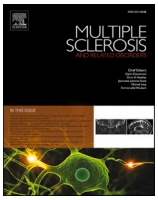
Multiple Sclerosis and Related Disorders 66 (2022) 104065

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard



X-linked genetic risk factors that promote autoimmunity and dampen remyelination are associated with multiple sclerosis susceptibility

Kirill Borziak *, Joseph Finkelstein

Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY 10029 United States

ARTICLE INFO

Keywords:
Multiple sclerosis
Genetic risk factors
X chromosome

ABSTRACT

Background: Multiple sclerosis (MS) is a chronic neurodegenerative disease, which has a strong genetic component and is more prevalent in women. MS is caused by an autoimmunity initiated inflammatory response which leads to axon demyelination, followed by axon loss, plaque formation and neurodegeneration. The goal of this article was to explore X-linked genetic factors that are associated with MS susceptibility.

Methods: Using UK Biobank microarray, we analyzed the prevalence of alleles on the X chromosome to identify variants potentially involved in MS. Overall, 488,225 patients across 18,857 markers were analyzed using PLINK.

Results: Our results identify 20 SNPs that are significantly more abundant in persons with MS. The genes associated with these SNPs belong to immunity (LAMP2, AVPR2, MTMR8, F8, BCOR, PORCN, and ELF4) and remyelination (NSDHL, HS6ST2, RBM10, TAZ, and AR) pathways that are potentially of great significance for understanding the onset and progression of multiple sclerosis. We further identified a significant 20-fold increase in incidence of MS cases in women with co-occurrences of SNPs associated with myelination and immunity functions.

Conclusions: Our analysis provides novel insights into the roles of X-linked genes in the onset and presentation of multiple sclerosis, identifying 20 SNPs in 14 genes involved primarily in immunity and myelination functions that are significantly more abundant in persons with MS. Our co-occurrence analysis suggests that concurrent disruption of both myelination and immune systems significantly increases the risk of MS onset in women.

1. Introduction

The pathogenesis of multiple sclerosis (MS) involves inflammatory mediators, which leads to apoptosis of oligodendrocytes and damage the axon myelin sheath (Lassmann and van Horssen, 2011). In MS, infiltration of immune cells through the blood-brain barrier (BBB) and their release of inflammatory cytokines represent the earliest cerebrovascular abnormalities seen in MS brains (Ortiz et al., 2014). The BBB consists of cerebral endothelial cells, pericytes and their basal lamina. Disruption of the BBB, in pathological conditions such as MS, allows T lymphocytes activated in the periphery to infiltrate the central nervous system to trigger the immune responses responsible for myelin damage (Ortiz et al., 2014). The infiltrating lymphocytes release cytotoxic factors including pro-inflammatory cytokines, proteases, and reactive oxygen species, initiate microglia and astrocytes, and recruit macrophages and other lymphocytes (Golden and Voskuhl, 2017). These pro-inflammatory processes then lead to demyelination and axonal death (Kamm et al., 2014).

Reduced conduction ability, due to decreased myelination, causes the deficiencies in sensation, movement, and cognition typically associated with MS (Garg and Smith, 2015). The myelin repair, remyelination process, does occur and is able to reverse the damage due to inflammation; however, repeated attacks result in less effective remyelination and the formation of plaques around the damaged axon (Lassmann and van Horssen, 2011; Garg and Smith, 2015). Remyelination begins spontaneously within MS lesions and is then associated with functional recovery and clinical remittances of MS symptoms (Podbielska et al., 2013). The new myelin sheath acts as a protective physical barrier against damage from inflammatory molecules and restores trophic support to the axon. Despite this, the remyelination process becomes less efficient with progressive damage, leading to increased neurodegeneration (Podbielska et al., 2013). Since the lack of myelination is the proximate cause of the axonal death and neurodegeneration associated with MS, remyelination has been an important topic of research in the treatment and recovery of persons with MS (Podbielska et al., 2013).

* Corresponding author.
E-mail address: Kirill.Borziak@mountsinai.org (K. Borziak).

https://doi.org/10.1016/j.msard.2022.104065
Received 1 February 2022; Received in revised form 14 April 2022; Accepted 17 July 2022
Available online 18 July 2022

It has been shown that MS affects women and men differently in regards to the central nervous system and immune system (Whitacre et al., 1999). For instance, MS is three times more prevalent in women than man during the reproductive years, while men have worse disease progression (Golden and Voskuhl, 2017; Bove, 2016). Because of the impact of MS on demyelination and the sex-linked differences in persons with MS, we focused our analysis on the X chromosome. The X chromosome has long been under investigation in its role in MS (Greer and McCombe, 2011). Among the potential roles of the X chromosome in accounting for the sex-linked differences in persons with MS are the role of unbalanced X chromosome inactivation in predisposing women to MS (Knudsen et al., 2007), and the presence of susceptibility genes on the X chromosome (Greer and McCombe, 2011). Among these potential genes of interest is the IL2R γ chain, an important component of interleukin receptors which has been found in increased levels in MS-affected brain tissues (Peerlings et al., 2021). Other potential genes of interest included Foxp3, the master regulator of regulatory T cells, which decrease in persons with MS (Huan et al., 2005); and the CD40 ligand, CD154, which is over-expressed on CD4+ T cells in persons with MS (Balashov et al., 1997). Additionally, previous research has been able to identify a significant X-linked SNP associated with MS, using genome-wide analysis (ABC 2019). However, despite the presence of multiple immune and remyelination response genes on the X chromosome, the SNPs located within these genes have not been fully analyzed for their implication in the presentation of MS. In addition variant co-occurrence analysis has not been systematically conducted previously.

In this study, we focused on analyzing the large array data set from the UK Biobank repository to identify causal low frequency alleles (single nucleotide polymorphisms) on X chromosome affecting immune and remyelination responses. The UK Biobank is a particularly useful resource, since it is a population-based data repository, with a focus on middle and old age diseases (UK Biobank Coordinating Centre; Ollier et al., 2005). The repository currently contains data from over 500,000 participants, collected from 2010 onwards (C Sudlow et al., 2015). The goal of the study was to analyze over 18,000 SNPs on the X chromosome in order to identify their associations with MS occurrence.

2. Material and methods

Genomic data from the UK Biobank Axiom® Array (Bycroft et al., 2018) and biometric data was obtained from the UK Biobank (C Sudlow et al., 2015). The UK Biobank Axiom Array covers 820,967 SNP and indel markers across all chromosomes. The array contains rare coding variants, composed of 30,581 protein truncating variants and 80,581 missense variants. Additionally, the array contains 348,569 common variants genome-wide and 280,838 low frequency variants genome-wide. In total allele variant information was obtained from 488,225 patients. To identify persons with MS, we utilized the International Classification of Diseases 10th Revision (World Health O, 2004) code G35, which is specific for multiple sclerosis within demyelinating diseases of the central nervous system. Specifically, the following Summary Diagnosis data fields from the UK Biobank were searched for G35: 41,270 (Diagnoses - ICD10), 41,202 (Diagnoses - main ICD10), 41,204 (Diagnoses - secondary ICD10), and 41,201 (External causes - ICD10).

Genomic data for the X chromosome was obtained from the UK Biobank data field 22,418. Specifically, the PLINK binary biallelic genotype table file ukb22418_cX_b0_v2.bed, and it's associated PLINK sample information file ukb22418_cX_b0_v2_s488225.fam were downloaded using the gfetch utility. Additionally, the PLINK extended MAP file for chromosome X was obtained from UK Biobank Resource 1963. The previously converted MS diagnosis information was encoded as the phenotype variable for the PLINK analysis. PLINK was used to perform the association analysis, using the default parameters, for all 18,857 SNPs that were mapped to the X chromosome (Purcell et al., 2007). False Discovery Rate multi-testing correction was employed to control the number of false positives.

Manhattan plots and Q-Q plots for significant SNPs were generated using the qqman R package (Turner, 2018). Functional clustering of genes was performed using DAVID 6.8 Functional Annotation Tool (Huang da et al., 2009). The Benjamini multiple testing correction was employed on the resultant p-values. The full Homo sapiens gene set was used as the background gene set.

Co-occurrence analysis was performed by grouping significant SNPs from myelination and autoimmune implicated genes into two categories. For each individual, co-occurrence was defined a presence of at least one alternate allele from each group. The significance of the interaction between MS diagnosis and the co-occurrence of myelination and autoimmune implicated variants was then assessed using a Chi-squared test.

2.1. Ethical approval

UK Biobank had obtained ethics approval from the North West Multicenter Research Ethics Committee (approval number: 16/NW/0274). Informed consent from all participants was obtained by the UK Biobank. The UK Biobank approved an application for use of the data (ID 69,385). All data used in this analysis has been fully de-identified by the UK Biobank, following the de-identification protocol V2. Further, the received Participant Data was released to researchers with distinct encrypted random number identifiers.

3. Results

Using the publically available UK Biobank resource, we examined 488,225 patients available in the database. The mean age of recruitment for participants in the UK Biobank set was 56.53 (S.D. = 8.1). The mean age of participants with MS at recruitment was 57.48 (S.D. = 9.2). To understand the available cohort of persons with MS, we first examined the frequency of multiple sclerosis in the database. In contrast with previous estimates of multiple sclerosis prevalence of 203.4 per 100,000 population of the United Kingdom (Mackenzie et al., 2014), we observed a significantly higher ratio of 404.3 individuals per 100,000 ($\chi 2$ p-value = 2.04e-21). In total, 1974 diagnosed cases of MS as defined by ICD10 code G35 were identified from the UK Biobank cohort. Next, we examined the sex based frequency of MS in the UK Biobank cohort. 0.536% (1419/264,772) of women were diagnosed with MS as defined by the G35 ICD10 code. In contrast, only 0.248% (555/223,453) of men have an MS diagnosis. In accordance with the previously reported higher prevalence of MS among women of 2.3–3.5:1 (Compston and Coles, 2002), we observed a ratio of 2.16:1, with 71.9% of MS cases in the UK Biobank being diagnosed in women. This constitutes a significantly higher prevalence of MS in women in the UK Biobank cohort ($\chi 2$ p-value = 1.76e-55).

To identify SNPs significantly associated with MS, we performed PLINK association case-control analysis for the X chromosome. In total, 488,377 individuals were analyzed. Of these, 223,453 were encoded as men, 264,772 as women, and 152 as unspecified sex. Further, 1974 were mapped as MS cases, 486,251 as controls, with 152 possessing a missing phenotype. Thus, a total genotyping rate in remaining individuals is 0.98. In total, we analyzed 18,857 SNPs that were present on the X chromosome as part of the UK Biobank Axiom® Array (Bycroft et al., 2018) We first analyzed the Q-Q plot, which revealed that the PLINK analysis had sufficient power to identify significant X-linked SNPs associated with MS (Fig. 1). To identify significant SNPs, we performed False Discovery Rate multi-testing corrections on the resultant PLINK p-values. This analysis revealed 44 SNPs that were significant to a FDR < 0.01 level (Fig. 2). This included 20 SNP variants that were significantly more abundant in persons with MS, and 24 SNP variants that were significantly more abundant in control cases. In total, the 20 significant SNP variants positively associated with MS prevalence were mapped to 14 genes (Table 1). All 14 genes are located outside the pseudoautosomal regions, implying that all male cases are homozygous at these SNP

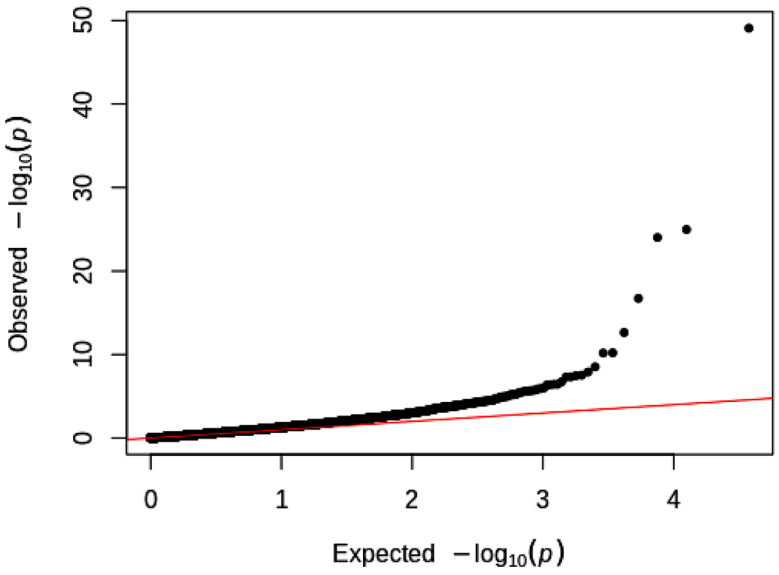


Fig. 1. Q-Q plot of the MS case-control association analysis for the 18,857 SNPs that were present on the X chromosome.

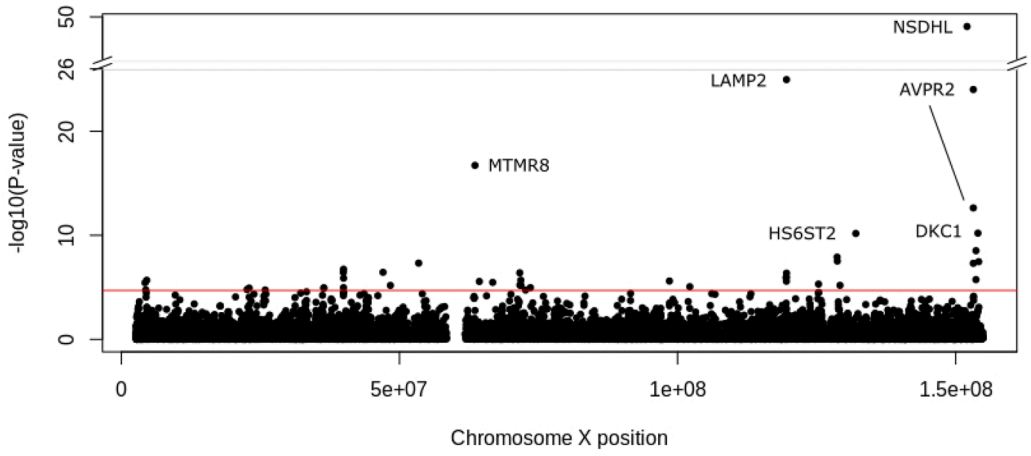


Fig. 2. Manhattan plot of the p-value from MS case-control association analysis for the 18,857 SNPs that were present on the X chromosome. The red line represents the FDR < 0.01 (P-value < 2.31E-5) significance line. The most significant SNPs (P-value < 1E-7) are labeled with their associated gene names.

locations. Suggesting the importance of these genes in proper homeostasis and metabolic functioning, the 14 genes were significantly enriched (Benjamini p-value = 3.2E-3) in UniProt keyword for Disease mutation (9/14). To get a better understanding of how these SNPs affect the function of the genes in which they reside, we next examined the individual SNPs.

We observed that the majority of genes with significant SNPs could be classified into two categories, widely implicated in MS onset and progression: myelination and immunity pathways. In total, we were able to identify 5 genes and 6 significant SNPs implicated in myelination functions, and 7 genes and 12 significant SNPs in immunity related functions (Table 1). Interestingly, the most significant SNP is implicated in myelination functionality.

3.1. Myelination implicated SNPs

The most significant SNP was rs797045835, which maps to the NAD (P) dependent steroid dehydrogenase-like (NSDHL) gene (FDR p-value = 1.57E-45, OR = NA). The alternative allele for this SNP is an 8 nt deletion, which results in a frameshift mutation affecting the tail 27 aa of the NSDHL protein. The NSDHL protein performs essential roles in the production of cholesterol. Further, the cholesterol pathway and NSDHL specifically have been recently implicated in the demyelination associated with MS (Ulrich et al., 2010). Specifically, downregulation of cholesterol biosynthesis was associated with increased demyelination, and it is possible to speculate that the frameshift mutation results in the dysregulation of the NSDHL protein.

The SNP rs950792996 (FDR p-value = 1.78E-7, OR = 55.28) represents an intron variant in heparan sulfate 6-O-sulfotransferase 2 (HS6ST2). Recent research has shown that heparan sulfate accumulation by oligodendrocyte cells slows demyelination and promotes remyelination associated with MS (Macchi et al., 2020). Thus our results suggest that HS6ST2 has a potential role in helping generate specific heparan sulfates responsible for promoting remyelination, which this variant potentially disrupts.

The gene tafazzin (TAZ) contain two significant SNPs: rs387907218 and Affx-89,017,095. rs387907218 (FDR p-value = 7.08E-6, OR = 44.2) is a missense variant (*G* > R) that is associated with infantile dilated X-linked cardiomyopathy (Bissler et al., 2002). Affx-89,017,095 (FDR p-value = 1.55E-3, OR = 27.6) also represents a missense mutation (*I* > N) in the putative acyl-acceptor binding pocket of the Lysophospholipid Acyltransferases (LPLATs) of Glycerophospholipid Biosynthesis: AGPAT-like domain, with the potential to disrupt or alter substrate binding. TAZ plays an important role in remodeling cardiolipin and by

Table 1
Genes with significant (FDR < 0.01) MS associated SNPs, classified by their putative roles in remyelination and autoimmunity.

Gene name	Gene symbol	SNPs	Putative role
androgen receptor	AR	rs367604031	myelination
arginine vasopressin receptor 2	AVPR2	Affx-89,012,620; Affx-89,008,152; Affx-89,010,658	immunity
BCL6 corepressor	BCOR	rs199676230	immunity
coagulation factor VIII	F8	rs369414658	immunity
dyskerin pseudouridine synthase 1	DKC1	rs121912302	other
E74 like ETS transcription factor 4	ELF4	rs373568641	immunity
heparan sulfate 6-O-sulfotransferase 2	HS6ST2	rs950792996	myelination
lysosomal associated membrane protein 2	LAMP2	rs1194422515; rs42895; rs41300191; rs42886	immunity
myotubularin related protein 8	MTMR8	rs766668643	immunity
NAD(P) dependent steroid dehydrogenase-like	NSDHL	rs797045835	myelination
porcupine O-acyltransferase	PORCN	rs1556974235	immunity
RIB43A domain with coiled-coils 1	RIBC1	rs782346908	other
RNA binding motif protein 10	RBM10	rs139585263	myelination
tafazzin	TAZ	rs387907218; Affx-89,017,095	myelination

extension the proper structure and function of mitochondria (Chin and Conway, 2020). It has been experimentally shown that cerebrospinal fluid of patients with progressive MS causes neuronal mitochondrial elongation, which is thought to contribute to the metabolic impairment of neuronal bioenergetics underlying neurodegeneration associated with MS (Wentling et al., 2019). Recently, increased expression of TAZ has been identified in MS populations relative to control groups (Khalilian et al., 2021). Further, inhibiting nuclear localization of TAZ, using dimethyl fumarate mediated inhibition of PI3-K/Akt1 pathway, reduces the relapse rate and disability progression in persons with MS (Toyama et al., 2018; Miclea et al., 2016). Finally, it has been shown that proper functioning of TAZ is crucial for Schwann cells to maintain nerve myelination in adult tissues (Grove et al., 2017).

RNA binding motif protein 10 (RBM10) contains the significant SNP rs139585263 (FDR p-value = 4.51E-4, OR = 31.57), which encodes a synonymous variant. RBM10 is an RNA-binding protein that regulates alternative splicing of DNA (cytosine-5)-methyltransferase 3b (DNMT3B) (Atsumi et al., 2017). DNMT3B regulates the activity of NF-κB-responsive promoters and consequently inflammation development. Further, increased demethylation activity of, in part, DNMT3B has been shown to coincide with hippocampal demyelination in persons with MS (Chomyk et al., 2017). If the observed synonymous variant is able to increase translational efficiency of RBM10, it has the downstream potential to promote demethylation and, in turn, the demyelination associated with MS.

A significant SNP rs367604031 (FDR p-value = 2.25E-3, OR = 13.39) is also located in the Androgen receptor gene (AR). It represents a missense (*E* > *Q*) variant, which is localized between two phosphorylation sites and adjacent to the transcription activation unit Tau-5 (Tan et al., 2015; Guo et al., 2006). A functional AR is required for promoting remyelination, as shown through the action of testosterone and 5αDHT treatment of experimental autoimmune encephalomyelitis (Hussain et al., 2013). Indeed, lower testosterone levels have been identified in men with MS and correlate with worsened scores of physical and cognitive disability (Chitnis, 2018).

3.2. Immunity implicated SNPs

The most significant immunity implicated SNP was rs1194422515, which maps to lysosomal associated membrane protein 2 (LAMP2) (FDR p-value = 1.00E-21, OR = 221.7). The alternative allele is a single nucleotide deletion, inducing a frameshift, which results in disruption of terminal 218 amino acid residues that encode the second lumenal domain and a protein binding site. In addition to rs1194422515, three additional SNPs are significantly associated with MS risk. Those SNPs are rs42895 (FDR p-value = 4.53E-4, OR = 1.194); rs42886 (FDR p-value = 1.83E-3, OR = 1.486); and rs41300191 (FDR p-value = 1.01E-3, OR = 1.177). While rs42895 and rs42886 are intron variants, rs41300191 is a 3′ UTR variant. Autophagy, in which LAMP2 participates, is tightly linked to autoimmune regulation, and directly participates in the progress of MS. Further, as inflammation and oxidative stress are increased in MS lesions, LAMP2 expression is reduced. As such the resultant frameshift mutation induced by the alternative rs1194422515 allele likely similarly reduces the abundance of LAMP2 and in turn promotes inflammation and oxidative stress.

The arginine vasopressin receptor 2 (AVPR2) gene has 3 significant SNPs with alternative alleles more prevalent in MS cases. These four SNPs are Affx-89,012,620 (FDR p-value = 6.03E-21, OR = 213.5); Affx-89,008,152 (FDR p-value = 8.71E-10, OR = 73.63); and Affx-89,010,658 (FDR p-value = 7.11E-5, OR = 36.69). Affx-89,012,620 is a 1 nucleotide insertion, resulting in a frameshift mutation, affecting the terminal 59 amino acid residues. This region also includes transmembrane helix 7 of the AVPR2 protein. Disruption of this transmembrane region by other SNP insertions, such as rs886040961, has been implicated in causing nephrogenic diabetes insipidus. Affx-89,008,152 and Affx-89,010,658 result in missense mutations. Vasopressin has been implicated in multiple sclerosis. Vasopressin (AVP) is released after brain injury and contributes to the inflammatory response. Previous research showed that blocking the AVPR2 receptor can decrease BBB permeability and affect MS progression (Viñuela-Berni et al., 2020). Thus the association of MS with these AVPR2 mutations seems to indicate increased function of the AVPR2 receptor in promoting BBB permeability.

Another significant SNP, rs766668643, (FDR p-value = 8.85E-14, OR = 110.4) encodes a stop gain variant in the myotubularin related protein 8 gene. The stop gain prematurely terminated the MTMR8 gene, removing the Myotubularin-like phosphatase domain. MTMR8 functions to dephosphorylate phosphatidylinositol 3-phosphate [PtdIns(3)P], which in turn decreases activity of autophagy processes (Zou et al., 2012). Given that autophagy is implicated in two of the main hallmarks of MS, neurodegeneration and inflammation (Misrielal et al., 2020), it has been recently shown in a cohort study that autophagic activity was increased in relapsing-remitting persons with MS (Hassanpour et al., 2020). We speculate that the stop gain mutation to MTMR8, which removes its catalytic domain, promotes increased PtdIns(3)P levels, which in turn promote increased autophagy as seen in relapsing-remitting persons with MS.

We also identify a significant SNP in the coagulation factor VIII gene. The SNP rs369414658 (FDR p-value = 5.87E-5, OR = 18.45) is a serine to threonine missense variant. Factor VIII has also been implicated in BBB permeability (Ziliotto et al., 2019). Previous research has also shown an association between MS and a factor VIII deficit in a family-based study (Capra et al., 1993).

The BCL6 corepressor gene (BCOR) presents an interesting case as it contains a SNP variant that is significantly more abundant in persons with MS (rs199676230) and also three SNP variants that are significantly more abundant in the control population (rs5963736, rs5963739, rs4076107). The rs199676230 SNP (FDR p-value = 4.24E-4, OR = 31.55) variant encodes a stop codon gain mutation. This mutation disrupts two ankyrin protein-protein binding domains, and the critical Polycomb Group Ring Finger 1 (PCGF1) binding domain. PCGF1 is an important factor in regulation of hematopoietic cell differentiation

(Ross et al., 2012). In turn, BCOR has been shown to be an important regulator of Th2 cell differentiation (Kusam et al., 2003). Further, reduction of Th2 derived cytokines, particularly CCL22, appear to play an important role in the pathogenesis of MS, especially in women (Jafarzadeh et al., 2014).

The porcupine O-acyltransferase (PORCN) gene contains a significant SNP, rs1556974235 (FDR p-value = 3.74E-3, OR = 24.54), which a missense (R > C) variant. PORCN acts as a stimulator of Wnt secretion, which has been implicated in promoting the proper formation of the BBB phenotype (Laksitorini et al., 2019).

Finally, ETS-related transcription factor Elf-4 (ELF4) contains the significant SNP rs373568641 (FDR p-value = 3.79E-3, OR = 12.68), which encodes a missense (*S* > P*)* variant. ELF4 functions to inhibit differentiation of CD4+ T cells into Th17 cells (Lee et al., 2014). Th17 cells play an important role in the pathogenesis of multiple sclerosis by promoting BBB disruption and promote central nervous system inflammation through CD4+ lymphocyte recruitment (Kebir et al., 2007).

3.3. Other significant SNPs

Another significant SNP, rs121912302, (FDR p-value = 1.94E-7, OR = 55.44) is located in the dyskerin pseudouridine synthase 1 (DKC1) gene. This is a missense mutation, which has previously been associated with X-linked dyskeratosis congenita (Knight et al., 1999). One of the primary roles of DKC1 is its essential function in telomere stability and in preventing telomere shortening (Parry et al., 2011). Research also suggest a possible role for accelerated telomere shortening in the progression and pathobiology of MS (Hecker et al., 2021).

The gene RIB43A domain with coiled-coils 1 (RIBC1) contains one significant SNP, rs782346908 (FDR p-value = 7.35E-5, OR = 36.8). It encodes a splice donor variant, which has the potential to affect protein structure through exclusion of exons or inclusion of intron sequences into the mature mRNA.

3.4. Co-occurrence analysis

Our results revealed the importance of mutations in myelination and immunity genes on the X-chromosome in the presentation of MS. Due to the importance of deficient remyelination and overactive autoimmunity (particularly with regard to BBB disruption driven inflammation) phenotypes in driving multiple sclerosis progression, we asked whether individuals with mutant variants from both functional classes (immunity and remyelination) were more likely to be present in MS (Table 2). To analyze the importance of the concurrent disruption of myelination and immune functions in driving the MS phenotypes, we performed a co-occurrence analysis. First, we observed that no men have been identified in possessing alternative alleles for significant SNPs associated with remyelination implicated genes. Further, of all women with these remyelination implicated SNPs (58 total), none possessed two copies of the alternant variants and were all heretozygous at these SNP positions. The co-occurrence analysis revealed that concurrent presence of both remyelination and autoimmunity SNPs was significantly enriched in women with MS (20.7X; $\chi 2$ p-value=0.0). This result implies that women possessing variant alleles in both groups (immunity and remyelination) have a 20 fold higher risk of developing MS. This effect was largely driven by remyelination SNPs that were also significantly enriched in women diagnosed with MS (22.8X; $\chi 2$ p-value=0.0). Our results indicate that individuals with co-occurring variant alleles in both X-linked remyelination implicated genes and X-linked immune functioning genes are over 20 times more likely to have MS.

Table 2
Co-occurrence of significant SNPs classified as myelination functioning and immunity functioning in the UK Biobank cohort as variant frequency per 100,000 people.

Category	Frequency of co-occurrence	Frequency of immune implicated SNPs	Frequency of remyelination implicated SNPs
Total population	9.22	54,059.50	13.11
All women	17.00	65,080.52	24.17
All men	0.00	41,000.57	0.00
Women with MS	352.36	70,472.16	493.31
Men with MS	0.00	46,126.13	0.00
Women without MS	15.19	65,051.47	21.64
Men without MS	0.00	40,987.81	0.00

4. Discussion

Here we present an analysis using a large, genomic, publically-available resource, the UK Biobank, to identify alleles which potentially contribute to the observed sex-bias in presentation of multiple sclerosis. Since MS is a progressive, autoimmune disease which presents through axonal demyelination and neuronal death, and presents almost three times more commonly in women, we examined the X chromosome for possible informative variant alleles. Using the genomic and biomedical information from 488,377 individuals available as part of the UK Biobank cohort, we performed chromosome-wide association analysis for MS occurrence with the 18,857 SNPs that were present on the X chromosome. Our analysis identified 20 significant SNPs, at an FDR level of less than 0.01, that were significantly associated with MS. These SNPs belong to 14 genes. Although many of these genes have been tangentially implicated in MS, as described in the results, our results present the first evidence of causal alleles within them that are significantly associated with MS occurrence. Among them are NSDHL, LAMP2, AVPR2, MTMR8, HS6ST2, DKC1, TAZ, and F8. These genes fall into two main categories: seven genes that are implicated in inflammatory responses, and five genes that are implicated in myelination functions.

Genes that promote inflammation do so through a variety of pathways. However, we observed that the majority (4) are implicated in BBB phenotypes: AVPR2, F8, PORCN, and ELF4. The breakdown of the BBB, in which the significant SNPs from these genes have been implicated, has also been implicated in allowing infiltration of lymphocytes which release pro-inflammatory cytokines, proteases, and reactive oxygen species, responsible for demyelination (Golden and Voskuhl, 2017; Kamm et al., 2014). Thus our results further support the role of a breakdown in BBB functionality in driving MS onset. The other immunity implicated genes, LAMP2 and MTMR8, regulate the inflammatory response, while BCOR regulates immune cell development. Autophagy, in which LAMP2 and MTMR8 have been functionally implicated, plays a major role in two of the main hallmarks of MS, neurodegeneration and inflammation, making it especially important to understand how this pathway contributes to MS manifestation and progression (Misrielal et al., 2020). This may be particularly impactful for LAMP2, where disrupted lysosome function has previously been implicated in neurodegenerative diseases, including Alzheimer's disease, amyotrophic lateral sclerosis and familial Parkinson's disease (Nixon, 2013).

The five genes that affect myelination are involved in diverse pathways. The most significant SNP from our analysis resides in the NSDHL gene, which regulates production of cholesterol by NSDHL, an important ingredient in myelin formation. High cholesterol levels have been shown as essential for myelin membrane growth (Saher et al., 2005). Other significant genes affect myelination through less direct pathways, including heparan sulfate production by HS6ST2, promotion of demethylation by RBM10, or proliferation regulation by TAZ. Of particular interest was the significant SNP located within the androgen receptor gene. AR has been shown to promote remyelination, through the action of testosterone and 5αDHT (Hussain et al., 2013). Indeed, testosterone shows both neuroprotective effects and protects against autoimmunity (Saher et al., 2005). This result provides new support to the

immunomodulatory role of testosterone and further suggests that low testosterone levels potentially predispose women to increased rates of MS. Highlighting the particularly important role the myelination genes play, none of the significant SNPs were found to be present in men or in two copies in women.

The immunity and myelination pathways have been previously shown to play important roles in the onset and progression of MS (Lassmann and van Horssen, 2011; Ortiz et al., 2014). However, the interplay between these pathways has not been fully explored. Due to the importance of immunity and myelination pathways in the onset and progression of MS, we asked if those individuals with concurrent mutations in significant SNPs to both the immunity genes and the myelination genes were more likely to be diagnosed with MS. For this purpose, we performed a co-occurrence analysis, which revealed that mutations in both myelination and autoimmune/blood brain barrier functionalities significantly increase the risks to developing MS, particularly in women. Women with mutations in both groups of genes were over 20 times more likely to have a MS diagnosis, compared to those with a mutations in only one or neither of these gene groups. These results continue to further strengthen the BBB permeability initiated inflammatory response and the remyelination response as the two critical forces that cause MS onset, axon loss, and subsequent progressive neurodegeneration that characterizes MS. In particular our results further support the important role the remyelination process appears to play in protecting axons from the aberrant autoimmune process, characteristic of MS.

Multiple previous studies have attempted to implicate SNPs in the onset or progression of MS using large genomic datasets (Parnell and Booth, 2017; Gresle et al., 2020; Cotsapas and Mitrovic, 2018; George et al., 2016). However, large genome-wide association studies (GWAS) have mostly focused on autosomal differences, at the expense of analyzing the potential effect of the X chromosome linked SNPs (Voskuhl, 2020). Previous research has been able to identify only one X-linked SNP associated with MS, as part of a genome-wide analysis (ABC, 2019). The reported SNP rs2807267 is missing from the UK Biobank Axiom array, so we were unable to further validate its significance. We believe our work is the first to focus specifically on the X chromosome as the potential source SNPs, which can be implicated in the onset and progression of MS. Our results shed further light on how mutation to the inflammatory and remyelination pathways are able to promote MS occurrence and progression. In particular, our identification of mutations of genes involved in BBB permeability (AVPR2, F8, PORCN, and ELF4) sheds additional light on the mechanisms which can promote infiltration of T-helper cells, which release cytokines responsible for demyelinated lesions associated with MS (Golden and Voskuhl, 2017; Kamm et al., 2014).

Although our research provides new insights into SNPs causal for MS, some limitations attenuate the power of the identified SNPs in predicting the likelihood of MS onset or the disease severity. One major limitation is the utilization of the UK Biobank Axiom Array, which covers only 18,857 SNPs across the X chromosome out of the potential 156 million total base pairs. To address this limitation, we will carry out this analysis using the UK Biobank exome sequencing dataset, once it becomes fully available in the future. The utilization of full protein coding regions will allow us to verify the SNPs identified here and to create a much fuller representation of the protein coding changes across the entire X chromosome which might be causal for MS onset. An additional limitation is the gross aggregation of MS cases. Given the differential disease presentation and prognosis across the four different types of MS, particularly distinctions between relapsing-remitting and progressive MS types, our study is limited in classifying the identified SNPs by potential disease severity. Although the UK Biobank resource does not have MS type designation, in future follow-up research we will aim to utilize the brain MRI images available through UK Biobank to classify patients by MS type, using unsupervised machine learning approaches such as Eshagi et al. (Eshaghi et al., 2021). To further the utility of the identified SNPs in predicting MS onset or progression, we also plan to analyze the differences in distributions of the significant SNPs we identified between women and men. This will allow us to better understand which of the significant genes are causal for MS, specifically in women.

5. Conclusion

Our analysis provides a novel insight into the roles of X-linked genes in the onset and presentation of multiple sclerosis. We identify 20 SNPs in 14 genes involved primarily in immunity and myelination functions that are significantly more abundant in persons with MS. The immunity genes primarily function in maintaining the blood-brain barrier, the disruption of which allows for the onset of autoimmune mediated inflammation and demyelination. The implicated myelination genes highlight the importance of a properly functioning myelination system to help prevent the onset of neurodegeneration characteristic of MS. Finally, our co-occurrence analysis revealed that concurrent disruption of both myelination and immune systems significantly increases the risk of MS onset in women by 20 fold.

Disclosure

The author reports no competing interests to declare with regards to this work.

CRediT authorship contribution statement

Kirill Borziak: Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Joseph Finkelstein:** Conceptualization, Validation, Writing – review & editing, Supervision, Project administration, Resources.

Acknowledgements

Funding was provided by NIH grant number UL1TR001433.

References

Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science 365 (6460), 2019.
Atsumi, T., Suzuki, H., Jiang, J.J., et al., 2017. Rbm10 regulates inflammation development via alternative splicing of Dnmt3b. Int. Immunol. 29 (12), 581–591.
Balashov, K.E., Smith, D.R., Khoury, S.J., Hafler, D.A., Weiner, H.L., 1997. Increased interleukin 12 production in progressive multiple sclerosis: induction by activated CD4+ T cells via CD40 ligand. Proc. Natl. Acad. Sci. U. S. A. 94 (2), 599–603.
Bissler, J.J., Tsoras, M., Göring, H.H., et al., 2002. Infantile dilated X-linked cardiomyopathy, G4.5 mutations, altered lipids, and ultrastructural malformations of mitochondria in heart, liver, and skeletal muscle. Lab. Invest. 82 (3), 335–344.
Bove, R., 2016. Women's Issues in Multiple Sclerosis. Semin. Neurol. 36 (2), 154–162.
Bycroft, C., Freeman, C., Petkova, D., et al., 2018. The UK Biobank resource with deep phenotyping and genomic data. Nature 562 (7726), 203–209.
Capra, R., Mattioli, F., Kalman, B., Marcianò, N., Berenzi, A., Benetti, A., 1993. Two sisters with multiple sclerosis, lamellar ichthyosis, beta thalassaemia minor and a deficiency of factor VIII. J. Neurol. 240 (6), 336–338.
Chin, M.T., Conway, S.J., 2020. Role of Tafazzin in Mitochondrial Function, Development and Disease. J. Dev. Biol. 8 (2).
Chitnis, T., 2018. The role of testosterone in MS risk and course. Mult. Scler. 24 (1), 36–41.
Chomyk, A.M., Volsko, C., Tripathi, A., et al., 2017. DNA methylation in demyelinated multiple sclerosis hippocampus. Sci. Rep. 7 (1), 8696.
Compston, A., Coles, A., 2002. Multiple sclerosis. Lancet 359 (9313), 1221–1231.
Cotsapas, C., Mitrovic, M., 2018. Genome-wide association studies of multiple sclerosis. Clin. Transl. Immunol. 7 (6), e1018.
Eshaghi, A., Young, A.L., Wijeratne, P.A., et al., 2021. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. Nat. Commun. 12 (1), 2078.
Garg, N., Smith, T.W., 2015. An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. Brain Behav. 5 (9), e00362.
George, M.F., Briggs, F.B., Shao, X., et al., 2016. Multiple sclerosis risk loci and disease severity in 7,125 individuals from 10 studies. Neurol. Genet. 2 (4), e87.
Golden, L.C., Voskuhl, R., 2017. The importance of studying sex differences in disease: the example of multiple sclerosis. J. Neurosci. Res. 95 (1–2), 633–643.
Greer, J.M., McCombe, P.A., 2011. Role of gender in multiple sclerosis: clinical effects and potential molecular mechanisms. J. Neuroimmunol. 234 (1–2), 7–18.

Gresle, M.M., Jordan, M.A., Stankovich, J., et al., 2020. Multiple sclerosis risk variants regulate gene expression in innate and adaptive immune cells. Life Sci. Alliance 3 (7).
Grove, M., Kim, H., Santerre, M., et al., 2017. YAP/TAZ initiate and maintain Schwann cell myelination. Elife 6.
Guo, Z., Dai, B., Jiang, T., et al., 2006. Regulation of androgen receptor activity by tyrosine phosphorylation. Cancer Cell 10 (4), 309–319.
Hassanpour, M., Hajihassani, F., Hiradfar, A., et al., 2020. Real-state of autophagy signaling pathway in neurodegenerative disease; focus on multiple sclerosis. J. Inflamm. (Lond.) 17, 6.
Hecker, M., Fitzner, B., Jäger, K., et al., 2021. Leukocyte Telomere Length in Patients with Multiple Sclerosis and Its Association with Clinical Phenotypes. Mol. Neurobiol. 58 (6), 2886–2896.
Huan, J., Culbertson, N., Spencer, L., et al., 2005. Decreased FOXP3 levels in multiple sclerosis patients. J. Neurosci. Res. 81 (1), 45–52.
Huang da, W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4 (1), 44–57.
Hussain, R., Ghoumari, A.M., Bielecki, B., et al., Jan 2013. The neural androgen receptor: a therapeutic target for myelin repair in chronic demyelination. Brain 136 (Pt 1), 132–146.
Jafarzadeh, A., Ebrahimi, H.A., Bagherzadeh, S., et al., 2014. Lower serum levels of Th2-related chemokine CCL22 in women patients with multiple sclerosis: a comparison between patients and healthy women. Inflammation 37 (2), 604–610.
Kamm, C.P., Uitdehaag, B.M., Polman, C.H., 2014. Multiple sclerosis: current knowledge and future outlook. Eur. Neurol. 72 (3–4), 132–141.
Kebir, H., Kreymborg, K., Ifergan, I., et al., 2007. Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. Nat. Med. 13 (10), 1173–1175.
Khalilian, S., Hojati, Z., Dehghanian, F., et al., 2021. Gene expression profiles of YAP1, TAZ, CRB3, and VDR in familial and sporadic multiple sclerosis among an Iranian population. Sci. Rep. 11 (1), 7713.
Knight, S.W., Heiss, N.S., Vulliamy, T.J., et al., 1999. X-linked dyskeratosis congenita is predominantly caused by missense mutations in the DKC1 gene. Am. J. Hum. Genet. 65 (1), 50–58.
Knudsen, G.P., Harbo, H.F., Smestad, C., et al., 2007. X chromosome inactivation in females with multiple sclerosis. Eur. J. Neurol. 14 (12), 1392–1396.
Kusam, S., Toney, L.M., Sato, H., Dent, A.L., 2003. Inhibition of Th2 differentiation and GATA-3 expression by BCL-6. J. Immunol. 170 (5), 2435–2441.
Laksitorini, M.D., Yathindranath, V., Xiong, W., Hombach-Klonisch, S., Miller, D.W., 2019. Modulation of Wnt/β-catenin signaling promotes blood-brain barrier phenotype in cultured brain endothelial cells. Sci. Rep. 9 (1), 19718.
Lassmann, H., van Horssen, J., 2011. The molecular basis of neurodegeneration in multiple sclerosis. FEBS Lett. 585 (23), 3715–3723.
Lee, P.H., Puppi, M., Schluns, K.S., Yu-Lee, L.Y., Dong, C., Lacorazza, H.D, 2014. The transcription factor E74-like factor 4 suppresses differentiation of proliferating CD4 + T cells to the Th17 lineage. J. Immunol. 192 (1), 178–188.
Macchi, M., Magalon, K., Zimmer, C., et al., 2020. Mature oligodendrocytes bordering lesions limit demyelination and favor myelin repair via heparan sulfate production. Elife 9.
Mackenzie, I.S., Morant, S.V., Bloomfield, G.A., MacDonald, T.M., Riordan, J., 2014. Incidence and prevalence of multiple sclerosis in the UK 1990–2010: a descriptive study in the General Practice Research Database. Journal of Neurology, Neurosurgery & Psychiatry 85 (1), 76.
Miclea, A., Leussink, V.I., Hartung, H.P., Gold, R., Hoepner, R., 2016. Safety and efficacy of dimethyl fumarate in multiple sclerosis: a multi-center observational study. J. Neurol. 263 (8), 1626–1632.
Misrielal, C., Mauthe, M., Reggiori, F., Eggen, B.J.L., 2020. Autophagy in Multiple Sclerosis: two Sides of the Same Coin. Front. Cell. Neurosci. 14, 603710.
Nixon, R.A., 2013. The role of autophagy in neurodegenerative disease. Nat. Med. 19 (8), 983–997.
Ollier, W., Sprosen, T., Peakman, T., 2005. UK Biobank: from concept to reality. Pharmacogenomics 6 (6), 639–646.
Ortiz, G.G., Pacheco-Moisés, F.P., Macías-Islas, M., et al., 2014. Role of the blood-brain barrier in multiple sclerosis. Arch. Med. Res. 45 (8), 687–697.
Parnell, G.P., Booth, D.R., 2017. The Multiple Sclerosis (MS) Genetic Risk Factors Indicate both Acquired and Innate Immune Cell Subsets Contribute to MS Pathogenesis and Identify Novel Therapeutic Opportunities. Front. Immunol. 8, 425.
Parry, E.M., Alder, J.K., Lee, S.S., et al., 2011. Decreased dyskerin levels as a mechanism of telomere shortening in X-linked dyskeratosis congenita. J. Med. Genet. 48 (5), 327–333.
Peerlings, D., Mimpen, M., Damoiseaux, J., 2021. The IL-2 - IL-2 receptor pathway: key to understanding multiple sclerosis. J. Transl. Autoimmun. 4, 100123.
Podbielska, M., Banik, N.L., Kurowska, E., Hogan, E.L., 2013. Myelin recovery in multiple sclerosis: the challenge of remyelination. Brain Sci. 3 (3), 1282–1324.
Purcell, S., Neale, B., Todd-Brown, K., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.
Ross, K., Sedello, A.K., Todd, G.P., et al., 2012. Polycomb group ring finger 1 cooperates with Runx1 in regulating differentiation and self-renewal of hematopoietic cells. Blood 119 (18), 4152–4161.
Saher, G., Brügger, B., Lappe-Siefke, C., et al., 2005. High cholesterol level is essential for myelin membrane growth. Nat. Neurosci. 8 (4), 468–475.
Sudlow, C., Gallacher, J., Allen, N., et al., 2015a. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12 (3), e1001779.
Sudlow, C., Gallacher, J., Allen, N., et al., 2015b. UK Biobank: an Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. 12 (3), e1001779.
Tan, M.H., Li, J., Xu, H.E., Melcher, K., Yong, E.L, 2015. Androgen receptor: structure, role in prostate cancer and drug discovery. Acta Pharmacol. Sin. 36 (1), 3–23.
Toyama, T., Looney, A.P., Baker, B.M., et al., 2018. Therapeutic Targeting of TAZ and YAP by Dimethyl Fumarate in Systemic Sclerosis Fibrosis. J. Invest. Dermatol. 138 (1), 78–88.
Turner, S.D., 2018. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. 3 (25), 731.
UK Biobank Coordinating Centre. Protocol for a large-scale prospective epidemiological resource. https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf.
Ulrich, R., Kalkuhl, A., Deschl, U., Baumgärtner, W., 2010. Machine learning approach identifies new pathways associated with demyelination in a viral model of multiple sclerosis. J. Cell. Mol. Med. 14 (1–2), 434–448.
Viñuela-Berni, V., Gómez-González, B., Quintanar-Stephano, A., 2020. Blockade of Arginine Vasopressin receptors prevents blood-brain barrier breakdown in Experimental Autoimmune Encephalomyelitis. Sci. Rep. 10 (1), 467.
Voskuhl, R.R., 2020. The effect of sex on multiple sclerosis risk and disease progression. Mult. Scler. 26 (5), 554–560.
Wentling, M., Lopez-Gomez, C., Park, H.J., et al., 2019. A metabolic perspective on CSF-mediated neurodegeneration in multiple sclerosis. Brain 142 (9), 2756–2774.
Whitacre, C.C., Reingold, S.C., O'Looney, P.A, 1999. A gender gap in autoimmunity. Science 283 (5406), 1277–1278.
World Health O, 2004. ICD-10: International Statistical Classification of Diseases and Related Health problems: Tenth Revision, 2nd ed. World Health Organization, Geneva.
Ziliotto, N., Bernardi, F., Jakimovski, D., Zivadinov, R., 2019. Coagulation Pathways in Neurological Diseases: multiple Sclerosis. Front. Neurol. 10, 409.
Zou, J., Zhang, C., Marjanovic, J., Kisseleva, M.V., Majerus, P.W., Wilson, M.P., 2012. Myotubularin-related protein (MTMR) 9 determines the enzymatic activity, substrate specificity, and role in autophagy of MTMR8. Proc. Natl. Acad. Sci. U. S. A. 109 (24), 9539–9544.