

## Whole exome sequencing in multi-incident families identifies novel candidate genes for multiple sclerosis

J. Horjus, B.Sc. <sup>1,\*</sup>, T. Banda, M.Sc. <sup>1,\*</sup>, M. Heerings, M.A.N.P. <sup>1,\*</sup>, M. Hakobjan, B.A.S. <sup>1</sup>, W. De Witte, B.Sc. <sup>1</sup>, D.J. Heersema, M.D., Ph.D. <sup>2</sup>, A.J. Jansen, M.D. <sup>3</sup>, E.M.M. Strijbis, M.D., Ph.D. <sup>4</sup>, B.A. de Jong, M.D., Ph.D. <sup>5</sup>, A.E.J. Slettenaar, M.A.N.P. <sup>6</sup>, E.M.P.E. Zeinstra, M.D., Ph.D. <sup>7</sup>, E.L.J. Hoogervorst, M.D., Ph.D. <sup>8</sup>, B. Franke, Ph.D. <sup>1,9,10</sup>, W. Kruijer, Ph.D. <sup>11</sup>, P.J. Jongen, M.D., Ph.D. <sup>12,13</sup>, L. Visser, M.D., Ph.D. <sup>14,15</sup>, G. Poelmans, M.D., Ph.D. <sup>1</sup>

<sup>1</sup>Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>2</sup>Department of Neurology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

<sup>3</sup>Department of Neurology, Bravis Hospital, Bergen op Zoom, The Netherlands

<sup>4</sup>Department of Neurology, Amsterdam UMC, location VUmc, Amsterdam, The Netherlands

<sup>5</sup>Department of Neurology, Amsterdam UMC, Vrije Universiteit Amsterdam, MS Center Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands

<sup>6</sup>Department of Neurology, Medisch Spectrum Twente, Enschede, The Netherlands

<sup>7</sup>Department of Neurology, Isala Hospital, Meppel, The Netherlands

<sup>8</sup>Department of Neurology, St Antoniusziekenhuis, Utrecht/Nieuwegein, the Netherlands

<sup>9</sup>Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

<sup>10</sup>Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>11</sup>Independent Life Science Consultant, Leusden, The Netherlands

<sup>12</sup>MS4 Research Institute, Nijmegen, The Netherlands

<sup>13</sup>Department of Community & Occupational Medicine, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands

<sup>14</sup>Department of Neurology, St. Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands

<sup>15</sup>Department of Care Ethics, University of Humanistic Studies, Utrecht, The Netherlands

\* equal contribution

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

## **Funding**

This work was supported by the National Multiple Sclerosis Foundation, The Netherlands.

## **Conflicts of interest**

P.J. Jongen received honoraria from Bayer Netherlands for consulting activities. G. Poelmans is director of Drug Target ID, Ltd. All other authors do not report any conflicts of interest.

## Abstract

Multiple sclerosis (MS) is a degenerative disease of the central nervous system in which auto-immunity-induced demyelination occurs. MS is thought to be caused by a complex interplay of environmental and genetic risk factors. While most genetic studies have focused on identifying common genetic variants for MS through genome-wide association studies, the objective of the present study was to identify rare genetic variants contributing to MS susceptibility. We used whole exome sequencing (WES) followed by co-segregation analyses in nine multi-incident families with 2 to 4 affected individuals. WES was performed in 31 family members with and without MS. After applying a suite of selection criteria, co-segregation analyses for a number of rare variants selected from the WES results were performed, adding 24 family members. This approach resulted in 12 exonic rare variants that showed acceptable co-segregation with MS within the nine families, implicating the genes *MBP*, *PLK1*, *MECP2*, *MTMR7*, *TOX3*, *CPT1A*, *SORCS1*, *TRIM66*, *ITPR3*, *TTC28*, *CACNA1F*, and *PRAM1*. Of these, three genes (*MBP*, *MECP2*, and *CPT1A*) have been previously reported as carrying MS-related rare variants. Six additional genes (*MTMR7*, *TOX3*, *SORCS1*, *ITPR3*, *TTC28*, and *PRAM1*) have also been implicated in MS through common genetic variants. The proteins encoded by all twelve genes containing rare variants interact in a molecular framework that points to biological processes involved in (de-/re-)myelination and auto-immunity. Our approach provides clues to possible molecular mechanisms underlying MS that should be further studied in cellular and/or animal models.

## Introduction

Multiple Sclerosis (MS) is one of the most common neurological diseases, affecting over two million people globally and showing an increasing incidence <sup>1</sup>. Auto-immunity pathways play an important role in the pathogenesis of MS <sup>2</sup>. The focal areas of inflammation associated with MS are often characterized by reduced integrity of the blood-brain barrier (BBB), leading to infiltration of several types of immune cells, including myelin-specific autoreactive T-cells and macrophages <sup>3,4</sup>. This auto-immune response results in the activation of microglia, which in turn promotes further neuroinflammation and demyelination through the release of proinflammatory cytokines <sup>5</sup>. Additionally, proinflammatory mediators released by activated T-cells and microglia cause damage to oligodendrocytes (ODCs) that produce and maintain the myelin sheaths around axons in the central nervous system (CNS) <sup>6</sup>. The MS-associated demyelinating events in the CNS and the inability to remyelinate the resulting lesions eventually lead to axonal and neuronal degeneration <sup>7</sup>. Depending on the extent and location of the MS lesions, a variety of symptoms can occur, such as motor, sensory, cognitive, and visual symptoms and, eventually, severe overall disability.

In case of remyelination, which is a normal process in healthy individuals, the underlying axons remain intact and neurodegeneration is prevented. Remyelination is achieved through the differentiation of progenitor cells into mature ODCs <sup>8</sup>. Microglia and macrophages, which are also able to release anti-inflammatory cytokines in addition to the abovementioned proinflammatory cytokines <sup>9</sup>, have been found to stimulate ODC differentiation from progenitor cells, therefore contributing to remyelination <sup>10</sup>. Remyelination can occur in the remission phases in the relapsing-remitting form of MS (RRMS), the (initial) subtype found in approximately 90% of MS cases <sup>11</sup>. In RRMS, relapses characterized by (various) neurological symptoms due to inflammatory and active demyelinating lesions alternate with periods in which these symptoms are partially or completely absent. In 15-80% of MS cases, the disease course becomes progressive within 10 to 15 years, which is called secondary progressive MS (SPMS) <sup>12,13</sup>. The small group of patients with primary progressive MS (PPMS) experiences uninterrupted progression from disease onset.

Progressive forms of MS are pathologically characterized by brain atrophy, which can be linked to axonal and neuronal degeneration due to the lack of remyelination <sup>14</sup>.

MS is thought to be caused by a complex interplay of multiple environmental and genetic risk factors that contribute to disease onset and progression <sup>15</sup>. Environmental risk factors for MS include Epstein-Barr virus infection <sup>16</sup>, vitamin D deficiency, and cigarette smoking <sup>17</sup>. Females are affected up to three times more often than males, and this is assumed to be largely due to environmental factors <sup>18</sup>. Furthermore, first-degree relatives of patients have a 20–40-fold risk of developing MS compared to the general population, and this risk is 300-fold for monozygotic twins <sup>19</sup>. The genetic contribution to MS susceptibility seems to be in part polygenic, i.e. a large number of common variants within multiple genes each have a small contribution to overall disease risk <sup>20</sup>. Large-scale genome-wide association studies (GWASs) have shown that the common genetic risk for MS is for the largest part due to variance in the Major Histocompatibility Complex (MHC) gene cluster, including multiple human leucocyte antigen (*HLA*) genes <sup>21</sup>. These genes encode proteins with important immunological functions, which is in line with the characterization of MS as an auto-immune disease. A total of 20–30% of MS heritability can be explained by common variants within *HLA* genes <sup>22,23</sup>. The most recent and largest meta-analysis of GWASs of MS analysed 8,278,136 common genetics variants (i.e., single nucleotide polymorphisms, SNPs) in 47,429 MS cases and 68,374 controls. This meta-GWAS identified 233 SNPs that showed genome-wide significant association with MS ( $P < 5.00E-08$ ), including 200 autosomal SNPs in genes outside the MHC cluster and one SNP on the X-chromosome <sup>23</sup>.

In addition to common genetic variants, rare variants have been shown to contribute to MS risk. Next-generation sequencing methods, especially whole exome or whole genome sequencing (WES or WGS), are considered to be the elementary technology for the discovery of rare variants involved in diseases, but exome chips are also frequently used. In this respect, a large exome chip study of 120,991 coding low-frequency variants with minor allele frequencies (MAFs) between 0.01% and 5% compared 32,367 MS cases and 36,012 controls <sup>24</sup>.

The results of this study showed that up to 5% of MS heritability is explained by low-frequency variants. Some of these variants were found to impact genes that have not yet been observed in GWASs and that encode proteins involved in diverse immunity-related processes, such as regulatory T cell homeostasis, IFN $\gamma$  biology, and NF $\kappa$ B signaling<sup>24</sup>.

WES and WGS studies of rare variants in MS generally use family designs. When multiple family members are affected within a family, rare variants contributing to disease risk can be considered to be present at higher frequencies compared to the general population<sup>25</sup>. In addition, accumulation of rare susceptibility variants is observed in multi-incident MS families compared to single-case families<sup>26</sup>. This makes multi-incident MS families of interest for identifying and studying rare variants involved in the disease. In the past 10 years, a number of studies have reported rare, putatively pathogenic genetic variants in MS patients from multi-incident families through WES<sup>22,27-32</sup>. Most of these family studies, however, only involved a small number of families, analyses were sometimes targeted exclusively at known MS loci, and the vast majority of the identified variants were not studied in replication studies or did not replicate<sup>27,29,33-36</sup>. Therefore, more robust WES/WGS studies and integration across studies of common and rare genetic variant findings are still needed to better understand the rare genetic component of MS and provide additional insight into the molecular pathways underlying the disease.

In the present study, WES and co-segregation analyses were applied in nine Dutch MS families with 2 to 4 affected family members, in order to identify rare, exonic variants co-segregating with the disease in these families. Our findings were integrated with the results from large rare and common variant studies of MS. Subsequently, based on the proteins encoded by the genes in which we identified rare variants, we constructed a molecular framework that provides further insights into the mechanisms underlying MS as a first step towards validation studies in cellular and/or animal models.

## Methods

### Participants

Multi-incident MS families of Dutch descent were recruited through an online advertisement on the website of the Dutch National MS Foundation (that also funded this study). In order to participate in this study, at least two people with MS needed to be present in a family, preferably across two generations. Participants had to be minimally 20 years of age and have a diagnosis of MS - made by a neurologist - according to the McDonald criteria for at least one year<sup>37</sup>. After online application, families were contacted by phone to plan a home visit. At this home visit, blood or saliva for DNA isolation was collected from the MS cases and as many unaffected family members as possible. In addition, informed consent was obtained and a questionnaire was administered to collect information about disease type and duration, MS-related treatments and more general information such as (previous) infectious and other diseases in both MS cases and unaffected family members. Lastly, a pedigree of the family was drawn to obtain more information about the disease penetrance. A total of nine multi-incident MS families were included. The 'Medische Ethische Toetsingscommissie (METC) Brabant' – i.e., the relevant Dutch Medical Ethical Committee – approved this study, and written informed consent was provided by each participant (approved research study number: NL62481.028.17 / P1738).

### Whole exome sequencing and analysis

DNA was available for a total of 55 members across the nine families, of which 26 were diagnosed with MS and 29 were unaffected family members. Genomic DNA was isolated from either peripheral venous blood or saliva according to standard procedures of the department of Human Genetics of the Radboud University Medical Center, Nijmegen, The Netherlands. Whole exome sequencing was performed by BGI (Copenhagen, Denmark) on at least two MS cases per family and one unaffected family member. Exome capture was performed by using the SureSelect<sup>XT</sup> Human All Exon V5 enrichment kit (Agilent Technologies, Santa Clara, CA, USA).

Raw image files were generated on HiSeq 4000 platform (Illumina, San Diego, CA, USA), making use of the 2x 100 bp paired-end module with a minimal median target coverage of 50X and an average Phred-like consensus quality of 30 for at least 80% of called bases. This implies that 80% of called bases have an accuracy of 99.9%. Base calling was implemented through the default parameters of the Illumina software. Sequencing reads were then aligned against the human reference genome GRCh37/hg19 using the Burrows-Wheeler Aligner (BWA). Variants were subsequently called by the Genome Analysis Toolkit (GATK) HaplotypeCaller. Annotation of variants was performed according to in-house pipelines of the department of Human Genetics. Data on 31 sequenced exomes were available from this procedure.

### **Selection of candidate single nucleotide variants**

For each family, rare single nucleotide variants (SNVs) in the exome were selected that were present in all sequenced MS cases but absent in the sequenced unaffected family member(s), making use of the in-house tool HitPanda. In order to identify rare variants most likely involved in the pathogenesis of MS in each family, additional selection procedures were applied.

Firstly, SNVs occurring in at least two families were selected. Of these SNVs, only those identified in  $\leq 1\%$  of the general population were selected. Both non-synonymous substitutions – 'missense mutations' – and other (putative) loss-of-function variants were included. In addition, variants with a Combined Annotation Dependent Depletion (CADD) score below 20 were omitted. The CADD score provides a measurement of the deleteriousness of a variant; a score of 20 implies that the variant belongs to the 1% most deleterious substitutions in the human genome<sup>38</sup>.

As a second selection strategy, a selection was made among variants within single families. Only rare SNVs identified in  $\leq 0.05\%$  of the general population were included, using the same criteria as described above. In addition to missense variants, all putative loss-of-function variants fulfilling these criteria were included. Among the selected missense variants, a further selection was made based on prior evidence.



Variants within genes for which literature evidence of a clear link with MS could be found were selected. Additionally, variants were selected based on evidence of association in the abovementioned large exome chip <sup>24</sup> and GWAS <sup>23</sup> analyses from the International Multiple Sclerosis Genetics Consortium (IMSGC). Using the summary statistics of the GWAS, downloaded from the website of the IMSGC (<https://imsgc.net/>), we performed a gene-wide association analysis using the Multimarker Analysis of GenoMic Annotation (MAGMA) software <sup>39</sup>. From this, a weighted P-value for the combination of all SNPs within a gene was calculated (see **Supplementary Table 1** for all genes with uncorrected P-values <0.05). We required variants from the exome study and genes from the gene-wide GWAS to show association at uncorrected P<0.05 for selection of a WES-derived variant for further analysis.

### **Validation of SNVs**

Validation was performed by Sanger sequencing of the region of interest for each variant in the respective family. The region of interest was sequenced in all family members with DNA available, including the 24 family members in which WES had not been performed. Primer pairs spanning the variant of interest were designed according to in-house pipelines, making use of the University of California Santa Cruz (UCSC) Genome Browser and Primer3Plus online software. Primer optimization was achieved by running a gradient polymerase chain reaction (PCR) on test DNA to determine the optimal annealing temperature for each primer pair. Primer sequences and annealing temperatures can be found in **Supplementary Table 2**. The purified PCR product from the participating family members was then sequenced at the sequencing facility of the department of Human Genetics. Sequences were analyzed using Vector NTI software (Invitrogen, Carlsbad, CA, USA) and Alamut Visual v2.10 software (Interactive Biosoftware, Rouen, France).

Co-segregation analysis was then performed to assess to which extent MS cases in a given family carried the identified rare variant, while unaffected family members did not.

We designated those variants that were present in all MS cases and <33% of unaffected family members as showing ‘acceptable’ co-segregation.

### **Construction of molecular framework**

Subsequently, we constructed a molecular framework based on the proteins encoded by genes containing rare variants that met our abovementioned selection criteria and showed acceptable co-segregation. In short, we conducted an elaborate literature analysis of the (putative) function of these proteins, their interactions and evidence of these genes/proteins being implicated in MS. In addition, we placed some other proteins that have been implicated in MS and functionally link the proteins from our WES in the framework. Based on much larger data sets, we have previously applied a similar approach to build so-called 'molecular landscapes' for other neurological disorders, such as Parkinson's disease<sup>40</sup> and amyotrophic lateral sclerosis<sup>41</sup>.

## **Results**

Nine families with 2 to 4 MS cases were included in this study. The characteristics of the 55 participants (26 MS patients and 29 unaffected family members) are presented in **Table 1**. The mean age at DNA collection was  $45.6 \pm 15.2$  years for MS cases and  $51.2 \pm 16.0$  years for the unaffected family members. The average age at MS onset was  $31.6 \pm 8.5$  years. The male to female ratio was 1:2.9 for MS cases and 1:1.2 for unaffected family members. All forms of MS were observed within and across the families (RRMS, n=13; SPMS, n=6, PPMS, n=2, diagnosis made by a neurologist), while for five cases, the MS form/disease course was not possible to determine.

Whole exome sequencing of 31 individuals from nine MS families resulted in 6.9 gigabases of mapped reads. Average target coverage was 80-fold per exome. An average of 122,969 variants was called per individual (range 115,583 – 130,153). Of these variants, 94.1% was annotated in dbSNP (v150).

Selection of the variants that were present in all sequenced MS cases of a family and absent in the unaffected sequenced family members resulted in an average total of 17,050 variants per family (range 7,636 – 23,881). Applying the described selection criteria, we obtained 15 missense variants and 3 deletions across all families, of which five missense variants were present in two out of nine families. In **Table 2**, the 18 variants are listed with frequencies, nucleotide and amino acid changes, and CADD scores. Rare missense variants in the genes *PLK1*, *PRAM1*, *CACNA1F*, *CPT1A*, and *TRIM66* were identified in two families. Among the 18 variants, missense variants were found in three genes that have been previously linked to MS: *MBP*<sup>42,43</sup>, *MECP2*<sup>44,45</sup>, and *CPT1A*<sup>46</sup>. The missense variants in *PRAM1* and *USH2A* showed nominally significant association with MS in the exome chip study of the IMSGC<sup>24</sup>. Furthermore, at least nominally significant gene-wide P-values were found in the data from the GWAS of the IMSGC for additional genes carrying rare variants in our study, i.e. *ZNF641*, *EXOC2*, *CCNI*, *ITPR3*, *SORCS1*, *PRAM1*, *MTMR7*, *TTC28*, and *TOX3*<sup>23</sup>. An overview of all 18 variants meeting the selection criteria with accompanying P-values (where applicable) can be found in **Supplementary Table 3**.

Sanger sequencing was performed in all family members with available DNA in order to validate the WES results and to genotype the 24 family members in whom WES had not been performed. As shown in **Table 3** and **Supplementary Figure 1** - containing the pedigrees of the nine families with all 18 variants; as this figure may contain identifying information, it is only available upon request from the corresponding author - subsequent co-segregation analyses of each variant in the respective families indicated that the variants within five genes completely co-segregated with MS: *PLK1*, *MBP*, *SORCS1*, *MTMR7*, and *TOX3*. The identified variants in *PRAM1*, *CPT1A*, *TRIM66*, *CACNA1F*, *MECP2*, *TTC28*, and *ITPR3* were present in all MS cases in the respective families, but were also present in 20-33% of unaffected family members. The missense variants in *EXOC2* and *CCNI*, and the deletions in *CNGB1* and *MANBA*, were present in 50-66% of the unaffected individuals. Lastly, the missense variants in *ZNF641* and *USH2A* were not present in all MS cases and present in 25% and 17% of unaffected family members, respectively.

Given the complex genetic etiology of MS, we did not expect a monogenic etiology in the families investigated here. Therefore, we argue that variants showing incomplete segregation are also of interest for further study. Investigating the function of the proteins encoded by the 12 genes that showed ‘acceptable’ co-segregation (presence in all MS cases and <33% of unaffected individuals) – *PLK1*, *MBP*, *ITPR3*, *SORCS1*, *PRAM1*, *CACNA1F*, *MECP2*, *CPT1A*, *MTMR7*, *TRIM66*, *TTC28*, and *TOX3* – we found that the proteins encoded by these genes could be linked to (de/re)myelination and/or the auto-immune response, as shown in the molecular framework in **Figure 1**.

## Discussion

In the current study, we set out to identify rare genetic variants involved in MS through WES in multi-incident MS families. Among nine families, we identified rare variants in 12 genes that showed at least acceptable co-segregation with the disease. For each of these 12 genes, a link was identified with (de-/re)myelination processes and/or auto-immune responses that are characteristic of MS, and the encoded proteins could be placed in a molecular framework (**Figure 1**). Below, we will discuss how these 12 proteins – *MBP*, *PLK1*, *MECP2*, *MTMR7*, *TOX3*, *CPT1A*, *SORCS1*, *TRIM66*, *ITPR3*, *TTC28*, *CACNA1F*, and *PRAM1* – and a number of additional molecules/proteins (most of which have also been linked to MS) operate in the framework.

MS is characterized by an auto-immune response against the myelin sheath that surrounds axons and is produced by the ODCs. Several genes in the molecular framework encode proteins that regulate myelination. Firstly, myelin basic protein (*MBP*) is the most important structural component of the myelin sheath. The most commonly used animal model of MS, experimental autoimmune encephalomyelitis (EAE), can be established through immunization (production of antibodies) of mice against *Mbp*<sup>47</sup>. In this respect, antibodies against *MBP* have been found in EAE mice and a proportion of MS patients<sup>48</sup>. In humans, anti-*MBP* antibodies were detected during the conversion of clinically isolated syndrome (CIS, a prodromal form of the disease) to clinical MS<sup>42</sup>. Moreover, common genetic variation within *MBP* was found to predict the clinical course of MS<sup>43</sup>.

Interestingly, in the most recently reported WES study of MS, a missense variant in *MBP* was found in four MS patients from the same family<sup>32</sup>, further pointing towards a role of rare variants in *MBP* in MS susceptibility. The product of *PLK1* is a kinase that phosphorylates and regulates the activity of MTHFR, an enzyme involved in the synthesis of methionine and its product S-adenosylmethionine (SAM)<sup>49,50</sup>. SAM methylates MBP at arginine sites within the protein, which protects MBP against degradation and is hence important for maintaining the integrity of the myelin sheath<sup>51</sup>. Common genetic variants in *MTHFR* have already been associated with the risk for MS development in multiple populations<sup>52-55</sup>. *PLK1* also inhibits the protein complex MTORC1, which in turn is an important positive regulator of (re)myelination by ODCs<sup>56,57</sup>. The molecular framework gene *MECP2* encodes a transcriptional regulator that downregulates the expression of MBP in ODCs<sup>45,58</sup>. *Mecp2* has also been linked to negatively regulating (re)myelination in the EAE mouse model<sup>44,59</sup>. *MTMR7* encodes a member of the myotubularin protein family of phosphatases and is upregulated after ODC differentiation, which may point to a role in (re)myelination<sup>60</sup>. *TOX3* encodes a transcription factor that positively regulates the survival of ODCs after injury and may therefore be involved in remyelination<sup>61</sup>. Another important protein in the framework is the product of the *CPT1A* gene, an enzyme located in the mitochondrial membrane that catalyzes the rate-limiting step in fatty acid oxidation. When oxidized, fatty acids cause oxidative stress and cannot be used for myelin production, which has a negative effect on (re)myelination by ODCs<sup>46,62</sup>. In two isolated, indigenous populations from northern Canada - the Inuit and Hutterites - two different rare genetic variants in *CPT1A* occur in the majority of the population. As people carrying either one of these variants have a much smaller chance to develop MS compared to the nonindigenous population of Canada, these variants seem to protect against the disease. The two variants result in 78% decreased activity and complete inactivity of *CPT1A* in the Inuit and Hutterite populations, respectively<sup>46</sup>. This study also showed that in a mouse model of the low-activity Inuit-variant in *CPT1A*, EAE could not be induced, further providing evidence for a protective role for this variant.

The molecular framework gene *SORCS1* encodes a membrane protein that is highly expressed in neurons and is involved in downregulating the expression of agouti-related peptide (AGRP)<sup>63,64</sup>. AGRP has been linked to MS and is a secreted molecule that binds to and is an antagonist of the melanocortin-4 receptor (MC4R)<sup>65-67</sup>. MC4Rs are highly expressed in ODCs and their precursor cells. When bound by their main agonist, the melanocortin peptide ACTH (not shown in **Figure 1**), MC4Rs have been shown to protect ODCs from damaging insults and increase proliferation of their precursor cells, hence contributing to (re)myelination<sup>68</sup>. TRIM66 in the framework is a transcription factor that is highly expressed in both neurons and macrophages. Interestingly, two rare variants in *TRIM66* have been associated with MS severity<sup>69</sup>. In neurons, TRIM66 regulates the activity of SIRT6, a histone-deacetylating enzyme that is involved in epigenetic processes and has been recently suggested as a novel drug target for MS<sup>70,71</sup>. In (peripheral) macrophages, TRIM66 is involved in upregulating the expression of endogenous interferon beta (IFNB1)<sup>72</sup>, which is interesting as recombinant, exogenous IFNB1 is one of the earliest and still used disease-modifying treatments for MS<sup>73</sup>. Lastly, ITPR3 is a protein in the endoplasmic reticulum (ER) membrane that increases the release of calcium from the ER to the cytoplasm of neurons. This triggers a downstream signaling cascade that can eventually result in neuronal death/neurodegeneration in both the EAE mouse model of MS<sup>74</sup> and stem cell-derived neurons from MS patients<sup>75</sup>, as sustained high calcium levels in cells/neurons are highly toxic.

In addition to TRIM66 - which functions both in neurons and peripheral macrophages - three proteins encoded by genes with MS-related rare variants found in our study operate in the framework outside of the CNS. First, TTC28 is a cytoplasmic protein located in vascular endothelial cells that constitute the blood-brain barrier (BBB) between the periphery and the CNS. TTC28 binds and functionally interacts with ICAM1<sup>76</sup>, a membrane receptor that has been consistently linked to the pathogenesis of MS, as it promotes transcellular diapedesis of inflammatory T-cells from the periphery to the brain cells across the BBB, a key pathological process underlying MS<sup>77</sup>. Moreover, exogenous IFNB1 administration leads to downregulation of ICAM1 in vascular endothelial cells of the BBB<sup>78</sup>.

The two other identified genes encode proteins with functions in peripheral T-cells. *CACNA1F* encodes an L-type voltage dependent calcium channel that transports calcium into T-cells. A deficiency of *CACNA1F* leads to reduced calcium influx and ultimately disturbances in T-cell function that are characteristic of MS<sup>79-81</sup>. *PRAM1* is a membrane protein that binds and interacts with *TRIM27*, a cytoplasmic protein that negatively regulates calcium influx into T-cells, with a similar effect on T-cell function as *CACNA1F* deficiency<sup>82,83</sup>.

In summary, 9 of the 12 genes with rare variants associated with MS in our study encode proteins that are (potentially) involved in regulating (de-/re-)myelination (*MBP*, *PLK1*, *MECP2*, *MTMR7*, *TOX3*, *CPT1A*, *SORCS1*, *TRIM66*, and *ITPR3*), while 4 of the 12 genes (*TRIM66*, *TTC28*, *CACNA1F*, and *PRAM1*) are linked to (auto-)immune processes. Although they need to be replicated/validated (see below), our findings may imply that rare variants predominantly affect genes involved in (de-/re-)myelination and, eventually, neurodegeneration rather than the acute auto-immune reaction that is associated with disease relapses.

Our study should be viewed in the context of a number of strengths and limitations. A strength is the fact that many of the genes found affected are supported by evidence from previous studies of rare and common genetic variants. One important limitation is the small number of family members for which co-segregation analyses of the identified variants could be performed. Larger families and a larger numbers of participants would yield power for statistical analysis of these co-segregation analyses. Future research should therefore be aimed at including more family members in this type of studies. Moreover, 7 of the 12 identified variants did not show complete co-segregation with MS. For instance, unaffected subject I:2 in family 1 (**Supplementary Figure 1**) carries the identified missense variant in *EXOC2*, which led us to not further consider this variant for the framework, as it is found in 1 out of 2 tested unaffected family members. However, no variants in *PLK1* and *MBP* are found in subject I:2 that are present – together with the *EXOC2* variant - in all three MS patients from this family: II:5, II:11 and II:13.

Further, none of the three variants were found in the second unaffected member of the family, II:4. Another example is the identified missense variant in *CACNA1F* that was found in families 3 and 7 (**Supplementary Figure 1**); in both these families, one unaffected family member also carried this variant (i.e. III:3 and II:9, respectively). This means that, across both families, 2 out of 7 unaffected family members (29%) carry the *CACNA1F* variant. However, individual III:3 from family 3 may still be at risk of developing MS, as she is only in her 30s. These patterns of incomplete co-segregation suggest an oligogenic inheritance of MS with the effects of a small(er) number of variants still needed for the MS phenotype to be fully expressed.

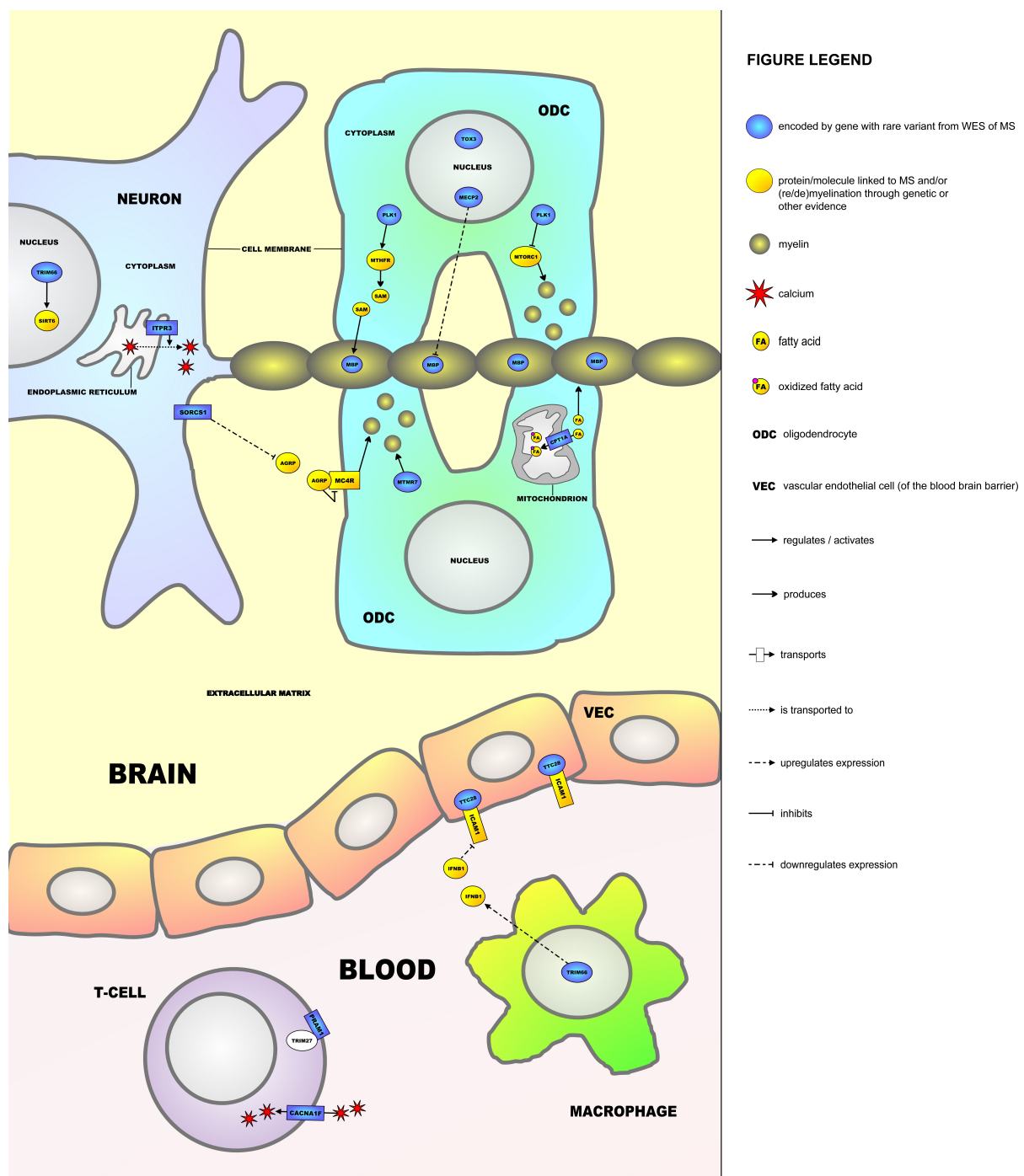
As indicated above, a number of genes that were identified in this study have been previously linked to MS, and it would be of interest to perform functional experiments involving these or other identified genes to further elucidate the (pathogenic) molecular mechanisms underlying MS. For instance, in the study by Mørkholt et al., it was shown that a rare variant in *Cpt1a* results in ineffective EAE inducement in mice <sup>46</sup>. In this respect, it could be speculated that the *CPT1A* variant that was identified in this study leads to a gain-of-function of the CPT1A enzyme, and it could be tested whether this is indeed the case through *in vitro* studies.

In conclusion, through WES and subsequent Sanger sequencing in multi-incident MS families, rare genetic variants in 12 genes were identified that show an acceptable co-segregation pattern with the disease. The proteins encoded by these genes interact in a molecular framework that contains both (de/re)myelination and auto-immunity-related processes. Our findings provide further insights into the molecular mechanisms underlying MS that should be studied in cellular and/or animal models, and that could be leveraged for the development of novel treatment strategies for the disease, and this specifically aimed at promoting remyelination.



## **Acknowledgements**

We would like to thank all the MS patients and their unaffected family members who have contributed to this study. In addition, we are grateful to professor Christian Gilissen and his research group at the Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands, for their assistance with the analysis and interpretation of the WES data.



**Figure 1.** A molecular framework containing 12 interacting proteins encoded by genes harboring rare variants that were identified through whole exome sequencing (WES) in nine multi-incident multiple sclerosis (MS) families. MBP, PLK1, MECP2, MTMR7, TOX3, CPT1A, SORCS1, TRIM66, ITPR3, TTC28, CACNA1F, and PRAM1 are all linked to (de-/re)myelination and/or the auto-immune response characteristic of MS.

**Table 1.** An overview of the clinical characteristics of the participants included in the whole exome sequencing (WES) study of MS (26 MS patients and 29 unaffected family members)/ F: female; M: male; NA: MS type not available; RRMS: relapsing-remitting MS; SPMS: secondary progressive MS; PPMS: primary progressive MS.

Family	Participant	Gender	MS diagnosis	MS subtype	Age range at diagnosis (years)	Age range at DNA collection (years)	WES performed
1	II.1	F	No	-	-	61-65	No
1	II.4	F	No	-	-	36-40	Yes
1	II.5	M	Yes	RRMS	26-30	36-40	Yes
1	II.11	F	Yes	RRMS	26-30	31-35	No
1	II.13	F	Yes	RRMS	26-30	26-30	Yes
2	II.1	M	No	-	-	81-85	Yes
2	III.1	M	No	-	-	56-60	No
2	III.2	F	Yes	SPMS	46-50	56-60	Yes
2	III.10	F	No	-	-	51-55	No
2	III.12	F	Yes	SPMS	16-20	41-45	Yes
2	IV.1	F	No	-	-	26-30	No
2	IV.2	M	Yes	RRMS	21-25	21-25	No
3	III.1	M	Yes	RRMS	21-25	31-35	Yes
3	II.1	M	No	-	-	56-60	Yes
3	II.2	F	Yes	RRMS	31-35	56-60	Yes
3	II.3	M	No	-	-	51-55	No
3	III.3	F	No	-	-	31-35	No
4	III.2	M	Yes	SPMS	51-55	71-75	Yes
4	III.3	F	No	-	-	66-70	Yes
4	III.6	F	No	-	-	61-65	No
4	III.8	F	Yes	NA	21-25	56-60	No
4	IV.1	F	No	-	-	31-35	No
4	IV.2	F	Yes	RRMS	26-30	31-35	Yes
4	IV.4	F	Yes	RRMS	31-35	31-35	Yes
4	IV.5	M	No	-	-	46-50	No
4	IV.6	M	No	-	-	26-30	No
4	IV.7	F	No	-	-	21-25	No
5	I.1	M	Yes	SPMS	26-30	66-70	Yes
5	I.2	F	No	-	-	61-65	No
5	II.1	F	No	-	-	36-40	No
5	II.2	M	Yes	RRMS	36-40	36-40	Yes
5	II.3	F	Yes	RRMS	26-30	36-40	Yes
5	II.4	M	No	-	-	36-40	Yes
6	I.1	F	No	-	-	61-65	Yes
6	II.1	F	Yes	RRMS	26-30	31-35	Yes
6	II.2	F	No	-	-	26-30	No
6	II.3	F	Yes	RRMS	21-25	26-30	Yes

<b>Table 1 – continued.</b>							
<b>Family</b>	<b>Participant</b>	<b>Gender</b>	<b>MS diagnosis</b>	<b>MS subtype</b>	<b>Age range at diagnosis (years)</b>	<b>Age range at DNA collection (years)</b>	<b>WES performed</b>
7	II.2	F	Yes	SPMS	36-40	66-70	Yes
7	II.5	M	No	-	-	61-65	No
7	II.7	F	Yes	SPMS	31-35	61-65	No
7	II.9	F	No	-	-	61-65	No
7	III.1	F	Yes	PPMS	36-40	36-40	Yes
7	III.2	F	Yes	NA	21-25	36-40	Yes
7	III.3	F	No	-	-	36-40	Yes
7	II.10	M	No	-	-	56-60	No
8	I.1	M	No	-		66-70	No
8	I.2	F	No	-		66-70	No
8	II.1	M	Yes	NA	31-35	41-45	Yes
8	II.2	F	Yes	NA	21-25	36-40	Yes
8	II.3	M	No	-		36-40	Yes
8	II.4	F	Yes	RRMS	31-35	36-40	Yes
9	II.7	F	Yes	NA	41-45	76-80	Yes
9	II.8	F	Yes	PPMS	51-55	71-75	Yes
9	III.1	M	No	-	-	56-60	No
9	III.2	M	No	-	-	51-55	Yes

**Table 2. Rare genetic variants selected for validation.** An overview of the variants that were selected according to our criteria for nine multi-incident MS families. All variants are non-synonymous substitutions, resulting in missense variants, apart from three loss-of-function variants constituting a deletion (DEL) of 7-15 nucleotides. Variants with a reference SNP identifier (rs ID) are known to NCBI's dbSNP database (v150). For these variants, minor allele frequencies (MAFs) are mentioned. If a variant is novel, frequencies are indicated as NA (not available). Both nucleotide and amino acid changes are provided for each identified variant (where applicable), as well as the Combined Annotation Dependent Depletion score (CADD). The criteria for selection of each of the individual variants are shown in **Supplementary Table 3**.

Family	Chromosome	Gene	Variant ID	MAF (%)	Nucleotide change	Amino acid change	CADD
1	chr6	<i>EXOC2</i>	rs760365995	≤ 0.05	G>A	p.Pro134Leu	25.0
1	chr16	<i>PLK1</i>	rs35056440	≤ 1	C>T	p.Leu261Phe	22.9
1	chr18	<i>MBP</i>	18:74696828	NA	C>A	p.Gly151Val	25.4
2	chr12	<i>ZNF641</i>	rs200502528	≤ 0.05	G>A	p.His342Tyr	27.7
3	chr4	<i>CCNI</i>	rs139547927	≤ 0.05	T>C	p.Tyr334Cys	29.1
3	chr 6	<i>ITPR3</i>	6: 33648383 - 33648394	NA	DEL (a)	-	26.6
3	chr10	<i>SORCS1</i>	10:108439392	NA	G>A	p.Ser554Leu	34.0
3	chr16	<i>PLK1</i>	rs35056440	≤ 1	C>A	p.Leu261Phe	22.9
3	chr19	<i>PRAM1</i>	rs138042924	≤ 1	A>G	p.Ser387Pro	22.0
3	chrX	<i>CACNA1F</i>	X:49067910	NA	T>A	p.Met1324Leu	23.4
3	chrX	<i>MECP2</i>	rs61751445	≤ 0.05	G>A	p.Thr311Met	26.1

Table 2 – continued.							
Family	Chromosome	Gene	Variant ID	MAF (%)	Nucleotide change	Amino acid change	CADD
4	chr1	<i>USH2A</i>	rs200802261	≤ 0.05	G>A	p.Pro2870Leu	31.0
5	chr11	<i>CPT1A</i>	rs140958507	≤ 1	C>T	p.Arg288Gln	23.7
6	chr8	<i>MTMR7</i>	rs760803217	≤ 0.05	G>A	p.Thr6Met	22.5
6	chr11	<i>TRIM66</i>	rs138444298	≤ 1	G>A	p.Arg1037Trp	21.5
6	chr16	<i>CNGB1</i>	rs776392588	≤ 0.05	DEL (b)	-	51.0
7	chrX	<i>CACNA1F</i>	X:49067910	NA	T>A	p.Met1324Leu	23.4
8	chr22	<i>TTC28</i>	22:28501278	NA	T>C	p.Tyr1099Cys	32.0
9	chr4	<i>MANBA</i>	rs1283991082	≤ 0.05	DEL (c)	-	30.0
9	chr11	<i>TRIM66</i>	rs138444298	≤ 1	G>A	p.Arg1037Trp	21.5
9	chr11	<i>CPT1A</i>	rs140958507	≤ 1	C>T	p.Arg288Gln	23.7
9	chr16	<i>TOX3</i>	16:52580565	NA	T>C	p.Tyr24Cys	23.8
9	chr19	<i>PRAM1</i>	rs138042924	≤ 1	A>G	p.Ser387Pro	22.0

- (a) Deletion of 12 nucleotides: GTGCTGAGCGTT > -
- (b) Deletion of 7 nucleotides: AAGTAGT > -
- (c) Deletion of 15 nucleotides: GAGCAAAGAAATTCT > -

**Table 3. Co-segregation of the 18 selected rare genetic variants with MS.** Complete co-segregation is indicated as the variant being present in all MS cases and absent in all controls within one or two families, while acceptable co-segregation implies that the variant is present in all MS cases and is present in a maximum of 33% of controls. In total, 12 rare genetic variants show complete or acceptable co-segregation.

Gene	Variant ID	Amino acid change	Co-segregation MS cases / controls (%)
<i>EXOC2</i>	rs760365995	p.Pro134Leu	100/50
<i>PLK1</i>	rs35056440	p.Leu261Phe	100/0
<i>MBP</i>	18:74696828	p.Gly151Val	100/0
<i>ZNF641</i>	rs200502528	p.His342Tyr	66/25
<i>CCNI</i>	rs139547927	p.Tyr334Cys	100/66
<i>ITPR3</i>	6: 33648383 - 33648394	-(DEL (a))	100/33
<i>SORCS1</i>	10:108439392	p.Ser554Leu	100/0
<i>PRAM1</i>	rs138042924	p.Ser387Pro	100/20
<i>CACNA1F</i>	X:49067910	p.Met1324Leu	100/29
<i>MECP2</i>	rs61751445	p.Thr311Met	100/33
<i>USH2A</i>	rs200802261	p.Pro2870Leu	75/17
<i>CPT1A</i>	rs140958507	p.Arg288Gln	100/20
<i>MTMR7</i>	rs760803217	p.Thr6Met	100/0

<b>Table 3 – continued.</b>			
<b>Gene</b>	<b>Variant ID</b>	<b>Amino acid change</b>	<b>Co-segregation MS cases / controls (%)</b>
<i>TRIM66</i>	rs138444298	p.Arg1037Trp	100/25
<i>CNGB1</i>	rs776392588	- (DEL (b))	100/50
<i>TTC28</i>	22:28501278	p.Tyr1099Cys	100/33
<i>MANBA</i>	rs1283991082	- (DEL (c))	100/50
<i>TOX3</i>	16:52580565	p.Tyr24Cys	100/0

- (a) Deletion of 12 nucleotides: GTGCTGAGCGTT > -
- (b) Deletion of 7 nucleotides: AAGTAGT > -
- (c) Deletion of 15 nucleotides: GAGCAAAGAAATTCT > -



<b>Supplementary Table 2.</b> An overview of the sequences of the forward and reverse primers that were used for the Sanger sequencing and their optimal annealing temperatures in order to validate the 18 variants in the family members of the multi-incident MS families.						
Family	Chr	Gene	Variant ID	Forward primer	Reverse primer	Optimal annealing temperature
1	chr6	<i>EXOC2</i>	rs760365995	5' TTCCTGGGATGTTTCTGGAG	5' CCAACAGGCATTTTGTATTAGG	59 °C
1	chr16	<i>PLK1</i>	rs35056440	5' AGGATGCCTGACCTTTGTTT	5' GGTGGGATAGACCTATTCTGG	59 °C
1	chr18	<i>MBP</i>	18:74696828	5' GAAATCCACGCGTAAATGC	5' TGAACCACTACCCAGTCCAG	59 °C
2	chr12	<i>ZNF641</i>	rs200502528	5' TCAGCCAGTGTCTACAAGG	5' GGCGAAGACATCACCTCATC	59 °C
3	chr4	<i>CCNI</i>	rs139547927	5' AGAAACAGGCTGCAAAGGTG	5' GCCCAGACTTCTCCAAGGAC	59 °C
3	chr6	<i>ITPR3</i>	6: 33648383 - 33648394	5' AGGAGGTGTGGGAGTGGAG 5' AGGAGGTGTGGGAGTGGAG	5' ACCTGCAGGGAAGTGCTG 5' CTAGGCAAGTCGGCCCTG	56 °C 58 °C
3	chr10	<i>SORCS1</i>	10:108439392	5' GGTTTTCTGCAGACTTGCTTC	5' TATTTGGGGATGGAAGATG	59 °C
3	chr16	<i>PLK1</i>	rs35056440	5' AGGATGCCTGACCTTTGTTT	5' GGTGGGATAGACCTATTCTGG	59 °C
3	chr19	<i>PRAM1</i>	rs138042924	5' CACCCACCTATGGATGCTG	5' AGTCAGCGTGCTTCCAAG	59 °C
3	chrX	<i>CACNA1F</i>	X:49067910	5' GTGGTGGTTGTGAGGAAATG	5' ATCTCCTCATTCCCCAGTCC	59 °C
3	chrX	<i>MECP2</i>	rs61751445	5' TCTGAGTGGTGGTATGGTG	5' CAGGTCATGGTATCAAACG	59 °C
4	chr1	<i>USH2A</i>	rs200802261	5' CGCTCACACAATGAAGACAC	5' GTGCGAAACTGGCAGTAAAG	59 °C
5	chr11	<i>CPT1A</i>	rs140958507	5' CCCGGCCACAACTGTATATT	5' CAAAGTCAAAGTGGAGAAAATCA	59 °C
6	chr8	<i>MTMR7</i>	rs760803217	5' CAGCAGCAGCTAAGGCTAGG	5' CAGCCTGTACTCCCCATCTC	59 °C
6	chr11	<i>TRIM66</i>	rs138444298	5' CAAGCCAGAGGAAAAGTGC	5' CCAAGCCTGTGGCTAGAAAG	59 °C
6	chr16	<i>CNGB1</i>	rs776392588	5' TCCGATCATCACAGAGAAAGC	5' TCTGTCTTTGCACAACATGC	58 °C

Supplementary Table 2 – continued.						
Family	Chr	Gene	Variant ID	Forward primer	Reverse primer	Optimal annealing temperature
7	chrX	<i>CACNA1F</i>	X:49067910	5' GTGGTGGTTGTGAGGAAATG	5' ATCTCCTCATTCCCCAGTCC	59 °C
8	chr22	<i>TTC28</i>	22:28501278	5' CTGAGACCCTTACTCCACACAG	5' CCTTGAACGCCAGCTGAAC	59 °C
9	chr4	<i>MANBA</i>	rs1283991082	5' CTTGATCTTAGAAAAAATCC	5' GCCTCTGACTACTTGACC	56 °C
9	chr11	<i>TRIM66</i>	rs138444298	5' CAAGCCAGAGGAAAAGTGC	5' CCAAGCCTGTGGCTAGAAAG	59 °C
9	chr11	<i>CPT1A</i>	rs140958507	5' CCCGGCCACAACTGTATATT	5' CAAAGTCAAAGTGGAGAAAATCA	59 °C
9	chr16	<i>TOX3</i>	16:52580565	5' GCCAAAGAAAAGTGCAGCAG	5' CTCGACGGTGGATCTAGCG	59 °C
9	chr19	<i>PRAM1</i>	rs138042924	5' CACCCACCTATGGATGCTG	5' AGTCAGCGTGCTTCCAAG	59 °C

**Supplementary Table 3.** An overview of the selection criteria that each of the 18 rare genetic variants meets is provided. If a variant or gene had a significant P-value in either the exome chip study<sup>24</sup> or based on the results from the genome-wide association study (GWAS)<sup>23</sup> – i.e. the gene-wide P-value – by the International Multiple Sclerosis Genetics Consortium (IMSGC), said P-value is provided.

Gene	Variant ID	Amino acid change	Present in 2/9 families	Literature link(s) with MS	P-value exome chip	Gene-wide P-value
<i>EXOC2</i>	rs760365995	p.Pro134Leu				2.50E-03
<i>PLK1</i>	rs35056440	p.Leu261Phe	✓			
<i>MBP</i>	18:74696828	p.Gly151Val		✓ <sup>42,43</sup>		
<i>ZNF641</i>	rs200502528	p.His342Tyr				3.76E-02
<i>CCNI</i>	rs139547927	p.Tyr334Cys				1.46E-03
<i>ITPR3</i>	6: 33648383 - 33648394	- (DEL (a))				2.65E-05
<i>SORCS1</i>	10:108439392	p.Ser554Leu				3.74E-02
<i>PRAM1</i>	rs138042924	p.Ser387Pro	✓		3.26E-02	3.55E-04
<i>CACNA1F</i>	X:49067910	p.Met1324Leu	✓			
<i>MECP2</i>	rs61751445	p.Thr311Met		✓ <sup>44,45</sup>		
<i>USH2A</i>	rs200802261	p.Pro2870Leu			2.76E-02	
<i>CPT1A</i>	rs140958507	p.Arg288Gln	✓	✓ <sup>46</sup>		

<b>Supplementary Table 3 – continued.</b>						
<b>Gene</b>	<b>Variant ID</b>	<b>Amino acid change</b>	<b>Present in 2/9 families</b>	<b>Literature link(s) with MS</b>	<b>P-value exome chip</b>	<b>Gene-wide P-value</b>
<i>MTMR7</i>	rs760803217	p.Thr6Met				2.83E-02
<i>TRIM66</i>	rs138444298	p.Arg1037Trp	✓			
<i>CNGB1</i>	rs776392588	- (DEL (b))				
<i>TTC28</i>	22:28501278	p.Tyr1099Cys				3.15E-07
<i>MANBA</i>	rs1283991082	- (DEL (c))				
<i>TOX3</i>	16:52580565	p.Tyr24Cys				3.07E-02

- (a) Deletion of 12 nucleotides: GTGCTGAGCGTT > -
- (b) Deletion of 7 nucleotides: AAGTAGT > -
- (c) Deletion of 15 nucleotides: GAGCAAAGAAATTCT > -

## References

- 1 Wallin, M.T. *et al.* Global, regional, and national burden of multiple sclerosis 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet. Neurology* **18**, 269-285 (2019).
- 2 McFarland, H. F. & Martin, R. Multiple sclerosis: a complicated picture of autoimmunity. *Nature immunology* **8**, 913-919 (2007).
- 3 Hemmer, B., Archelos, J. J. & Hartung, H. P. New concepts in the immunopathogenesis of multiple sclerosis. *Nature reviews. Neuroscience* **3**, 291-301 (2002).
- 4 Ciccarelli, O. *et al.* Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. *The Lancet. Neurology* **13** (2014).
- 5 Bogie, J. F., Stinissen, P. & Hendriks, J. J. Macrophage subsets and microglia in multiple sclerosis. *Acta neuropathologica* **128**, 191-213 (2014).
- 6 Duffy, S. S., Lees, J. G. & Moalem-Taylor, G. The contribution of immune and glial cell types in experimental autoimmune encephalomyelitis and multiple sclerosis. *Multiple sclerosis international* **2014**, 285245 (2014).
- 7 Trapp, B. D. & Nave, K. A. Multiple sclerosis: an immune or neurodegenerative disorder? *Annual review of neuroscience* **31**, 247-269 (2008).
- 8 Franklin, R. J. M. & Ffrench-Constant, C. Regenerating CNS myelin - from mechanisms to experimental medicines. *Nature reviews. Neuroscience* **18**, 753-769 (2017).
- 9 Chu, F. *et al.* The roles of macrophages and microglia in multiple sclerosis and experimental autoimmune encephalomyelitis. *Journal of neuroimmunology* **318**, 1-7 (2018).
- 10 Miron, V. E. *et al.* M2 microglia and macrophages drive oligodendrocyte differentiation during CNS remyelination. *Nature neuroscience* **16**, 1211-1218 (2013).
- 11 Zettl, U. K., Stüve, O. & Patejdl, R. Immune-mediated CNS diseases: a review on nosological classification and clinical features. *Autoimmunity reviews* **11**, 167-173 (2012).
- 12 Karussis, D. The diagnosis of multiple sclerosis and the various related demyelinating syndromes: a critical review. *Journal of autoimmunity* **48-49**, 134-142 (2014).
- 13 Thompson, A. J., Baranzini, S. E., Geurts, J., Hemmer, B. & Ciccarelli, O. Multiple sclerosis. *Lancet (London, England)* **391**, 1622-1636 (2018).
- 14 Mahad, D. H., Trapp, B. D. & Lassmann, H. Pathological mechanisms in progressive multiple sclerosis. *The Lancet. Neurology* **14**, 183-193 (2015).
- 15 Goodin, D. S., Khankhanian, P., Gourraud, P. A. & Vince, N. The nature of genetic and environmental susceptibility to multiple sclerosis. *PLoS One* **16**, e0246157 (2021).
- 16 Bjornevik, K. *et al.* Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science (New York, N.Y.)* **375**, 296-301 (2022).
- 17 Amato, M. P. *et al.* Environmental modifiable risk factors for multiple sclerosis: Report from the 2016ECTRIMS focused workshop. *Multiple sclerosis (Houndmills, Basingstoke, England)* **24**, 590-603 (2018).
- 18 Orton, S. M. *et al.* Sex ratio of multiple sclerosis in Canada: a longitudinal study. *The Lancet. Neurology* **5**, 932-936 (2006).
- 19 Sadovnick, A. D., Ebers, G. C., Dymont, D. A. & Risch, N. J. Evidence for genetic basis of multiple sclerosis. The Canadian Collaborative Study Group. *Lancet (London, England)* **347**, 1728-1730 (1996).
- 20 Canto, E. & Oksenberg, J. R. Multiple sclerosis genetics. *Multiple sclerosis (Houndmills, Basingstoke, England)* **24**, 75-79 (2018).
- 21 Hollenbach, J. A. & Oksenberg, J. R. The immunogenetics of multiple sclerosis: A comprehensive review. *Journal of autoimmunity* **64**, 13-25 (2015).
- 22 Mescheriakova, J. Y. *et al.* Linkage analysis and whole exome sequencing identify a novel candidate gene in a Dutch multiple sclerosis family. *Multiple sclerosis (Houndmills, Basingstoke, England)* **25**, 909-917(2019).

- 23 International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science (New York, N.Y.)* **365**, eaav7188 (2019).
- 24 International Multiple Sclerosis Genetics Consortium. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. *Cell* **175**, 1679-1687.e7 (2018).
- 25 Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753 (2009).
- 26 Gourraud, P. A. *et al.* Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. *Annals of neurology* **69**, 65-74 (2011).
- 27 Ramagopalan, S. V. *et al.* Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Annals of neurology* **70**, 881-886 (2011).
- 28 Dyment, D. A. *et al.* Exome sequencing identifies a novel multiple sclerosis susceptibility variant in the TYK2 gene. *Neurology* **79**, 406-411 (2012).
- 29 Wang, Z. *et al.* Nuclear Receptor NR1H3 in Familial Multiple Sclerosis. *Neuron* **90**, 948-954 (2016).
- 30 Maver, A. *et al.* Identification of rare genetic variation of NLRP1 gene in familial multiple sclerosis. *Scientific reports* **7**, 3715 (2017).
- 31 Sadovnick, A. D. *et al.* Purinergic receptors P2RX4 and P2RX7 in familial multiple sclerosis. *Human mutation* **38**, 736-744 (2017).
- 32 Zrzavy, T. *et al.* Exome-Sequence Analyses of Four Multi-Incident Multiple Sclerosis Families. *Genes* **11**, 988 (2020).
- 33 Ban, M. *et al.* No evidence of association between mutant alleles of the CYP27B1 gene and multiple sclerosis. *Annals of neurology* **73**, 430-432 (2013).
- 34 Barizzone, N. *et al.* No evidence for a role of rare CYP27B1 functional variations in multiple sclerosis. *Annals of neurology* **73**, 433-437 (2013).
- 35 International Multiple Sclerosis Genetics Consortium. NR1H3 p.Arg415Gln Is Not Associated to Multiple Sclerosis Risk. *Neuron* **92**, 333-335 (2016).
- 36 Minikel, E. V. & MacArthur, D. G. Publicly Available Data Provide Evidence against NR1H3 R415Q Causing Multiple Sclerosis. *Neuron* **92**, 336-338 (2016).
- 37 de Seze, J. & Bigaut, K. Multiple sclerosis diagnostic criteria: From poser to the 2017 revised McDonald criteria. *Presse Med* **50**, 104089 (2021).
- 38 Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nature genetics* **46**, 310-315 (2014).
- 39 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS computational biology* **11**, e1004219 (2015).
- 40 Klemann, C. *et al.* Integrated molecular landscape of Parkinson's disease. *NPJ Parkinsons Dis* **3**, 14 (2017).
- 41 Klemann, C., Visser, J. E., Van Den Bosch, L., Martens, G. J. M. & Poelmans, G. Integrated molecular landscape of amyotrophic lateral sclerosis provides insights into disease etiology. *Brain Pathol* **28**, 203-211 (2018).
- 42 Berger, T. *et al.* Antimyelin antibodies as a predictor of clinically definite multiple sclerosis after a first demyelinating event. *The New England journal of medicine* **349**, 139-145 (2003).
- 43 Zhou, Y. *et al.* Variation within MBP gene predicts disease course in multiple sclerosis. *Brain and behavior* **7**, e00670 (2017).
- 44 Khorshid Ahmad, T. *et al.* Experimental Autoimmune Encephalomyelitis (EAE)-Induced Elevated Expression of the E1 Isoform of Methyl CpG Binding Protein 2 (MeCP2E1): Implications in Multiple Sclerosis (MS)-Induced Neurological Disability and Associated Myelin Damage. *International journal of molecular sciences* **18**, 1254 (2017).
- 45 Sharma, K., Singh, J. & Pillai, P. P. MeCP2 Differentially Regulate the Myelin MBP and PLP Protein Expression in Oligodendrocytes and C6 Glioma. *Journal of molecular neuroscience* **65**, 343-350 (2018).

- 46 Mørkholt, A. S. *et al.* CPT1A plays a key role in the development and treatment of multiple sclerosis and experimental autoimmune encephalomyelitis. *Scientific reports* **9**, 13299 (2019).
- 47 Terry, R. L., Ifergan, I. & Miller, S. D. Experimental Autoimmune Encephalomyelitis in Mice. *Methods in molecular biology (Clifton, N.J.)* **1304**, 145-160 (2016).
- 48 Ponomarenko, N. A. *et al.* Autoantibodies to myelin basic protein catalyze site-specific degradation of their antigen. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 281-286 (2006).
- 49 Li, X. *et al.* Polo-like kinase 1 (PLK1)-dependent phosphorylation of methylenetetrahydrofolate reductase (MTHFR) regulates replication via histone methylation. *Cell cycle* **16**, 1933-1942 (2017).
- 50 Froese, D. S. *et al.* Structural basis for the regulation of human 5,10-methylenetetrahydrofolate reductase by phosphorylation and S-adenosylmethionine inhibition. *Nature communications* **9**, 2261 (2018).
- 51 Kim, S., Lim, I. K., Park, G. H. & Paik, W. K. Biological methylation of myelin basic protein: enzymology and biological significance. *The international journal of biochemistry & cell biology* **29**, 743-751 (1997).
- 52 Klotz, L. *et al.* The variant methylenetetrahydrofolate reductase c.1298A>C (p.E429A) is associated with multiple sclerosis in a German case-control study. *Neuroscience letters* **468**, 183-185 (2010).
- 53 Ineichen, B. V. *et al.* Genetic variants of homocysteine metabolism and multiple sclerosis: a case-control study. *Neuroscience letters* **562**, 75-78 (2014).
- 54 Cevik, B. *et al.* Association of methylenetetrahydrofolate reductase gene C677T polymorphism with multiple sclerosis in Turkish patients. *Journal of investigative medicine : the official publication of the American Federation for Clinical Research* **62**, 980-984 (2014).
- 55 Naghibalhossaini, F., Ehyakonandeh, H., Nikseresht, A. & Kamali, E. Association Between MTHFR Genetic Variants and Multiple Sclerosis in a Southern Iranian Population. *International journal of molecular and cellular medicine* **4**, 87-93 (2015).
- 56 Lebrun-Julien, F. *et al.* Balanced mTORC1 activity in oligodendrocytes is required for accurate CNS myelination. *The Journal of* **34**, 8432-8448 (2014).
- 57 Ruf, S. *et al.* PLK1 (polo like kinase 1) inhibits MTOR complex 1 and promotes autophagy. *Autophagy* **13**, 486-505 (2017).
- 58 Sharma, K., Singh, J., Pillai, P. P. & Frost, E. E. Involvement of MeCP2 in Regulation of Myelin-Related Gene Expression in Cultured Rat Oligodendrocytes. *Journal of molecular neuroscience* **57**, 176-184 (2015).
- 59 KhorshidAhmad, T. *et al.* Transcriptional Regulation of Brain-Derived Neurotrophic Factor (BDNF) by Methyl CpG Binding Protein 2 (MeCP2): a Novel Mechanism for Re-Myelination and/or Myelin Repair Involved in the Treatment of Multiple Sclerosis (MS). *Molecular neurobiology* **53**, 1092-1107 (2016).
- 60 Kippert, A., Trajkovic, K., Fitzner, D., Opitz, L. & Simons, M. Identification of Tmem10/Opalin as a novel marker for oligodendrocytes using gene expression profiling. *BMC neuroscience* **9**, 40, (2008).
- 61 Bastien, D. *et al.* IL-1 $\alpha$  Gene Deletion Protects Oligodendrocytes after Spinal Cord Injury through Upregulation of the Survival Factor Tox3. *The Journal of neuroscience* **35**, 10715-10730 (2015).
- 62 Sedel, F., Bernard, D., Mock, D. M. & Tourbah, A. Targeting demyelination and virtual hypoxia with high-dose biotin as a treatment for progressive multiple sclerosis. *Neuropharmacology* **110**, 644-653 (2016).
- 63 Hermey, G. *et al.* SorCS1, a member of the novel sorting receptor family, is localized in somata and dendrites of neurons throughout the murine brain. *Neuroscience letters* **313**, 83-87 (2001).
- 64 Subkhangulova, A. *et al.* SORCS1 and SORCS3 control energy balance and orexigenic peptide production. *EMBO reports* **19**, e44810 (2018).

- 65 Steinman, L. & Zamvil, S. Transcriptional analysis of targets in multiple sclerosis. *Nature Reviews Immunology* **3**, 483-492 (2003).
- 66 Matarese, G., Carrieri, P. B., Montella, S., De Rosa, V. & La Cava, A. Leptin as a metabolic link to multiple sclerosis. *Nature reviews. Neurology* **6**, 455-461 (2010).
- 67 Zachmann, J. *et al.* Combined Computational and Structural Approach into Understanding the Role of Peptide Binding and Activation of Melanocortin Receptor 4. *Journal of chemical information and modeling* **60**, 1461-1468 (2020).
- 68 Benjamins, J. A., Nedelkoska, L. & Lisak, R. P. Melanocortin receptor subtypes are expressed on cells in the oligodendroglial lineage and signal ACTH protection. *Journal of neuroscience research* **96**, 427-435 (2018).
- 69 Sadovnick, A. D. *et al.* Genetic modifiers of multiple sclerosis progression, severity and onset. *Clinical immunology (Orlando, Fla.)* **180**, 100-105 (2017).
- 70 Chen, J. *et al.* TRIM66 reads unmodified H3R2K4 and H3K56ac to respond to DNA damage in embryonic stem cells. *Nature communications* **10**, 4273 (2019).
- 71 Foolad, F., Khodagholi, F. & Javan, M. Sirtuins in Multiple Sclerosis: The crossroad of neurodegeneration, autoimmunity and metabolism. *Multiple sclerosis and related disorders* **34**, 47-58 (2019).
- 72 Versteeg, G. A. *et al.* The E3-ligase TRIM family of proteins regulates signaling pathways triggered by innate immune pattern-recognition receptors. *Immunity* **38**, 384-398 (2013).
- 73 Cohan, S. L. *et al.* Interferons and Multiple Sclerosis: Lessons from 25 Years of Clinical and Real-World Experience with Intramuscular Interferon Beta-1a (Avonex). *CNS drugs* **35**, 743-767 (2021).
- 74 Mo, M., Hoang, H. T., Schmidt, S., Clark, R. B. & Ehrlich, B. E. The role of chromogranin B in an animal model of multiple sclerosis. *Mol Cell Neurosci* **56**, 102-114 (2013).
- 75 Woo, M. S. *et al.* Neuronal metabotropic glutamate receptor 8 protects against neurodegeneration in CNS inflammation. *J Exp Med* **218**, e20201290 (2021).
- 76 Huttlin, E. L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505-509 (2017).
- 77 Abadier, M. *et al.* Cell surface levels of endothelial ICAM-1 influence the transcellular or paracellular T-cell diapedesis across the blood-brain barrier. *European journal of immunology* **45**, 1043-1058 (2015).
- 78 Defazio, G. *et al.* Interferon beta-1a downregulates TNF $\alpha$ -induced intercellular adhesion molecule 1 expression on brain microvascular endothelial cells through a tyrosine kinase-dependent pathway. *Brain research* **881**, 227-230 (2000).
- 79 Omilusik, K. *et al.* The Ca(v)1.4 calcium channel is a critical regulator of T cell receptor signaling and naive T cell homeostasis. *Immunity* **35**, 349-360 (2011).
- 80 Baecher-Allan, C., Kaskow, B. J. & Weiner, H. L. Multiple Sclerosis: Mechanisms and Immunotherapy. *Neuron* **97**, 742-768 (2018).
- 81 Fenninger, F. *et al.* Mutation of an L-Type Calcium Channel Gene Leads to T Lymphocyte Dysfunction. *Frontiers in immunology* **10**, 2473 (2019).
- 82 Cao, T., Duprez, E., Borden, K. L., Freemont, P. S. & Etkin, L. D. Ret finger protein is a normal component of PML nuclear bodies and interacts directly with PML. *Journal of cell science* **111**, 1319-1329 (1998).
- 83 Cai, X. *et al.* Tripartite motif containing protein 27 negatively regulates CD4 T cells by ubiquitinating and inhibiting the class II PI3K-C2 $\beta$ . *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20072-20077 (2011).