NA = not available or not applicable.*

There was modest population stratification (inflation factor 1.07). However, exclusion of outliers regarding genomic kinship had no influence on the results. Univariate ORs of the 57 risk SNPs for our patients with pediatric-onset and adult-onset MS are presented in supplementary table1. We calculated the wGRS to investigate whether there is implication of the 57 risk SNPs in pediatric-onset MS risk, and if so, whether it is similarly implicated as in adults. The mean wGRS was similar between pediatric-onset MS and adult-onset MS

(7.32 ± 0.53 vs. 7.40 ± 0.52, p=0.29). The mean wGRS differed significantly between both MS groups and general population controls (p<0.01). We found a significantly higher mean wGRS in pediatric-onset MS patients as compared to children with monophasic ADS (7.32 ± 0.53 vs. 7.10 ± 0.47, p=0.01) and as compared to controls (7.32 ± 0.53 vs. 7.11 ± 0.53, p<0.01). In contrast, there was no difference in mean wGRS of participants with monophasic ADS (7.10 ± 0.47) and controls (7.11 ± 0.53). We did not find differences in mean GRS between children from the Canadian and the Dutch cohort. As an exploratory analysis, we were also interested in whether mean wGRS differed between children with an ADEM (n =52) presentation at ADS and controls or children with pediatric-onset MS. Since ADEM reflects a specific ADS presentation with encephalopathy and typical MRI features and often is post-infectious.[9] No difference was found in mean wGRS between children with ADEM (7.06 ± 0.45) and controls, while the mean wGRS was higher in children with MS as compared to the children with ADEM (7.32 ± 0.53 vs. 7.06 ± 0.45, p<0.01). Similar results were found for all comparisons using the unweighted GRS. In Figure 1A the distribution of the number of risk alleles in pediatric-onset MS patients and children with monophasic ADS is presented.

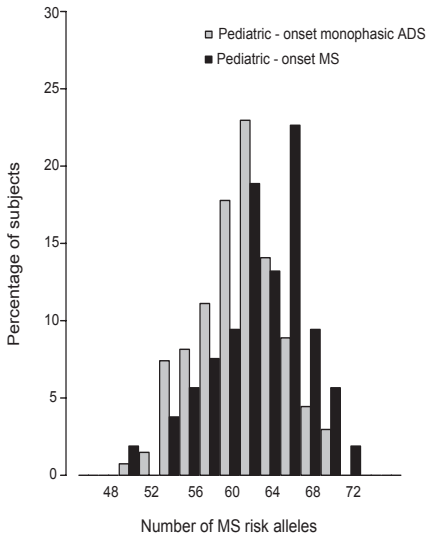


Figure 1A: Distribution of the number of risk alleles in children with MS (black) and monophasic ADS (gray).

It should be noted that patients can harbor 0, 1 or 2 alleles for each risk SNP.
ADS = Acquired Demyelinating Syndrome, MS = Multiple Sclerosis.

Figure 1B displays the distribution of wGRS values as plot boxes for children with MS, children with monophasic ADS, adults with MS and controls from the general population.

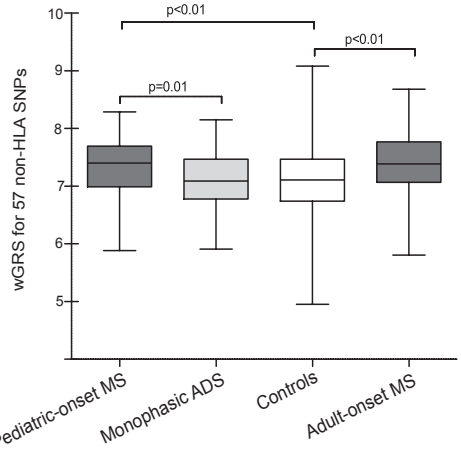


Figure 1B: Box plots presenting the distribution of wGRS values in patients and controls.

The distribution of wGRS is presented for children with MS, children with monophasic ADS, adults with MS and controls from the general population using box plots. We found a significantly higher mean wGRS in pediatric-onset MS patients as compared to children with monophasic ADS (7.32 ± 0.53 vs. 7.10 ± 0.47, p=0.01) and as compared to controls (7.32 ± 0.53 vs. 7.11 ± 0.53, p<0.01). The mean wGRS in adult-onset MS patients was significantly higher compared with controls (7.40 ± 0.52 vs. 7.11 ± 0.53, p<0.01). The mean wGRS was similar between pediatric-onset MS and adult-onset MS (7.32 ± 0.53 vs. 7.40 ± 0.52, p=0.29).
wGRS = weighted Genetic Risk Score, MS = Multiple Sclerosis, ADS = Acquired Demyelinating Syndrome, SNP = Single Nucleotide Polymorphism.

We next examined how well the wGRS values were able to discriminate between children with MS and monophasic ADS or controls, as well as between adults with MS and controls. To do this, we created ROC-curves and calculated AUC-values, including the effects of *HLA-DRB1*15* and sex. AUC values and their confidence intervals are presented in Table 2.

The ability of the wGRS for the 57 non-HLA risk SNPs to discriminate between children with MS and those with monophasic ADS was moderate (AUC =0.64), but improved with the addition of sex and *HLA-DRB1*15* (AUC =0.70). Comparable AUCs of the wGRS for the 57 risk SNPs (AUC =0.66) and for the 57 risk SNPs combined with sex and *HLA-DRB1*15* (AUC =0.73) were found in our adult-onset MS patients when compared with controls. The combined effects of the 57 risk SNPs exceeded the effect of *HLA-DRB1*15* alone in both models.

Table 2: AUC values for risk models for pediatric-onset MS versus monophasic ADS and adult-onset MS versus controls from the general population.

	Pediatric-onset MS		Adult-onset MS	
	AUC	*95% CI*	*AUC*	*95% CI*
HLA-DRB1*15	0.60	0.52 - 0.67	0.63	0.60 - 0.65
***HLA-DRB1*15* and sex**	0.66	0.58 - 0.74	0.67	0.64 - 0.69
57 non-HLA SNPs	0.64	0.54 - 0.73	0.66	0.63 - 0.68
57 non-HLA SNPs and *HLA-DRB1*15*	0.66	0.58 - 0.75	0.71	0.69 - 0.74
57 non-HLA SNPs, *HLA-DRB1*15* and sex	0.70	0.62 - 0.79	0.73	0.71 - 0.76

In Figure 2, we present the ROC-curves for our model predicting pediatric-onset MS in children with ADS. In contrast, the same model using wGRS of the 57 non-HLA SNPs had no ability to discriminate between individuals with monophasic ADS and controls (AUC =0.50).

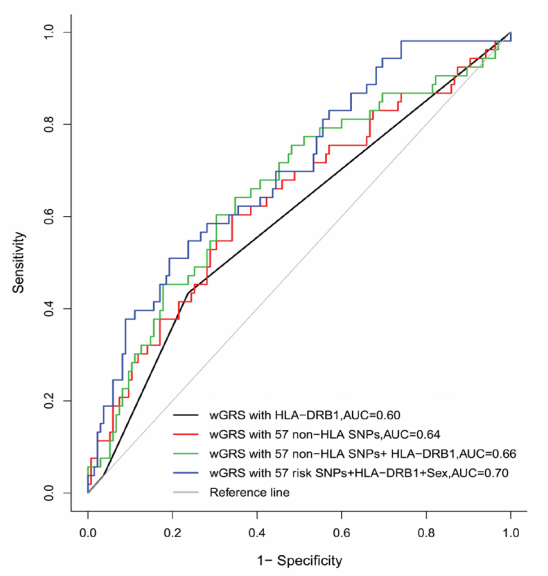


Figure 2: Receiver operating characteristic (ROC) curve for models identifying diagnosis of pediatric-onset MS among children with ADS.

*The results for four separate models to identify pediatric-onset MS cases among children with ADS: wGRS with HLA-DRB1*15 (black); wGRS with 57 non-HLA risk loci (red); wGRS with HLA-DRB1*15 and 57 non-HLA risk loci (green), and wGRS with HLA-DRB1*15, including 57 non-HLA risk loci and sex (blue).*

AUC = Area Under the Curve, HLA = Human Leukocyte Antigen, ROC = Receiver Operating Characteristic, SNP = Single Nucleotide Polymorphism, wGRS = weighted Genetic Risk Score.

DISCUSSION

We report a unique analysis of the 57 non-HLA SNPs recently found to confer risk for adult-onset MS, in a large prospective pediatric ADS cohort including children ascertained to have MS and children with monophasic ADS. Using a compound weighted genetic risk score of the 57 SNPs, we found that mean wGRS significantly differs between pediatric-onset MS patients and controls and between children with MS and those with monophasic ADS. Our results indicate that the 57 non-HLA risk SNPs implicated in adult-onset MS, also contribute to risk of MS in children. These SNPs do not appear to confer a general risk of CNS inflammation in children since wGRS of children with monophasic ADS and controls did not differ.

Disease onset during childhood may represent a heightened genetic susceptibility (a greater "genetic load") or a particularly powerful interaction between genetic factors and childhood environmental risk exposures. We found no significant differences between the GRS of the 57 non-HLA risk SNPs in children and adults diagnosed with MS, suggesting a similar cumulative genetic contribution to disease risk in both pediatric-onset and adult-onset disease. However, whether the very same loci make the same contributions to the pathophysiology of pediatric- and adult-onset MS remains to be fully elucidated, and would require large sample sizes to distinguish individual SNP contributions.

While mean wGRS was higher in the pediatric-onset MS group as compared to both the monophasic ADS and control group, our AUC modeling indicated only a modest ability to discriminate between children with MS and monophasic ADS (AUC =0.70 for final adjusted model). The same model applied to adult-onset disease was comparable in its ability to distinguish between MS patients and controls (AUC =0.73 for final adjusted model) and its discriminatory ability was similar to other published models using a compound genetic risk score in adult-onset MS.[8,16] For comparison, the AUC of LDL-cholesterol as a risk predictor of coronary heart disease was 0.74 in men and 0.77 in woman in a large prospective study.[23]

As we expected, presence of *HLA-DRB1*15* alone had a high contribution to the overall predictive ability of the model (AUC =0.60). However, a surprising finding of our study was that the predictive value as reflected in AUCs of the 57 non-HLA risk SNPs together

(AUC =0.64) was larger than the predictive value of the major MS risk allele *HLA-DRB1*15* (AUC =0.60) alone. We found a similar result in our adult MS patients.

There are several limitations in our study. While our pediatric MS cohort is relatively large given the rarity of this condition, our overall numbers still limit comparisons between groups. Future studies would be aided by large-scale multinational collaborations to facilitate the inclusion of more patients. Despite a mean duration of follow-up of 66.0 months (range 24 – 160), it remains possible that some of the children currently classified as having monophasic disease will be diagnosed with MS in the future. We do not expect that this number will be very high, since pediatric MS studies have demonstrated a high early relapse rate[24] and given that the time interval between incident attack and second event is typically less than 12 months.[2,25] In order to study a genetically homogenous group as possible, our study focused on individuals of European ancestry. Replication studies including individuals of mixed ethnicities will be valuable, though the field is currently hampered by differences in the distribution and linkage disequilibrium of the genetic variants. Other areas of future study include the generation of more complex prediction models that incorporate not only genetic susceptibility but also known environmental factors such as serum 25-hydroxyvitamin D levels and viral exposures.[26] As has been seen in models for adult-onset MS and other autoimmune diseases, it is likely that the incorporation of non-genetic risk factors to the current genetic risk model will lead to improved predictive ability for pediatric-onset MS.[27]

REFERENCES

1. Banwell, B. *et al.* Multiple sclerosis in children: clinical diagnosis, therapeutic strategies, and future directions. Lancet Neurol. 2007;6:887-902.
2. Banwell, B. *et al.* Clinical, environmental, and genetic determinants of multiple sclerosis in children with acute demyelination: a prospective national cohort study. Lancet Neurol. 2011;10:436-445.
3. Neuteboom, R. F. *et al.* Prognostic factors after a first attack of inflammatory CNS demyelination in children. Neurology. 2008;71:967-973.
4. Miller, D. H. *et al.* Clinically isolated syndromes. Lancet Neurol. 2012;11:157-169.
5. Disanto G, Magalhaes S, Handel AE, et al. HLA-DRB1 confers increased risk of pediatric-onset MS in children with acquired demyelination. Neurology 2011;76:781-786.
6. Ramagopalan, S. V. *et al.* Multiple sclerosis and the major histocompatibility complex. Curr Opin Neurol. 2009;22:219-225.
7. International Multiple Sclerosis Genetics Consortium. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476:214-219.
8. De Jager, P. L. *et al.* Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol. 2009;8:1111-1119.
9. Krupp, L. B. *et al*. Consensus definitions proposed for pediatric multiple sclerosis and related disorders. Neurology. 2007;68:S7-12.
10. Polman, C. H. *et al.* Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol. 2005;58:840-846.
11. Hofman, A. *et al.* Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol. 1991;7:403-422.
12. Miller, S. A. *et al.* A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.
13. Johnson, A. D. *et al.* SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24:2938-2939.
14. de Bakker, P. I. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet. 2006;38:1166-1172.
15. Field, J. *et al.*. A polymorphism in the HLA-DPB1 gene is associated with susceptibility to multiple sclerosis. PLoS One. 2010;5:e13454.
16. Jafari, N. *et al.* Perspectives on the use of multiple sclerosis risk genes for prediction. PLoS One. 2011;6:e26493.
17. Wallin, M. T. *et al.* Multiple sclerosis in US veterans of the Vietnam era and later military service: race, sex, and geography. Ann Neurol. 2004;55:65-71.
18. Orton, S. M. *et al.* Sex ratio of multiple sclerosis in Canada: a longitudinal study. Lancet Neurol. 2006;5:932-936.
19. Janssens, A. C. *et al.* Strengthening the reporting of Genetic RIsk Prediction Studies: the GRIPS Statement. PLoS Med. 2011;8:e1000420.
20. The R project for statistical computing [online]. Available at: www.R-project.org. Accessed October 8, 2012.
21. Kundu, S. *et al.* PredictABEL: an R package for the assessment of risk prediction models. Eur J Epidemiol. 2011;26:261-264.
22. Sing, T, *et al.* ROCR: visualizing classifier performance in R. Bioinformatics. 2005;21:3940-3941.
23. Wilson, P. W. *et al.* Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97:1837-1847.

24. Gorman, M. P. *et al.* Increased relapse rate in pediatric-onset compared with adult-onset multiple sclerosis. Arch Neurol. 2009;66:54-59.

25. Mikaeloff, Y. *et al.* First episode of acute CNS inflammatory demyelination in childhood: prognostic factors for multiple sclerosis and disability. J Pediatr. 2004;144:246-252.

26. Handel, A. E. *et al.* Environmental factors and their timing in adult-onset multiple sclerosis. Nat Rev Neurol. 2010;6:156-166.

27. Simon, K. C. *et al*. Combined effects of smoking, anti-EBNA antibodies, and HLA-DRB1*1501 on multiple sclerosis risk. Neurology. 2010;74:1365-1371.

SUPPLEMENTARY DATA

Table 1: Overview of the 57 GWAS-implicated risk SNPs and their odds ratios (ORs), and the univariate ORs for patients with adult-onset and pediatric-onset MS.

				Risk	GWAS	(n=9772)	Adult-onset MS (n=466)		Pediatric-onset MS (n=53)	
Chr	dbSNP rs-number	Position	Gene	Allele	OR *	95%CI	OR	95%CI	OR	95%CI
1	rs4648356	2699024	*MMEL1(TNFRSF14)*	C	1·14	1·12-1·16	1·26	1·08-1·48	1·55	0·98-2·45
1	rs11810217	92920965	*EVI5*	A	1·15	1·13-1·16	0·99	0·84-1·16	1·68	1·12-2·51
1	rs11581062	101180107	*VCAM1*	G	1·12	1·1-1·13	1·08	0·93-1·27	1·13	0·74-1·72
1	rs1335532	116902480	*CD58*	A	1·22	1·19-1·24	1·12	0·89-1·41	0·62	0·37-1·02
1	rs1323292	190807644	*RGS1*	A	1·12	1·1-1·14	1·13	0·94-1·36	1·50	0·86-2·6
1	rs7522462	199148218	*C1orf106(KIF21B)*	G	1·11	1·1-1·13	1·25	1·07-1·47	2·04	1·23-3·41
2	rs12466022	43212565	*No gene*	C	1·11	1·1-1·13	1·24	1·05-1·46	1·06	0·68-1·64
2	rs7595037	68500599	*PLEK*	A	1·11	1·1-1·12	1·17	1·02-1·35	1·59	1·06-2·38
2	rs17174870	112381672	*MERTK*	G	1·11	1·09-1·13	1	0·85-1·19	0·94	0·6-1·47
2	rs10201872	230814968	*SP140*	A	1·14	1·12-1·16	1·23	1·03-1·47	0·97	0·58-1·62
3	rs11129295	27763784	*EOMES*	A	1·11	1·09-1·12	1·12	0·97-1·29	1·18	0·8-1·74
3	rs669607	28046448	*No gene*	C	1·13	1·12-1·15	1·17	1·02-1·35	1·30	0·88-1·91
3	rs2028597	107041527	*CBLB*	G	1·13	1·06-1·21	1·14	0·85-1·53	0·80	0·4-1·6
3	rs2293370	120702624	*TMEM39A(CD80)*	G	1·13	1·11-1·15	1·17	0·97-1·42	1·13	0·68-1·9
3	rs9282641	123279458	*CD86*	G	1·21	1·18-1·24	0·96	0·74-1·25	1·47	0·64-3·37
3	rs2243123	161192345	*IL-12A*	G	1·08	1·06-1·1	1·03	0·88-1·2	1·18	0·79-1·78
4	rs228614	103797685	*NFKB1(MANBA)*	G	1·09	1·07-1·1	1·08	0·94-1·24	1·24	0·84-1·84

5	rs6897932	35910332	*IL7R*	G	1·11	1·09-1·13	1·22	1·03-1·43	1·53	0·95-2·46
5	rs4613763	40428485	*PTGER4*	G	1·2	1·8-1·22	1·18	0·97-1·44	1·10	0·64-1·88
5	rs2546890	158692478	*IL12B*	A	1·11	1·1-1·13	1·02	0·89-1·18	0·90	0·61-1·32
6	rs12212193	91053490	*BACH2*	G	1·09	1·08-1·1	1·01	0·87-1·16	1·38	0·94-2·03
6	rs802734	128320491	*THEMIS*	A	1·1	1·09-1·12	1·16	0·99-1·37	0·91	0·6-1·38
6	rs11154801	135781048	*MYB(AHI1)*	A	1·13	1·11-1·15	1·11	0·96-1·28	1·07	0·72-1·58
6	rs17066096	137494601	*IL22RA2*	G	1·14	1·12-1·15	1·23	1·04-1·44	1·42	0·93-2·17
6	rs13192841	138008907	*No gene*	A	1·1	1·09-1·12	1·24	1·06-1·45	1·06	0·69-1·64
6	rs1738074	159385965	*TAGAP*	G	1·13	1·12-1·15	1·09	0·94-1·26	0·97	0·65-1·44
7	rs354033	148920397	*ZNF746*	G	1·11	1·1-1·13	1·05	0·89-1·24	0·89	0·57-1·38
8	rs1520333	79563593	*IL7*	G	1·1	1·08-1·11	1·11	0·94-1·3	1·15	0·75-1·76
8	rs4410871	128884211	*MYC*	G	1·11	1·09-1·12	1·15	0·97-1·35	1·01	0·65-1·56
8	rs2019960	129261453	*PVT1*	G	1·12	1·1-1·13	1·05	0·89-1·25	0·95	0·59-1·53
10	rs3118470	6141719	*IL2RA*	G	1·12	1·1-1·13	1·19	1·03-1·38	0·82	0·54-1·26
10	rs1250550	80730323	*ZMIZ1*	A	1·1	1·09-1·12	1·12	0·97-1·3	1·03	0·69-1·55
10	rs7923837	94471897	*HHEX*	G	1·1	1·08-1·11	1·11	0·95-1·28	1·00	0·67-1·5
11	rs650258	60588858	*CD6*	G	1·12	1·1-1·13	1·16	1-1·35	0·84	0·56-1·24
11	rs630923	118259563	*CXCR5*	C	1·12	1·1-1·14	1·28	1·05-1·58	1·57	0·86-2·88
12	rs1800693	6310270	*TNFRSF1A*	G	1·12	1·11-1·14	1·14	0·98-1·31	1·06	0·72-1·56
12	rs10466829	9767358	*CLECL1*	A	1·09	1·08-1·11	1·13	0·98-1·3	1·14	0·78-1·68
12	rs12368653	56419523	*CYP27B1*	A	1·1	1·09-1·12	1·05	0·91-1·21	0·97	0·66-1·43
12	rs949143	122161116	*ARL6IP4*	G	1·08	1·04-1·12	1·18	1·02-1·38	1·06	0·7-1·61
14	rs4902647	68323944	*ZFP36L1*	G	1·11	1·1-1·13	1·12	0·97-1·29	0·95	0·64-1·39

Table 1: Continued

14	rs2300603	75075310	*BATF*	A	1·11	1·09-1·12	0·97	0·82-1·14	0·86	0·56-1·31
14	rs2119704	87557442	*GALC(GPR65)*	C	1·22	1·19-1·25	1·01	0·77-1·32	2·12	0·78-5·79
16	rs2744148	1013553	*SOX8*	G	1·12	1·1-1·14	1·12	0·92-1·36	0·88	0·5-1·55
16	rs7200786	11085302	*CLEC16A(CIITA)*	A	1·15	1·13-1·16	1·19	1·04-1·38	1·09	0·74-1·6
16	rs13333054	84568534	*IRF8*	A	1·11	1·1-1·13	0·98	0·83-1·16	0·94	0·59-1·49
17	rs9891119	37761506	*STAT3*	C	1·11	1·09-1·12	1·23	1·06-1·42	0·94	0·63-1·4
17	rs180515	55379057	*RPS6KB1*	G	1·09	1·08-1·11	1·04	0·9-1·21	1·13	0·77-1·68
18	rs7238078	54535172	*MALT1*	A	1·12	1·1-1·14	0·99	0·84-1·17	0·78	0·51-1·19
19	rs1077667	6619972	*TNFSF14*	G	1·16	1·14-1·18	1·35	1·12-1·63	1·27	0·77-2·1
19	rs8112449	10381064	*TYK2(ICAM3)*	G	1·08	1·07-1·1	1·15	0·98-1·35	0·91	0·6-1·37
19	rs874628	18165700	*MPV17L2(IL12RB1)*	A	1·11	1·09-1·12	1·16	0·99-1·37	1·43	0·9-2·28
19	rs2303759	54560863	*DKKL1(CD37)*	C	1·11	1·09-1·13	1·3	1·11-1·52	1·18	0·77-1·8
20	rs2425752	44135527	*CD40*	A	1·11	1·1-1·13	1·17	1-1·37	0·76	0·48-1·22
20	rs2248359	52224925	*CYP24A1*	G	1·12	1·1-1·13	1·11	0·95-1·28	0·76	0·52-1·13
20	rs6062314	61880157	*TNFRSF6B*	A	1·16	1·14-1·19	1·23	0·92-1·66	0·55	0·31-1
22	rs2283792	20461125	*MAPK1*	C	1·1	1·08-1·11	1·11	0·97-1·29	1·41	0·95-2·09
22	rs140522	49318132	*SCO2*	A	1·1	1·09-1·12	1·08	0·93-1·25	1·45	0·98-2·16
6	rs9271366**	32694832	*HLA-DRB1*15*	G	3.1		2.54	2.15-3.01	2.04	1.29-3.22

** Weighted genetic risk score (wGRS) was calculated by multiplying the number of risk alleles for each SNP with the effect size (log OR) obtained from the GWAS.*

*** Field et al., PlosOne 2010*

PART III

RARE GENETIC RISK FACTORS IN MS FAMILIES

Chapter 4.1

LINKAGE ANALYSIS AND WHOLE EXOME SEQUENCING IDENTIFY A NOVEL CANDIDATE GENE IN A DUTCH MULTIPLE SCLEROSIS FAMILY

J.Y. Mescheriakova, A.J.M.H. Verkerk, N. Amin, A.G. Uitterlinden, C.M. van Duijn and R.Q. Hintzen.

Multiple Sclerosis Journal, 2019

ABSTRACT

Background

Multiple sclerosis (MS) is a complex disease resulting from the joint effect of many genes. It has been speculated that rare variants might explain part of the missing heritability of MS.

Objective

To identify rare coding genetic variants by analyzing a large MS pedigree with 11 affected individuals in several generations.

Methods

Genome-wide linkage screen and whole exome sequencing (WES) were performed to identify novel coding variants in the shared region(s) and in the known 110 MS risk loci. The candidate variants were then assessed in 591 MS patients and 3169 controls.

Results

Suggestive evidence for linkage was obtained to 7q11.22-q11.23. In WES data a rare missense variant p.R183C in *FKBP6* was identified that segregated with the disease in this family. The minor allele frequency was higher in an independent cohort of MS patients than in healthy controls (1.27% versus 0.95%), but not significant (OR=1.33 [95% CI 0.8-2.4], p=0.31).

Conclusion

The rare missense variant in *FKBP6* was identified in a large Dutch MS family segregating with the disease. This association to MS was not found in an independent MS cohort. Overall, genome-wide studies in larger cohorts are needed to adequately investigate the role of rare variants in MS risk.

INTRODUCTION

Multiple sclerosis (MS) is a complex demyelinating disorder of the central nervous system. While the cause of MS is still unknown, there is overwhelming evidence that genetic factors are involved.[1, 2]Family studies show that 1 in 5 individuals with MS have an affected family member, and that the chances of developing MS are 20 times higher for a first-degree relative than for an individual from the general population.[3, 4]

The main genetic locus of MS risk in the North-European population is the human leukocyte antigen (HLA) class II region (the classical *HLA-DRB1*15:01* allele).[5] Additionally, since the development of genome-wide association studies (GWAS), many non-HLA genetic variants have been identified.[6-9] Associated genetic variants both within and outside HLA regions are seen more frequently in familial MS than in MS patients with no family history of MS.[10, 11]

Despite the success in the discovery of new associated loci, the variants in the HLA-class II region and the variants identified by GWAS can to date only explain about 27% of the heritability of MS.[9]More importantly, they do not explain MS in families in which the disease segregates as a major locus with a large effect on the disease risk. It has been speculated that rare variants might explain part of the missing heritability of MS, in particular in atypical familial cases.[12, 13]Although families with multiple affected individuals are extremely rare in MS[14], these families may be relevant to investigate and can help in the search to identify new pathways that until now remained undetected in GWAS. Indeed, there are a few reports of whole exome sequencing (WES) successfully identifying rare variants in several MSfamilies.[15-19]

To identify new rare and coding variants involved in the familial aggregation of MS, we performed WES in a large Dutch MS family with 11 affected individuals. We also investigated whether the variants identified in this family could be found in a cohort of unrelated MS patients and healthy controls.

MATERIALS AND METHODS

Study subjects

A multi-incident MS family was ascertained in 2003 through the still ongoing longitudinal study on Genes and Environment interaction in MS (GEMS) in the Netherlands, and has been evaluated for the incidence of new MS cases over the past 13 years. The family originated from the North-East part of the Netherlands, and there was no consanguinity. The number of MS cases was 11 (Figure 1). The MS diagnosis was evaluated according to the standard diagnostic criteria. [20, 21]The medical ethics committee of Erasmus Medical Centre (EMC) approved this study. Written informed consent was obtained from all participants. Findings were assessed in a cohort of 591 Dutch patients and 3169 unrelated Dutch controls. The clinical characteristics of the MS patients are described elsewhere. [11]In the patient group 251 patients were identified as sporadic cases and 340 patients were identified as multiplex cases from 158 MS multiplex families. Healthy controls were enrolled in the longitudinal Rotterdam Study [22].

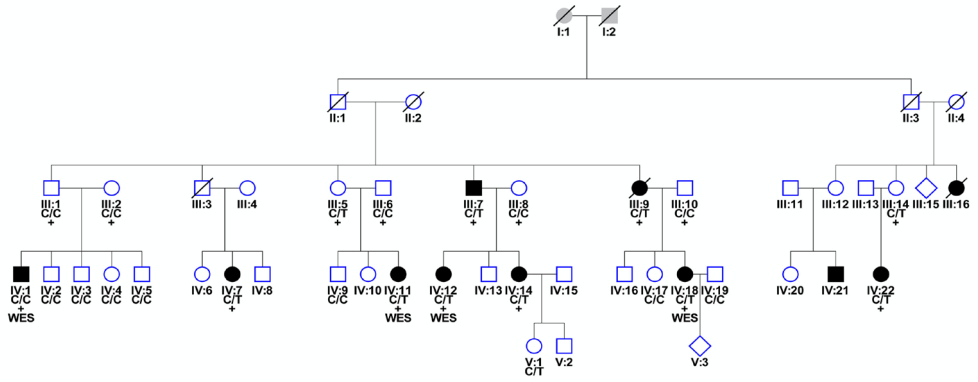


Figure 1: Dutch MS family with a high prevalence of multiple sclerosis.

Filled symbols represent affected individuals. Squares represent males, circles represent females and diamonds represent individuals of an unknown gender. Symbols with a diagonal line represent deceased individuals.

** = individuals included in the linkage analysis*

WES = individuals included in whole exome sequencing analysis

Genetic analysis

DNA was available from 24 family members and was isolated by a standardized method. [23] Twenty-two individuals were genotyped on IlluminaCytoSNP-12 array and two individuals (III:14 and IV:22, Figure 1) on IlluminaCytoSNP-850K array, as they

were included at a later stage. The list of overlapping markers from both arrays was used to calculate LOD scores with Allegro implemented in the EasyLinkage Plus.[24]Single nucleotide polymorphisms (SNPs) with a call rate <95% and those showing Mendelian inconsistencies were excluded from the calculations. MS patients were treated as affected, non-carrier family members as healthy, and unaffected mutation carriers with affected children were defined as having an unknown disease status. A multipoint parametric linkage analysis was performed with Allegro with genotypes from 16 individuals (figure 1) with a SNP spacing of 0.2 and 0.5 cM. LOD scores in sets of 100 markers were calculated assuming the disease in this family to be an autosomal dominant disorder with a population risk allele frequency of 0.001, a penetrance of 70%and a phenocopy rate of 0.001.[14]Regions showing a LOD score >2.0 were used as candidate regions. Flanking SNP markers were used as borders of the haplotype-sharing regions.

Four distantly related individuals were whole exome sequenced (Figure1) on the IlluminaHiSeq2000 platform using Agilent's SureSelectAllExon Kit. Variants with a minor allele frequency of ≤ 1% were searched (see supplementary data for details). Candidate single nucleotide variants (SNVs) from the whole exome sequencing (WES) analysis were then validated in all available family members using iPLEXGoldassays on a MassARRAY system (AgenaBioScience, San Diego, CA, USA). Samples with variant call rate lower than 50% were discarded. The average genotype call rate was 98.9%. Array based genotypes of independent Dutch MS cases and healthy controls from the Rotterdam study were available for a secondary analysis.

RESULTS

Characteristics of the affected family members

Clinical characteristics of MS patients are presented in Table 1. MS patients were 47.9 ± 16.3 years old. Their mean age at the disease onset was 31.3 ± 14.0 years, and the male to female ratio was 1: 3.5. The median disease duration was 16.6 years (ranged from 1.5 to 28.8 years). The median Expanded Disability Status Scale (EDSS) score was 3.5 (mean 4.7 SD ± 3.13) and the mean MS Severity Score (MSSS) was 5.3 (SD ± 3.26). MS with a bout-onset was predominantly represented (i.e. relapsing remitting MS n= 5, secondary progressive MS n=3, two patients had an unknown disease course and 1 patient had primary progressive MS). We interviewed seemingly healthy individuals by phone for disease symptoms, none of these showed evidence for MS based on signs and symptoms.

Table 1. Clinical characteristics of MS patients from the MS family.

ID	Gender	Disease course	Age at disease onset (years)	Age at last examination (years)	EDSS	Disease duration in years (till EDSS)	MSSS
III:7	M	RRMS	63	76	3.5	13	3.94
III:9	F	SPMS	40	69	9.0	29	9.75
III:16	F	MS	NA	NA	NA	NA	NA
IV:1	M	SPMS	23	40	7.5	17	8.38
IV:7	F	RRMS	23	25	1.5	1.5	4.30
IV:11	F	RRMS	14	35	4.0	22	2.97
IV:12	F	RRMS	33	37	2.0	4	4.83
IV:14	F	PPMS	33	48	2.5	15	2.33
IV:18	F	RRMS	25	48	2.5	23	1.26
IV:21	M	MS	NA	NA	NA	NA	NA
IV:22	F	SPMS	27	53	9.5	25	9.98

F = female; M = male; RR = relapsing-remitting; SP = secondary progressive; PP = primary progressive; NA = not available; EDSS = Expanded Disability Status Scale; MSSS = Multiple Sclerosis Severity Score.

Genetic analysis

Multi-point parametric linkage analysis of 16 patients (of which 9 with MS) resulted in a maximum LOD-score of 3.0 on chromosome 7 (Supplementary Figure 1). Analysis of chromosome 7 revealed a shared haplotype region between flanking SNPs rs11972782 and rs6959538 (Human GRCh 37/hg19, chr7:69877261-73139762) in all but one affected individual (IV:1, figure 1). Four distantly related individuals were then whole exome sequenced. WES generated 5.1 Gb of reads per individual. The mean coverage was 44x, and at least 20x of coverage was achieved for 71% of the exome for four samples. On average 35313variants (range 33840 – 36087) were called per individual after quality control. To reduce the number of variants, we filtered the WES data by narrowing the genomic region to the haplotype shared region found in the linkage analysis and setting the minor allele frequency (MAF) at≤ 1% in 1000 Genomes Project and ESP database**.** Further analysis of exonic and splice variants containing non-synonymous and stop variants revealed only one rare non-synonymous SNV rs147213094 (C/T, located on chr7:72745738MAF in ESP= 0.0041, gnomAD =0.002962, ExAC= 0.003155) in the *FKBP6* gene, shared by 3 of the 4 sequenced individuals. Other chromosomes were analysed in WES with the above-mentioned criteria, and MAF between 1 and 5 % were

also used. Several potential candidate variants in the following genes shared by 3 and 4sequenced individuals were found: *FAM184B* (rs61741403), *ACOT12* (rs34607174), *PPP1R9A* (rs61737465), *ADAMTS13* (rs28503257), *SPTY2D1* (rs35411689), *ZNF750* (rs35283702), *MED13L* (rs113890513) and *ABCC3* (rs11568591). By means of sequenom analysis the variants were assessed in 4 WES sequenced individuals and genotyped in the remaining family members. Only the SNP rs147213094 in the *FKBP6* gene located in the linkage region segregated with MS in this family, and was present in 8/9 (88.9%; one affected individual (IV:1) that did not share the haplotype also did not carry the *FKBP6* risk variant) of affected individuals (with DNA available). There were 3/15 (20%) carriers of this variant among unaffected individuals (with DNA available). A healthy female (V:1) carried the risk allele and was according to her age still at risk for MS. Two other unaffected individuals were parents and obligate carriers (III:5 and III:14) who transmitted the allele to their affected offspring.

There were no differences in the HLA-class II risk allele frequency between cases and unaffected family members (p=0.23, Table 2). After correcting for the *HLA-DRB1* status, MS patients carried significantly more often the T-allele than the healthy individuals in this family (OR = 6.9 [95% CI 1.2-39.8, p=0.03). Furthermore, MS patients with the T-allele (minor allele of *FKBP6)* did not differ in their *HLA-DRB1* status compared to MS patients that did not carry the *FKBP6* minor allele (p=0.06).

The frequency of rs147213094 was further analyzed in 591MS cases and 3169 healthy controls. The MAF was slightly higher in all MS patients than in healthy controls (respectively 15/1182alleles (MAF=1.27%) versus 60/6338 alleles (MAF=0.95%)), but there was no significant association with the disease risk (OR= 1.33 [95% CI 0.8-2.4], p=0.31). In the patient group 251 patients were identified as sporadic cases and 340 patients were identified as multiplex cases from 158 MS multiplex families. Sporadic MS patients had MAF of 0.8% (4/502 alleles) and multiplex patients had MAF of 1.62 % (11/680 alleles) of rs147213094.Taking only the probands from the MS multiplex families (n=158, 8/316 with MAF= 2.5%) revealed a marginal significant association with the disease risk (OR= 2.7 [95%CI 1.3-5.7], p=0.01 with a MAF in patients of 2.5% compared to 0.95% in general Dutch population). This is however a a-posteriori analysis that lacks genome-wide significance and adjustments for sex, population structure and MS polygenic risk score based on approximately 200 common susceptibility loci in and outside the HLA-class II region. The variant was enriched in Dutch controls compared to the low frequency in gnomAD (MAF= 0.002962) and ExAc database (MAF=0.003155).

Frameshift insertions, deletions, splicesite mutations, stopgain and stoploss variants located in the linkage region and genome-wide, shared by sequenced individuals were inspected but no potential candidates were identified.

Table 2. Genotypes of MS patients and their family members.

ID	Gender	Disease course	FKBP6 rs147213094, C/T (T- allele is the risk allele)	HLA-DRB1*1501 rs3129868 A/C (r^2 =0.974 with rs9271366; A-allele is the risk allele
III:1	M	NA	0/0	0/1
III:2	F	NA	0/0	0/0
III:5	F	NA	0/1	0/1
III:6	M	NA	0/0	0/0
III:7	M	RRMS	0/1	0/1
III:8	F	NA	0/0	0/0
III:9	F	SPMS	0/1	0/1
III:10	M	NA	0/0	0/1
III:14	F	NA	0/1	1/1
III:16	F	MS	unknown	unknown
IV:1	M	SPMS	0/0	0/1
IV: 2	M	NA	0/0	0/0
IV:3	M	NA	0/0	0/1
IV:4	F	NA	0/0	0/1
IV:5	M	NA	0/0	0/1
IV:7	F	RRMS	0/1	0/0
IV:9	M	NA	0/0	0/0
IV:11	F	RRMS	0/1	0/1
IV:12	F	RRMS	0/1	0/0
IV:14	F	PPMS	0/1	0/1
IV:17	F	NA	0/0	0/0
IV:18	F	RRMS	0/1	1/1
IV:19	M	NA	0/0	unknown
IV:21	M	MS	unknown	unknown
IV:22	F	SPMS	0/1	1/1
V:1	F	NA	0/1	0/1

F = female; M = male; RR = relapsing-remitting; SP = secondary progressive; PP = primary progressive; NA = not applicable; EDSS = Expanded Disability Status Scale; MSSS = Multiple Sclerosis Severity Score.

There were no rare damaging variants in the 110 MS risk genes identified by the recent MS GWAS[9] shared by sequenced individuals.

In earlier few WES studies in MS families several rare variants were reported: a p.A53T variant in *TYK2* (rs55762744)[15], a p.R389H variant in *CYP27B1*(rs118204009)[16, 17], a p.G420D variant in *PLG* (rs139071351)[19], and a p.Arg415G in variant in *NR1H3*(rs61731956)[18]. All variants were found in the Canadian MS patients. None could be found in our WES data.

The SNV rs147213094 (C/T) encodes a missense mutation in exon 5 of *FKBP6* that changes arginine in cysteine (p.R183C). The variant overlaps several transcripts and in many instances tagged as a damaging variant by multiple algorithms including PolyPhen, SIFT, MetalLR, MetaSVM, Mutation Taster, PROVEAN, but not by InterVar that classifies this variant as 'benign' (according to 2015 American College of Medical Genetics Criteria).The computational tools for measurement of evolutionary constraint showed an elevated level of conservation for the arginine residue in mammals (GERP++ = 5.46; PhyloP = 2.35; SiPhy= 8.64), indicating that this amino acid is important for protein function (Supplementary figure 2). Additional support is provided by the CADD score (19.6) indicating that the amino acid substitution belongs to the 10% of the most deleterious substitutions genome-wide.

The mRNA of FKBP6 gene is broadly expressed in all tissues, being the highest in the testes (Figure 2).

Analysis of protein-protein interactions by STRING (software version 10.0) highlighted the interaction of FKBP6 with Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and heat shock protein (Hsp)90 (Figure 3). [25]The included pathways are: regulation of nitric oxide biosynthesis, estrogen signaling pathway and MHC-class II protein binding.

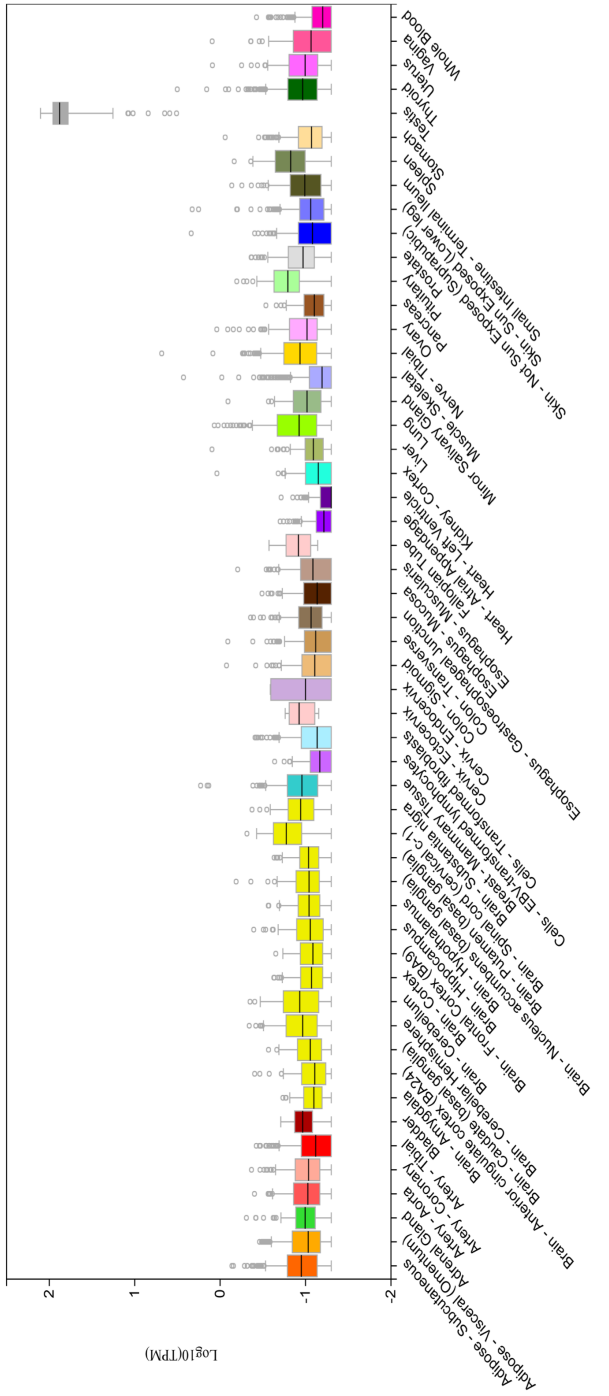


Figure 2: Expression of the *FKBP6* gene in human tissues.

Expression of the FKBP6 from the GTEX database with highest expression in the testes.

Source: https://www.gtexportal.org/home/

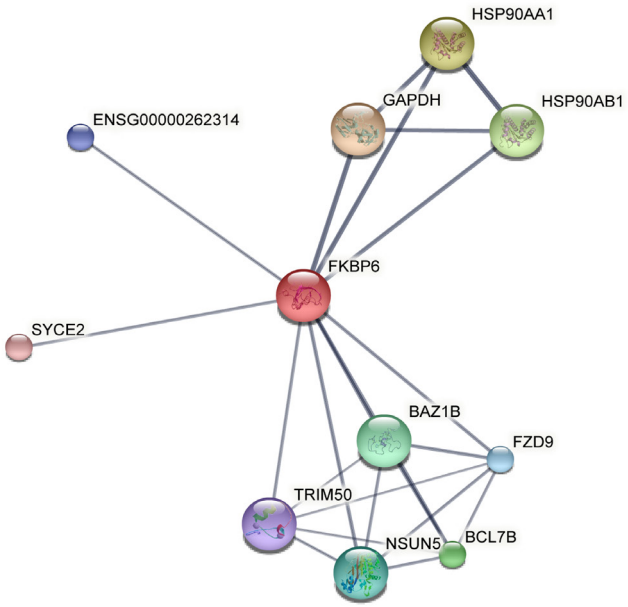


Figure 3: The FKBP6 related protein-protein interactions.

Analysis performed by STRING software (version 10.0). Thicker lines indicate more confident protein-protein interactions.

DISCUSSION

We found a rare missense variant(rs147213094) in the *FKBP6* gene in a large Dutch MS family which resides in the linkage area in the chromosomal region 7q11.22-q11.23. We identified this variant following a classical linkage analysis that revealed a suggestive evidence for linkage to the 7q11 region. This variant was shared by 3 of the 4 relatives that were exome sequenced. In the same WES data searching for variants shared by 3 and 4 sequenced individuals other eight variants emerged, however only *FKBP6* was found to segregate with MS when extending the family to 9 cases (8 of 9 were carriers). One individual with MS did not carry the risk variant. Probably other, non-genetic factors, are involved in this individual, because MS has a multi-factorial nature. While the variant showed good segregation pattern in the family, its association outside the family was not significant in a moderate-sized validation cohort of cases and controls. In a secondary analysis, an increased frequency of the variant was found in the probands and in the total multiplex group, but not in sporadic patients. This observation however

lacks genome-wide significance in this moderate-sized population and statistical adjustments have not been possible for sex, population structure and MS polygenic risk score. Despite the marginal enrichment of this variant in the multiplex MS, the hypothesis that this rare variant in FKBP6 gene is more representative for multiplex MS and not for sporadic MS per se in Dutch population cannot be claimed here. Validation in other and larger cohorts of multiplex MS families will be needed.

The *FKBP6* gene is located on chromosome 7 and belongs to the immunophilins FK-506 binding protein (FKBP) family of highly conserved proteins which possess binding abilities to immunosuppressive drugs. *FKBP6* was first described in the context of Williams-Beuren syndrome (WBS), a genomic disorder with congenital cardiovascular defects, dysmorphic facial features, mental retardation, azoospermie and hypercalcemia caused by a hemizygous contiguous gene deletion on chromosome 7q11.23 encompassing 28 genes in the largest deletions, including *FKBP6*.[26] But also smaller deletions are found in patients excluding *FKBP6.*[27] *FKBP6* was identified as a component of the synaptonemal complex, that forms between two homologous chromosomes during meiosis and mediates chromosome pairing, synapsis and genetic recombination. Four missense variants in *FKBP6* have been associated with human male infertility in a study in idiopathic infertile men.[28] A recent paper identified that the co-chaperone FKBP6 acts as a host factor required for hepatitis C virus (HCV) replication, i.e. HCV replication was completely suppressed in FKBP6-knockout hepatoma cell lines, while the expression of FKBP6 restored HCV replication in FKBP6-knockout cells.[29]However, we speculate that FKBP6 might have as yet undiscovered functions.

The *FKBP6* gene is not among the known MS risk loci. There is some evidence for its possible role in MS susceptibility. The chromosomal region 7q11-q21 has previously displayed a suggestive linkage to MS with a LOD=1.14 in 52 multiplex families (31% with multiple affected generations and/or affected collateral relatives) of European descent.[30]In the experimental autoimmune encephalomyelitis (EAE) model in rats there was a significant linkage in the chromosomal region analogous to the human region containing *FKBP6* with the incidence and duration of EAE, and also the maximum EAE score. [31] A case report described a co-occurrence of WBS and MS in an adolescent patient. [32]

To further explore how *FKBP6* might be linked to MS, we studied pathways and interactions of FKPB6 protein with other proteins. Analysis of protein-protein

interactions by STRING highlighted the interaction of FKBP6 with GAPDH and Hsp90. Studies have shown that FKBP6 (also known as FKBP36) can inhibit GAPDH activity and expression.[33]When GAPDH enzyme activity is inhibited, neurons display chromatin condensation, internucleosomal DNA cleavage, and cytoplasmic shrinking[34], and this could result in widespread neuroaxonal apoptosis and degeneration. FKBP6 also interacts with Hsp90 by being its co-chaperone, and this complex can interact with different proteins.

The rs147213094 in *FKBP6* encodes a missense mutation that changes arginine to cysteine (R183C). The variant is located in exon 5 and expected to affect the tetratricopeptide (TPR) repeat of the FKBP6 protein. The TPR repeat domains are critical for trans-membrane protein-protein interactions and the formation of multiprotein complexes. [35] It is as yet not clear how the identified polymorphism might affect the protein-protein interaction process, and be involved in MS susceptibility.

We failed to identify new rare variants in the 110 MS GWAS loci and previously reported rare variants in *TYK2, CYP27B1, PLG* and *NR1H3* in the exome sequenced patients. [15, 16, 18, 19] Concerning the rare p.R389H mutation in *CYP27B1* found in Canadian families*[16]*, only one study could replicate this variant in a multi-incident MS family of the Canadian descent, but not in the general Canadian MS population[17]. Other studies failed to replicate this finding in ethnically different populations.[36, 37]There are no studies we are aware of that tried to replicate the rare variants in *TYK2*[15] and *PLG*[19]. A recent publication of a rare variant in *NR1H3,* which claimed that this variant was causal for familial MS and a common variant in the same gene was associated with primary progressive MS, caused a discussion in the scientific community[38, 39]with several parties debating the arguments pro and against its role in MS. It is very likely that the rare variants are family- and population specific and have therefore different distributions in different populations compared to common variants.[40] Rare variants can also show higher levels of stratification and therefore lack replication in association studies.[40]Systematic interrogation of genome-wide rare variants in larger cohorts is needed to address this issue properly.

There were several limitations to our study. Firstly, by finding a suggestive linkage on chromosome 7 and combining these data with WES, we were able to point to a SNV with medium increased risk that turned out to be segregating with the disease in this family. The search for rare variants in other regions and the MS GWAS loci relied on

the analysis of only 4 exome sequenced individuals. If there were more of rare SNVs with medium increasing MS risk, these would probably be missed because of only few individuals WES analysed. Secondly, because we are dealing with a rare variant, the number of patients in the validation cohort should be preferably in the thousands. The major limitation of this study is lack of power in the validation cohort due to limited patient resources.

In summary, we here report a rare variant in the *FKBP6* gene found in a Dutch multi-incident MS family using linkage analysis and whole exome sequencing technology. Evidence for its involvement in familial MS in general is not sufficiently supported here. The few studies thus far that claimed the influence of rare genetic variants on MS susceptibility are still devoid of replication.[36-39]Systematic interrogation of genome-wide rare variants in larger cohorts is needed to answer the question if and to what extend rare variants contribute to MS risk.

REFERENCES

1. Kenealy, S. J. *et al.* The genetic epidemiology of multiple sclerosis. J Neuroimmunol. 2003;143:7-12.
2. Sadovnick, A. D. & Ebers, G. C. Genetics of multiple sclerosis. Neurol Clin. 1995;13:99-118.
3. O'Gorman, C. *et al.* Modelling genetic susceptibility to multiple sclerosis with family data. Neuroepidemiology. 2013;40:1-12.
4. Ebers, G. C. *et al.* The natural history of multiple sclerosis: a geographically based study: 8: familial multiple sclerosis. Brain. 2000;123: 641-9.
5. Patsopoulos, N. A. *et al.* Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013;9: e1003926.
6. Australia and New Zealand Multiple Sclerosis Genetics Consortium. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet. 2009;41: 824-8.
7. International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med. 2007;357:851-62.
8. International Multiple Sclerosis Genetics Consortium. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476:214-9.
9. International Multiple Sclerosis Genetics Consortium. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353-60.
10. Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann Neurol. 2011;69:65-74.
11. Mescheriakova, J. Y. *et al.* Burden of genetic risk variants in multiple sclerosis families in the Netherlands. MSJ-ETC. 2016:1-8.
12. Bodmer, W. *et al.* Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet. 2008;40:695-701.
13. Schork, N. J. *et al.* Common vs. rare allele hypotheses for complex diseases. Curr Opin Genet Dev. 2009;19:212-9.
14. Dyment, D. A. *et al.* A genome scan in a single pedigree with a high prevalence of multiple sclerosis. J Neurol Neurosurg Psychiatry. 2008;79:158-62.
15. Dyment, D. A. *et al.* Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. Neurology. 2012;79:406-11.
16. Ramagopalan, S. V. *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. Ann Neurol. 2011;70:881-6.
17. Ross, J. P. *et al*. Analysis of CYP27B1 in multiple sclerosis. J Neuroimmunol. 2014;266:64-6.
18. Wang, Z. *et al.* Nuclearreceptor NR1H3 in familial multiplesclerosis. Neuron. 2016;90:948-54.
19. Sadovnick, A. D. *et al*. Analysis of Plasminogen Genetic Variants in Multiple Sclerosis Patients. G3. 2016;6:2073-9.
20. Poser, C. M. *et al.* New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol. 1983;13:227-31.
21. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50:121-7.
22. Hofman, A. *et al.* The Rotterdam Study: 2012 objectives and design update. Eur J Epidemiol. 2011;26:657-86.
23. Miller, S. A. *et al.* A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.
24. Hoffmann, K. & Lindner, T. H. easyLINKAGE-Plus-automated linkage analyses using

large-scale SNP data. Bioinformatics. 2005;21:3565-7.

25. Jensen, L. J. *et al.* STRING 8-a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 2009;37:D412-6.
26. Schubert, C. The genomic basis of the Williams-Beuren syndrome. Cell Mol Life Sci. 2009;66:1178-97.
27. Fusco, C. *et al.* Smaller and larger deletions of the Williams Beuren syndrome region implicate genes involved in mild facial phenotype, epilepsy and autistic traits. Eur J Hum Genet. 2014;22:64-70.
28. Zhang, W. *et al.* Mutation screening of the FKBP6 gene and its association study with spermatogenic impairment in idiopathic infertile men. Reproduction. 2007;133:511-6.
29. Kasai, H. *et al.* Involvement of FKBP6 in hepatitis C virus replication. Sci Rep. 2015;5:16699.
30. Haines, J. L. *et al.* A complete genomic screen for multiple sclerosis underscores a role for the major histocompatability complex. The Multiple Sclerosis Genetics Group. Nat Genet. 1996;13:469-71.
31. Becanovic, K. *et al.* Advanced intercross line mapping of Eae5 reveals Ncf-1 and CLDN4 as candidate genes for experimental autoimmune encephalomyelitis. J Immunol. 2006; 176:6055-64.
32. Papageorgiou, E. *et al.* Multiple sclerosis in association with Williams-Beuren syndrome. J Paediatr Child Health. 2010;46:612-4.
33. Jarczowski, F. *et al.* FKBP36 is an inherent multifunctional glyceraldehyde-3-phosphate dehydrogenase inhibitor. J Biol Chem. 2009;284:766-73.
34. Nakazawa, M. *et al.* Koningic acid (a potent glyceraldehyde-3-phosphate dehydrogenase inhibitor)-induced fragmentation and condensation of DNA in NG108-15 cells. J Neurochem. 1997;68:2493-9.
35. Allan, R. K. & Ratajczak ,T. Versatile TPR domains accommodate different modes of target protein recognition and function. Cell Stress Chaperones. 2011;16:353-67.
36. Ban, M. *et al.* No evidence of association between mutant alleles of the CYP27B1 gene and multiple sclerosis. Ann Neurol. 2013;73:430-2.
37. Barizzone, N. *et al.* No evidence for a role of rare CYP27B1 functional variations in multiple sclerosis. Ann Neurol. 2013;73:433-7.
38. Minikel, E. V. *et al.* Publicly Available Data Provide Evidence against NR1H3 R415Q Causing Multiple Sclerosis. Neuron. 2016;92:336-8.
39. International Multiple Sclerosis Genetics Consortium. NR1H3 p.Arg415Gln Is Not Associated to Multiple Sclerosis Risk. Neuron. 2016;92:333-5.
40. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in spatially structured populations. Nat Genet. 2012;44:243-6.

SUPPLEMENTARY MATERIAL

Supplementary data on Methods and Materials

Whole exome sequencing

Four distantly related individuals were selected for WES (indicated in Figure1). Exome capture was performed using Agilent's SureSelect AllExon Kit. Paired-end sequencing (2x100 base pair reads) was performed on the IlluminaHiSeq2000 platform according to the manufacturers' instructions. Reads were mapped to the human reference genome sequence (assemblyGRCh37/hg19) using the Burrows-Wheeler Alignment Tool.[1] The identified genetic variants were called with the Genome analysis Tool Kit (GATK) using the following quality criteria: Phred-like consensus quality of ≥ 30, quality by depth ≥ 5, coverage of ≥ 5 and strandbias< 0.75.[2] All variants were annotated by ANNOVAR.[3] For identification and filtering of the variants with a minor allele frequency of ≤ 1%, we used dbSNP138 (http://www.ncbi.nlm.nih.gov/projects/SNP/), the 1000 Genome Project (http://www.1000genomes.org/), and the National Heart Lung and Blood Institute (NHLBI) Grand Opportunity Exome Sequencing Project (ESP; https://esp.gs.washington.edu/drupal/). For interpretation of functional and deleterious variants, the following tools were used and added to the annotation file: 1) scores from 8 prediction algorithms (PolyPhen2[4], Sorting Intolerant from Tolerant (SIFT) [5], LRT[6], MutationTaster[7], Mutation Assessor [8], FATHMM[9], Radial SVM[10]and LR[11]; 2) a scaled Combined Annotation-Dependent Depletion (CADD) score which integrates many diverse annotations into a single measure of potential functional impact for each genetic variant[12]; a scaled CADD- score of ≥10 indicates that variants are predicted to be the 10% most deleterious substitutions in the human genome, a score of greater or equal 20 indicates the 1% most deleterious and so on; and 3) prediction scores for evolutionary conservation (GERP++ RS[13], PhyloP[14]andSiPhy[15]).

Validation by SNP genotyping

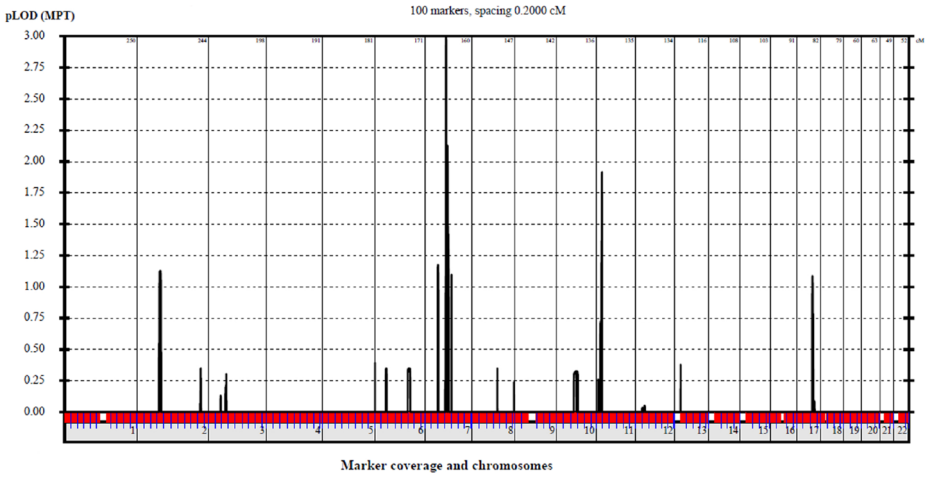
Candidate Single nucleotide variants (SNVs) from WES analysis were validated in all available family members from the pedigree using iPLEX Gold assays on a MassARRAY system (AgenaBioScience, San Diego, CA, USA). The assays were based on multiplex PCR and were designed with online available Assay Design Suite software from the same company. Clustering was called using TyperAnalyzer 4.0.22.67 software (AgenaBioScience). To ensure genotyping quality, SNV calling was verified visually to ensure distinctly formed clusters. Additionally, samples with SNV call rate lower than 50% were discarded. The average genotype call rate was 98.9%. Array based genotypes

of independent Dutch MS cases and healthy controls from the Rotterdam study were available for a secondary analysis.

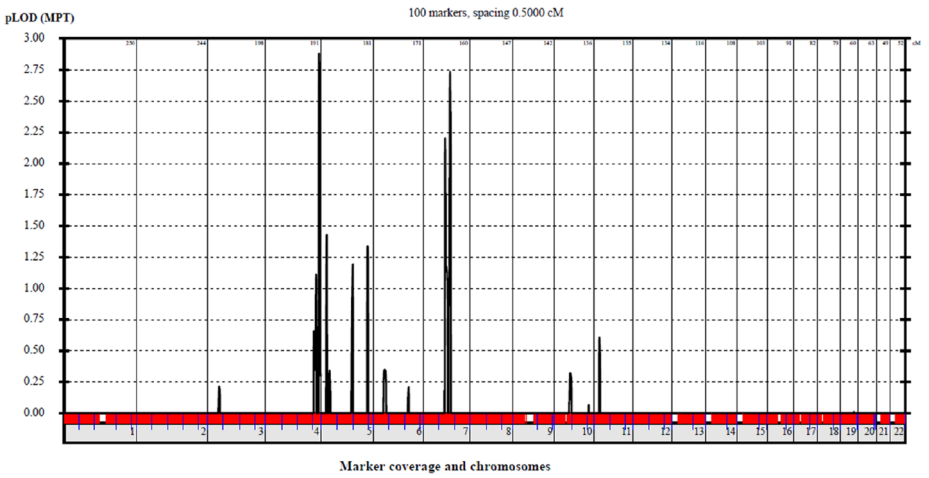
SUPPLEMENTARY REFERENCES

1. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-60.
2. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297-303.
3. Wang, K. *et al.* ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.
4. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248-9.
5. Ng, P. C. & Henikoff, S. Predicting deleterious amino acid substitutions. Genome Res. 2001;11:863-74.
6. Chun, S. & Fay, J. C. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19:1553-61.
7. Schwarz, J. M. *et al.* MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11:361-2.
8. Reva, B. *et al.* Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011;39:e118.
9. Shihab, H. A. *et al.* Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. 2013;34:57-65.
10. Cortes, C. & Vapnik, V. Support-vector networks. Mach Learn. 1995; 20: 273-97.
11. Agrestie, A. Categorical Data Analysis. 2002.
12. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310-5.
13. Davydov, E. V. *et al.* Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010;6:e1001025.
14. Cooper, G. M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 2005;15:901-13.
15. Garber, M. *et al.* Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics. 2009;25:54-62.

A

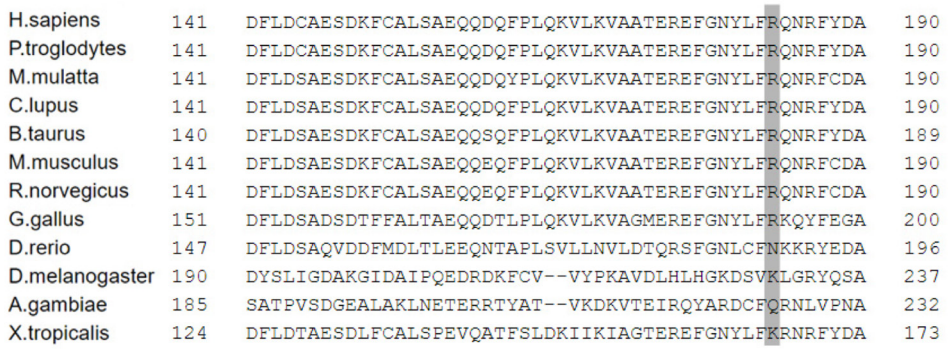


B



Supplementary figure 1. Plots of the LOD scores.

Plots of the LOD scores using Allegro. A multipoint linkage analysis was performed with a SNP spacing of 0.2 (A) and 0.5 (B) cM. Regions on the chromosome 7 with a LOD score >2.0 were used as candidate regions.



Supplementary figure 2. Sequence alignment of the FKBP6 protein across the species.

Multiple alignments of the FKBP6 protein across the different species. The amino acid arginine is indicated in grey and is conserved in several mammals and chicken.

Chapter 4.2

RESPONSE TO VILARIÑO-GÜELL ET AL.

J.Y. Mescheriakova, A.J.M.H. Verkerk, N. Amin, A.G. Uitterlinden, C.M. van Duijn and R.Q. Hintzen.

Multiple Sclerosis Journal, 2019

In response to our paper[1] Vilariño-Güell et al.[2] genotyped *FKBP6* p.R183C in 2383 MS patients (1952 with familial MS and 431 with sporadic MS) and 1055 controls from the Canadian population. They found a minor allele frequency (MAF) of 0.31% in all MS patients and MAF of 0.28% in controls, but the difference was not significant. They also showed that MAF of this variant was 1.4 times higher (we found 2x higher) in familial MS compared to sporadic MS (MAF 0.33% vs 0.23% respectively), although not statistically significant. *FKBP6* p.R183C did not segregate in their Canadian families.

It is important to emphasize that rare alleles have different genetic structures in ethnically different populations. [3] The frequency of this variant was in fact 3x higher in the Dutch control population (0.95%) compared to Canadian controls (0.28%) and the Exome Aggregation database (0.32%). In this light, one should be cautious with one by one comparison of the populations, even seemingly comparable ones.

Furthermore, rare variants can be private to a specific family of a specific origin. The lack of segregation in Canadian families does not simply eliminate the possibility that *FKBP6* p.R183C is a potentially disease causing variant for this specific Dutch MS family. As MS is a heterogenous disorder with a high degree of variability between individuals and ethnicities in terms of clinical features, genetics, pathology and immunological phenotypes[4], it is possible that other factors or a combination of factors have the upper hand in the Canadian families in terms of MS susceptibility. Moreover, the lack of association in Canadian case-control analysis may be due to the lack of power resulting from the rarity of the variant in this population.

The linkage area on chromosome 7 (chr7:69877261-73139762) contains about ~40 genes. We revisited raw exome sequencing data and inspected the alignments in this area using Integrative Genomics Viewer (IGV) [8]. No inter-chromosomal rearrangements were observed in 4 sequenced individuals. As stated in the paper[1] frameshift insertions, deletions, splicesite mutations, stopgain, and stoploss variants located in the linkage region shared by sequenced individuals were inspected, but no potential candidates have been identified. The authors further comment on the possibility that the true pathogenic variant could have been overlooked because of variable degree of coverage. A sufficient coverage on a single base resolution of at least 20x is usually necessary for reliable detection of sequence variations .[5] At least 10x of coverage was achieved for 90.1% and 20x for 71% of the whole exome for four samples. Similar coverage was found for chromosome 7, including the linkage area. It is a general fact that not all

targeted regions are captured at the same efficiency in whole exome sequencing (WES). Even not every kit covers the same exome regions.[6] The possibility of missing out the 'actual disease causing variant' thus cannot be ruled out here. However, even with deep coverage, models for SNP calling can sometime lead to considerable false positive rates[7], thus not only the measure of coverage but other stringent filtering parameters are also important. Our sequencing data went through a well-designed quality control pipeline to ensure the quality of the called variants.

In our paper, we made an effort to identify a risk variant by means of WES which was located in the linkage area and segregated with MS in the Dutch family. Our data analysis met the criteria needed for reliable detection. Despite this, it is possible that *FKBP6* p.R183C may not be the actual causal variant but rather be in LD with another risk variant not covered by exome-sequencing, as WES might miss out on variants in inadequately targeted regions and is unsuitable for finding potential candidates in cryptic exons or non-coding regions.

REFERENCES

1. Mescheriakova, J. Y. *et al.* Linkage analysis and whole exome sequencing identify a novel candidate gene in a Dutch multiple sclerosis family. Mult Scler. 2018:1352458518777202.
2. Vilarino-Guell, C. *et al.* Analysis of Canadian multiple sclerosis patients does not support a role for FKBP6 in disease. Mult Scler. 2018;25:1011-1013.
3. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in spatially structured populations. Nat Genet. 2012;44:243-6.
4. Disanto, G. *et al.* Heterogeneity in multiple sclerosis: scratching the surface of a complex disease. Autoimmune Dis. 2010;2011:932351.
5. Mertes, F. *et al.* Targeted enrichment of genomic DNA regions for next-generation sequencing. Brief Funct Genomics. 2011; 10:374-86.
6. Harismendy, O. *et al.* Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol. 2009; 10:R32.
7. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491-8.
8. Robinson, J. T. *et al.* Integrative genomics viewer. Nat Biotechnol. 2011;29:24-6.

PART IV

ENVIRONMENTAL RISK FACTORS IN MS FAMILIES

Chapter 5

EBNA-1 TITER GRADIENT IN MULTIPLE SCLEROSIS FAMILIES INDICATES A GENETIC CONTRIBUTION

J.Y. Mescheriakova, G.P. van Nierop, A.A. van der Eijk, K.L. Kreft and R.Q. Hintzen.

Neurology: Neuroimmunology & Neuroinflammation, 2020

ABSTRACT

Objective

In multiplex multiple sclerosis (MS) families, we determined the humoral immune response to Epstein–Barr virus (EBV) nuclear antigen-1 (EBNA-1)-specific immunoglobulin γ (IgG) titers in MS patients, their healthy siblings and biologically unrelated healthy spouses, and investigated the role of specific genetic loci on the antiviral IgG titers.

Methods

IgG levels against EBNA-1 and Varicella Zoster Virus (VZV) as control were measured. *HLA-DRB1*1501,* and *HLA-A*02* tagging single nucleotide polymorphisms (SNPs) were genotyped. We assessed associations between these SNPs and antiviral IgG titers.

Results

Odds ratio (OR) for abundant EBNA-1 IgG was highest in MS patients and intermediate in their siblings compared to spouses. We confirmed that *HLA-DRB1*1501* is associated with abundant EBNA-1 IgG. After stratification for *HLA-DRB1*1501*, the EBNA-1 IgG gradient was still significant in MS patients and young siblings compared to spouses. *HLA-A*02* was not explanatory for EBNA-1 IgG titer gradient. No associations for VZV IgG were found.

Conclusions

In families with MS, the EBNA-1 IgG gradient, being the highest in MS patients, intermediate in their siblings and lowest in biologically unrelated spouses indicates a genetic contribution to EBNA-1 IgG levels, that is only partially explained by *HLA-DRB1*1501* carriership.

INTRODUCTION

Familial clustering in multiple sclerosis (MS) is supportive for strong genetic determinants in MS etiology. The *HLA-DRB1*1501* containing haplotype is the strongest genetic MS-associated risk factor, while *HLA-A*02* has a protective effect on MS.[1,2] Additionally, more than 200 non-HLA MS susceptibility loci with modest odds ratios (ORs) have been identified. [3] The associated genetic factors are seen more often in familial MS than in non-familial MS.[4,5]

Besides genetic factors, environmental factors contribute to the risk of developing MS[6]. A recent meta-analysis of twin studies showed that environmental influences contribute for 21% of MS liability variance.[7] The major environmental risk factor is an infection with the herpesviridae family member Epstein-Barr virus (EBV).[8,9] Furthermore, immunoglobulin γ (IgG) response to EBV nuclear antigen 1 (EBNA-1) is heritable for 22-43%, suggesting that host-genetic factors are important in the immune response to EBV.[10–12] In relation to EBV antibody titers in MS patients and their twins and siblings, only a few small-sized studies were conducted that showed somewhat variable results.[13–17]

The aim of our study was to determine the influence of genetic factors on humoral immune response towards EBNA-1 in multiplex MS families, siblings and controls. We hypothesized that due to shared genetic pool of MS patients, their healthy siblings might have an increased IgG response to EBNA-1 compared to unrelated controls. Therefore, we determined serum EBNA-1 and varicella zoster virus IgG as a control herpesvirus not associated with development of MS[18] in these three groups and assessed the influence of *HLA-DRB1*1501* and *HLA-A*02* on antiviral titers.

MATERIALS AND METHODS

Study participants

The majority of the participants (257 MS patients and 173 unaffected siblings from 136 multiplex MS families, and their 135 unrelated healthy spouses) were included from the still ongoing study on gene-environment interaction in MS (GEMS) in the Netherlands. In this study, multiplex MS families are included, in which at least one first- or second-degree relative of an affected proband was also diagnosed with MS. The

remaining participants (44 MS patients, 25 unaffected siblings and 39 healthy spouses) were included from the Genetic Research in Isolated Populations (GRIP) study. Details of ascertainment are described elsewhere.[19] The diagnosis of MS in all patients was evaluated according to the standard diagnostic criteria.[20,21]

Serological testing

Sera samples were collected and stored at -80˚C. Serum EBNA-1 IgG and VZV IgG levels were determined using well-validated chemoluminescent assays (Liaison XL, Diasorin) according to the manufacturers' instruction. In samples negative for EBNA-1, anti-virus capsid antigen (VCA) IgG (Diasorin) was measured to ascertain EBV seroprevalence. If antibody levels were above the threshold of the assay, samples were diluted 20-fold using sample diluent (Diasorin) and re-analyzed. EBNA-1 and VCA double seronegative and VZV seronegative individuals were omitted from further analyses to prevent bias.

Genotyping

Genomic DNA was isolated using standardized methods.[22] MS-associated single nucleotide polymorphisms (SNPs; table e-1) were genotyped using the Sequenom platform according to manufacturers' instruction. The average genotype call rate for both SNPs was 99 %.

Statistical analysis

Data was analyzed using SPSS version 25.0 (SPSS Inc), and GraphPad Prism5 (GraphPad) was used to construct the graphs. Cases with missing data were omitted. (EBNA-1 and VZV IgG titers were not normally distributed, also not after log transformation (both $p<0.001$, Kolmogorov-Smirnov test). Therefore, IgG levels were dichotomized as above or below the 75th percentile of the levels of the spouses. We used 2-tailed t-test or non-parametric Mann-Whitney U test to compare continuous variables. χ^2-test or Fisher exact test were used to analyze nominal data. Generalized linear models (GLM) were used for pairwise comparison of the study groups in relation to EBNA-1 and VZV titers, adjusted for gender, and household (i.e. to which family the samples belong to). Dichotomized IgG levels were also used to assess the odds ratios (OR) for the MS-risk SNPs. ORs and associated confidence intervals (CI) were calculated using logistic regression. The number of homozygous SNP carriers for a minor allele in both SNPs was < 50, therefore the homozygous group was pooled with heterozygous allele carriers (table e-2) to improve power. Two-sided p-values less than 0.05 were considered

significant. Significance is indicated in the figures as * p<0.05, ** p<0.01, *** p<0.001. Bonferroni correction was applied to correct for multiple testing.

Standard protocol approvals, registrations and patient consents

Written informed consent was obtained from all participants with approval from the medical ethical committee of the Erasmus MC (Rotterdam, the Netherlands).

Data availability

All data are available from the corresponding author.

RESULTS

This study included a total of 301 MS patients, their 198 unaffected siblings and 174 unrelated healthy spouse controls (see Figure 1 for the study flowchart). Table 1 shows the clinical and demographic characteristics of the study population.

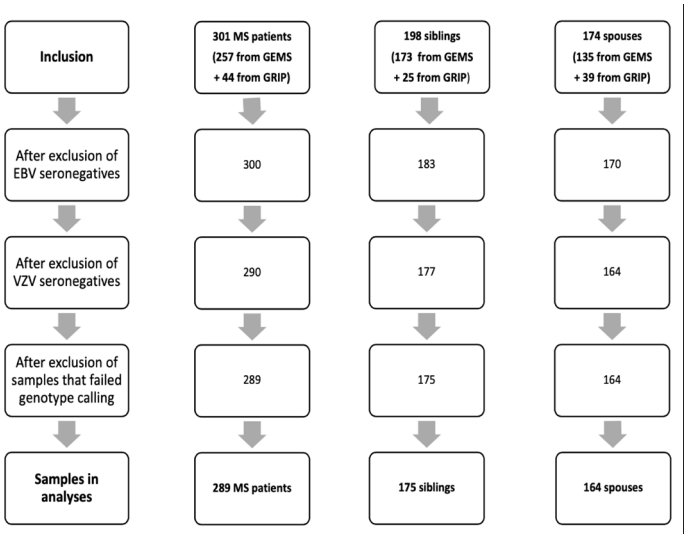


Figure 1. Flowchart of the study.

Only genotyped Epstein-Barr virus (EBV, determined as EBV nuclear antigen-1 or viral capsid antigen IgG positive individuals) and VZV IgG seropositive individuals were included in this study. GEMS, study on gene-environment interaction in MS; GRIP, Genetic Research in Isolated Populations program.

EBNA-1 IgG titers were inversely correlated with age

EBV seroprevalence in MS patients was higher than in their spouses, and their siblings (Table 1). No differences were found in VZV seroprevalence between all study groups.

A one-way analysis of variance showed that the effect of age at sampling was significant ($F(2,662)=11.3$, $p=1.5x10^{-5}$). MS patients were significantly younger than siblings and spouses (Table 1). Age was inversely correlated with EBNA-1 IgG titers (Spearmans'ρ= -0.1, $p=1.3x10^{-2}$) and age at sampling was lower in the group with high EBNA-1 titers (> 75th percentile) (mean age 48.3y, 95% CI 46.8 – 49.8), compared to the group with low EBNA-1 IgG titers (< 75th percentile) (mean age at sampling 50.6y, 95% CI 49.4 – 51.9, adjusted $p=2.0x10^{-2}$). Therefore, we stratified all data into two age groups by using the median age, i.e. under 50 (young, <50y) and above 50 years of age at sampling (old, >50y). Young MS patients had higher EBNA-1 IgG titers compared to elderly MS patients (OR = 1.7, 95% CI 1.0 – 2.7, adjusted $p=3.5x10^{-2}$).

EBNA-1 IgG titers were highest in MS patients, intermediate in siblings, and lowest in their spouses

Young MS patients (<50y) had an increased risk for high EBNA-1 IgG titers (> 75th percentile) compared to spouses (Figure 2A). Also, young siblings had an increased risk for high EBNA-1 IgG titers compared to spouses (Figure 2A). In this age group, MS patients were more likely to have high EBNA-1 IgG titers compared to siblings (OR=2.7, 95% CI 1.5 – 5.0, adjusted $p=2.0x10^{-3}$). Also, elderly MS patients (>50y) had an increased risk for high EBNA-1 titers compared to spouses, and their old siblings (Figure 2A). There was no association between high VZV antibody titers and the study groups (Figure 2B). EBNA-1 and VZV antibody titers were further not associated with gender, clinical disease course or disease duration (data not shown).

Table 1. Clinical and demographic characteristics of MS patients, siblings and spouses.

	MS patients N=301	Siblings N=198	Spouses N=174	p-value
Gender (N) **female:male ratio**	206:95 2.2:1	93:105 0.9:1	68:106 0.6:1	$1.8x10^{-10}$
Age at sampling (y)	46.9±12.6	51.7±12.2	51.1±12.1	Between groups: $1.5x10^{-5}$ MS vs Sibs: $6.8x10^{-5}$ MS vs Spouses: $0.1x10^{-2}$
Age at disease onset (y)	33±10	-	-	-
EDSS (median)	3.5	-	-	-
MSSS	4.7± 2.7	-	-	-
Disease duration (y)	15±11	-	-	-
RR-MS **CIS** **SP-MS** **PP-MS**	207 (68.8%) 15 (5.0%) 41 (13.6%) 38 (12.6%)	-	-	-
EBV seropositive	300 (99.7%; 95% CI 99.0 – 100.0)	183 (92.4%; 95% CI 88.7 – 96.1)	170 (97.7%; 95% CI 95.5 – 99.9)	$1.6x10^{-5}$
VZV seropositive	291 (96.7%; 95% CI 94.7 – 98.7)	190 (96.0%; 95% CI 93.2 – 98.7)	168 (96.6%; 95% CI 93.8 – 99.3)	0.9

N, number; EDSS, Expanded Disability Status Scale; MSSS, Multiple Sclerosis Severity Score; RR, relapsing-remitting; CIS, clinically isolated syndrome; SP, secondary progressive; PP, primary progressive; EBV, Epstein-Barr virus; VZV, varicella zoster virus. Means and standard deviations are shown unless otherwise specified.

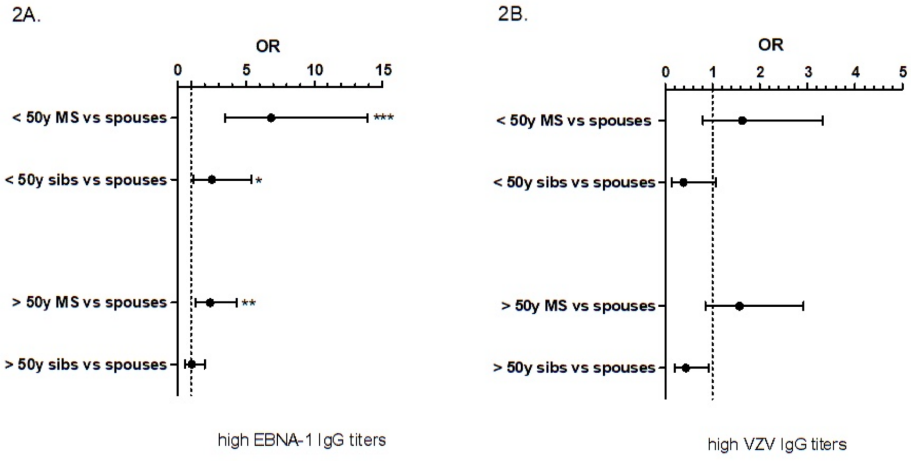


Figure 2. MS patients and their siblings have increased risk of high EBNA-1 titers compared to spouses.

(A) Odds ratio (OR) for high EBNA-1 IgG titers (> 75th percentile) is increased in MS patients in both age categories and siblings younger than 50 years of age compared with spouses, suggesting that MS patients and their siblings are more likely to have higher EBNA-1 IgG titers compared to spouses. (B) No differences in OR for high VZV IgG titers were found between MS patients and siblings in comparison to the spouses. ORs were calculated by means of logistic regression analysis and generalized linear models (GLM) were used for between-group comparisons with Bonferroni corrected and adjusted p-values for gender, disease status and household. The whiskers indicate 95% confidence interval. Dashed line in each graph represents the reference group (spouses).

HLA-DRB1*1501 carriership is associated with high EBNA-1 IgG in older MS patients and their siblings

The risk allele (G allele) frequency of the *HLA-DRB1*1501* tagging SNP (rs9271366) was higher in MS patients (39%; OR = 3.1 95% CI 2.2 – 4.4) and siblings (28%; OR=1.9 95% CI 1.3 – 2.8) when compared to spouses (17%). *HLA-DRB1*1501* was significantly associated with MS (OR=4.2, 95%CI 2.7 – 6.6, adjusted p= $7.4x10^{-10}$). In both MS and non-MS subjects, *HLA-DRB1*1501* was associated with elevated EBNA-1 IgG levels (respectively OR=1.6 95% CI interval 1.0 – 2.6, adjusted p=$4.4x10^{-2}$ and OR=2.2, 95% CI 1.4 – 3.6 adjusted p=$2.0x10^{-3}$). In contrast, VZV IgG titers were not associated with *HLA-DRB1*1501.*

*HLA-DRB1*1501* risk genotype was associated with high EBNA-1 IgG titers in older MS patients and older siblings (Figure 3A).

OR for high EBNA-1 IgG titers was increased in young MS patients versus spouses irrespective of *HLA-DRB1*1501* carriership (Figure 3B). In older MS patients compared to controls, OR for high EBNA-1 IgG titers is only increased in *HLA-DRB1*1501*-positive individuals (Figure 3A and 3B).

In the *HLA-DRB1*1501* negative group (AA genotype), young MS patients had higher EBNA-1 IgG titers than elderly MS patients (OR= 2.7, 95% CI 1.1 – 6.1, adjusted $p=1.4x10^{-2}$). Also, young MS patients had higher EBNA-1 IgG titers than siblings (OR=4.4, 95% CI 1.7 – 11.2, adjusted $p=2.0x10^{-3}$) and spouses (OR=7.4, 95% CI 3.0 – 18.3, adjusted $p=1.4x10^{-5}$).

There was no association between *HLA-DRB1*1501* and high VZV IgG titers (> 75th percentile) in all study groups in both age categories (Figure 3C).

HLA-A*02 is not associated with MS and EBNA-1 IgG titers

The *HLA-A*02* tagging SNP (rs6457110) was not associated with MS in our study (OR = 0.9 95% CI 0.6 – 1.2, p=0.4 adjusted for age, gender, household and *HLA-DRB1*1501*). The *HLA-A*02* tagging SNP was also not associated with EBNA-1 IgG and VZV titers, stratified for all study groups and age categories (data not shown).

Risk for high EBNA-1 IgG titers is increased in MS patients, HLA-DRB1*1501 carriers and young siblings of MS patients

Lastly, we performed a multivariate analysis to assess the individual effect of having MS or being a family member of a MS patient, and having *HLA-DRB1*1501*, stratified for age. In both age categories, having MS is an independent risk factor for high EBNA-1 IgG levels, and this effect is independent of *HLA-DRB1*1501*. Also, being a young family member of a MS patient increases the risk of abundant EBNA-1 IgG (figure 4). These results underscore the importance of the host genetics on the humoral immune response against EBV. Moreover, it indicates the importance of other genetic factors, for example other polymorphisms and epigenetic changes over time, contributing to the humoral immune response against EBV in MS.

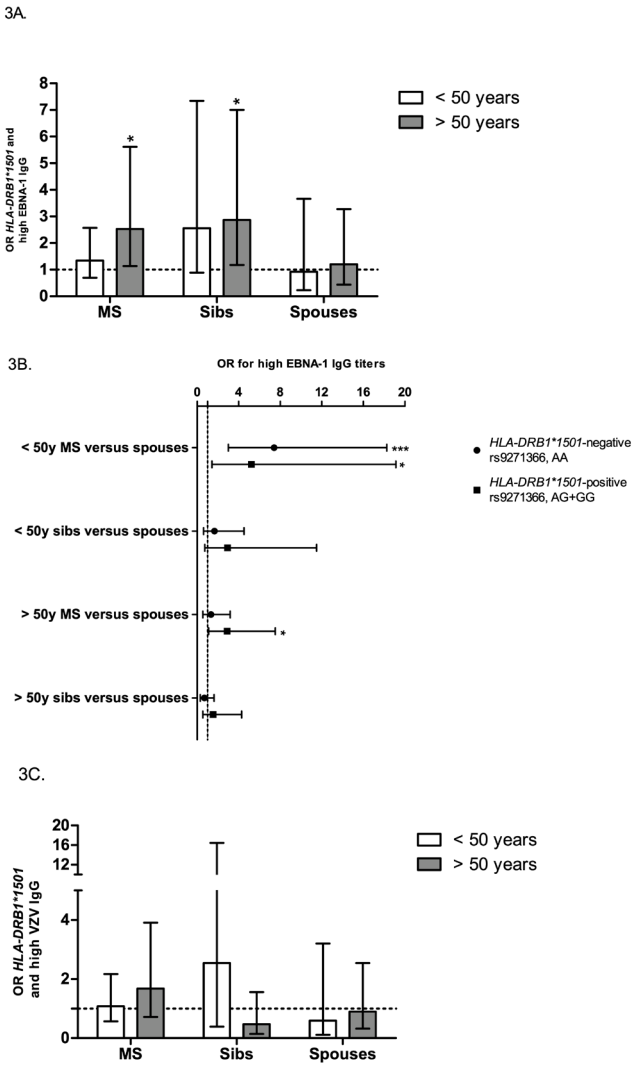


Figure 3. *HLA-DRB1*1501* carriership is associated with high EBNA-1 IgG titers.

*(A) ORs for the association between HLA-DRB1*1501 carriership and high EBNA-1 IgG titers were calculated for MS patients, siblings and spouses in two age categories. HLA-DRB1*1501 carriership is associated with high EBNA-1 IgG titers in old MS patients and siblings. (B) OR for high EBNA-1 IgG titers (> 75th percentile) is increased in young MS patients versus spouses in both HLA-DRB1*1501 positive and negative individuals. In older MS patients compared to controls, OR for high EBNA-1 IgG titers is only increased in HLA-DRB1*1501-positive individuals. (C) HLA-DRB1*1501 is not associated with high VZV IgG titers in all study groups. Logistic regression analyses were performed to calculate ORs and generalized linear models (GLM) were used for between-group comparisons with Bonferroni corrected and adjusted p-values for gender, disease status and household). The whiskers indicate 95% confidence interval.*

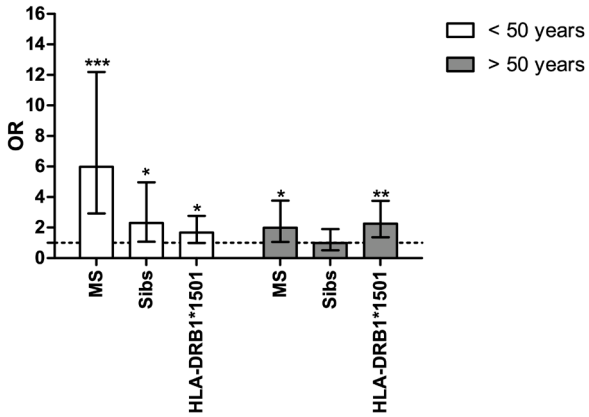


Figure 4. The risk of high EBNA-1 IgG is associated with age, *HLA-DRB1*1501* carriership, having MS or being a family member of an MS patient.

*The risk of high EBNA-1 IgG is associated with age, having MS or being a family member of an MS patient, and HLA-DRB1*1501 carriership. OR, odds ratio.*

DISCUSSION

We showed a gradient in EBNA-1 IgG being the highest in MS patients, intermediate in their siblings and the lowest in spouses. Congruent with our results, Comabella et al.[14] found an increased IgG response to EBNA-1 in MS patients compared to healthy siblings without a biologically unrelated control group. A very small study using a non-linear EBNA-1 IgG quantification method assessing MS patients and their siblings in comparison to unrelated controls, showed results comparable to our study.[17]

The observed EBNA-1 IgG titer gradient in our study suggests a genetic contribution of MS-related SNPs to EBNA-1 IgG titers. Indeed, in previous studies first-degree family members of MS patients showed to have higher genetic load for MS-associated risk SNPs compared to spouses.[4,5] We confirmed that *HLA-DRB1*1501* carriership significantly associates with high EBNA-1 IgG titers. A recent study showed that *HLA-DRB1*1501* carriership also in healthy controls, fully unrelated to MS patients, is associated with enhanced EBNA-1 IgG levels.[23] The high prevalence of *HLA-DRB1*1501* in MS patients, intermediate in siblings and low in spouses could partly explains the gradient observed in EBNA-1 IgG titers in the three groups. After stratification for *HLA-DRB1*1501*, the observed gradient was still present in the young MS patients, siblings and spouses. This

implies that *HLA-DRB1*1501* is not the only player for the generation of the anti EBV immune response in young MS patients. The *HLA-A*02* tagging SNP was not associated with the EBNA-1 IgG titers. Likely, other genetic factors besides *HLA-DRB1*1501* contribute to increased EBNA-1 IgG response. [24,25] Next to shared genetics, shared environment between MS patients and their siblings in early life, when EBV infection typically occurs, might partly be responsible for the found EBNA-1 IgG gradient.

Moreover, we found that high EBNA-1 titers were associated with low age. When stratified into study groups, EBNA-1 IgG was higher in young MS patients compared to elderly MS patients. This could be due to a particular phenomenon called immunosenescence, i.e. a gradual negative dysregulation of the immune system in the elderly.[26] There is overwhelming evidence that the amount of antibodies induced after an immunization response is strongly reduced with physiological ageing, and that titers decline occur more rapidly in elderly. [27,28] In our study we found that particularly in elderly patients and their elderly siblings, high EBNA-1 IgG titers were associated with the presence of the *HLA-DRB1*1501*. This may suggest that a certain genetic make-up is needed to react adequately to the EBV infection in elderly. Indeed, *HLA-DRB1*1501* is strongly correlated with EBNA-1 IgG.[29]

There was a limitation to our study. Due to a moderate sample size, we limited ourselves to testing only *HLA-DRB1*1501* and *HLA-A*02* loci in the genetic association analyses in relation to viral titers. Because of a suggestive genetic contribution on the EBNA-1 IgG response, it would be interesting to assess how non-HLA MS risk SNPs are involved in this process.

In summary, our study showed that EBNA-1 IgG titers were highest in MS patients, intermediate in siblings and low in spouses which suggests a strong genetic contribution on the EBNA-1 response that is partially associated with *HLA-DRB1*1501*. Correcting for *HLA-DRB1*1501* did not abrogate this association which suggests that additional genetic factors may contribute to this gradient. Further large genetic studies assessing MS genetics and EBNA-1 IgG responses are needed to determine which MS risk SNP are contributing to the enhanced EBNA-1 IgG response in MS. Also, EBNA-1 IgG titers were higher in young MS patients compared to elderly MS patients. In order to study the complex interaction between environmental and genetic factors, age needs to be considered. Potentially, this relates to the predominant relapsing-remitting disease progression in young, compared to the more progressive disease in older MS patients.

REFERENCES

1. Patsopoulos, N. A. *et al.* Fine-Mapping the Genetic Association of the Major Histocompatibility Complex in Multiple Sclerosis: HLA and Non-HLA Effects. PLoS Genet. 2013;9:e1003926
2. Fogdell-Hahn, A. *et al*. Multiple sclerosis: A modifying influence of HLA class I genes in an HLA class II associated autoimmune disease. Tissue Antigens. 2000;55:140-8.
3. Patsopoulos, N. A. *et al.* Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019;365:eaav7188.
4. Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann. Neurol. 2011;69:65-74.
5. Mescheriakova, J. Y. *et al.* Burden of genetic risk variants in multiple sclerosis families in the Netherlands. Mult Scler J Exp Transl Clin. 2016; 6;2:2055217316648721.
6. Belbasis, L. *et al.* Environmental risk factors and multiple sclerosis: An umbrella review of systematic reviews and meta-analyses. Lancet Neurol. 2015;14:263-73.
7. Fagnani, C. *et al.* Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. Mult Scler. 2015;21:1404-13.
8. Thacker, E. L. *et al*. Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis. Annals of Neurology. 2006;59:499-503.
9. Ascherio, A. *et al.* Epstein-Barr virus antibodies and risk of multiple sclerosis: A prospective study. J Am Med Assoc. 2001;286:3083-8.
10. Rubicz, R. *et al.* A Genome-Wide Integrative Genomic Study Localizes Genetic Factors Influencing Antibodies against Epstein-Barr Virus Nuclear Antigen 1 (EBNA-1). PLoS Genet. 2013;9:e1003147.
11. Rubicz, R. *et al.* Genetic factors influence serological measures of common infections. Hum Hered. 2011;72:133-41.
12. Sallah, N. *et al.* Whole-genome association study of antibody response to Epstein-Barr virus in an African population: a pilot. Glob Heal Epidemiol Genomics. 2017;2:e18.
13. Bergkvist, M. & Sandberg-Wollheim, M. Serological differences in monozygotic twin pairs discordant for multiple sclerosis. Acta Neurol Scand. 2001;104:262–5.
14. Comabella, M. *et al.* Antiviral immune response in patients with multiple sclerosis and healthy siblings. Mult Scler. 2010;16:355-8.
15. Kinnunen, E. *et al.* Viral Antibodies in Multiple Sclerosis: A Nationwide Co-Twin Study. Arch Neurol. 1990;47:743-6.
16. Ristori, G. *et al.* Antiviral immune response in patients with multiple sclerosis, healthy siblings and twins. Mult Scler. 2010;16:1527-8.
17. Sumaya, C. V. *et al.* Increased prevalence and titer of Epstein-Barr virus antibodies in patients with multiple sclerosis. Ann Neurol. 1985;17:371-7.
18. Burgoon, M. P. *et al.* Varicella zoster virus is not a disease-relevant antigen in multiple sclerosis. Ann Neurol. 2009; 65:474-9.
19. Hoppenbrouwers, I. A. *et al.* Familial clustering of multiple sclerosis in a Dutch genetic isolate. Mult Scler. 2007;13:17-24.
20. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Ann Neurol. 2001;50:121-7.
21. Poser, C. M. *et al.* New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. Ann Neurol. 1983;13:227-31.
22. Miller, S. A. *et al.* A simple salting out procedure for extracting DNA from

human nucleated cells. Nucleic Acids Res. 1988;16:1215.

23. Hedström, A. K. *et al.* High Levels of Epstein–Barr Virus nuclear antigen-1-specific antibodies and infectious mononucleosis act both independently and synergistically to increase multiple sclerosis risk. Front Neurol. 2020; 10:1368.

24. Kreft, K. L. *et al.* Elevated EBNA-1 IgG in MS is associated with genetic MS risk variants. Neurol. Neuroimmunol. NeuroInflammation. 2017;4:e406.

25. Zhou, Y. *et al.* Genetic loci for Epstein-Barr virus nuclear antigen-1 are associated with risk of multiple sclerosis. Mult Scler. 2016;22:1655-1664.

26. DeWitt, J. C. & Luebke, R. W. Immunological Aging. In *Comprehensive* Toxicology: Third Edition. 2018.

27. Weinberger, B. *et al.* Biology of immune responses to vaccines in elderly persons. Clin Infect Dis. 2008;46:1078-84.

28. Frasca, D. & Blomberg, B. B. Effects of aging on B cell function. Cur Opin Immunol. 2009;21:425-30.

29. Simon, K. C. *et al.* Combined effects of smoking, anti-EBNA antibodies, and HLA-DRB1*1501 on multiple sclerosis risk. Neurology. 2010;74:1365-71.

SUPPLEMENTAL MATERIAL

Supplemental table 1.

SNPs included in the study.

Chr	Gene	SNP	Risk Allele for MS	OR from original study	Ref	HWE
6	*HLA-DRB1*1501*	rs9271366	G	3.1	[1]	NS
6	*HLA-A*02*	rs6457110	T	1.47	[2]	NS

Chr, chromosome; SNP, single nucleotide polymorphism; OR, odds ratio; Ref, reference; HWE, Hardy-Weinberg equilibrium.

1. de Bakker, P. I. *et al*. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet. 2006;38:1166-72.
2. Jafari, N. *et al.* Infectious mononucleosis-linked HLA class I single nucleotide polymorphism is associated with multiple sclerosis. Mult Scler. 2010;16:1303-7.

Supplemental table 2.

Pooled homozygous SNP carriers with heterozygous SNP carriers as indicated in the table. This was done in all study groups (MS, sibs and spouses) in order to improve power of calculations in logistic regression analysis.
ORs of tagging SNPs (HLA-DBR1*1501 and HLA-A*02) were calculated in relation to high EBNA-1 IgG in healthy controls using a univariate logistic regression analysis adjusted for gender, age, and household. We did this for the both cohorts together (ORall) and separately (OR(gems) and OR(erf)).

SNPs		ORall	OR(gems)	OR(erf)
Rs9271366 *(HLA-DRB1*1501)*	AA | AG+ GG[1]	1.2 [0.5-2.8] p=0.6	1.6 [0.6-4.4] p=0.3	0.7 [0.2-3.1] p=0.7
Rs6457110 *(HLA-A*02)*	TT | AA+AT	0.8 [0.4-1.7] p=0.6	0.9 [0.4-2.3] p=0.9	0.7 [0.2-2.5] p=0.6

[1] *The risk allele has been underlined.*

PART V

MS AND OTHER AUTOIMMUNE DISEASES IN MS FAMILIES

Chapter 6

NO EXCESS OF AUTOIMMUNE DISEASES IN MULTIPLE SCLEROSIS FAMILIES FROM THE NETHERLANDS

J.Y. Mescheriakova and R.Q. Hintzen.

Acta Neurologica Scandinavica, 2018

ABSTRACT

Objectives

Autoimmune diseases (AIDs) cluster in families, however to what extent AIDs co-occur in MS multiplex families with two or more affected individuals is still controversial. The study aimed to evaluate co-existing AIDs in this type of families from the Netherlands.

Materials and Methods

155 MS multiplex families (155 MS probands, 959 first-degree relatives and 212 spouses) were characterized for a history of 11 AIDs by means of a self-administered questionnaire.

Results

In 43.2% of MS multiplex families at least one AID was present in the first-degree relatives. Overall frequency of AIDs was not significantly different between MS patients (11%), their first-degree family members (11%) and controls (5.2%). After correction for age at inclusion and gender, the odds ratios (OR) for AIDs were not significant for MS patients (OR=1.8 [0.77-4.34], p=0.17) and first-degree family members (OR=2.0 [0.98-4.10], p=0.06) when both compared to spouses. The frequency of AIDs in mothers did not differ from that in fathers after correction for sex bias (19% vs 8%, p=0.51). A presence of AID was more often reported in maternal than paternal second-degree relatives (23% versus 10%, p=0.0020).

Conclusion

Although nearly half of the Dutch MS multiplex families reported an AID, no excess of AIDs was present in MS patients from multiplex families or their first-degree family members compared to the spouses.

INTRODUCTION

Autoimmune diseases (AIDs) as a group affect approximately 8.5% of individuals world-wide.[1] Although AIDs include a broad range of phenotypic manifestations and severity, several features such as female predominance, production of antibodies and clustering of several AIDs in families, suggest that they share common etiological and genetic factors.

Multiple sclerosis (MS) is a complex inflammatory disease that affects the central nervous system (CNS), including the brain, the spinal cord and optic nerves. MS is considered an autoimmune disease because of the involvement of the T and B cells in its pathophysiology, and in the Northern Europeans MS has been associated with an extended MHC haplotype containing the *HLA-DRB1*1501* locus.[2] Network-based analysis of genome-wide association studies has shown a significant sharing of genetic loci between MS and other autoimmune diseases such as type 1 diabetes, rheumatoid arthritis, and Crohn's disease.[3] With a cross-phenotype meta-analysis Cotsapas et al. have shown that nearly 44% of immune disease-risk single nucleotide polymorphisms (SNPs) are associated to multiple, but not all AIDs.[4] Co-occurrence of several AIDs in one patient is therefore not surprising. Some studies suggest that first-degree relatives of probands with MS could be at a greater risk of autoimmune diseases other than MS[5-10] while some studies do not support these findings.[11,12] To what extent MS and other AIDs co-occur in MS multiplex families with two or more affected individuals with MS, also remain controversial.[5,10,11,13]

The aim of our study was to describe co-existing autoimmune diseases in MS patients and their first-degree relatives from the Dutch MS multiplex families, and compare them to spousal controls.

METHODS

The dataset consisted of 175 multiplex families of Northern-European ancestry, which were part of the study on gene environment interaction in MS (GEMS) conducted at the MS centre ErasMS of the Erasmus Medical Centre (EMC) in the Netherlands. This study started in 2002 and is still ongoing. Multiplex MS families are included in this study, in which at least one first- or second-degree relative of an affected proband also has

clinically characterized MS. The diagnosis of MS in all patients was evaluated according to standard diagnostic criteria.[14,15] The following clinical information was also collected: gender, age at disease onset, disease duration, clinical course, Expanded Disability Status Scale score (EDSS) and MS severity score (MSSS). Written informed consent was obtained from all participants with approval from the medical ethics committee of the EMC.

A personally administered questionnaire was used to collect epidemiological and environmental data about MS patients, their spouses and their families. All MS patients participating in GEMS filled in the questionnaire that was sent by mail. The questionnaire required information about family structure, family history of MS and information about autoimmune disorders in the index cases, their first- and second-degree family members, and the spouses. Information about the offspring was only requested if these individuals were above 18 years of age and informed consent could be obtained. The rate of AIDs in childhood is relatively low so this group would be less informative. We included the most common AIDs in the questionnaire, such as autoimmune thyroid disease (AITD) (Hashimoto's thyroiditis and Grave's disease), rheumatoid arthritis (RA), ulcerative colitis (UC), Crohn's disease (CD), psoriasis, pernicious anaemia (PA), type 1 diabetes (T1D), vitiligo, Sjögren's syndrome (SS) and systemic lupus erythematosus (SLE). We separately included sarcoidosis. To differentiate between type 1 and type 2 diabetes, age of onset (< 30 years) was added to the description. Patients were instructed to ask their relatives directly about these conditions and to fill in the information from memory on deceased relatives. Participants were invited to EMC to discuss the questionnaire. Investigators further clarified self-reported diagnoses by asking specific questions about each reported disease, including age of onset, whether specific lab tests were obtained, and whether specific medication was used. In case of doubt about a particular diagnosis, the reply was assumed to be negative to prevent false-positive diagnoses.

Statistical analysis

The analysis of categorical values was conducted by chi-square test, and odds ratios (OR) and corresponding 95% confidence intervals (CI) were computed. Continuous variables were analysed by non-parametric Mann-Whitney U-test. Two-sided p-values less than 0.05 were considered significant. We applied correction for multiple testing using the Benjamini-Hochberg procedure for controlling the false-positive rate. The analyses were performed utilizing SPSS software (IBM Company, version 21).

RESULTS

Description of the nuclear families

One hundred and seventy-five families were included in the study. Twenty families were excluded because of incomplete information about the first-degree relatives or because they declined further participation. The included 155 families were multiplex MS families by nature of this study, and 79% of MS index cases had a first-degree family member with MS. Of 155 families there were 72 families (46.5%) with more than 2 affected family members with MS (and a maximum of 9 affected individuals). The affected individuals were not restricted to one nuclear family. For the purposes of this study and to reduce recall-bias only one nuclear family per pedigree was selected. Data have been collected for 155 MS probands, 959 first-degree relatives (170 with MS and 789 without MS) including 453 siblings, 196 offspring, 310 parents, and 212 spouses of MS patients as controls. The clinical and demographic data are shown in Table 1**.**

Table 1. Clinical and demographic characteristics of the study participants.

	MS probands	First-degree relatives	First-degree affected relatives	First- degree unaffected relatives	Spousal controls
N	155	959	170	789	212
Age (years)	47.1 (12.5)	55.1 (23.2)	52.1 (17.6)	55.8 (24.2)	49.8 (12.5)
% Women	71.6%	50.78%	67.0%	47.3%	33.0%
Sex ratio (women:men)	2.5:1	1.03:1	2.0:1	0.9:1	0.49:1
Age at onset	33.5(9.99)	-	33.56 (10.6)	-	-
EDSS (median)	3.5	-	4.0	-	-
MSSS	4.6 (2.7)	-	3.7 (3.5)	-	-
Disease duration (years)	14.4 (10.5)	-	10.9 (12.3)	-	-
Proportion with RR-onset	86.5%	-	76.8%	-	-

N= number; EDSS= Expanded Disability Status Scale; MSSS= Multiple Sclerosis Severity Score; RR: relapsing-remitting. Means (SD) are shown unless otherwise specified.

Clinical characteristics of MS patients and first-degree family members

First-degree relatives were slightly older than MS probands and spousal controls (both $p < 0.01$). MS probands had the highest proportion of women, compared to their first-degree relatives and spouses (both $p<0.001$). There were no differences between MS probands and their first-degree affected family members in terms of age and gender distribution, age at disease onset and EDSS score. Although, MS probands had a longer disease duration ($p<0.001$) and a significantly higher MSSS ($p=0.004$) than their first-degree affected family members.

Autoimmune diseases in multiplex families

AID was reported in 17 of 155 MS probands (11%): 16 cases reported 1 AID and one case (0.6%) reported two AIDs (CD and RA). When all MS patients (MS probands and their first-degree affected family members; n=325) were inspected, similar results could be found (AIDs were reported in 11.4% of all MS patients).

Figure 1 shows the frequencies of each individual AID in all (n=325) MS patients from the nuclear families. Results for 155 MS probands were similar. The most common diseases present in probands and all MS cases were respectively AITD (2.6% and 3.7%), psoriasis (1.9% and 3.1%), RA (2.6% and 2.2%), and PA (1.9% and 1.2%).

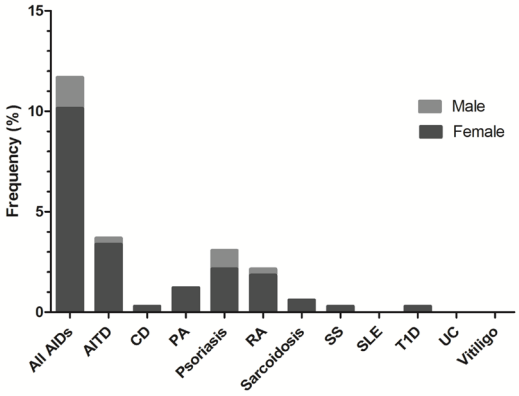


Figure 1. Frequency of autoimmune diseases in all MS patients from the nuclear families stratified by sex.

AIDs= autoimmune diseases; AITD = autoimmune thyroid disease; CD = Crohn's disease; PA = pernicious anaemia; RA = rheumatoid arthritis; SS = Sjögren's syndrome; SLE = systemic lupus erythematosus; T1D = type 1 diabetes; UC = ulcerative colitis

There were no differences in clinical characteristics such as age at disease onset, EDSS, MSSS, disease duration, and MS type between MS patients with and without a co-existing autoimmune disease. We observed that MS patients who reported an AID were more often women compared to MS patients without an AID (86.1% vs 67.1%, p=0.02).

In 67 of 155 MS multiplex families (43.2%) one or more AIDs other than MS were present in the first-degree relatives. There were 45 families with one AID, 19 families with two AIDs, and three families with three AIDs in the first-degree relatives. About 65 of 93 (69.9%) first-degree relatives with an AID were females.

Figure 2 and table 2 show frequencies of autoimmune diseases in MS probands, their first-degree relatives and spousal controls. There were no differences in the distribution of AIDs in MS probands (11%), their first-degree family members (11%) compared to spousal controls (5.2%) (respectively p=0.039, and p=0.012, but not significant after correction according to Benjamini-Hochberg procedure). Stratification on an individual AID, and gender did not harbour any significant result.

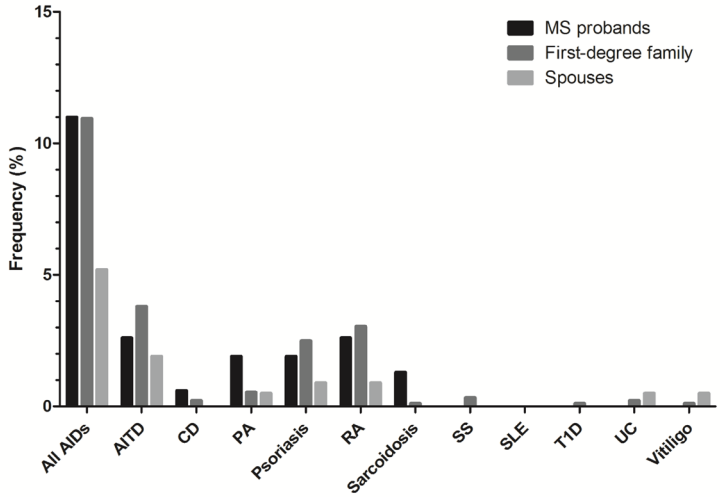


Figure 2. Frequencies of autoimmune diseases in MS probands, their first-degree family members and spouse controls.

AIDs= autoimmune diseases; AITD = autoimmune thyroid disease; CD = Crohn's disease; PA = pernicious anaemia; RA = rheumatoid arthritis; SS = Sjögren's syndrome; SLE = systemic lupus erythematosus; T1D = type 1 diabetes; UC = ulcerative colitis.

Table 2. Frequencies of autoimmune diseases in MS probands, their first-degree family members and spousal controls.

Autoimmune disease	MS probands (n=155)	First-degree relatives (n=922)	Spousal controls (n=212)	p-value Probands vs Spouses	p-value First-degree relatives vs Spouses	p-value Probands vs First-degree relatives
All AIDs	17(11%)	101(11%)	11 (5.2%)	0.039; NS after BH correction	0.012; NS after BH correction	0.98
AITD	4 (2.6%)	35 (3.8%)	4 (1.9%)	0.73	0.21	0.64
CD	1 (0.6%)	2 (0.2%)	0	0.42	0.50	0.37
PA	3 (1.9%)	5 (0.5%)	1 (0.5%)	0.18	0.90	0.09
Psoriasis	3 (1.9%)	23 (2.5%)	2 (0.9%)	0.65	0.20	0.68
RA	4 (2.6%)	28 (3.0%)	2 (0.9%)	0.25	0.10	0.77
Sarcoïdosis	2 (1.3%)	1 (0.1%)	0	0.18	0.63	0.06
SS	0	3 (0.3%)	0	-	0.41	0.48
SLE	0	0	0	-	-	-
T1D	0	1 (0.1%)	0	-	0.63	0.68
UC	0	2 (0.2%)	1 (0.5%)	0.39	0.46	0.56
Vitiligo	0	1 (0.1%)	1 (0.5%)	0.39	0.34	0.68

AIDs= autoimmune diseases; AIDs= autoimmune diseases; AITD = autoimmune thyroid disease; CD = Crohn's disease; PA = pernicious anaemia; RA = rheumatoid arthritis; SS = Sjögren's syndrome; SLE = systemic lupus erythematosus; T1D = type 1 diabetes; UC = ulcerative colitis; NS = not significant; BH correction = Benjamini-Hochberg correction.

We compared the frequency of AIDs in the spouse controls to the population prevalence of AIDs in the Netherlands. The overall prevalence of the AIDs was comparable between the general Dutch population (3.6 – 6.9 %, depending on whether the upper or the lower limit of the prevalence of SS is considered), [16-25] and our control group (5.2%).

The odds ratio (OR) for autoimmune diseases was 2.6 (95% CI [1.12 – 6.04], p= 0.027) for MS probands and 2.5 (95% CI [1.25 – 5.08], p= 0.01) for first degree relatives compared to the spousal controls. After correction for age at inclusion and gender the ORs were no longer significant (OR =1.8 for MS probands versus spouse controls (95% CI [0.77-4.34], p=0.17), and OR = 2.0 for first-degree family members versus spouse controls (95% CI [0.98-4.10], p = 0.06)). The presence of an autoimmune disease was influenced by gender (OR=2.9 [1.93-4.49], $p< 0.1x10^{-3}$) and not by the MS affection status (OR=1.35, 95% CI [0.89-2.04], p=0.16). This was especially the case in AITD (OR=4.6, 95% CI [2.11-9.94], $p=1.23x10^{-4}$) and RA (OR=2.85, 95% CI [1.32-6.16], $p=0.8x10^{-2}$).

Autoimmune diseases in maternal and paternal families

The frequency of MS did not differ between mothers and fathers of MS probands (respectively 16% vs 10%, p= 0.80 after correction for sex bias in MS). The frequency of other autoimmune diseases also did not differ between mothers and fathers of MS probands (respectively 19% vs 8%, sex bias corrected p-value =0.51). Moreover, we counted the number of maternal and paternal nuclear families (by looking at parents and sibs of mothers and fathers of MS index cases) that were positive for a history of an AID. MS was equally reported in maternal and paternal relatives (11% vs 10%, p=0.85). For mothers there were 35 nuclear families with a positive history of AID other than MS (22.6%) and for fathers there were 15 nuclear families with a positive history of AID other than MS (9.7%, p=0.002). The gender ratio (women:men) for persons who reported an AID in these families was 6:1 in maternal and 2:1 paternal families. For both maternal and paternal families, RA was reported in approximately 50% of AID cases, AITD in 30%, and psoriasis in 15% of maternal and 8% of paternal AID cases. The rest group consisted of SLE (n=1), SS (n=1) and T1D (n=1). The sibs:parents ratio of cases with an AID was 40:60 for maternal and 50:50 for paternal families.

DISCUSSION

The purpose of our study was to describe commonly studied autoimmune diseases in multiplex MS families. We did not find an excess of auto-immunity in the MS patients. These results accord with some,[5,6,11]but not all studies.[7,9,12] About 43% (67 of 155) of the Dutch MS multiplex families reported one or more auto-immune diseases in the first-degree relatives. This number lies in the range of the reported frequencies (9-64%)

for AIDs in the first-degree relatives worldwide,[5,7,13,26] being closer to the frequencies reported for MS multiplex families in particular (27-64%).[5,13]

We did not find an increased risk of auto-immune diseases in the first-degree relatives of MS probands, which was in line with previous studies.[11,12] In contrast, some other studies found an increased risk of auto-immunity in the first-degree relatives of the MS patients.[5-7,9,10] We had more than 95% power to detect a difference between first-degree relatives and controls (odds ratio of 1.2 and higher at the α –level of 0.05). Perhaps, the largest difference between the studies that found this association compared to studies that did not, was the control group. In our study as well as in the largest population-based cohort of more than 30.000 Canadians with MS,[11] spouses were used as controls, while studies that found positive association with AIDs mostly used population controls. Spouse controls are more likely than population controls to match for potentially confounding factors such as ethnic origin, age, socioeconomic background, education level and environmental factors.[11] However, if spouses choose their partner on the basis of one or more specific traits, the familial clustering of diseases can be underestimated.[27] Furthermore, by study design our control group consisted of more males than females compared to MS probands and their first-degree relatives. As it has been shown earlier that males tend to underreport diseases more often than females,[11] there is a potential risk of skew. However, our results are well in line with the overall reported frequency of autoimmune diseases in the Dutch population. There is a number of factors that could potentially cause conflicting results between our study and previously published reports, and these are: 1) differences in the methodology such as choosing a community dwelling versus a hospital-based study population, and hereby introducing a selection bias; 2) use of questionnaires with possible recall bias versus a register-based cohort design where the information can be misclassified or incomplete; 3) choice of a control group (spouses versus population controls as previously explained); 4) number and type of AIDs studied with a different prevalence per AID and a variation in cumulative prevalence.

HLA locus is the major genetic locus for the most autoimmune diseases. Table 3 shows the HLA loci associated with MS susceptibility that also play a role in the susceptibility of the studied AIDs. [7,28-37]

Table 3. HLA associations of MS with the studied autoimmune diseases.

Studied AIDs	HLA loci associated with MS susceptibility[2] that are also associated with studied AIDs	References
AITD	DRB1*0301 DRB1*04	28,29
CD	DRB1*04 DRB1*1401	30,31
PA	DRB1*15 DRB1*04	32
Psoriasis	HLA-DQA1	33
RA	DRB1*0401 DRB1*0404	31
Sarcoidosis	DRB1*1501 *DRB1*1401* *DRB1*0401 (protective in sarcoidosis)* *DQB1*0602 (protective in sarcoidosis)*	34
SS	**DRB1*0301**	35
SLE	DRB1*1501 **DRB1*0301** *DQB1*0602*	36
T1D	**DRB1*0301** DRB1*0401 DRB1*0404 *DRB1*1501 (protective in T1D)* DQB1*0602 *(protective in T1D)* DQA1*0102 *(protective in T1D)*	7,31
UC	**No evident HLA loci shared with MS**	31
Vitiligo	HLA-DQA1 HLA-A*0201 (protective in MS)	37

AIDs= autoimmune diseases; AITD = autoimmune thyroid disease; CD = Crohn's disease; PA = pernicious anaemia; RA = rheumatoid arthritis; SS = Sjögren's syndrome; SLE = systemic lupus erythematosus; T1D = type 1 diabetes; UC = ulcerative colitis

Next to the known HLA genes, genome-wide association (GWA) studies have identified hundreds of genetic loci outside the HLA region for a variety of AIDs. Some of these loci are also being shared across these diseases[4] and are involved in the certain biological pathways.[38] Studies have shown that several genetic loci can have opposite effects in different AIDs.[39] For studies like our study, it is thus very essential to look at what kind of AIDs are studied in relation to MS, because some AIDs will and some will not cluster

with each other. Moreover, whether a certain genetic profile will lead to an expression of one or more expected phenotypes in a patient and in his/her related family members (where sharing of the same genetic pool is assumed), depends also on an interplay with certain environmental and other genetics factors.[39] These interactions vary per family and per population studied.

In our study the presence of autoimmune diseases was significantly influenced by gender. This is not surprising because most autoimmune diseases occur significantly more often in women than men. The reasons for gender differences in certain autoimmune diseases remain unknown, but have been attributed to sex hormone influence, fetal microchimerism, X chromosome inactivation, and X chromosome abnormalities.[40]
Although mothers of MS probands did not differ from fathers in reported AID frequency after correction for sex bias, maternal families reported more often an AID than paternal, and in maternal families the proportion of women who reported an AID was higher compared to paternal families. Of note, comparable results were reported by Ketelslegers et al. for Dutch children with acquired demyelinating syndrome where increased frequency of AIDs was also observed in maternal families.[41] A maternal transmission of autoimmunity has been proposed for some autoimmune diseases such as juvenile idiopathic arthritis,[42] SS[43]and T1D.[44] The question whether a maternal parent-of-origin effect is present here, cannot be answered because the study has not been designed and powered to appropriately address this issue.

There were several limitations to this study. First, our study is a questionnaire based medium sized study and it is not free from potential recall-bias. Second, women are more aware of the diseases in their family than men[11] leading to a possible under-reporting of diseases in male subjects and their families. And third, perhaps the most important limitation is the absence of stringent medical confirmation of reported diagnoses.

In summary, our study did not find evidence for an association of MS with other autoimmune diseases in MS multiplex families. From GWA studies in autoimmune diseases it became clear that some but not all genetic risk loci (majority) are shared across various autoimmune diseases, and that risk alleles in one autoimmune disease can be protective for an another autoimmune disease.[39]

REFERENCES

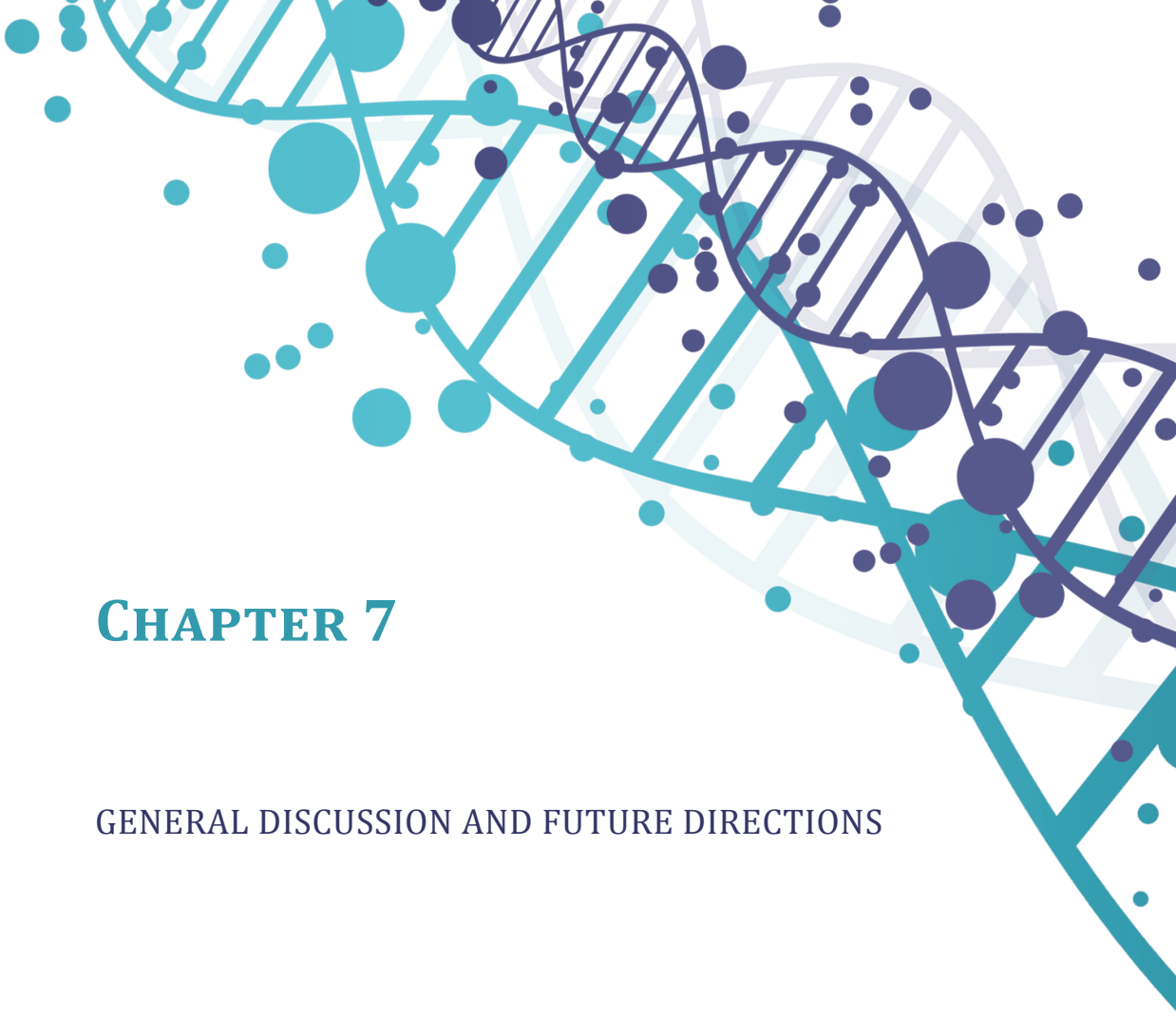
1. Cooper, G. S. *et al*. Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. J Autoimmun. 2009;33:197-207.
2. Patsopoulos, N. A. *et al.* Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013;9:e1003926.
3. Baranzini, S. E. The genetics of autoimmune diseases: a networked perspective. Curr Opin Immunol. 2009;21:596-605.
4. Cotsapas, C. *et al.* Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 2011;7:e1002254.
5. Broadley, S. A. *et al.* Autoimmune disease in first-degree relatives of patients with multiple sclerosis. A UK survey. Brain. 2000;123:1102-1111.
6. Henderson, R.D. *et al.* The occurrence of autoimmune diseases in patients with multiple sclerosis and their families. J Clin Neurosci. 2000;7;434-437.
7. Laroni, A. *et al.* Multiple sclerosis and autoimmune diseases: epidemiology and HLA-DR association in North-east Italy. J Neurol. 2006;253:636-639.
8. Criswell, L. A. *et al.* Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. Am J Hum Genet. 2005;76:561-571.
9. Nielsen, N. M. *et al.* Autoimmune diseases in patients with multiple sclerosis and their first-degree relatives: a nationwide cohort study in Denmark. Mult Scler. 2008;14:823-829.
10. Deretzi, G. *et al.* Familial prevalence of autoimmune disorders in multiple sclerosis in Northern Greece. Mult Scler. 2010;16:1091-1101.
11. Ramagopalan, S. V. *et al.* Autoimmune disease in families with multiple sclerosis: a population-based study. Lancet Neurol. 2007;6:604-610.
12. Midgard, R. *et al.* Multiple sclerosis and chronic inflammatory diseases. A case-control study. Acta Neurol Scand. 1996;93:322-328.
13. Barcellos, L. F. *et al.* Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. Lancet Neurol. 2006;5:924-931.
14. Poser, C. M. *et al.* New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol. 1983;13:227-231.
15. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50:121-127.
16. van der Linden, M. *et al.* Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk: klachten en aandoeningen in de bevolking en in de huisartspraktijk. Utrecht, Bilthoven: NIVEL, Rijksinstituut voor Volksgezondheid en Milieu; 2004.
17. de Groof, E. J. *et al.* Burden of disease and increasing prevalence of inflammatory bowel disease in a population-based cohort in the Netherlands. Eur J Gastroenterol Hepatol. 2016;28:1065-1072.
18. Andres, E. *&* Serraj, K. Optimal management of pernicious anemia. J Blood Med. 2012;3:97-103.
19. Van Peet, P. G. *et al.* NHG-Standaard Psoriasis(derde herziening). Huisarts Wet 2014;57:128-135.
20. Sloot, R. *et al.* Reumatische aandoeningen in Nederland: Ervaringen en kerngetallen. In: NIVEL, ed. Utrecht 2016.
21. Drent, M. *et al.* Richtlijn sarcoïdose en infliximab (Remicade). Maastricht: ILD

care (Maastricht) en Centrum interstitiële longziekten (Nieuwegein). Augustus 2007.

22. Tomsic, M. *et al.* Prevalence of Sjogren's syndrome in Slovenia. Rheumatology. 1999;38:164-170.
23. Danchenko, N. *et al.* Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. Lupus. 2006;15:308-318.
24. Spaans, E. A. *et al.* The incidence of type 1 diabetes is still increasing in the Netherlands, but has stabilised in children under five (Young DUDEs-1). Acta Paediatr. 2015;104:626-629.
25. Zhang, Y. *et al.* The Prevalence of Vitiligo: A Meta-Analysis. PLoS One. 2016;11:e0163806.
26. Heinzlef, O. *et al.* Autoimmune diseases in families of French patients with multiple sclerosis. Acta Neurol Scand. 2000;101:36-40.
27. Marrie, R. A. Autoimmune disease and multiple sclerosis: methods, methods, methods. Lancet Neurol. 2007;6:575-576.
28. Gough, S. C. *et al*. The HLA region and autoimmune disease: Associations and mechanisms of action. Curr Genomics. 2007;8:453-465.
29. Petrone, A. *et al.* Association of DRB1*04-DQB1*0301 haplotype and lack of association of two polymorphic sites at CTLA-4 gene with Hashimoto's thyroiditis in an Italian population. Thyroid. 2001;11:171-175.
30. Stokkers, P. C. *et al.* HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis. Gut. 1999;45:395-401.
31. Fernando, M. M. *et al.* Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genet. 2008;4:e1000024.
32. Ungar, B. *et al.* HLA-DR patterns in pernicious anaemia. Br Med J. 1981;282:768-770.
33. Prinz JC. Autoimmune aspects of psoriasis: Heritability and autoantigens. Autoimmun Rev. 2017;16:970-979.
34. Rossman, M. D. *et al.* HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. Am J Hum Genet. 2003;73:720-735.
35. Teos, L. Y. *et al.* Genetics of Sjogren's syndrome. Clin Immunol. 2017;182:41-47.
36. Graham, R. R. *et al.* Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. Am J Hum Genet. 2002;71:543-553.
37. Cavalli, G. *et al.* MHC class II super-enhancer increases surface expression of HLA-DR and HLA-DQ and affects cytokine production in autoimmune vitiligo. Proc Natl Acad Sci U S A. 2016;113:1363-1368.
38. Cotsapas, C. *et al.* Immune-mediated disease genetics: the shared basis of pathogenesis. Trends Immunol. 2013;34:22-26.
39. Sirota, M. *et al.* Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet. 2009;5:e1000792.
40. Ngo, S. T. *et al*. Gender differences in autoimmune disease. Front Neuroendocrinol. 2014;35:347-369.
41. Ketelslegers, I. A. *et al.* Incidence of acquired demyelinating syndromes of the CNS in Dutch children: a nationwide study. J Neurol. 2012;259:1929-1935.
42. Zeft, A. *et al.* Familial autoimmunity: maternal parent-of-origin effect in juvenile idiopathic arthritis. Clin Rheumatol. 2008;27:241-244.
43. Anaya, J. M. *et al.* Autoimmune disease aggregation in families with primary Sjogren's syndrome. J Rheumatol. 2006;33:2227-2234.
44. Samuelsson, U. *et al.* A fourfold difference in the incidence of type 1 diabetes between Sweden and Lithuania but similar prevalence of autoimmunity. Diabetes Res Clin Pract. 2004;66:173-181.

PART VI

GENERAL DISCUSSION



Chapter 7

GENERAL DISCUSSION AND FUTURE DIRECTIONS

Multiple sclerosis (MS) is a complex inflammatory and degenerative disease of the central nervous system. MS clusters in families, and first-degree relatives have a higher risk of developing MS themselves compared to general population. Studying MS families helps to understand the role of established and new risk factors (environmental and genetic) in disease pathogenesis, susceptibility and its trajectory. This chapter summarizes the main findings of the studies presented in this thesis and places them in a broader perspective. Finally, the directions for further research are presented and discussed.

MAIN FINDINGS

- Cumulative genetic risk score (103 genetic SNPs including the HLA-DRB1*1501) was higher in MS multiplex patients than in sporadic cases and healthy controls indicating a higher genetic burden in the familial cases. The difference in genetic burden was primarily driven by a higher HLA-DRB1*1501 allele frequency in familial cases (**chapter 2**).

- 57 risk loci for adult-onset MS were associated with childhood-onset MS, but not with monophasic acquired demyelinating syndrome in children. The joint effect of 57 risk loci exceeded the effect of the HLA-DRB1*1501 alone in both pediatric- and adult-onset MS models (**chapter 3**).

- By means of whole exome sequencing (WES) and linkage analysis a rare missense variant (rs147213094) in the FKBP6 gene (7q11.22– q11.23) was found in a large Dutch MS family. This SNP segregated with the disease in the family, but was not significantly associated with MS in a larger cohort of unrelated MS patients and controls indicating that rare private genetic variants do not contribute to the main risk of MS in the general population (**chapter 4.1**).

- There was a gradient in EBNA-1 IgG titers, being the highest in MS patients, intermediate in their unaffected siblings and the lowest in unrelated spouses suggesting a strong genetic contribution in the host's immune response to EBV. Indeed, HLA-DRB1*1501 was associated with elevated EBNA-1 IgG titres, but only partially explained the found gradient suggesting that additional genetic factors may also contribute. Additionally, EBNA-1 IgG titers were higher in young MS patients compared to old MS patients, and hereby pointing towards the immune decline with an increasing age (**chapter 5**).

- No excess of autoimmunity was found in MS patients from multiplex families compared to healthy controls. Although nearly 43% of MS multiplex families reported one or more autoimmune diseases in the first-degree relatives, no increased risk of autoimmune diseases was found in the first-degree relatives of MS probands. As some genetic risk loci are predisposing for some autoimmune diseases, and being protective in other autoimmune diseases, it really matters what autoimmune diseases are studied together (**chapter 6**).

COMMON GENETIC VARIANTS AND DISEASE PREDICTION

For nearly 40 years the only genetic factor associated with susceptibility to MS was the HLA region. Subsequent advancements in genotyping platforms and the development of more effective statistical methods allowed genome-wide association studies (GWAS) for the identification of several susceptibility loci.[1–12] Up until now, 233 independent genome-wide significant associations across the genome are associated with MS susceptibility, including 32 independent effects within the HLA region, and one in chromosome X. The next step is to determine the biology underlying these genetic associations. This is quite challenging because many identified genetic variants could be the tags for functionally relevant variants rather than having a meaning themselves due to extensive linkage disequilibrium. Furthermore, the majority of the variants discovered in the GWAS map to non-coding regulatory regions[4] and are likely to influence the expression of genes that lie hundreds to thousands base pairs away. The genes that lie closest to these genetic variants can be grouped together in networks and pathways by use of publicly available databases such as Encyclopedia of DNA elements (ENCODE) to better understand how the expression of these genes is altered, what cell types and tissues are involved, and to what pathways these gene/proteins belong. Recently, International Multiple Sclerosis Genetic Consortium (IMSGC) explored these variants in terms of biological relevance. Using cell-specific protein network analyses, IMSGC have shown that significant genetic MS associations from GWAS are associated with genes and gene products that are involved in the cellular pathways in monocytes, T and B cells. [13] The molecular functions of the GWAS risk genes [13] are depicted in figure 1.

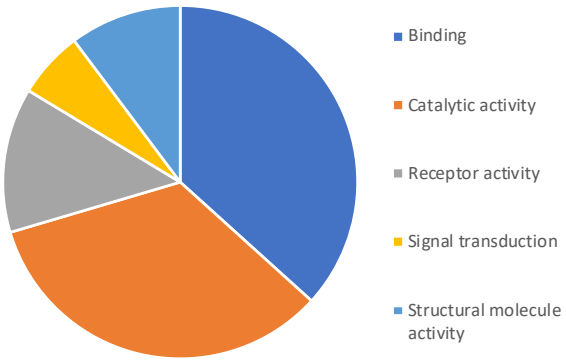


Figure 1. Molecular functions of the GWAS risk genes.

Another goal of understanding genetic variants is to investigate their utility in disease prediction, for better counseling of patients, such as prediction of conversion from CIS to MS, prediction of disease, and also prediction of MS in familial cases, specifically the unaffected family members of patients with MS. A better understanding of which individuals are at risk of developing MS, will create an opportunity to deploy early interventions such as life-style changes (i.e. implementing vitamin D into the diet, cessation of smoking, weight loss).

The complex polygenic architecture of MS with hundreds of causal genetic variants with small effects implies that no single variant on its own could be used to predict disease susceptibility. In order to take the complex genetic predisposition into account, a MS genetic burden score (a genetic risk score (GRS)) can be used, which includes genome-wide susceptibility loci. [14] To ascertain how well GRS discriminates between individuals with demyelinating disease and controls, a receiver operating characteristic (ROC) curve is generated by plotting the sensitivity of the continuous GRS score against 1–specificity, and calculating the area-under- the-curve (AUC). One of the first GRS studies in MS included only 16 MS risk loci to predict MS in an independent case-control dataset and had an AUC of 0.70. The risk score could not predict conversion of a clinically isolated syndrome (CIS) to MS.[15]

In this thesis we applied the GRS method in 2 types of population to test its performance in the disease prediction: in sporadic and familial MS (**chapter 2**) and pediatric-onset MS (**chapter 3**). In **chapter 2,** GRS based on 102 non-human leukocyte antigen loci and *HLA-DRB1*1501* was calculated in patients from multiplex MS families, sporadic MS cases and healthy controls. We showed that the cumulative genetic risk score based on 103 genetic risk factors (including the *HLA-DRB1*1501*) was significantly higher in MS multiplex patients than in sporadic cases and healthy controls in the Dutch population. The difference in genetic burden was primarily driven by a higher *HLA-DRB1*1501* allele frequency in familial cases. Higher genetic load in familial MS compared to sporadic MS was found in most other studies. [16–19] Two studies, that were less powered because of lower numbers of multiplex cases, found no difference in genetic load between multiplex and sporadic MS.[20,21] Population-related differences in allele frequencies could result in a distinct genetic load in those studies, as has been the case in rare variant studies of other disorders such as Alzheimer's disease. [22]

As for MS prediction, to differentiate between healthy controls and MS patients, GRS model based on 102 genetic risk loci and the *HLA-DRB1*1501* was modest in discriminating MS from controls (AUC 0.72); it performed better in discriminating multiplex MS from healthy controls (AUC 0.77) than sporadic MS from healthy controls (AUC 0.69), suggesting an enrichment of the risk genes in the familial cases. In one of the most advanced research fields in genetics, the field of cardiovascular disease (CVD), usage of risk factors for disease prediction and decision-making has a long history. In the late 90's the famous Framingham risk score (based on clinical parameters such as age, sex, diabetes, smoking status, total cholesterol, high-density lipoprotein cholesterol, and blood pressure) was developed for the prediction of CVD where an AUC of about 0.77 was used. [23] Although the AUC for MS found in our study is approaching the value of the Framingham risk score, a value with a higher specificity is required for a relatively rare condition such as MS. In order to obtain higher AUC a considerable number of additional common variants and/or stronger associated (rare) variants with higher odds ratios are needed.[24] Additional clinical parameters and environmental factors in the model could be beneficial in disease prediction, although risk estimates will probably not change that much when considering putative environmental factors in MS.[25] However, the combination of genome wide significant and marginally significant variants in a polygenic risk scores may yield an interesting alternative. Although it has been difficult to add predictive power to the Framingham risk score[26], in the field of CVD more recent studies based on hundreds of thousands of people and millions of genetic variants indicate that genetic risk scores can now outperform traditional risk factors (the Framingham score) in risk prediction.[27] Development of genome-wide polygenic scores makes it possible to identify individuals with clinically significantly increased disease risk. In coronary artery disease (CAD), the utility of polygenic risk score (in > 120.000 individuals of British ancestry and based on > 6.6 million genetic variants) has shown to successfully identify not only individuals at higher disease risk, but also individuals at a 3-fold, 4-fold and 5-fold higher risk; these are risk estimates comparable to levels conferred by rare monogenic mutations. [28] For these patients in the highest percentile there is a battery of preventive interventions available targeting dyslipidemia, hypertension, diabetes among others and healthy lifestyle which can substantially reduce CVD risk (i.e. these patients benefit most).[29,30]An important question that will be answered in the near future, is whether and how these polygenic scores will be used in clinical cardiology practice and GP based prevention.

Utility of a polygenic risk score in MS is an interesting concept as MS susceptibility is polygenic by nature, and the identified risk variants are just the tip of an iceberg

of all risk variants involved.[31,32] Before polygenic risk score can be applied in MS for the disease prediction, higher patient numbers (desirably above 100.000 individuals per study arm) of similar ancestry are needed to better understand the polygenic architecture of MS. This is a major challenge as such large sample sizes are difficult to achieve even in multicenter collaborations, because MS is a rare disease.

At present, the current knowledge of susceptibility genes does not (yet) contribute to clinical application in MS, in terms of diagnosis, prognosis, prediction of treatment response, or genetic counseling in MS families. At the moment, the greatest contribution of GWAS is the insight into MS biology and its altered pathways. Several identified genetic variants in the past have been successfully studied in functional studies. [33–39] However, further functional studies are needed to demonstrate that the found genetic variants have a causal functional change and to further understand the mechanisms that lead to MS susceptibility. One of the highlights of this GWAS is the genetic evidence of the X chromosome involvement in MS, the disease with a 3:1 female preponderance.

PEDIATRIC-ONSET MS

About 5-10% of all MS patients, had their first attack during childhood (before the age of 18 years).[40] Despite existing parallels in childhood-onset and adulthood-onset MS, there are fundamental differences in presentation and disease course. Pediatric-onset MS has a relapsing-remitting course in about 98% of all cases compared with around 84% of adult cases[41]; an initial progressive disease course is extremely rare in children with MS. Children with MS tend to have a higher relapse rate[42–44] and MRI lesion load compared to adults with MS.[45,46] On the contrary, progression of disability is slower in children with MS. [41,47] Despite longer disease duration to the secondary progressive disease phase, pediatric-onset MS patients will reach disability milestones at younger age than adult-onset MS patients. [41,47]

Pediatric-onset MS has been proposed as a unique time window for the study of early MS disease mechanisms, as the putative disease window of disease susceptibility is shorter in children than adults, and probably starts from (pre)conception. The research question is whether children with MS have a different genetic profile compared to adult-onset MS patients.

In **chapter 3**, we describe the involvement of adult-onset MS associated genetic variants in childhood-onset MS, and investigate their ability of disease prediction. At the time of our study only 57 genome-wide significant non-HLA SNPs with modest effects were known to be associated with MS risk in adult population. In **chapter 3**, we included children with ADS, adult-onset MS patients, and adult controls, and calculated GRS based on all known non-HLA SNPs and *HLA-DRB1*1501.* This study showed that next to *HLA-DRB1*1501,* 57 risk SNPs found in adults with MS also conferred susceptibility to MS in children. GRS of pediatric MS patients was higher than in children with monophasic ADS and healthy adult controls (both $p<0.01$). Mean GRS did not differ between childhood-onset MS and adult-onset MS ($p=0.29$). The genetic risk score of children with monophasic ADS did not differ from that of controls, implying that these SNPs do not confer a general risk of CNS inflammation in children. The discriminative ability of this risk score (of the full model) to differentiate between children with MS and monophasic ADS was moderate (AUC 0.70), and was comparable to that of adult MS versus healthy controls (AUC 0.73). The combined effect of 57 SNPs exceeded the effect of *HLA- DRB1*1501* alone in the risk prediction models for pediatric- and adult-onset MS.

As childhood-onset MS is a rare condition that is usually treated in tertiary clinics, few small to moderate sized studies (few hundreds of patients) were published about pediatric-onset MS in relation to genetic risk variants over the past years.[48–52] We now know that pediatric-onset MS indeed associates with similar genetic risk alleles as have been documented for adult-onset MS. This is the case for both *HLA-DRB1*1501*[49,53–55] as for non-HLA risk alleles[49,54]. The genetic burden seems to be higher in pediatric-onset MS compared to adult-onset MS[38] - a finding however not supported by our own smaller-sized study exploring a smaller set of genetic risk variants[54]. Furthermore, two other GRS models are independently associated with pediatric-onset MS susceptibility, that is GRS score based on vitamin D- associated SNPs (OR 0.72 95% CI 0.55-0.94) and GRS based on BMI-associated SNPs (OR 1.17 95% CI 1.05-1.30) (both after adjusting for sex, *HLA-DRB1*15*, GRS based on 110 non-HLA SNPs, and genetic ancestry).[48] In term of prediction of disease course, both *HLA-DRB1*15* carriership and GRS based on 110 non-HLA SNPs are not associated with relapse rate in children. However, increase of vitamin D level in *HLA-DRB1*15* carriers, results in decreased relapse rate, which is not the case in *HLA-DRB1*15* negative individuals. [51] Earlier, the presence of a vitamin D response element (VDRE) in the *HLA-DRB1* promotor region capable of binding a recombinant vitamin D receptor (VDR) has been described; increased cell surface expression of

HLA-DRB1 upon stimulation with 1,25-dihydroxyvitamin D3 has also been reported. [56] The precise mechanism and the timing of the *HLA-DBR1*- vitamin D interaction (could hypothetically be *in utero* or early childhood) still needs to be elucidated.

An important lesson from our studies, is that the disease prediction of both adult-onset and pediatric-onset MS is still moderate and cannot be implicated in the clinical practice. One of the explanations for this is the simplicity of the model used, as only common MS risk variants were included. Utility of a polygenic risk score encompassing genome-wide and marginally significant results might be more suitable. Inclusion of the genetic variants in combination with the environmental factors such as smoking status, vitamin D levels, infections, and variables for the interaction between genetic and environmental factors, could increase the utility of GRS and improve the model. Improvement of the disease prediction and the disease course is essential for better patient counselling, since it is now widely accepted that early treatment of MS patients results in a more favorable disease course.[57]

RARE GENETIC VARIANTS

The MS associated 233 common variants found in the recent GWAS study [32] explained up to 39% of the heritability of the disease (the proportion of the total phenotypic variation that is due to additive genetic factors) leaving open the question of the so-called 'missing heritability' of common traits. Such phenomenon can be explained by different mechanisms, such as hidden epigenetic effects, gene by gene and gene by environment interactions, and rare penetrant semi-private genetic variants. Recently, interest is increasing in rare genetic variants, in particular the ones with large(r) effects. Family-based studies offer a powerful setting to search for rare variants because of a higher prevalence of the disease. These families are also the setting for personalized interventions, i.e., interventions that are not justified in the general population due to the rarity of these genetic variants.

With the completion of the human reference genome, and automation and technical improvements of sequencing techniques, it became possible to sequence a genome quickly and affordably. Sequencing allows the identification of genetic variants that are private to patients that have developed the disease and affect heritable phenotypes conveying a high risk, including disease-causing mutations. The most widely used

targeted sequencing method clinically is whole exome sequencing (WES) that targets the protein-coding regions of the human genome which represent less than 2% of the genome, but contain about 85% of known disease-related variants. In **chapter 4.1** we describe a Dutch multiplex MS family with 11 affected individuals in several generations, in which linkage analysis and WES were applied to search for rare variants. We found a rare missense variant (rs147213094) in the *FKBP6* gene which resided in the linkage area in the chromosomal region 7q11.22– q11.23. While this genetic variant showed significant segregation pattern with the disease in the family, its association outside the family was not proven in the validation cohort of 591 Dutch MS cases and 3169 controls. When the patient group was divided in sporadic and multiplex cases, an increased frequency of this variant was found in the MS cases from multiplex families, but not in sporadic MS patients. *FKBP6* gene is not listed among MS risk genes from the latest GWAS study.[32] Although there is some evidence of its possible association with MS,[58–60] it is unknown how it might be related to the mechanisms involved in MS pathogenesis. The chromosomal region 7q11– q21 containing *FKBP6* has shown a suggestive linkage to MS with a LOD = 1.14 in 52 multiplex families of European descent. [58] A case report described a co-occurrence of Williams – Beuren syndrome and MS in an adolescent patient.[60] Furthermore, in the experimental autoimmune encephalomyelitis (EAE) model in rats, a significant linkage in the chromosomal region analogous to the human region containing *FKBP6* was found with the incidence and duration of EAE, and also the maximum EAE score.[59]

The *FKBP6* gene is located on chromosome 7 and belongs to the immunophilins FK-506 binding protein (FKBP) family of highly conserved proteins that possess binding abilities to immunosuppressive drugs. *FKBP6* was first described in the context of Williams–Beuren syndrome, a dominant autosomal disorder caused by a hemizygous contiguous gene deletion on chromosome 7q11.23 containing 28 genes in the largest deletions, including *FKBP6*.[61] Smaller deletions are also found in these patients excluding *FKBP6*. [62] *FKBP6* is part of the synaptonemal complex that forms between two homologous chromosomes during meiosis, and mediates chromosome pairing, synapsis, and genetic recombination. In a study in idiopathic infertile men four missense variants in *FKBP6* were associated with male infertility.[63] A recent paper identified that the co-chaperone FKBP6 acts as a host factor required for hepatitis C virus (HCV) replication.[64] HCV replication was completely suppressed in FKBP6-knockout hepatoma cell lines, while the expression of FKBP6 restored HCV replication in FKBP6-knockout cells.[64] We speculate that *FKBP6* gene might have yet undiscovered functions.

Up to date several other rare variants in MS susceptibility genes were found using WES: *CYP27B1*[65], *TYK2*[66], *NR1H3*[67], *NLRP1*[68] , *PLG*[69], *P2RX4/P2RX7*[70,71] and *GALR2*[72]. Most of these studies are still devoid of replication.[73–77] When making a replication effort of rare variants, a sufficient power should be generated by using a high enough number of cases and controls due to rarity of the variants in most populations. This was however lacking in the most replication studies. More recently, a GWAS study done by IMSGC demonstrated the role of rare variants (2 in known MS risk genes and 4 in new genes) in a large international cohort of patients and controls using the exome-chip array. These variants were however not found in the family studies. It is important to emphasize that rare alleles have different frequency in ethnically different populations, and could be population- or even family-specific.[78] In the aforementioned GWAS, many samples from different populations with different minor allele frequencies (MAF) are taken together causing the dilution of the allele frequencies of such variants, and some variants can therefore be overlooked.

It is exciting that in a complex disease such as MS, WES has been successful in the identification of rare variants in family-based studies. On the other hand, the big question remains what this actually means for MS in general, as these variants have very low MAF, were found in single families and seem to play a limited role in sporadic cases, which represent a broader spectrum of MS population. Yet, these variants may reveal important information on the pathogenesis of MS. Although the role of rare variants with large effects is still debated in MS, in other complex diseases such as schizophrenia[79,80] and Alzheimer's disease[81–83], the contribution of rare variants to the disease risk has been established more conveniently. To better understand the role of the variants identified, functional studies are needed. Functional studies in rare variants have been proven difficult for MS.

As explained in **chapter 4.2**, there are some critical notes on the utility of WES. It is a fact that not all targeted regions are captured at the same efficiency in WES. Not every kit covers the same exome regions[84] and the causal variant may rather be in a non-protein coding region, also in familial patients. The possibility of missing out the 'actual disease-causing variant' is thus present. WES is limited in to detecting large insertions/deletions/duplications, large rearrangements, mitochondrial genome mutations and it does not detect changes in methylation. Furthermore, WES might miss out variants in inadequately targeted regions and is unsuitable for finding potential candidates in cryptic exons or non-coding regions.[85]

As WES focuses only on the coding regions and has its limitations, the question remains whether we should continue to look for rare coding variants or switch the tactics to other genetic regions, as many GWAS signals are from intronic and intergenic sites of the genome. As we recently started to understand, these regions belong to important regulatory regions[86,87] that may be highly relevant for middle-aged and late-aged disorders that do not express at birth. Detailed investigations using whole-genome sequencing (WGS), which also capture the regulatory regions, have not yet been performed in MS. As technology keeps evolving and will hopefully become cheaper in the near future, the utility of WGS on a large scale could provide convenient answers to the question about the role of rare variants and the 'missing heritability' of MS.

IMMUNE RESPONSE TO EPSTEIN-BARR VIRUS INFECTION IN MS FAMILIES

Besides MS genetics, involving over 230 risk loci, also environmental factors play an important role in MS susceptibility, of which Epstein–Barr virus (EBV) is the most well-known and consequently replicated one.[88,89] Almost all MS patients have antibodies against EBV [90], while EBV antibodies are present in 90%–95% of healthy controls.[91,92] Infectious mononucleosis (IM) is a common manifestation of EBV infection in adolescence and adulthood.[93] The risk for MS is increased in individuals with a history of IM compared with individuals without such a history.[94] The most consistent finding in MS is an increase of antibodies against EBV nuclear antigen 1 (EBNA-1). In fact, elevated EBNA-1 immunoglobulin G (IgG) levels increase the risk of developing MS later in life.[95–97] There are several mechanisms and theories by which EBV infection could increase the risk of MS. One theory is the molecular mimicry theory, where the immune response to EBV infection in individuals with a specific genetic susceptibility cross-reacts with myelin antigens in which both cross-reacting T lymphocytes[98,99] and antibodies[100] are involved. This is supported by the finding that MS patients have an increased frequency and broadened specificity of CD4+ T cells recognizing EBNA-1[101], and the presence of increased immunoreactivity to two EBV peptides, one of which is EBNA-1, in both serum and CSF of MS patients[102]. Another theory is that EBV could be an initiator of a pathological process as MS tends to occur years after IM[103], and that other factors are needed to trigger the disease onset.

The IgG response to EBNA-1 has a heritability of 22-43%, suggesting that genetic factors are important in the immune activation.[104–106] The strongest genetic risk factor for MS,

*HLA-DRB1*15*, is associated with enhanced EBNA-1IgG levels. Also, SNPs in *VCAM1, EVI5, ILDR1/CD86, EOMES, MYB, CARD11, IL2RA, PRDX5, SOX8, CLEC16A* were associated with increased EBNA-1 IgG levels.[107,108] Although a formal Mendelian Randomisation experiment was not conducted, this suggests that increased EBNA-1 IgG levels are part of the pathogenesis of MS.

Research in this thesis is primarily focussed on MS families, and as explained previously, first-degree family members are at higher risk of developing MS themselves compared to general population. This familial clustering is partly explained by a shared genetic pool, that is transmitted from generation to generation, but may also be the consequence of the shared exposure to viral infections such as EBV. A recent meta-analysis of twin studies showed that shared environmental influences contribute for 21% (95% CI 10-30%) of MS liability variance indicating that shared environmental factors during the early stages of life must be of some influence on MS susceptibility in MS families. [109] This may be related to infections with EBV, but also other factors such as exposure to (passive) smoking.

In **chapter 5** we assessed the humoral immune response (IgG) to EBNA-1 in MS patients, their healthy siblings and unrelated healthy spouses. We found an EBNA-1 IgG titer gradient, being the highest in MS patients, intermediate in siblings and the lowest in spouse controls. Because in our previous study first-degree family members were shown to have higher genetic load compared to unrelated controls[110], we validated the role of *HLA-DRB1*1501*, the major genetic locus for MS risk, on enhanced humoral immune responses against EBV. *HLA-DRB1*1501* carriership was associated with elevated EBNA-1 IgG titers, but it only partially explained the elevated EBNA-1 IgG levels, as the frequency of *HLA-DBR1*1501* risk allele mirrored the EBNA-1 IgG gradient in the study groups. In the *HLA-DRB1*1501* negative group, the significant difference was still present between EBNA-1 IgG titers of young MS patients and spouses. This suggests that next to *HLA-DRB1*1501* also other genetic factors are involved in the anti EBV immune response in young MS patients. Furthermore, we also tested an *HLA-A*02* tagging SNP, but this one was not associated with elevated EBNA-1 IgG titers. Because of moderate study size, we limited our genetic association analysis to only *HLA-DBR1*1501* and *HLA-A*02* loci. We may therefore miss the contribution of other yet unexplored genetic factors or their interactions in relation to EBV and MS risk. In fact, in another yet unpublished study, the preliminary results of 165 MS risk SNPs from the most recent MS GWAS[32] showed that at least 15 of these SNPs are correlated with EBNA-1 IgG titers

in MS patients, in both *HLA-DRB1*1501* risk allele carriers and non-carriers (*personal communications with K. Kreft).*

Next, we found that high EBNA-1 titers were associated with younger age. EBNA-1 IgG titers were higher in young MS patients (< 50 years of age) compared to older MS patients. (> 50 years of age). This could be due to a particular phenomenon called immunosenescence (less productive immune system with increasing age).[111] There is overwhelming evidence that the amount of antibodies induced after an immunisation response is strongly reduced with physiological ageing, and that titers decline more rapidly in elderly people.[112,113] Furthermore, in our study particularly in old MS patients and old siblings, elevated EBNA-1 IgG titers were associated with the presence of the *HLA-DRB1*1501* risk genotype. This suggest that in older people, in order to react adequately to the EBV infection, a certain genetic make-up is needed.

MS AND CO-OCCURRENCE OF OTHER AUTOIMMUNE DISEASES

An overlap between MS-associated genetic variants and those associated with other autoimmune diseases (AIDs) is an interesting finding that may reveal common pathological pathways.[114,115] Studying the presence of additional autoimmune diseases in both people with MS, and their first-degree relatives has been pursued over many years, with studies employing a variety of designs and yielding conflicting results.[116]

In chapter 6 we explored whether there is a clustering of autoimmune diseases in patients with MS and their first-degree relatives in the Dutch MS multiplex families. A personally administered questionnaire was used to collect epidemiological and environmental data. The following autoimmune diseases were included in the questionnaire: autoimmune thyroid disease (Hashimoto's thyroiditis and Grave's disease), rheumatoid arthritis, ulcerative colitis, Crohn's disease, psoriasis, pernicious anemia, type 1 diabetes, vitiligo, Sjögren's syndrome and systemic lupus erythematosus. We did not find an excess of autoimmunity in the patients with MS compared to controls, supported by several other studies in this field.[117–119] About 43% of the Dutch MS multiplex families reported one or more autoimmune diseases in the first-degree relatives. This number was in the range of the reported numbers worldwide (27%-64%).[117,120] However, there was no increased risk of autoimmune diseases in the first-degree relatives of MS probands, which was in line with some[119,121] but not all previous studies[117,122–124]. There are several factors that

could potentially cause conflicting results between our study and previously published studies, such as a different methodology (community dwelling vs a hospital-based study population); usage of questionnaires with possible recall bias vs a register-based cohort design where the information can be misclassified or incomplete; number and type of AIDs studied. From GWA studies in autoimmune diseases, it became clear that some but not all genetic risk loci (majority) are shared across various autoimmune diseases.[115] Some genetic variants can be associated with several traits, but it should be noted that some have opposite effects in these diseases, thus leading to 'non-clustering' of AIDs. It is thus very important to look at what kind of AIDs are studied together in relation to MS, because some AIDs will and some will not cluster with each other.

FUTURE RESEARCH

GWAS studies in MS have been proven incredibly successful with the most recent discovery of more than 200 genetic risk loci. Further increasing the samples sizes and utility of larger SNP arrays will only increase the number of newly associated SNPs with very modest effect sizes on the disease risk. Similar to other complex diseases, the polygenic risk scores may be able to identify a small group of patients at a very high risk of developing disease. A major challenge will be the question how to prevent the disease such as in these individuals. It will remain challenging for the future to further uncover the functional meaning of the genetic variants and translate the gathered knowledge into the understanding of MS pathophysiology. The latter might even be the most challenging step, as each of the genetic loci will have to be followed-up experimentally in multiple conditions. The next step will also be the discovery and implementation of the large-scale and fast-throughput fully-automated testing facilities to investigate each of the loci more time- and cost-effectively.

Rare variants in multiplex families

New research directions are needed to further unravel MS genetic architecture and help clarify the missing heritability. It has been suggested that at least part of the genetic component of this complex phenotype may be due to rare variants – the so-called "common disease/rare variant hypothesis".[125] This approach may identify those at a high genetic risk, a group that is eligible for very targeted personalized medicine and interventions. From a pharmacological perspective, targeting few rare variants will be 'easier' than targeting the joint effect of many polygenic genes and the proteins they encode for. Up until now the focus has been on the rare coding genetic variants found by means of the whole exome sequencing. Most of the rare variants in MS were found in single families. Searches for rare variants in a case-control setting were fruitful, but not that spectacular because only a limited number of rare variants were detected[126], similar to the findings for other complex disorders. The next step in answering the question about the role of rare variants in MS will need a different approach, i.e. it will need international collaboration to make the aggregation of larger numbers of multiplex MS families from different ethnicities possible. Families with multiple MS patients in different generations will probably have the highest yield in terms of rare variants. This will be challenging because large MS families are quite rare.

Other variants

Investigations of variants in the regulatory regions and structural variants, copy-number variants, insertions/deletions, by means of whole genome sequencing, might help to clarify the missing pieces of the heritability puzzle. A part of missing heritability could also be explained by genetic interactions or epistasis of the already known genetic variants.[127] Demonstration of epistasis in human populations studies is however very difficult and would require huge samples.[128] Thus, the magnitude of the contribution of epistasis to the missing heritability of complex diseases such as MS still remains to be determined.

Signature genes

Gene expression studies might offer valuable insights into MS pathophysiology and help solve the heritability puzzle. Changes in expression of a group of genes, called signatures, could potentially capture disease-specific alterations. One of the first studies characterizing such signatures, showed that T cell specific genes were overexpressed in MS patients compared to controls, consistent with postulated pathogenesis of MS. [129] Another study had shown that gene signatures have a potential to identify subgroup of MS patients with a more active disease, characterized by an increased risk of a relapse while on first-line disease-modifying treatments.[130] This is an interesting finding, and when validated, might help to divide MS patients in meaningful subsets and allow a more personalized patient care such as an earlier switch of first-line treatment to a more aggressive treatment.

Single-cell RNA sequencing

Recent advances in molecular biology and nanotechnology have given rise to single-cell sequencing technologies that measure a broad range of cellular parameters. Of these technologies, single-cell RNA sequencing is the most widely used one.[131] The recent single-cell studies in post mortem MS brain tissue have identified an altered oligodendroglial heterogeneity[132], selective cortical neuron damage and glial activation associated with lineage- and region-specific transcriptomic changes. [133] Another study showed that in MS a transcriptional diversity in blood is increased, while in CNS a cell type diversity is increased, hereby providing evidence for compartmentalized mechanisms which drive human autoimmunity in the brain. [134] Integration of these newly acquired data with the known genetic knowledge, will add important pieces to the MS puzzle.

Epigenetics

A very interesting avenue of research in MS genetics will be epigenetics, which are usually non-inherited modifications to the DNA that regulate gene expression. Modification in DNA methylation has been reported in blood, CD4+, and CD8+ T cells as well as in pathology-free brain regions from MS patients.[135–138] Moreover, genetic variants in the loci encoding epigenetic machinery genes have been associated with MS.[139,140] For the two important susceptibility factors in MS, HLA and smoking, research has shown that DNA methylation mediates impact of the HLA locus on MS susceptibility[141], and that smoking alters DNA methylation of blood cells in MS patients.[142] Perhaps the most vivid example of the involvement of the epigenetic mechanisms is the pregnancy-related modification of disease activity in MS, where the rising level of estrogens modulates T-cell activity through complex epigenetic regulation mechanisms.[143] More and more evidence is available that supports the role of epigenetic processes in the pathogenesis of MS. Yet, methodological limitations delay the understanding of functional relevance of the detected alterations, and make interpretation and translation of epigenetic changes into biology still very challenging. [138]

Mendelian randomisation

Mendelian randomisation (MR) is based on the notion that genetic factors can explain the cause of non-genetic factors, such as smoking or high body mass index. MR is an analytical method that uses genetic variants as instrumental variables to investigate the causal relationship between modifiable risk factors and disease. However, MR only works if the information of the genetics of a disease is sufficient. In the past ten years, many papers have been published using MR because GWAS have produced more that hundred thousand of reliable genetic variants that strongly associate with modifiable risk factors and behaviors. Because MR uses genetic variants, which are fixed at conception, it can overcome some types of confounding that are inevitable in observational studies, such as reversed causation. [144] In MS, MR has been used to assess causal effects of environmental risk factors on the disease (including pediatric-onset MS), such as lowered vitamin D[48,145,146] and obesity[48,147–149]. Considering the high prevalence of these exposures, prevention can now be considered. In other medical fields such as coronary artery disease, MR has gone beyond exposure-outcome association, and is used to find worthy targets for development of new drugs.[150] Because MR studies have the power to determine whether a particular risk factor/ drug target is a causative risk factor or just simply a biomarker of a disease[150], MR can help prioritize clinical trials and drug development, and reduce the time, cost and failures associated with clinical trials. [144]

Gut microbiota

Another interesting and developing MS research field is the gut microbiota. It appears that gut bacteria are causally related to brain autoimmunity. Studies in a germ-free transgenic mouse model of spontaneous experimental autoimmune encephalomyelitis (EAE; a well-validated animal model of MS) have shown that these mice were resistant to developing EAE, which was reversed by fecal microbiota transplant from normal mice.[151] In other studies, when EAE mice were transplanted with MS twin-derived microbiota, it induced a significantly higher incidence of autoimmunity than the healthy twin-derived microbiota.[152] Germ-free mice studies showed that gut microbiota is crucial in regulating the blood-brain barrier (BBB), as loss of integrity of the BBB is a hallmark of MS.[153] Multiple studies support the concept that gut microbiota is a key regulator of neuroinflammation.[154–157] Further research is needed to understand the mechanisms by which microbiota is involved in MS pathophysiology. Future studies will also have to show what dietary recommendations can be given to MS patients in order to modulate the inflammatory cascade that might act through microbiota to diminish the MS disease activity.

Prediction of disease

MS has a significant impact on families, often creating psychological stress in each family member, influencing their wellbeing and quality of life.[158] First-degree family members of MS patients have in addition higher risks of developing MS themselves. Several attempts have been made to predict MS in the family members using simple models with genetic markers. The predictions scores are moderate and still cannot be used in the clinical practice. Because of the multifactorial nature of MS, the prediction models should incorporate detailed information on epidemiological, clinical, immunological, genetic and imaging biomarkers in sufficiently large number of subjects to increase the prediction utility. Improvement of clinical prediction of the disease and disease course is essential for a better patient and family counselling. Identification of these high-risk individuals will create an opportunity for the clinical trials of primary prevention of MS, based on, for example, life-style changes (i.e. healthy diet with an addition of vitamin D, smoking cessation, weight reduction) or early interventions with application of immunomodulating therapies .[19]

Personalized medicine

Personalized medicine has the aim to streamline clinical decision making by using biological information, for example a genetic test or biomarker, in regard to the prevention,

diagnosis, and treatment of a disease.[159] The concept of personalized medicine lies in the fact that there is a great variety in inter-individual differences with respect to mechanisms, contributing factors and effects of disease processes on individuals. The contemporary focus of personalized medicine is rooted in the findings of genetic studies that show that individuals vary widely because they possess subsets of genetic variants that exist in the human population as a whole.[160] In addition, some subsets of these genetic variants are *de novo* mutations, and may be unique to an individual. These genetic differences between individuals, in combination with environmental exposures and epigenetic changes explain for the most part why there is so much variability with respect to disease susceptibility and response to interventions, and all need to be taken into account when determining the optimal way for disease prevention, detection and treatment.[160] In order to make early disease prediction possible, it is important to find mutations with large enough effects and to understand the polygenic effects. However, in MS mutations with large effect are very rare. The functional effect of these mutations can be studied in cellular experiments using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 technologies. CRISPR/Cas9 technique allows the creation of cell lines with and without a specific mutation, and the comparison of these two would give an insight into the mutation effects controlling for the genetic background associated with the patients' genome. Also, it is now possible to harvest cells from individual patient and use pluripotency induction (induced pluripotent stem cell or 'iPSC') method on these cells in order to generate cell types of relevance for a specific condition, hereby creating a 'disease in a dish' cellular model of a patients' condition. [160] iPSC will help to overcome limitations associated with autoimmune-mediated demyelinating animal models used for MS, as iPSC will make human cell lines accessible, especially those in the brain tissues. Based on the GWAS the monocytes-macrophage-microglial cells are of interest in MS.[13] MS-iPSC that differentiate into the target cells, with the specific genetic and epigenetic MS background, may be further used for the functional studies to better understand the functional effects of the MS-associated genetic variants and epigenetic changes, but also for drug screening and toxicology studies. [161] Although our understanding of MS immunopathology and genetic architecture is moving forward, there are still so many questions that need be answered, and the utility of iPSC could be helpful in elucidating key pathways in the initiation and development of MS.

MS registries

Current treatment strategy in MS, especially in patients with an active disease course and accumulating disability, is to start treatment as early as possible to prevent further disease progression and neurodegeneration. To ensure a proper treatment for an individual patient with MS, it requires clear profiling of this individual patient through a better characterization of clinical and immunopathological subtypes of the disease, and a definition of clinical criteria for responsiveness and/or treatment failure. Better profiling can be achieved through a collection of large amounts of real-world data in MS patient care. MS registries are a very meaningful tool for collecting long-term data of a large number of patients (and their relatives) over a very long period of time. Here, high quality data is documented in a standardized way and can be retrieved easily and efficiently for different purposes, [162] something that a standard clinical study is not able to produce due to its limited scope. MS registers enable correlations of disease progression with epidemiological and clinical characteristics. Big data collections like these overcome the limitations in statistical power that characterize most clinical studies (only if data is properly recorded).[163] There are several MS registers worldwide that capture different characteristics of the disease. Some are physician-oriented, some are patient-oriented and some others are a combination of both with different advantages and disadvantages in each case.[164] For MS registries to be successful, better coordination in standardization of data collection between registries is still needed in order to increase and leverage existing and future efforts.[165] In the near future, with the development of bioinformatical computation systems and machine learning algorithms, it will become possible to better integrate this MS big data into intelligent management systems where the data is not only collected, but can be used to predict and analyze treatment response patterns on the basis of patients' clinical profiling, genomic make-up, epigenetic tendencies, and environmental data, and allow personalized disease predictions and treatments.

REFERENCES

1. Comabella, M. *et al.* Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. PLoS One. 2008; 3:e3490.
2. Bahlo, M. *et al.* Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet. 2009; 41:824-8.
3. International Multiple Sclerosis Genetics Consortium (IMSGC) *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353-60.
4. Patsopoulos, N. A. *et al.* Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019; 365:eaav7188.
5. Baranzini, S. E. *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet. 2009;18:767-78.
6. De Jager, P. L. *et al.* Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet. 2009;41:776-82.
7. International Multiple Sclerosis Genetic Consortium. Risk Alleles for Multiple Sclerosis identified by a genomewide study. N Engl J Med. 2009; 357:851-62.
8. Sanna, S. *et al.* Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. Nat Genet. 2010;42:495-7.
9. Nischwitz, S. *et al.* Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. J Neuroimmunol. 2010;227:162-6.
10. Jakkula, E. *et al.* Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. Am J Hum Genet. 2010;86:285-91.
11. Sawcer, S. *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. 2011;476:214-9.
12. Patsopoulos, N. A. *et al.* Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol. 2011;70:897-912.
13. International Multiple Sclerosis Genetics Consortium. A systems biology approach uncovers cell-specific gene regulatory effects of genetic associations in multiple sclerosis. Nat Commun. 2019;10:2236.
14. Wray, N. R. *et al.* Prediction of individual genetic risk to disease from genome-wide association studies. Genome Res. 2007;17:1520-8.
15. De Jager, P. L. *et al.* Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol. 2009;8:1111-9.
16. D'Netto, M. J. *et al.* Risk alleles for multiple sclerosis in multiplex families. Neurology. 2009;72:1984-8.
17. Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann Neurol. 2011;69:65-74.
18. Isobe, N. *et al.* Genetic burden in multiple sclerosis families. Genes Immun. 2013; 14:434-440.
19. Xia, Z. *et al.* Genes and Environment in Multiple Sclerosis project: A platform to investigate multiple sclerosis risk. Ann Neurol. 2016;79:178-89.
20. Esposito, F. *et al.* Impact of MS genetic loci on familial aggregation, clinical phenotype, and disease prediction. Neurol Neuroimmunol Neuroinflamm. 2015;2:e129.
21. Harbo, H. F. *et al.* Oligoclonal bands and age at onset correlate with genetic risk score in multiple sclerosis. Mult Scler. 2014;20:660-8.

22. Van Der Lee, S. J. *et al.* PLD3 variants in population studies. Nature. 2015;520:E2-3.

23. Wilson, P. W. F. *et al.* Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97:1837-47.

24. Jafari, N. *et al.* Perspectives on the use of multiple sclerosis risk genes for prediction. PLoS One. 2011;6:e26493.

25. Sawcer, S. & Wason, J. Risk in complex genetics: 'all models are wrong but some are useful'. Ann Neurol. 2012;72:502-9.

26. Tzoulaki, I. *et al.* Assessment of claims of improved prediction beyond the Framingham risk score. JAMA. 2009;302:2345-52.

27. Knowles, J. W. & Ashley, E. A. Cardiovascular disease: The rise of the genetic risk score. PLoS Med. 2018;15:e1002546.

28. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics 2018;50:1219-1224.

29. Mega, J. L. *et al.* Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: An analysis of primary and secondary prevention trials. Lancet. 2015;385:2264-2271.

30. Khera, A. V. *et al.* Genetic risk, adherence to a healthy lifestyle, and coronary disease. N Engl J Med. 2016;375:2349-2358.

31. The International Multiple Sclerosis Genetics Consortium. Evidence for polygenic susceptibility to multiple zclerosis-the shape of things to come. Am J Hum Genet. 2010;86:621-5.

32. Patsopoulos, N. A. *et al.* Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science. 2019;365:eaav7188.

33. Van Luijn, M. M. *et al.* Multiple sclerosis-associated CLEC16A controls HLA class II expression via late endosome biogenesis. Brain. 2015;138:1531-47

34. De Jager, P. L. *et al.* The role of the CD58 locus in multiple sclerosis. Proc Natl Acad Sci. 2009;106:5264-9.

35. Gregory, S. G. *et al.* Interleukin 7 receptor α chain (IL7R) shows allelic and functional association with multiple sclerosis. Nat Genet. 2007;39:1083-91.

36. Gregory, A. P. *et al.* TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. Nature. 2012;488:508-511.

37. Maier, L. M. *et al.* Soluble IL-2RA levels in multiple sclerosis subjects and the effect of soluble IL-2RA on immune responses. J Immunol. 2009;182:1541-7.

38. Raj, T. *et al.* Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. Science. 2014; 344:519-23.

39. Kreft, K. L. *et al.* Decreased systemic IL-7 and soluble IL-7Rα in multiple sclerosis patients. Genes Immun. 2012;13:587-92.

40. Banwell, B. L. Multiple sclerosis in children. Handb Clin Neurol. 2014;122:427-41.

41. Renoux, C. *et al.* The natural history of multiple sclerosis with childhood onset. Clin Neurol Neurosurg. 2008;110:897-904.

42. Boiko, A. *et al.* Early onset multiple sclerosis: A longitudinal study. Neurology. 2002;59:1006-10.

43. Gorman, M. P. *et al.* Increased relapse rate in pediatric-onset compared with adult-onset multiple sclerosis. Arch Neurol. 2009;66:54-9.

44. Benson, L. A. *et al.* Elevated relapse rates in pediatric compared to adult MS persist for at least 6 years. Mult Scler Relat Disord. 2014;3:186-93.

45. Yeh, E. A. *et al.* Magnetic resonance imaging characteristics of children and adults with paediatric-onset multiple sclerosis. Brain. 2009;132:3392-400.

46. Waubant, E. *et al.* Difference in disease burden and activity in pediatric patients on brain magnetic resonance imaging at time

of multiple sclerosis onset vs adults. Arch Neurol. 2009;66:967-71.

47. Simone, I. L. *et al.* Course and prognosis in early-onset MS: Comparison with adult-onset forms. Neurology. 2002; 59:1922-8.

48. Gianfrancesco, M. A. *et al.* Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. Neurology. 2017;88:1623-1629.

49. Gianfrancesco, M. A. *et al.* Genetic risk factors for pediatric-onset multiple sclerosis. Mult Scler. 2018;24:1825-1834.

50. Graves, J. S. *et al.* The multiple sclerosis risk allele within the AHI1 gene is associated with relapses in children and adults. Mult Scler Relat Disord. 2018; 19:161-165.

51. Graves, J. S. *et al.* Genetic predictors of relapse rate in pediatric MS. Mult Scler. 2016;22:1528-1535.

52. van Rensburg, S. J. *et al.* Identification of an iron-responsive subtype in two children diagnosed with relapsing-remitting multiple sclerosis using whole exome sequencing. Mol Genet Metab Reports. 2019;19:100465.

53. Banwell, B. *et al.* Clinical, environmental, and genetic determinants of multiple sclerosis in children with acute demyelination: A prospective national cohort study. Lancet Neurol. 2011; 10:436-45.

54. Daniëlle van Pelt, E. *et al.* Risk genes associated with pediatric-Onset MS but not with monophasic acquired CNS demyelination. Neurology. 2013;81:1996-2001.

55. Disanto, G. *et al.* HLA-DRB1 confers increased risk of pediatric-onset MS in children with acquired demyelination. Neurology. 2011;76:781-6.

56. Ramagopalan, S. V. *et al.* Expression of the multiple sclerosis-associated MHC class II allele HLA-DRB1*1501 is regulated by vitamin D. PLoS Genet. 2009;5:e1000369.

57. Stüve, O. *et al.* Pharmacological treatment of early multiple sclerosis. Drugs. 2008;68:73-83.

58. Haines, J. L. *et al.* A complete genomic screen for multiple sclerosis underscores a role for the major histocompatibility complex. Nat Genet. 1996;13:469-71.

59. Becanovic, K. *et al.* Advanced intercross line mapping of Eae5 reveals Ncf-1 and CLDN4 as candidate genes for experimental autoimmune encephalomyelitis. J Immunol. 2014;176:6055-64.

60. Papageorgiou, E. *et al.* Multiple sclerosis in association with Williams-Beuren syndrome. J Paediatr Child Heal. 2010;46:612–614.

61. Schubert, C. The genomic basis of the Williams - Beuren syndrome. Cell Mol Life Sci. 2009;66:1178-97.

62. Fusco, C. *et al.* Smaller and larger deletions of the Williams Beuren syndrome region implicate genes involved in mild facial phenotype, epilepsy and autistic traits. Eur J Hum Genet. 2014;22:64-70.

63. Zhang, W. *et al.* Mutation screening of the FKBP6 gene and its association study with spermatogenic impairment in idiopathic infertile men. Reproduction 2007;133:511-6.

64. Kasai, H. *et al.* Involvement of FKBP6 in hepatitis C virus replication. Sci Rep. 2015;5:16699.

65. Ramagopalan, S. V. *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. Ann Neurol. 2011;70:881-6.

66. Dyment, D. A. *et al.* Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. Neurology. 2012;79:406-11.

67. Wang, Z. *et al.* Nuclear Receptor NR1H3 in Familial Multiple Sclerosis. Neuron. 2016;90:948-54.

68. Maver, A. *et al.* Identification of rare genetic variation of NLRP1 gene in familial multiple sclerosis. Sci Rep. 2017;7:3715.

69. Sadovnick, A. D. *et al.* Analysis of Plasminogen Genetic Variants in Multiple Sclerosis Patients. G3 (Bethesda). 2016;6:2073-9.

70. Sadovnick, A. D. *et al.* Purinergic receptors P2RX4 and P2RX7 in familial multiple sclerosis. Hum Mutat. 2017;38:736-744.

71. Zrzavy, T. *et al.* A rare P2RX7 variant in a Hungarian family with multiple sclerosis. Mult Scler Relat Disord. 2019;27:340-341.

72. Garcia-Rosa, S. *et al.* A non-functional galanin receptor-2 in a multiple sclerosis patient. Pharmacogenomics J. 2019;19:72-82.

73. Ban, M. *et al.* No evidence of association between mutant alleles of the CYP27B1 gene and multiple sclerosis. Ann Neurol. 2013;73:430-2.

74. Barizzone, N. *et al.* No evidence for a role of rare CYP27B1 functional variations in multiple sclerosis. Ann Neurol. 2013;73:433-7.

75. Minikel, E. V. & MacArthur, D. G. Publicly available data provide evidence against NR1H3 R415Q causing multiple sclerosis. Neuron. 2016;92:336-338.

76. International Multiple Sclerosis Genetics Consortium. NR1H3 p.Arg415Gln is not associated to multiple sclerosis risk. Neuron. 2016;92:333-335.

77. Vilariño-Güell, C. *et al.* Analysis of Canadian multiple sclerosis patients does not support a role for FKBP6 in disease. Mult Scler. 2018;25:1011-1013.

78. Mathieson, I. & McVean, G. Differential confounding of rare and common variants in spatially structured populations. Nat Genet. 2012;44:243-6.

79. Singh, T. *et al.* The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. Nat Genet. 2017;49:1167-1173.

80. Coelewij, L. & Curtis, D. Mini-review: Update on the genetics of schizophrenia. Ann Hum Genet. 2018;82:239-243.

81. Vardarajan, B. N. *et al.* Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Ann Neurol. 2015;78:487-98.

82. Guerreiro, R. *et al.* TREM2 variants in Alzheimer's disease. N Engl J Med. 2013; 368:117-27.

83. Cruchaga, C. *et al.* Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature. 2014; 505:550-554.

84. Harismendy, O. *et al.* Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol. 2009;10:R32.

85. Seaby, E. G. *et al.* Exome sequencing explained: A practical guide to its clinical application. Brief Funct Genomics. 2016; 15:374-84.

86. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57-74.

87. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317-30.

88. Munger, K. L. *et al.* Anti-Epstein-Barr virus antibodies as serological markers of multiple sclerosis: A prospective study among United States military personnel. Mult Scler. 2011;17:1185-93.

89. Ascherio, A. & Munger, K. L. Epstein-barr virus infection and multiple sclerosis: A review. J Neuroimmune Pharmacol. 2010;5:271-7.

90. Ascherio, A. & Munger, K. L. Environmental risk factors for multiple sclerosis. Part I: The role of infection. Ann Neurol. 2007;61:288-99.

91. Almohmeed, Y. H. *et al.* Systematic review and meta-analysis of the sero-epidemiological association between Epstein Barr virus and multiple sclerosis. PLoS One. 2013;8:e61110.

92. Levin, L. I. *et al.* Primary infection with the Epstein-Barr virus and risk of multiple sclerosis. Ann Neurol. 2010; 67:824-30.

93. Cohen, J.I. Epstein-Barr virus infection. N Engl J Med. 2000;343:481-92.

94. Thacker, E. L. *et al*. Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis. Ann Neurol. 2006;59:499-503.

95. Thacker, E. L. *et al.* Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis. Ann Neurol. 2006;59:499-503.

96. Ascherio, A. *et al.* Epstein-Barr virus antibodies and risk of multiple sclerosis: A prospective study. JAMA 2001; 286:3083-8.

97. Levin, L. I. *et al.* Temporal relationship between elevation of Epstein-Barr virus antibody titers and initial onset of neurological symptoms in multiple sclerosis. JAMA. 2005;293:2496-500.

98. Lang, H. L. E. *et al.* A functional and structural basis for TCR cross-reactivity in multiple sclerosis. Nat Immunol. 2002; 3:940-3.

99. Holmøy, T. & Vartdal, F. Cerebrospinal fluid T cells from multiple sclerosis patients recognize autologous Epstein-Barr virus-transformed B cells. J. Neurovirol. 2004;10:52-6.

100. Rand, K. H. *et al.* Epstein-Barr virus nuclear antigen-1 (EBNA-1) associated oligoclonal bands in patients with multiple sclerosis. J Neurol Sci. 2000; 173:32-9.

101. Lünemann, J. D. *et al.* Increased frequency and broadened specificity of latent EBV nuclear antigen-1-specific T cells in multiple sclerosis. Brain. 2006; 129:1493-506.

102. Cepok, S. *et al.* Identification of Epstein-Barr virus proteins as putative targets of the immune response in multiple sclerosis. J Clin Invest.2005;115:1352-60.

103. Goldacre, M. J. *et al.* Multiple sclerosis after infectious mononucleosis: Record linkage study. J Epidemiol Community Health. 2004;58:1032-5.

104. Rubicz, R. *et al.* A Genome-wide integrative genomic study localizes genetic factors influencing antibodies against Epstein-Barr virus nuclear antigen 1 (EBNA-1). PLoS Genet. 2013; 9:e1003147.

105. Rubicz, R. *et al.* Genetic factors influence serological measures of common infections. Hum Hered. 2011;72:133-41.

106. Sallah, N. *et al.* Whole-genome association study of antibody response to Epstein-Barr virus in an African population: a pilot. Glob Heal Epidemiol Genom. 2017;2:e18.

107. Zhou, Y. *et al.* Genetic loci for Epstein-Barr virus nuclear antigen-1 are associated with risk of multiple sclerosis. Mult Scler. 2016;22:1655-1664.

108. Kreft, K. L. *et al.* Elevated EBNA-1 IgG in MS is associated with genetic MS risk variants. Neurol. Neuroimmunol. NeuroInflamm. 2017;4:e406.

109. Fagnani, C. *et al.* Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. Mult Scler. 2015;21:1404-13.

110. Mescheriakova, J. Y. *et al.* Burden of genetic risk variants in multiple sclerosis families in the Netherlands. Mult Scler J – Exp Transl Clin. 2016; 2:2055217316648721.

111. DeWitt, J. C. & Luebke, R. W. Immunological Aging. in Comprehensive Toxicology: Third Edition. 2018.

112. Weinberger, B. *et al.* Biology of immune responses to vaccines in elderly persons. Clin Infect Dis. 2008;46:1078-84.

113. Frasca, D. & Blomberg, B. B. Effects of aging on B cell function. Curr Opin Immunol. 2009;21:425-30.

114. Baranzini, S. E. The genetics of autoimmune diseases: a networked perspective. Curr Opin Immunol. 2009; 21:596-605.

115. Cotsapas, C. *et al.* Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 2011;7:e1002254.

116. Dobson, R. & Giovannoni, G. Autoimmune disease in people with multiple sclerosis and their relatives: A systematic review and meta-analysis. J Neurol. 2013;260:1272-85.

117. Broadley, S. A. *et al.* Autoimmune disease in first-degree relatives of patients with

multiple sclerosis. A UK survey. Brain 2000;123:1102-11.

118. Henderson, R. D. *et al.* The occurrence of autoimmune diseases in patients with multiple sclerosis and their families. J Clin Neurosci. 2000;7:434-7.

119. Ramagopalan, S. V. *et al.* Autoimmune disease in families with multiple sclerosis: a population-based study. Lancet Neurol. 2007;6:604-10.

120. Barcellos, L. F. *et al.* Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. Lancet Neurol. 2006; 5:924-31.

121. Midgard, R. *et al.* Multiple sclerosis and chronic inflammatory diseases. A case-control study. Acta Neurol Scand. 1996;93:322–328.

122. Laroni, A. *et al.* Multiple sclerosis and autoimmune diseases: Epidemiology and HLA-DR association in North-east Italy. J. Neurol. 2006;253:636-9.

123. Nielsen, N. M. *et al.* Autoimmune diseases in patients with multiple sclerosis and their first-degree relatives: A nationwide cohort study in Denmark. Mult Scler. 2008;14:823-9.

124. Deretzi, G. *et al.* Familial prevalence of autoimmune disorders in multiple sclerosis in Northern Greece. Mult Scler. 2010;16:1091-101.

125. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature. 2010;467:1061-73.

126. International Multiple Sclerosis Genetics Consortium. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. Cell. 2018;175:1679-1687.e7.

127. Zuk, O. *et al.* The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci. 2012;109:1193-8.

128. Blanco-Gómez, A. *et al.* Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. BioEssays. 2016;38:664-73.

129. Gandhi, K. S. *et al.* The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dysregulation of specific T cell pathways in pathogenesis. Hum Mol Genet. 2010;19:2134-43.

130. Ottoboni, L. *et al.* An RNA profile identifies two subsets of multiple sclerosis patients differing in disease activity. Sci Transl Med. 2012;4:153ra131.

131. Stuart, T. & Satija, R. Integrative single-cell analysis. Nat Rev Genet. 2019;20:257-272.

132. Jäkel, S. *et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. Nature. 2019; 566:543-547.

133. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. Nature. 2019;573:75-82.

134. Schafflick, D. *et al.* Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis. Nat. Commun. 2020;11:247.

135. Graves, M. C. *et al.* Methylation differences at the HLA-DRB1 locus in CD4+ T-Cells are associated with multiple sclerosis. Mult Scler. 2014; 20:1033-41.

136. Huynh, J. L. *et al.* Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. Nat Neurosci. 2014;17:121-30.

137. Bos, S. D. *et al.* Genome-wide DNA methylation profiles indicate CD8+ T cell hypermethylation in multiple sclerosis. PLoS One. 2015;10:e0117403.

138. Zheleznyakova, G. Y. *et al.* Epigenetic research in multiple sclerosis: Progress, challenges, and opportunities. Physiol Genomics. 2017;49:447-461.

139. Andlauer, T. F. M. *et al.* Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. Sci Adv. 2016; 2:e1501678.

140. International Multiple Sclerosis Genetics Consortium *et al.* Analysis of immune-

related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45:1353-60.

141. Kular, L. *et al.* DNA methylation as a mediator of HLA-DRB1 15:01 and a protective variant in multiple sclerosis. Nat Commun. 2018;9:2397.

142. Marabita, F. *et al.* Smoking induces DNA methylation changes in Multiple Sclerosis patients with exposure-response relationship. Sci Rep. 2017;7:14589.

143. Maglione, A. *et al.* The adaptive immune system in multiple sclerosis: an estrogen-mediated point of view. Cells. 2019;8:1280.

144. Davies, N. M. *et al.* Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. BMJ. 2018; 362:k601.

145. Mokry, L. E. *et al.* Vitamin D and risk of multiple sclerosis: A Mendelian tandomization study. PLoS Med. 2015;12:e1001866.

146. Rhead, B. *et al.* Mendelian randomization shows a causal effect of low Vitamin D on multiple sclerosis risk. Neurol Genet. 2016;2:e97.

147. Mokry, L. E. *et al.* Obesity and multiple aclerosis: A Mendelian randomization study. PLoS Med. 2016;13:e1002053.

148. Devorak, J. *et al.* Large differences in adiponectin levels have no clear effect on multiple sclerosis risk: A Mendelian randomization study. Mult Scler. 2017;23:1461-1468.

149. Gianfrancesco, M.A. *et al.* Causal effect of genetic variants associated with body mass index on multiple sclerosis susceptibility. Am J Epidemiol. 2017;185:162-171.

150. Roberts, R. Mendelian randomization studies promise to ahorten the journey to FDA approval. JACC Basic to Transl. Sci. 2018;3:690-703.

151. Berer, K. *et al.* Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. Nature. 2011;479:538-41.

152. Berer, K. *et al.* Gut microbiota from multiple sclerosis patients enables spontaneous autoimmune encephalomyelitis in mice. Proc Natl Acad Sci U S A. 2017;114:10719-10724.

153. Braniste, V. *et al.* The gut microbiota influences blood-brain barrier permeability in mice. Sci Transl Med. 2014;6:263ra158.

154. Fransen, F. *et al.* Aged gut microbiota contributes to systemical inflammaging after transfer to germ-free mice. Front Immunol. 2017;8:1385

155. Rea, K. *et al.* The microbiome: A key regulator of stress and neuroinflammation. Neurobiol Stress. 2016;4:23-33.

156. Tankou, S. K. *et al.* A probiotic modulates the microbiome and immunity in multiple sclerosis. Ann Neurol. 2018;83:1147-1161.

157. Cekanaviciute, E. *et al.* Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. Proc Natl Acad Sci U S A. 2017;114:10713-10718.

158. Uccelli, M. M. The impact of multiple sclerosis on family members: a review of the literature. Neurodegener dis manag. 2014;4:177-85.

159. Vogenberg, F. R. *et al.* Personalized medicine - Part 1: Evolution and development into theranostics. *P T.* 2010;35:560-76.

160. Goetz, L. H. & Schork, N. J. Personalized medicine: motivation, challenges, and progress. Fertil Steril.2018;109:952-963.

161. Di Ruscio, A. *et al.* Multiple sclerosis: Getting personal with induced pluripotent stem cells. Cell Death Dis. 2015;6:e1806.

162. Ziemssen, T. *et al.* Multiple sclerosis: Clinical profiling and data collection as prerequisite for personalized medicine approach. BMC Neurol. 2016;16:124.

163. Rothenbacher, D. *et al.* New opportunities of real-world data from clinical routine settings in life-cycle management of drugs:

example of an integrative approach in multiple sclerosis. Curr Med Res Opin. 2015;31:953-65.

164. MS registries: putting people with MS at the heart of research. Available at: https://www.msif.org/ms-registries-putting-people-with-ms-at-the-heart-of-research/. (Accessed: 10th June 2019)

165. Bebo, B. F. *et al.* Landscape of MS patient cohorts and registries: Recommendations for maximizing impact. Mult Scler. 2018;24:579-586.

PART VII

SUMMARY / SAMENVATTING

SUMMARY

Multiple sclerosis (MS) is a complex neurological disease that clusters in families. In fact, 20% of MS patients have an affected family member. In addition, first-degree family members are at a higher risk of developing MS compared to the general population. Etiology of MS includes both genetic and environmental factors. The strongest and most studied genetic risk factor of MS susceptibility is the *HLA-DRB1*1501* allele. For several decades, HLA-class II region was the only known genetic risk factor for MS. With the improvement of technology and application of genome-wide association studies, more than 200 other risk loci were discovered. Despite discovery of genetic and several other biomarkers, MS heritability is not completely explained and prediction of MS in apparently healthy family members is still very limited.

The studies described in this thesis are executed in a unique collection of Dutch MS patients and their family members who belong to the multiplex MS families. Our aim was to explore genetic (known and unknown) and environmental factors in MS patients and their family members, and to study how these factors relate to MS risk in family members. Next to this, we also studied how risk genes of adult MS patients can be compared to risk genes in children with MS.

In **chapter 1**, a general introduction on current insights of genetic epidemiology of MS is given. In **chapter 2**, we studied the utility of combined genetic risk score (GRS) based on the common risk variants and *HLA-DRB1*1501* in familial and sporadic MS for disease prediction. GRS based on 103 genetic risk factors including the *HLA-DRB1*1501* was significantly higher in MS multiplex patients than in sporadic cases and Dutch healthy controls. This difference in genetic burden was primarily driven by a higher *HLA-DRB1*1501* allele frequency in familial cases. As for MS prediction, the cumulative GRS was modest in discriminating MS from controls; it performed better in discriminating multiplex MS from healthy controls compared to sporadic MS versus healthy controls, suggesting an enrichment of the risk genes in the familial cases.

In **chapter 3** we explored the utility of cumulative GRS for disease prediction in children with monophasic Acquired Demyelinating Syndrome (ADS) and MS. We showed that GRS was higher in children with MS than in children with monophasic ADS. The combined effect of 57 SNPs exceeded the effect of *HLA- DRB1*1501* alone in the risk models for pediatric- and adult-onset MS. The discriminative ability of this risk score

to differentiate between children with MS and monophasic ADS was moderate, and was comparable with disease prediction score for adult MS versus healthy controls.

In **chapter 4.1** we studied rare coding variants by means of linkage analysis and whole exome sequencing in a large Dutch MS family with multiple affected family members in different generations. We found a rare missense variant (rs147213094) in the *FKBP6* gene which resided in the linkage area in the chromosomal region 7q11.22– q11.23. This SNP showed a significant segregation pattern with the disease in the family, but was not significantly associated with MS in a larger cohort of unrelated MS patients and controls. When the case-control group was divided in sporadic and multiplex cases, there was an increased frequency of SNP in the MS multiplex families, but not in sporadic MS patients. Although literature shows some evidence of a possible association of *FKBP6* with MS, it is yet unknown how it might be related to MS pathogenesis.

Whole exome sequencing (WES) is the technique for detection of rare coding variants, however it can miss out the actual disease-causing variant. In **chapter 4.2** we explained why used WES technique has its pitfalls.

Next to the degree of relatedness, elevated EBNA-1 antibodies also increase the risk of MS. In **chapter 5** we examined the immune response to Epstein-Barr virus in MS patients, their healthy full-siblings and unrelated healthy spouses. We found a gradient in the EBNA-1 IgG titres, being highest in MS patients, intermediate in siblings and lowest in spouses. The gradient was not influence by *HLA-DRB1*1501*, that on itself was independently associated with high EBNA-1 IgG titres. An *HLA-A*02* tagging SNP was not associated with EBNA-1 IgG titers, and did not explain the gradient. Furthermore, our study showed that EBNA-1 IgG titres were higher in young MS patients compared to elderly MS patients, suggesting a role for immunosenescence. We found an association between *HLA-DRB1*1501* and high EBNA-1 IgG response in elderly MS patients and elderly siblings, suggesting that a certain genetic make-up is needed for the immune system to be able to maintain a proper immune response in elderly.

As MS has increased comorbidity with other autoimmune diseases, it has been shown that MS shares some of its genes with other autoimmune diseases. There are studies that suggested that first-degree relatives of MS patients could be at higher risk of autoimmune diseases other than MS due to shared genetics. In **chapter 6** we explored to what extent MS and other autoimmune diseases (AID) co-occur in MS patients and

their first-degree relatives from Dutch MS multiplex families. We showed that less than half of Dutch multiplex MS families reported one or more autoimmune diseases in the first-degree relatives. There were no differences in frequency of AID between MS patients, first-degree family members and spousal controls. Specifically, our study had shown that first-degree family members of MS patients are not at a higher risk of AID other than MS.

Main findings of this thesis are summarised and discussed in **chapter 7,** and suggestions are made for further research.

SAMENVATTING

Multiple sclerose (MS) is een complexe neurologische aandoening die zich in families clustert. Ongeveer 20% van MS-patiënten heeft een familielid met MS. Daarnaast hebben eerstegraads familieleden van MS-patiënten een hoger risico om zelf MS te ontwikkelen in vergelijking met de algemene bevolking. Bij de etiologie van MS zijn zowel genetische als omgevingsfactoren betrokken. Van alle risico genen wordt het dragerschap van het *HLA-DRB1*1501* (HLA-klasse II) allel het sterks geassocieerd met het risico op MS. Gedurende enkele decennia werd gedacht dat HLA-klasse II de enige bekende genetische risicofactor was voor MS. Met de technologische ontwikkelingen en door toepassing van genoomwijde associatie studies zijn meer dan 200 andere risico genen ontdekt. Ondanks deze ontwikkelingen en ontdekkingen van verschillende andere biomarkers, blijft het lastig om de erfelijkheid van MS volledig te verklaren en het ontstaan van MS te voorspellen bij ogenschijnlijk gezonde familieleden van MS-patiënten.

De in dit proefschrift beschreven onderzoeken zijn uitgevoerd in een unieke verzameling van Nederlandse MS-patiënten en hun familieleden die tot de multiplex MS-families behoren (dit zijn families waarin één of meer eerstegraads en/of tweedegraads familieleden zelf ook MS hebben). Wij hebben in deze families genetische (bekende en onbekende) factoren en omgevingsfactoren onderzocht, en gekeken hoe deze zich verhouden tot het risico op MS in familieleden. Daarnaast hebben we ook onderzocht hoe risicogenen van volwassen MS-patiënten kunnen worden vergeleken met risicogenen bij kinderen met MS.

In **hoofdstuk 1** worden de huidige inzichten in de genetische epidemiologie van MS besproken. In **hoofdstuk 2** wordt de bruikbaarheid van gecombineerde genetische risicoscore (GRS) voor de ziektevoorspelling bestudeerd in familiaire en sporadische MS; deze score is samengesteld uit niet-HLA risicogenen en *HLA-DRB1*1501*. GRS gebaseerd op 103 non-HLA risicogenen en *HLA-DRB1*1501* was significant hoger bij MS-patiënten uit multiplex families vergeleken met patiënten met sporadisch MS en gezonde controles. Dit verschil in 'genetische belastbaarheid' werd voornamelijk veroorzaakt door een hogere frequentie van *HLA-DRB1*1501* in familiaire MS. Wat betreft de voorspelling van de ziekte, scoorde de cumulatieve genetische risico score matig in het onderscheiden van MS-patiënten van gezonde mensen. Deze score presteerde beter bij het onderscheiden van familiaire MS van gezonde mensen, in

vergelijking met sporadische MS en gezonde controles. Dit gegeven suggereert dat er een verrijking is van de risicogenen in de familiaire MS.

In **hoofdstuk 3** werd onderzocht of cumulatieve GRS gebruikt kan worden voor de voorspelling van de ziekte in kinderen met monofasisch verworven demyeliniserend syndroom (Acquired Demyelinating Syndrome, ADS) en MS. We hebben aangetoond dat GRS gemiddeld hoger was bij kinderen met MS dan bij kinderen met monofasisch ADS, en dat GRS van monofasisch ADS was vergelijkbaar met gezonde controles. Het onderscheidend vermogen van deze risicoscore tussen kinderen met MS en monofasisch ADS was middelmatig en vergelijkbaar met de voorspellingsscore voor volwassen MS-patiënten en gezonde controles. Het gecombineerde effect van 57 risico genen overtrof het effect van *HLA-DRB1*1501* in het voorspellen van MS bij zowel kinderen als volwassenen.

In **hoofdstuk 4.1** hebben we zeldzame genetische varianten bestudeerd in een grote Nederlandse MS-familie waarin meerdere familieleden in verschillende generaties MS hadden. Dit werd gedaan middels linkage analyse en whole exome sequencing (WES). We vonden een zeldzaam genetisch variant (rs147213094) in het *FKBP6*-gen dat zich in het linkage gebied op chromosoom 7q11.22– q11.23 bevond. Dit genetische variant had een significant segregatiepatroon met de ziekte in de familie, maar was niet geassocieerd met MS in een grotere cohort van niet-verwante MS-patiënten en gezonde controles. Op het moment dat de case-controle groep verdeeld was in sporadische en familiare MS gevallen, was er een verhoogde frequentie van dit variant in familiaire MS, maar niet bij patiënten met sporadische MS. Hoewel er in de literatuur enig bewijs bestaat van een mogelijke associatie tussen *FKBP6* met MS, is het nog niet bekend op welke manier *FKBP6* betrokken is in de pathogenese van MS. WES kan gebruikt worden voor het opsporen van zeldzame genetische varianten, echter het heeft zijn valkuilen (**hoofdstuk 4.2**) en kan het werkelijke ziekte veroorzakende variant missen.

Naast de mate van verwantschap, zijn er omgevingsfactoren, zoals verhoogde EBNA-1 antilichamen, die het risico op MS verhogen. In **hoofdstuk 5** onderzochten we het immuunrespons op het Epstein-Barr-virus bij MS-patiënten, hun gezonde broers en zussen, en niet-verwante gezonde partners. We vonden een gradiënt in de EBNA-1 IgG-titers: deze was het hoogst bij MS-patiënten, gemiddeld bij broers en zussen, en het laagst bij partners. De gradiënt werd niet beïnvloed door de aan- of afwezigheid van *HLA-DRB1*1501,* die op zichzelf onafhankelijk geassocieerd was met hoge EBNA-1 IgG-

titers. EBNA-1 IgG-titers waren hoger bij jonge MS-patiënten vergeleken met oudere MS-patiënten, wijzend op een mogelijke rol voor immunosenescentie. We vonden ook een verband tussen *HLA-DRB1*1501* en hoog EBNA-1 IgG-respons bij oudere MS-patiënten en oudere broers en zussen.

MS heeft comorbiditeit met bepaalde auto-immuunziekten. Sommige genetische risico factoren die betrokken zijn bij MS, zijn ook geassocieerd met andere auto-immuun ziekten. Er zijn studies die suggereren dat eerstegraads familieleden van MS-patiënten daardoor een hoger risico hebben op andere auto-immuunziekten dan MS. In **hoofdstuk 6** is bekeken in hoeverre MS en andere auto-immuunziekten samen voorkomen bij MS-patiënten en hun eerstegraads familieleden in de Nederlandse MS-multiplex families. Minder dan de helft van de Nederlandse MS-families had een of meer auto-immuunziekten gemeld bij de eerstegraads familieleden. Er waren geen verschillen in frequentie van auto-immuun ziekten tussen MS-patiënten, eerstegraads familieleden en gezonde partners. Onze studie had aangetoond dat eerstegraads familieleden van MS-patiënten vooral geen hoger risico hebben op een auto-immuun aandoening anders dan MS vergeleken met normale populatie.

Tenslotte, de belangrijkste bevindingen van dit proefschrift worden samengevat en besproken in **hoofdstuk 7**, en worden er ideeën besproken voor verder onderzoek.

Epilogue

ABOUT THE AUTHOR

Julia Yevgenievna Mescheriakova was born on September 29th 1985, in Ashgabat, Turkmenistan (the former Soviet Union). In 1998 she moved with her family to the Netherlands. In her early years she attended secondary school at the Bernardinus College in Heerlen (Limburg). Hereafter she studied medicine at the Maastricht University Medical Center. After obtaining her medical degree in 2010, she worked at the department of neurology of Sint Franciscus Gasthuis in Rotterdam. In October 2011 she started her PhD research described in this thesis at the MS Centre ErasMS in Rotterdam under supervision of Prof. dr. R.Q. Hintzen (department of Neurology, Erasmus MC; head: Prof. dr. P.A.E. Sillevis Smitt). During her PhD research she coordinated the genetic family studies, and contributed to several national and international collaborations, such as the PROUD-study (Predicting the Outcome of demyelinating event study) and several international projects initiated by the International Multiple Sclerosis Genetics Consortium (IMSGC). Due to passing away of first promotor Prof. Hintzen, the supervision was taken over by Prof. dr. P.A. van Doorn (department of Neurology, Erasmus MC) and Prof. dr. ir. C.M. van Duijn (department of Epidemiology, Erasmus MC).

From January 2018 onwards, she works as a resident at the department of Rehabilitation at the Amsterdam UMC, location AMC in Amsterdam (head: Prof. dr. V. de Groot). She lives with her family in Leiden.

LIST OF PUBLICATIONS

Publications in this thesis

Mescheriakova JY, Kreft KL and Hintzen RQ. In: Multiple Sclerosis Immunology. A Foundation for Current and Future Treatments. Title: Genetics of multiple sclerosis.
Springer, 2013, ISBN 978-1-4614-7952-9, page 197-228.

Van Pelt E*, **Mescheriakova JY*,** Makani N, Ketelslegers IA, Neuteboom RF, Kundu S, Broer L, Janssens ACJW, Catsman-Berrevoets CE, van Duijn CM, Banwell B, Bar-Or A and Hintzen RQ. Risk genes associated with pediatric-onset MS but not with monophasic acquired CNS demyelination.
Neurology. 2013;81:1996-2001.

Mescheriakova JY, Broer L, Wahedi S, Uitterlinden AG, van Duijn CM and Hintzen RQ. Burden of genetic risk variants in multiple sclerosis families in the Netherlands.
Multiple Sclerosis Journal – Experimental, Translational and Clinical.
2016;2:2055217316648721.

Mescheriakova JY and Hintzen RQ. No excess of autoimmune diseases in multiple sclerosis families from the Netherlands.
Acta Neurologica Scandinavica. 2018;137:531-537.

Mescheriakova JY, Verkerk JMH, van Duijn CM and Hintzen RQ. Linkage analysis and exome sequencing identify a novel candidate gene in a Dutch multiple sclerosis family.
Multiple Sclerosis Journal. 2019;25:909-917.

Mescheriakova JY, Verkerk JMH, van Duijn CM and Hintzen RQ. Response to Vilariño-Güell et al.
Multiple Sclerosis Journal. 2019;25:1013-1015.

Mescheriakova JY, van Nierop GP, van der Eijk AA, Kreft KL and Hintzen RQ. EBNA-1 titer gradient in multiple sclerosis families indicates a genetic contribution.
Neurology: Neuroimmunology & Neuroinflammation. 2020;7:e872.

*** shared first author**

Other publications

van der Vuurst de Vries RM, **Mescheriakova JY**, Runia TF, Jafari N, Siepman TAM and Hintzen RQ. Soluble CD27 levels in CSF as a prognostic biomarker in clinically isolated syndrome. *JAMA Neurology. 2017;74:286-292.*

van der Vuurst de Vries RM*, van Pelt ED*, **Mescheriakova JY**, Wong YYM, Ketelslegers YA, Siepman TAM, Catsman CE, Neuteboom RF and Hintzen RQ. Disease course after CIS in children versus adults: a prospective cohort study. *European Journal of Neurology. 2017;24:315-321.*

van der Vuurst de Vries RM, **Mescheriakova JY**, Runia TF, Siepman TAM, Wokke BHA, Samijn JPA and Hintzen RQ. Smoking at time of CIS increases the risk of clinically definite multiple sclerosis. *Journal of Neurology. 2018;265:1010-1015.*

van der Vuurst de Vries RM, **Mescheriakova JY**, Wong YYM, Runia TF, Jafari N, Samijn JP, de Beukelaar JWK, Wokke BHA, Siepman TAM and Hintzen RQ. Application of the 2017 revised McDonald criteria for multiple sclerosis to patients with a typical clinically isolated syndrome. *JAMA Neurology. 2018;75:1392-1398.*

van der Vuurst de Vries RM, van den Dorpel JJA, **Mescheriakova JY**, TF, Jafari N, Siepman TAM and Hintzen RQ. Fatigue after a first attack of suspected multiple sclerosis. *Multiple Sclerosis Journal. 2018;24:974-981.*

van der Vuurst de Vries RM, Wong YYM, **Mescheriakova JY,** van Pelt ED, Runia TF, Jafari N, Siepman TA, Melief MJ, Wierenga-Wolf AF, van Luijn MM, Samijn JP, Neuteboom RF and Hintzen RQ. High neurofilament levels are associated with clinically definite multiple sclerosis in children and adults with clinically isolated syndrome. *Multiple Sclerosis Journal. 2019;25:958-967.*

Rijkers K*, **Mescheriakova J*,** Majoie HM, Lemmens EMP, van Wijk XMR, van Winden L, Schijns OEMG, Vles JSH and Hoogland G. Polymorphisms in CACNA1E and CamkIId are associated with seizure susceptibility of Sprague-Dawley rats. *Epilepsy Research. 2010;91:28-34.*

Lill CM, Luessi F, Alcina A, Sokolova EA, Ugidos N, de la Hera B, Guillot-Noël L, Malhotra, Reinthaler E, Schjeide BM, **Mescheriakova JY**, Mashychev A, Wohlers I, Akkad DA, Aktas O, Alloza I, Antigüedad A, Arroyo R, Astobiza I, Blaschke P, Boyko AN, Buttmann M, Chan A, Dörner T, Epplen JT, Favorova OO, Fedetz M, Fernández O, García-Martínez A, Gerdes LA, Graetz C, Hartung HP, Hoffjan S, Izquierdo G, Korobko DS, Kroner A, Kubisch C, Kümpfel T, Leyva L, Lohse P, Malkova NA, Montalban X, Popova EV, Rieckmann P, Rozhdestvenskii AS, Schmied C, Smagina IV, Tsareva EY, Winkelmann A, Zettl UK, Binder H, Cournu-Rebeix, Hintzen R, Zimprich A, Comabella M, Fontaine B, Urcelay E, Vandenbroeck K, Filipenko M, Matesanz F, Zipp F and Bertram L Genome-wide significant association with seven novel multiple sclerosis risk loci. *Journal of Medical Genetics. 2015;52:848-55.*

Lemmens E, Schijns O, Rijkers K, **Mescheriakova J,** van Wijk X, Beuls E and Hoogland G. Lange-termijn effecten van febriele convulsies in de neonatale rat. *Epilepsie: periodiek voor professionals. 2009;7:11-14.*

Co-authoring papers from International Multiple Sclerosis Genetics Consortium

International Multiple Sclerosis Genetics Consortium (IMSGC) *et al.* MANBA, CXCR5, SOX8, RPS6KB1, and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain. 2013;136:1778-82.*

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

International Multiple Sclerosis Genetics Consortium (UMSGC) *et al.* Low-frequency and rare-coding variation contributes to multiple sclerosis risk. *Cell. 2018;175:1679-1687.e7*

International Multiple Sclerosis Genetics Consortium (IMSGC) *et al.* Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science. 2019;365:eaav7188.*

PHD PORTFOLIO

1. PhD training	Year	Workload (ECTS)
Courses		
Genetics for dummies, MolMed Erasmus MC	2011	0.6
SNP VIII course, MolMed Erasmus MC	2011	2.0
BROK-course	2012	1.5
Advances in Genome-Wide Association Studies (GEO3, NIHES)	2012	1.4
Principles of Genetic Epidemiology (NIHES)	2012	0.7
Next Generation Sequencing data analysis (MolMed)	2012	1.0
Basic course on SPSS (MolMed)	2012	1.0
Basic Immunology Course, Department of Immunology EMC	2012	0.5
EDSS training	2014	0.2
Genetic Course in MS, ECTRIMS Summerschool (Talin, Estland)	2014	2.0
Biomedical English Writing Course (MolMed)	2016	4.0
Oral presentations		
Symposium 10-jarig bestaan ErasMS "Trots in Translatie"	2012	1.2
MS-goes Live, Sittard	2012	0.5
MS-goes Live, Rotterdam	2013	0.5
Referaat, dep. Neurology, Erasmus MC	2013	1.2
Annual MS meeting of the Dutch MS Research foundation (Hasselt, Belgium)	2013	1.2
Informatiebijeenkomst, MS-vereniging Gouda	2015	1.0
Informatiebijeenkomst, MS-vereniging Sassenheim	2016	1.0
Regionale MS Bijeenkomst voor de neurologen	2016, 2018	2.4
Poster presentations		
17the MolMed day, Erasmus MC (1 poster)	2013	1.0
ECTRIMS congress, Copenhagen Denmark (1 poster)	2013	1.0
ECTRIMS congress, Boston USA (1 poster)	2014	1.0
ECTRIMS congress, Paris, France (1 poster)	2017	1.0
(Inter)national conferences		
MS-symposium, VUMC Amsterdam	2016	0.5
ECTRIMS Congress (Amsterdam, Lyon, Copenhagen, Boston, Barcelona, Paris)	2011-2015, 2017	6.0
Annual MS meeting of the Dutch MS Research foundation (Hasselt, BE 2013; Oegstgeest, NL 2015)	2013, 2015	1.0
IMSGC-meeting Paris, France	2013	1.0
2. Teaching		
Supervising master thesis of a medical student	2015	3.0
Lectures for minor students	2012	0.5
3. Other		
Review of various papers for international peer-reviewed journals	2014-2016	0.6
Investigating doctor in clinical MS trials (FTY-720, Synergy, ORATORIO)	2011-2016	5.0

LIST OF ABBREVIATIONS

ADEM	acute disseminated encephalomyelitis
ADS	acquired demyelinating syndromes
AID	autoimmune diseases
AITD	autoimmune thyroid disease
AUC	area under the curve
BBB	blood-brain barrier
CAD	coronary artery disease
CD	Crohn's disease
CIS	clinically isolated syndrome
CNS	central nervous system
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CVD	cardiovascular disease
EAE	experimental autoimmune encephalomyelitis
EBNA-1	Epstein-Barr virus nuclear antigen 1
EBV	Epstein-Barr virus
EDSS	expanded disability status scale
ENCODE	Encyclopedia of DNA Elements
GEMS	study on gene-environment interaction in MS
GRS	genetic risk score
wGRS	weighted genetic risk score
uwGRS	unweighted genetic risk score
GWAS	genome wide association study
HCV	hepatitis C virus
HLA	Human Leukocyte Antigen
IgG	Immunoglobulin G
IM	infectious mononucleosis
IMSGC	International Multiple Sclerosis Genetics Consortium
iPSC	induced pluripotent stem cell
LOD	logarithm of odds
MAF	minor allele frequency
MHC	major histocompatibility complex
MOG	myelin oligodendrocyte glycoprotein
MR	mendelian randomization
MSSS	MS Severity Score

NMOSD	neuromyelitis optica spectrum disorders
ON	optic neuritis
OR	odds ratio
PA	pernicious anemia
PPI	protein-protein interactions
PPMS	primary progressive multiple sclerosis
QC	quality control
RA	rheumatoid arthritis
RNA	ribonucleic acid
ROC	receiver operating characteristic
RRMS	remitting-relapsing multiple sclerosis
RSIII	the longitudinal Rotterdam Study III
SLE	systemic lupus erythematosus
SNP	single nucleotide polymorphism
SPMS	secondary progressive multiple sclerosis
SS	Sjögren's syndrome
T1D	type 1 diabetes
TM	transverse myelitis
UC	ulcerative colitis
VDR	vitamin D receptor
VDRE	vitamin D response element
WES	whole exome sequencing
WGS	whole genome sequencing

DANKWOORD

Ten tijde van het schrijven van dit laatste hoofdstuk van mijn proefschrift, komt het besef dat mijn proefschrift eindelijk klaar is. Het was een zeer bewogen periode van mijn leven met zijn vele up- en enkele downmomenten. Veel mensen hebben een belangrijke rol gespeeld in het tot stand komen van dit proefschrift die ik graag wil bedanken.

Allereerst mijn promotor die mij voor het grootste deel van het PhD-traject heeft begeleid, Prof. dr. R.Q. Hintzen. Beste Rogier, ik wil je bedanken dat je me ooit een mogelijkheid hebt geboden om deel uit te maken van het MS-centrum en je onderzoeksteam, waar het voelde als thuiskomen. Ik heb altijd je passie voor het MS-vak bewonderd en je moed om tegen de geldende opinies in te gaan. Je was altijd heel goed op de hoogte van wat er speelde en wie je voor wat kon vragen. Door je zeer sociale instelling kende je veel mensen in de onderzoekswereld en erbuiten. Daarnaast was je een gezellig mens, die het belangrijk vond om ook kleine successen met elkaar te vieren. Dank voor je gezelligheid en je humor tijdens de congressen, borrels en allerlei uitjes. Tevens dank voor je begeleiding, begrip en steun op de momenten dat mijn (Phd)leven even anders liep. Ook jouw eigen leven is anders gelopen… Dank voor je zorgen dat de supervisie van mijn PhD in goede handen terecht kwam, namelijk die van Prof. dr. P.A. van Doorn en Prof. dr. ir. C.M. van Duijn. Het doet mij verdriet dat je niet aanwezig zal zijn op de voor mij speciale dag. Je wordt nog elke dag gemist door ons allen.

Mijn huidige eerste promotor, Prof. dr. P.A. van Doorn. Beste Pieter, wat was ik blij toen ik hoorde dat jij de supervisie van mijn PhD-traject wilde overnemen na het overlijden van Rogier. Bedankt voor je bereidheid om dit te doen. Ook dank voor je begeleiding en je sturing in dit laatste stukje van dit traject, je oog voor details, je heldere en kritische blik bij het lezen van dit proefschrift.

Veder gaat mijn dank uit naar mijn andere promotor, Prof. dr. ir. C.M. van Duijn. Beste Cornelia, met een snelle vaart heb jij de draad weer opgepakt waar Rogier het had moeten achterlaten. Dank voor je bereidheid om dit te doen, je kritisch blik en inhoudelijke bijdrage aan dit proefschrift. Jij en Pieter hebben mij net dat ene zetje gegeven om dit traject tot een succesvol einde te brengen.

Ik wil graag de leden van de leescommissie bedanken voor het kritisch lezen van het proefschrift: Prof. dr. J.D. Laman, Prof. dr. C.C.W. Klaver en Prof. dr. A.G. Uitterlinden.

Prof. dr. J.D. Laman, beste Jon, bedankt voor je enthousiasme en je supersnelle reactie en tips n.a.v. het lezen van mijn proefschrift. Daarnaast ook dank aan de overige commissieleden, namelijk: dr. J.J.F.M. Smolders, Prof. dr. C.J. van der Woude, Prof. dr. R. Willemsen, Prof. dr. M. D'hooghe, voor het plaats nemen in mijn promotiecommissie.

Dank aan alle patiënten en hun familieleden die hebben meegedaan aan de familiestudie. Zonder jullie inzet en bijdrage zou het onderzoek, beschreven in dit proefschrift, niet mogelijk zijn. In het bijzonder dank aan de familie Koster, welke financieel heeft bijgedragen aan het onderzoek in dit proefschrift.

Collegae van het MS-centrum: Dorine, Yuyi, Roos, Danielle, Tessel, Naghmeh, Immy, Rinze, Sarita, Hanne, Jacqueline, Josje, Yvonne en Leonieke. Ik ben erg trots dat ik destijds deel heb uitgemaakt van deze groep. Dank voor de fijne samenwerking en voor alle gezelligheid. Een paar mensen wil ik in het bijzonder bedanken.

Dorine: wat heb ik van jou onwijs veel geleerd over MS! Ik waardeer je behulpzaamheid, je geduld, je nauwkeurigheid, je laagdrempeligheid en al je tijd die je aan mij hebt besteed om mij een betere dokter te maken. Niet te vergeten je lieve attentie bij de geboorte van mijn zoontje. Dankjewel voor alles!

Tessel: ik ben blij dat jij er was toen ik startte met mijn promotieonderzoek. Dank dat ik op je terug kon vallen bij vragen over onderzoek en PROUD-patiënten. Dank voor je humor en gezelligheid bij uitstapjes en congressen.

Danielle: mijn kamergenootje, borrelmaatje op congressen, mede treating physician voor de studiepatiënten, coauteur op onze eerste grote publicatie waar we een fles bubbels op hebben gedronken. Dank voor je aanwezig in dit hele traject, je steun, je gezelligheid en een hele fijne samenwerking op meerdere vlakken binnen dit MS-team.

Roos: wat was ik blij toen jij bij ons onderzoek kwam doen, als opvolgster van Tessel. Je hebt met heel veel enthousiasme de PROUD-studie opgepakt. Het was heel fijn samenwerken met jou! Dank voor alle leuke, gezellige, leerzame momenten op de poli, tijdens congressen, etentjes en MS-evenementen. Dank voor je positiviteit, je steun en je goede raad wanneer ik dat nodig had.

YuYi: ik ben trots dat ik je heb leren kennen als mijn fijne MS-collega, en ook dichtbij als een heel prettig persoon die een ongelooflijk doorzettingsvermogen heeft, veel voor mekaar krijgt, een organisatietalent heeft (of het nou gaat om een etentje organiseren bij haar thuis of een babyshower), en een perfect oriënterend vermogen heeft (soms beter dan googlemaps). Samen met Joyce waren wij de leden van de 'cosy table', waar veel gesprekken over ons onderzoeksleven, privé, mannen, make-up en vooral het eten werden gevoerd. Vaak hebben Joyce en ik moeten lachen om je 'me, myself and I' momentjes waarbij die grote koptelefoon op je hoofd ging, en waarbij wij aan de snelheid en het geluid van jouw typkunsten konden afleiden of het wel of niet je dag was. Dames, dank voor jullie gezelligheid en humor! Ik denk er nog vaak met plezier aan terug.

Collegae van immunologie en virologie: Karim, Gijs, Marvin, Malou, Jamie, Liza, Marie-José en Annet. Dank voor het verwerken van de bloedsamples voor onze studies. Daarnaast ook voor jullie gezelligheid in de wandelgangen, borrels en evenementen i.k.v. MS-onderzoek. In het bijzonder dank aan Karim voor de fijne samenwerking, meeschrijven aan twee hoofdstukken van dit boekje, en voor de fijne aanvullingen na het kritisch lezen van mijn proefschrift.

Collegae op de afdeling epidemiologie en genetica: Jeanette, Andy, Michael, Pascal, en Ramazan. Dank voor het bewerken en opslaan van de DNA-samples.

Kamergenootjes op de 22e door de jaren heen: Stephan, Immy, Tessel, Yuyi, Joyce, Bianca, Alex, Carina, Christine, Christa, Esther, Laurike, Danielle v P, Danielle B, Marienke, Agnes, Roos, Marieke, Hanne. Dank voor de mooie en gezellige lunch- en koffiemomenten met lekkere taartjes en eigen baksels, viering van grote en kleine mijlpalen, limonchello en vodka met zure matjes- momentjes op vrijdagmiddagen, samen lopen naar de 22e als de lift weer eens uitviel, samen klagen als de vriezertelefoon weer eens afging midden in de nacht, na vertrek van Stephan alleen vrouwen in onze kamer gehad, waarbij er op een gegeven moment nog meer vrouwen bij kwamen en de kamer werd door een van de supervisoren (ik meen Rogier) omgedoopt tot 'een kippenhok'. Mijn spin/sportmaatjes (" the Ed groupies"): Bianca en Janneke, dank voor de hilarische momenten op de spinningfiets. Ik kijk met veel plezier hiernaar terug!
Mijn huidige opleiders binnen revalidatiegeneeskunde Herman, Rutger en Judith, dank voor de mogelijkheid die jullie mij hebben geboden binnen de opleiding om aan mijn proefschrift te werken.

Lieve familie, schoonfamilie en vrienden, dank voor jullie interesse in mijn onderzoek, steun en begrip afgelopen jaren. Mijn lieve ouders, mama en Rob, dank voor jullie onvoorwaardelijke liefde en jullie vertrouwen in mij. Mam, ik weet dat je heel trots op me bent en ik kan nu eindelijk van de daken schreeuwen 'het is zover'. Onze gezamenlijke droom wordt eindelijk werkelijkheid.

Peggy, mijn beste vriendin en vandaag als mijn paranimf aanwezig. Dank voor je eindeloze steun, je raad, je humor en onze gezellige samenzijn momenten afgelopen 11 jaar. Jij als niemand anders, begrijpt heel goed wat het doen van onderzoek inhoudt. Wat ben ik blij en dankbaar dat je vandaag naast me staat en dit belangrijke moment samen met mij viert.

Mijn kleine Gabriel (2019), the apple of my eye. Je brengt elke dag vreugde en positiviteit in mijn leven, wat mij geholpen heeft om door te zetten. Ik kijk ernaar uit om je te zien opgroeien en samen met jou onze 'journey' af te gaan leggen.

Lieve Stephen, hier sta je dan, naast mij vandaag, mijn steun en toeverlaat. Ik kan je niet vaak genoeg zeggen hoe blij ik ben dat jij er in mijn leven bent. Je dagelijkse portie humor, enthousiasme, geduld, relativeringsvermogen en liefde hebben mij geholpen om zover te komen. Dank voor al je hulp met het regelen van vele zaken rondom mijn promotie. Ik kijk uit naar de komst van onze klein meid in februari '22 en dat we samen een compleet gezinnetje gaan vormen.

GENETIC AND
ENVIRONMENTAL
RISK FACTORS IN

MULTIPLE SCLEROSIS MULTIPLEX FAMILIES

Multiple sclerosis (MS) is a complex neurological disease that clusters in families. First-degree family members of MS patients are at a higher risk of developing MS compared to the general population.

Despite advances in technology, MS heritability is not completely understood and prediction of MS in apparently healthy family members is still very limited.

The aim of this thesis was to explore genetic (known and unknown) and environmental factors in Dutch MS patients and their family members, and to study how these factors relate to MS risk in family members.

In addition, we studied how risk genes of adult MS patients can be compared to risk genes in children with MS.