

2022

# Rare variants and HLA haplotypes associated in patients with neuromyelitis optica spect...

*This work was made openly accessible by BU Faculty. Please [share](#) how this access benefits you. Your story matters.*

Version	Published version
Citation (published version):	I. Tabansky, A.J. Tanaka, J. Wang, G. Zhang, I. Dujmovic, S. Mader, V. Jeganathan, T. DeAngelis, M. Funaro, A. Harel, M. Messina, M. Shabbir, V. Nursey, W. DeGouvia, M. Laurent, K. Blitz, P. Jindra, M. Gudesblatt, Regeneron Genetics Center, A. King, J. Drulovic, E. Yunis, V. Brusic, Y. Shen, D.B. Keskin, S. Najjar, J.N.H. Stern. 2022. "Rare variants and HLA haplotypes associated in patients with neuromyelitis optica spectrum disorders." <i>Frontiers in Immunology</i> , Volume 13, pp.900605-. <a href="https://doi.org/10.3389/fimmu.2022.900605">https://doi.org/10.3389/fimmu.2022.900605</a>

<https://hdl.handle.net/2144/46948>

*Boston University*



## OPEN ACCESS

## EDITED BY

Linda L. Kusner,  
George Washington University,  
United States

## REVIEWED BY

Chris Bundell,  
Pathwest Laboratory Medicine,  
Australia  
Antonino Giordano,  
San Raffaele Scientific Institute  
(IRCCS), Italy

## \*CORRESPONDENCE

Joel N. H. Stern  
joel.stern@gmail.com;  
joel.n.h.stern@post.harvard.edu

<sup>†</sup>These authors have contributed  
equally to this work

## SPECIALTY SECTION

This article was submitted to  
Autoimmune and  
Autoinflammatory Disorders,  
a section of the journal  
Frontiers in Immunology

RECEIVED 20 March 2022

ACCEPTED 21 July 2022

PUBLISHED 04 October 2022

## CITATION

Tabansky I, Tanaka AJ, Wang J,  
Zhang G, Dujmovic I, Mader S,  
Jeganathan V, DeAngelis T, Funaro M,  
Harel A, Messina M, Shabbir M,  
Nurse V, DeGouvia W, Laurent M,  
Blitz K, Jindra P, Gudesblatt M,  
Regeneron Genetics Center, King A,  
Drulovic J, Yunis E, Brusica V, Shen Y,  
Keskin DB, Najjar S and Stern JNH  
(2022) Rare variants and HLA  
haplotypes associated in  
patients with neuromyelitis  
optica spectrum disorders.  
*Front. Immunol.* 13:900605.  
doi: 10.3389/fimmu.2022.900605

# Rare variants and HLA haplotypes associated in patients with neuromyelitis optica spectrum disorders

Inna Tabansky<sup>1,2,3,4,5,6†</sup>, Akemi J. Tanaka<sup>7†</sup>, Jiayao Wang<sup>7,8†</sup>,  
Guanglan Zhang<sup>9†</sup>, Irena Dujmovic<sup>10,11†</sup>, Simone Mader<sup>5,12†</sup>,  
Venkatesh Jeganathan<sup>5</sup>, Tracey DeAngelis<sup>13</sup>,  
Michael Funaro<sup>1,2,3,4,5</sup>, Asaff Harel<sup>14</sup>, Mark Messina<sup>1,2,3,4,5</sup>,  
Maya Shabbir<sup>5</sup>, Vishaan Nursey<sup>5</sup>, William DeGouvia<sup>1,2,3,4,5</sup>,  
Micheline Laurent<sup>1,2,3,4,5</sup>, Karen Blitz<sup>15</sup>, Peter Jindra<sup>16</sup>,  
Mark Gudesblatt<sup>12</sup>, Regeneron Genetics Center,  
Alejandra King<sup>17</sup>, Jelena Drulovic<sup>10</sup>, Edmond Yunis<sup>18</sup>,  
Vladimir Brusica<sup>19</sup>, Yufeng Shen<sup>8</sup>, Derin B. Keskin<sup>20,21</sup>,  
Souhel Najjar<sup>14</sup> and Joel N. H. Stern<sup>1,2,3,4,5\*</sup>

<sup>1</sup>Department of Neurology, Donald and Barbra Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States, <sup>2</sup>Department of Urology, Donald and Barbra Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States, <sup>3</sup>Department of Molecular Medicine, Donald and Barbra Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States, <sup>4</sup>Department of Science Education, Donald and Barbra Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States, <sup>5</sup>Institute of Molecular Medicine, The Feinstein Institutes for Medical Research, Manhasset, NY, United States, <sup>6</sup>Department of Neurobiology and Behavior, The Rockefeller University, New York, NY, United States, <sup>7</sup>Department of Pediatrics, Columbia University College of Physicians and Surgeons, New York, NY, United States, <sup>8</sup>Department of Biomedical Informatics and Department of Systems Biology, Columbia University, New York, NY, United States, <sup>9</sup>Department of Computer Science, Boston University, Boston, MA, United States, <sup>10</sup>Clinical Center of Serbia University School of Medicine, Belgrade, Serbia, <sup>11</sup>Department of Neurology, University of North Carolina School of Medicine, Chapel Hill, NC, United States, <sup>12</sup>Biomedical Center and University Hospitals, Ludwig Maximilian University Munich, Munich, Germany, <sup>13</sup>Department of Neurology, Neurological Associates of Long Island, New Hyde Park, NY, United States, <sup>14</sup>Department of Neurology, Lenox Hill Hospital, Northwell Health, New York, NY, United States, <sup>15</sup>Department of Neurology, South Shore Neurologic Associates, Patchogue, NY, United States, <sup>16</sup>Division of Abdominal Transplantation, Baylor College of Medicine, Baylor College of Medicine, Houston, TX, United States, <sup>17</sup>Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY, United States, <sup>18</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, United States, <sup>19</sup>School of Computer Science, University of Nottingham Ningbo China, Ningbo, China, <sup>20</sup>Department of Translational Immuno-Genomics for Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA, United States, <sup>21</sup>Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

Neuromyelitis optica spectrum disorders (NMOSD) are rare, debilitating autoimmune diseases of the central nervous system. Many NMOSD patients have antibodies to Aquaporin-4 (AQP4). Prior studies show associations of NMOSD with individual Human Leukocyte Antigen (HLA) alleles and with mutations in the complement pathway and potassium channels. HLA allele associations with NMOSD are inconsistent between populations, suggesting complex relationships between the identified alleles and risk of disease. We

used a retrospective case-control approach to identify contributing genetic variants in patients who met the diagnostic criteria for NMOSD and their unaffected family members. Potentially deleterious variants identified in NMOSD patients were compared to members of their families who do not have the disease and to existing databases of human genetic variation. HLA sequences from patients from Belgrade, Serbia, were compared to the frequency of HLA haplotypes in the general population in Belgrade. We analyzed exome sequencing on 40 NMOSD patients and identified rare inherited variants in the complement pathway and potassium channel genes. Haplotype analysis further detected two haplotypes, HLA-A\*01, B\*08, DRB1\*03 and HLA-A\*01, B\*08, C\*07, DRB1\*03, DQB1\*02, which were more prevalent in NMOSD patients than in unaffected individuals. *In silico* modeling indicates that HLA molecules within these haplotypes are predicted to bind AQP4 at several sites, potentially contributing to the development of autoimmunity. Our results point to possible autoimmune and neurodegenerative mechanisms that cause NMOSD, and can be used to investigate potential NMOSD drug targets.

#### KEYWORDS

neuromyelitis optica spectrum disorders, human leukocyte antigen, exome sequencing, autoimmunity, antibody mediated

## Introduction

Neuromyelitis optica spectrum disorders (NMOSD) are inflammatory disorders of the central nervous system marked by astrocyte damage, severe immune-mediated demyelination and axonal damage of the optic nerves and the spinal cord (1, 2). NMOSD has an estimated prevalence between 0.5 and 10 cases per 100,000 in the population (3–5), but patients may still be underdiagnosed due to the similarity of symptoms to other autoimmune neurodegenerative diseases, such as multiple sclerosis (MS), and the expanding spectrum of humorally mediated autoimmune syndromes of the CNS in NMOSD (6). The clinical distinction between NMOSD and MS, particularly in early disease stages, may be difficult because of a considerable overlap between the symptoms; however, accurate diagnosis is important as patients with NMOSD require different treatment (2, 7).

NMOSD was first described in a series of patients in 1894 (8), but a unique biomarker was not identified until 2004, when a disease-specific serum NMOSD-immunoglobulin G (IgG) antibody was detected in the vast majority of patients with NMOSD (9–12). It was not observed in individuals with other autoimmune neurodegenerative diseases (9–12). The NMOSD IgG antibody binds to astrocyte water channel protein Aquaporin-4 (AQP4): a member of the aquaporin water-selective membrane channel family of proteins. In humans, AQP4 is the predominant water channel protein in the brain. It is prevalent in astrocytes at the blood-brain barrier and has an essential role in brain water homeostasis (13). It has been demonstrated in several models that AQP4 autoantibodies

have a pathogenic role in NMSOD, and that they lead to tissue destruction primarily through complement-mediated cytotoxicity (13–18). NMOSD is characterized by astrocytic cell death and destruction of the blood-brain-barrier (BBB), while MS primarily affects oligodendrocytes (16, 19, 20). Destruction of the BBB in NMOSD may partially be mediated by AQP4-IgG antibodies (21) and by pathogenic autoreactive T cells (22–25). Neuronal cell death and oligodendrocyte damage are considered as secondary events (26).

In a subgroup of NMOSD patients with a diverse clinical phenotype, autoantibodies to a myelin specific glycoprotein (MOG) can be found at high titers (27–30). This group of patients is most recently defined as MOG associated diseases (MOGAD) (31, 32). It is becoming increasingly clear that patients with AQP4 IgG positive NMOSD and those with MOGAD manifest two separate CNS demyelinating diseases with different biological, clinical, and neuropathological characteristics. NMOSD most commonly presents as acute attacks of bilateral or rapidly sequential optic neuritis, or characteristic MRI findings of long spanning spinal cord lesions and/or brain lesions. Patients with variable presentation are also classified as having NMOSD (8). Patients with AQP4 IgG antibodies detected in blood have been associated with NMOSD whereas the presence of MOG antibodies leads to a diagnosis MOGAD (29).

NMOSD predominantly occurs in adults and is tenfold more prevalent in women than in men (33–35). Women are more frequently affected by relapsing disease (35). The average age of NMOSD onset is between 32 and 41 years of age, but it can be quite variable (36, 37), while the onset of MS commonly in the early

twenties. Based on limited population data, AQP4 seropositive NMOSD is more prevalent in populations of African and Afro-Caribbean descent in the US (5), while rates of MS are thought to be higher among people of European ancestry (38).

Previous studies have described an association with NMOSD and different alleles of the Human Leukocyte Antigen (HLA), which are variable between populations (39). Studies of genetic susceptibility of NMOSD have indicated an association with HLA DRB1\*03:01 in French, Brazilian, Mestizo Mexican and Afro-Caribbean populations, but not in Danish, Spanish, or Arab populations (40–45). Some of these patients were seronegative for AQP4 antibodies, but still met the criteria for NMOSD. In Japan, NMOSD was associated with HLA-DRB1\*1602 and DPB1\*0501 alleles (46). Genetic associations of NMOSD with alleles outside the HLA region are only starting to be explored. Genome wide association studies (GWAS) of NMOSD (43–46) have shown an association with single nucleotide polymorphisms (SNPs) located near genes encoding complement proteins and several candidate genes involved in NMOSD, including a potassium channel gene, KCMA1 (47–49).

In our present study we performed HLA haplotype analysis in addition to whole exome sequencing of individuals with NMOSD to identify candidate genes and pathways that underlie the pathogenesis of NMOSD disease.

NMOSD rarely affects multiple members of the same family. This could be due to environmental or other factors, including *de novo* and somatic mutations, which have recently been associated with multigenic inheritance in patients with neurodevelopmental and psychiatric conditions (50, 51). To understand whether *de novo* mutations could contribute to NMOSD, and to identify candidate genes and pathways, we used trio analysis of exome sequencing data in families in which at least one member was affected by NMOSD.

## Material and Methods

### Patients

The present study was approved by the Northwell Health Institutional Review Board. NMOSD patients (first and second degree) who met the revised diagnostic criteria for NMOSD were enrolled at the Clinic of Neurology Clinical Center of Serbia (n=26, referred to as Serbian patients) and at Northwell Health and Lenox Hill Hospital, Neurological Associates of Long Island, and South Shore Neurologic Associates, New York, USA (n=14). In total, 33 females and seven male patients were enrolled. Only NMOSD patients without any clinical evidence of other autoimmune diseases were included in this study.

Blood samples were collected from patients affected with NMOSD according to protocols approved by the Northwell Health Institutional Review Board (IRB# HS15-06-03). Written informed consent was obtained from all participants.

### Cell-based antibody assay

All NMOSD patients and relatives were screened for antibodies against AQP4 (M23 isoform) and human MOG using a cell-based assay as previously described (30, 52). Samples were tested using HEK 293T cells transfected with relevant epitopes, as well as untransfected negative controls. HEK 293T cells were used 24 hours after transfection with MOG, or 72 hours after transfection with AQP4. To test for antibody binding we first blocked with PBS/10% fetal calf serum (FCS) for an hour, then incubated with patient serum diluted with PBS/10% FCS at 4°C for an hour. To visualize antibody staining, cells were incubated with Alexa-594 conjugated goat anti-human IgG (Life Technologies) for 30 minutes at room temperature, then washed in PBS/10% FCS, stained with DAPI to assess the viability and visualized with a fluorescent microscope. Titer  $\geq$  1:160 were classified as MOG-IgG positive (30).

### High Resolution HLA typing

HLA typing was performed with genomic high resolution typing (second field) of HLA- A, B, C, DRB1, DRB3, DRB4, DRB5, DQA, DQB, DPA and DPB. All loci were analyzed by PCR-sequence specific oligonucleotide (PCR-SSO) probing using LABType SSO HD kits (One Lambda). A Luminex 200 (Luminex Corporation) flow analyzer was used to identify the fluorescence intensity of PE (phycoerythrin) of PCR products bound to fluorescently coded microspheres. The data were imported into HLA Fusion software version 4.4.0 (One Lambda) for analysis.

### Statistics

#### Analysis of HLA sequencing

Two datasets of HLA allele frequencies in Serbian populations from Allele Frequency Net Database (52) were used to compare HLA profiles of NMOSD patients to the general Serbian population. The dataset named “Serbia\_pop3” was derived from HLA typing of 1,992 volunteer bone marrow donors (53). The dataset named “Serbia\_Vojvodina” was derived from HLA typing of a sample population of 400 Serbians. Dataset Serbia\_pop3 was used for comparison to the patient population, as it was based on a much larger sample. Our HLA typing is of higher resolution than that of Serbia\_pop3, which is in 2-digit resolution. To make the results comparable, we computationally reduced our typing result to two decimal places, corresponding to nonsynonymous substitutions only (Supplementary Table 5). Using the Fisher exact test, we compared allele frequency in our Serbian patient cohort with the general Serbian population. The investigators were blinded to the clinical information of the patients in the analysis.

## Peptide binding predictions

NetMHC 4.0 (54, 55) and NetMHCII 2.2 servers (DTU Health Tech, Lyngby, Denmark) (56) were used for HLA class I and class II binding predictions, respectively. For class I binding prediction, thresholds used for strong and weak binders were top 1% rank and top 2% rank respectively. For class II binding prediction, thresholds used to define strong and weak binders were affinities of  $\leq 50\text{nM}$  and  $\leq 500\text{nM}$ , respectively.

## Visualization of HLA binding to AQP4

Pymol software (Schrodinger, Inc.) was used to model the protein structure of AQP4 (3GD8). The membrane spanning helices were identified based on the existing literature (22). Peptides were identified with the software based on the sequence. The intracellular region of AQP4 was identified based on the presence of N and C terminals. The peptide sequences closer to the N and C terminals could not be visualized because they had not been included in the crystal structure.

## Whole exome sequencing analysis

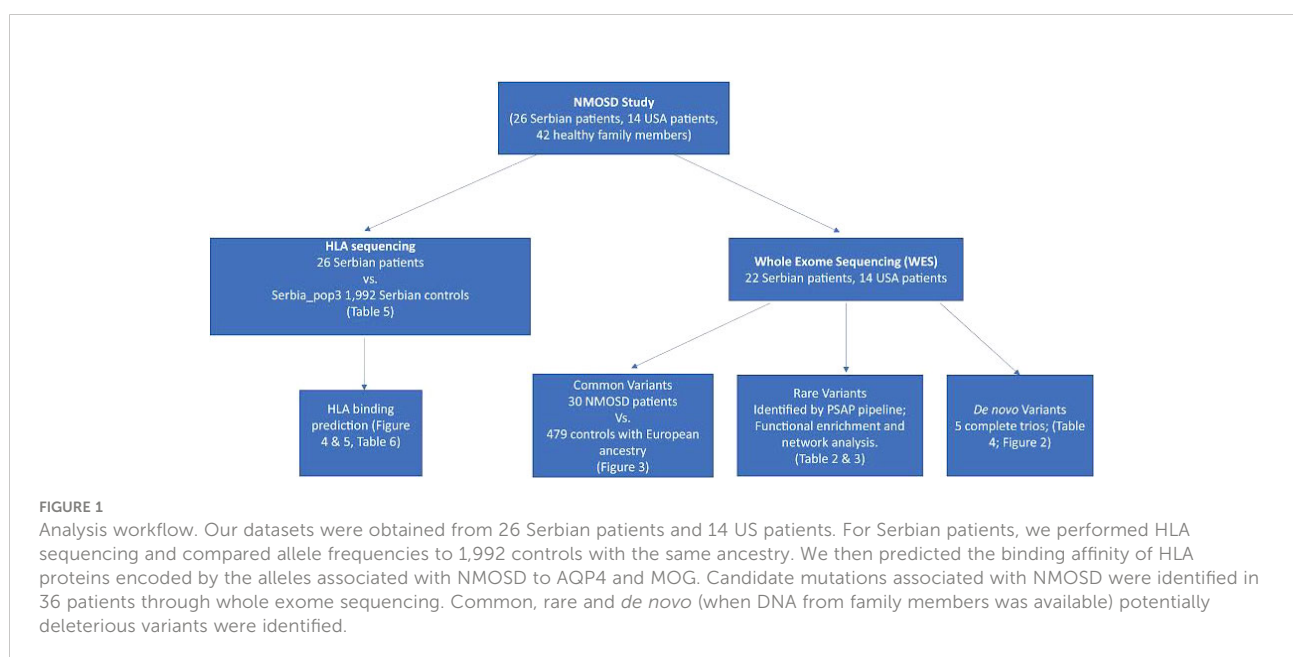
Whole exome sequencing was performed on 36 individuals, including 5 trios (Figure 1), 21 singletons, and 10 families with first and/or second degree relatives. Clinical characteristics of the patients are listed in Table 1. (Detailed patients information listed in Supplementary Table 4). Overall, 33 of 36 NMOSD

TABLE 1 Clinical and demographic characteristics of NMOSD Patients.

	NMOSD patients (n = 40)
Female (%)	80.8
Mean Age	47.6 (std=11.6)
AQP4-IgG Positive (%)	92.3
Clinical Presentations (%):	
Optical neuritis (ON)	65.4
Visual Loss	65.4
Myelopathy	92.3
Motor Weakness	92.3

patients where AQP4 IgG positive and none were MOG IgG positive. (Table 1). All relatives were tested negative for AQP4 IgG and MOG IgG. DNA for whole-exome sequencing was isolated using Qiagen DNA isolation kit. Whole exome sequencing was performed at Regeneron Genetics Center (Tarrytown, NY, USA). Exome sequence was captured by Xgen kit (Integrated DNA Technologies, Inc., Coralville, IA, USA) and sequenced on an Illumina Hi-seq 2500 (Illumina) with 78 base pair paired-end reads and read-depth coverage of at least 15X for 95% of target regions for all samples.

Data were processed by a bioinformatics pipeline based on Genome Analysis Toolkit (GATK) best practice. Reads were mapped to HG19 by BWA-mem algorithm (57). GATK-haplotype caller and GenotypeGVCF were used for variant discovery and joint genotyping the whole cohort. SmartPCA and in-house script based on a random forest algorithm were used for determining the populations of patients (58, 59). ANNOVAR was used to annotate the consequence of variants



and aggregate information about allele frequencies and *in silico* prediction algorithms of deleteriousness prediction (57).

All variants were filtered for Genome Aggregation Database (gnomAD) and ExAC allele frequency < 0.01, Genotype quality (GQ) > 60, minimal read depth of 5, minimal allele fraction of 0.1, Fisher allele balance test score (FS) < 25, Quilyt per depth (QD) > 2, segmental duplications < 0.95 and mappability  $\geq 1$ , which are more stringent than common variants, since rare variants usually have a higher false positive rate than common variants. We used Population Sampling Probability (PSAP) to identify potential rare variants as disease candidates (59) to address the small size of our cohort limiting the power of a case-control burden test. We ran PSAP on all patients, filtered for rare variants as Minor Allele Frequency (MAF) < 0.1%, and assess the p-value of how likely this variant was to be seen in the general population (ExAC) (60). More specifically, we used the following parameters: popscore < 1e-3, MAF in gnomAD and ExAC < 1e-3, Combined Annotation Dependent Depletion (CADD) score  $\geq 20$  to identify genes most likely to contribute to NMOSD. Variants were classified according to the ACMG variant interpretation guidelines (61). In addition to adhering to the ACMG criteria, our analysis specifically focused on variants that occurred in a gene with a known function, were shared among all affected individuals in our cohort, and predicted as deleterious by multiple *in silico* prediction algorithms including Combined Annotation Dependent Depletion (CADD) (62) and Rare Exome Variant Ensemble Learner (REVEL) (63). Functional enrichment of gene harbouring damaging mutations were performed with Enrichr (64). To identify whether these genes have interactions with each other to achieve same biological functions, we performed protein-protein interaction (PPI) network analysis with STRING database (<https://string-db.org/>, v10.5), and corresponding network were constructed using candidate genes selected by PSAP. Genes are represented by nodes and interaction between genes are represented by edges. Types of interactions were restricted based on experimental evidence, interaction databases and co-expression. The minimum requirement of evidence score was set to 0.9, which resulted in only the highest confidence of interaction.

### Common SNP association

To determine whether an association exists between NMOSD and common SNPs, we selected common variants (MAF > 1%) identified in the individuals. Control data were obtained from other disease cohorts, which were sequenced together with our NMOSD patients in the same sequencing batch. These cohorts are case-control studies of other diseases such as breast cancer, cardiomyopathy and obesity. We selected healthy family members from these cohorts as the controls, and potential technical batch effects that may impact our GWAS study were thus minimized. Genome-wide Complex Trait

Analysis (GCTA) (65) was used to check pedigrees and select index patients from control data. European cases and controls were used to perform association tests to avoid population stratification. Eigensoft (66) was used to perform principal component analysis (PCA) and an in house script based on random forest algorithm was used to select European individuals according to 1KG references. After selecting on population and relatedness, 30 NMOSD cases and 479 controls with European ancestry were selected. Wald test for logistic regression (b.wald) test implemented in Efficient and Parallelizable Association Container Toolbox (EPACTS) was used to perform statistical testing on selected SNPs and individuals. The first 10 eigen values from population PCA were used as covariates. For variant quality, filters on GQ > 30 and allele fraction in [0.3, 0.7] were applied. Variants that failed these filters were marked and not called.

## Results

### Identification of rare, deleterious variants by whole exome sequencing

Analysis of Whole Exome Sequencing (WES) data identified several rare variants (MAF < 1% of the general population) in genes potentially associated with NMOSD. Rare variants are increasingly becoming understood to play an important role in disease susceptibility to multiple disorders, including autoimmune neurodegenerative diseases, such as multiple sclerosis (MS) (67). After applying the filters for high-confidence rare-variant screening (see Method), we observed 346 statistically significant variants in 338 genes. The majority of the alleles were heterozygous missense variants (Supplementary Table 1). Two immune related pathways, Class I MHC mediated antigen processing and MHC class II antigen (Supplementary Table 2), were identified to be enriched with selected rare damaging missense variants. Mutations in a potassium channel have been previously proposed to confer susceptibility to NMOSD (49), and mutations in potassium channels were also identified to be enriched with rare damaging missense variants in our patient cohort (68). The PPI network analysis gave us three groups of highly interconnected genes with predicted damaging mutations in our exome sequencing patient cohort. The first set of genes, including UBE2Q1, HUWE1, UBE2H, ASB8, ZNRF1 and TRIM37, are pairwise connected in network and form the function of class I MHC mediated antigen processing. The second gene group consisted of MHC class II antigen genes: DCTN1, CANX, DCTN3 and OSBPL1A. The third network showed DCTN1, CLTC, DRD4, OSBPL1A and CCR4, which are genes involved in neuroactive ligand-receptor interaction, G alpha (i) signaling events and G Protein Coupled Receptors (GPCRs) downstream signaling.

## Rare Variants Associated with Immune Function in NMOSD patients

Since NMOSD involves an autoimmune component, we focused on variants specifically involved in immune function. Several variant candidates were identified in immune genes (69), including NOTCH1, SPINK5, GUSB, IL6ST, CCR4 and FN1 (fibronectin 1). These genes were all previously documented with deleterious effects of loss of function in heterozygotes, or dose-dependent effects in other disease processes (Table 2).

A recent study has indicated that mutations in the complement cascade, specifically complement 4A, can increase genetic susceptibility to NMOSD (47). We observed a marginally significant enrichment of rare damaging missense mutations ( $MAF < 1e-3$ ,  $REVEL > 0.5$ ,  $Enrichment = 3.6$ ,  $p\text{-value} = 0.054$ , binomial test) in the complement pathway. We identified two predicted pathogenic variants, one in C7 and one in C8B, in addition to multiple variants predicted deleterious to protein structure and function by *in silico* algorithms (Table 3). A variant documented to cause C8B deficiency, type II was also detected (70, 71). Two variants were detected in C1S, which is associated with Ehlers-Danlos Syndrome, Periodontal Type, 2

and Complement Component C1s Deficiency. The C7, C8B and C9 complement genes are all downstream of the complement C4A gene, which has previously been described as a potential risk factor for NMOSD (47). We did not identify C4A mutations in our NMOSD patient cohort.

## De novo Mutations in NMOSD patients

While it is possible for multiple cases of NMOSD to occur within a single family, the disease usually arises in a single family member. It is therefore possible that the disorder arises from multigenic factors, including *de novo* mutations that contribute to disease susceptibility. We identified three *de novo* mutations in patients with NMOSD by whole exome sequencing of trios, including the proband and the proband's mother and father (Table 4). Pedigrees of individuals in this study are illustrated in Figure 2.

In patient NMO-51, we detected a c.1645G>A (p.Glu549Lys) missense mutation in gene BIN2 (Bridging integrator2). BIN2 is a membrane remodeling adaptor for SH3 domain containing proteins located at dynamic actin-remodeling adhesive structures

TABLE 2 Selected candidate inherited mutations for NMOSD.

Gene	Chr	Pos	Ref	Alt	func	revel	cadd	gnomadAF	Samples	Other disorders
CCR4	3	32995947	T	C	Dmis	0.374	25.8	3.23E-05	NMO-63	Steven Johnson Syndrome
FN1	2	2.16E+08	G	A	Dmis	0.287	24.6	0.0002	NMO-38; NMO-54	glomerulopathy and spondylometaphyseal dysplasia
FN1	2	2.16E+08	A	G	Dmis	0.252	26.5	3.23E-05	NMO-62	glomerulopathy and spondylometaphyseal dysplasia
FN1	2	2.16E+08	G	A	Dmis	0.448	34	0.0002	NMO-38; NMO-54	glomerulopathy and spondylometaphyseal dysplasia
GUSB	7	65439334	A	T	Dmis	0.914	25.6	0	NMO-31	mucopolysaccharidosis type VII
IL6ST	5	55259272	T	C	Dmis	0.153	22.5	0.0002	NMO-66	arthritis-like autoimmune disease
IL6ST	5	55243391	C	T	Dmis	0.201	33	0	NMO-37	arthritis-like autoimmune disease
NOTCH1	9	1.39E+08	C	T	Dmis	0.169	24.5	0.0004	NMO-27	aortic valve disease, Adams-Oliver syndrome, T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, and head and neck squamous cell carcinoma.
NOTCH1	9	1.39E+08	G	A	Dmis	0.325	24.6	6.47E-05	NMO-55	aortic valve disease, Adams-Oliver syndrome, T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, and head and neck squamous cell carcinoma.
SPINK5	5	1.47E+08	G	T	Dmis	0.197	24	0	NMO-5	Netherton syndrome

TABLE 3 List of damaging rare genetic variants identified by exome sequencing in the genes encoding proteins that are part of the complement system.

Gene	Chr	Pos	Ref	Alt	func	revel	cadd	gnomadAF	Samples
C1S	12	7169873	A	G	Dmis	0.038	22.4	0.0006	NMO-32
C1S	12	7174371	C	T	Dmis	0.151	24.6	0	NMO-31
C3	19	6709721	T	C	Dmis	0.181	25.6	3.24E-05	NMO-69
C7	5	40931185	T	G	Dmis	0.554	26.6	0	NMO-41
C8B	1	57420448	G	C	Dmis	0.868	27.4	3.23E-05	NMO-42
C9	5	39341348	C	T	Dmis	0.579	34	0.0002	NMO-69

TABLE 4 *De novo* variants identified in patients with NMOSD.

Proband	Chr	Pos	Ref	Alt	Gene	GeneName	VarType	AAChange	gnomAD_ genome_MAF	MetaSVM	CADD	pLI
NMOSD-31	3	1.25E+08	G	T	SNX4	sorting nexin 4	missense	SNX4:NM_003794:exon13:c.C1251A;p.D417E	0	T	21	0.01
NMOSD-43	4	4E+07	G	A	PDS5A	PDS5 cohesin associated factor A	missense	PDS5A:NM_001100399:exon25:c.C2935T;p.R979C	0	T	34	1
NMOSD-61	12	5.2E+07	C	T	BIN2	bridging integrator 2	missense	BIN2:NM_001290008:exon10:c.G1549A;p.E517K-BIN2:NM_001290009:exon10:c.G1177A;p.E393K-BIN2:NM_001290007:exon11:c.G1471A;p.E491K-BIN2:NM_016293:exon11:c.G1645A;p.E549K	0.0003	D	17.79	0

of immune cells (72). It promotes cell motility and migration, probably *via* its interaction with the cell membrane and with podosome proteins that mediate interaction with the cytoskeleton. In patient NMO-32, we observed a missense c.1251C>A (p.Asp417Glu) mutation in Sorting Nexin 4 (SNX4) gene. SNX4 has been associated with Alzheimer's disease (AD) (73). SNX4 regulates  $\beta$ -site amyloid precursor protein-cleaving enzyme 1 (BACE1) trafficking, overexpression of SNX4 significantly increased the levels of BACE1 and  $\beta$ -amyloid (A $\beta$ ) in AD. In patient NMO-43, a missense mutation c.2935C>T (p.Arg979Cys) in the *PDS5A* gene, the protein encoded by this gene binds to the cohesin complex and associates with chromatin through most of the cell cycle. The encoded protein may play a role in regulating sister chromatid cohesion during mitosis.

## Common variants associations

In addition to analyzing rare and *de novo* mutations in patients, we also tested 200,292 common variants that were found in our exome sequencing dataset for associations with NMOSD. A Manhattan plot is shown in Figure 3A and Figure 3B shows the GWAS signal on chromosome 6. While there is a peak associated with the variant, it does not rise to the level of statistical significance, which is likely due to the small sample size of our study. Figure 3C shows a QQ plot of the test results. While multiple common variants were observed to be more prevalent in NMOSD patients than in the general population, none of them reached statistical significance.

We further tested whether our population sample could validate the results of a recent study (47) of NMOSD. The previous study used GWAS that identified 2 common SNPs as associated with NMOSD. These SNPs are rs28383224 (MAF = 0.42, odds ratio = 2.24, p-value =  $5.8 \times 10^{-12}$ ) mapped 21.5 kb downstream of *HLADQA1*; and rs1150757 (MAF = 0.10, odds ratio = 4.66, p-value =  $3.33 \times 10^{-16}$ ). Our analysis included only one of the previously identified SNPs, as rs28383224 occurs outside

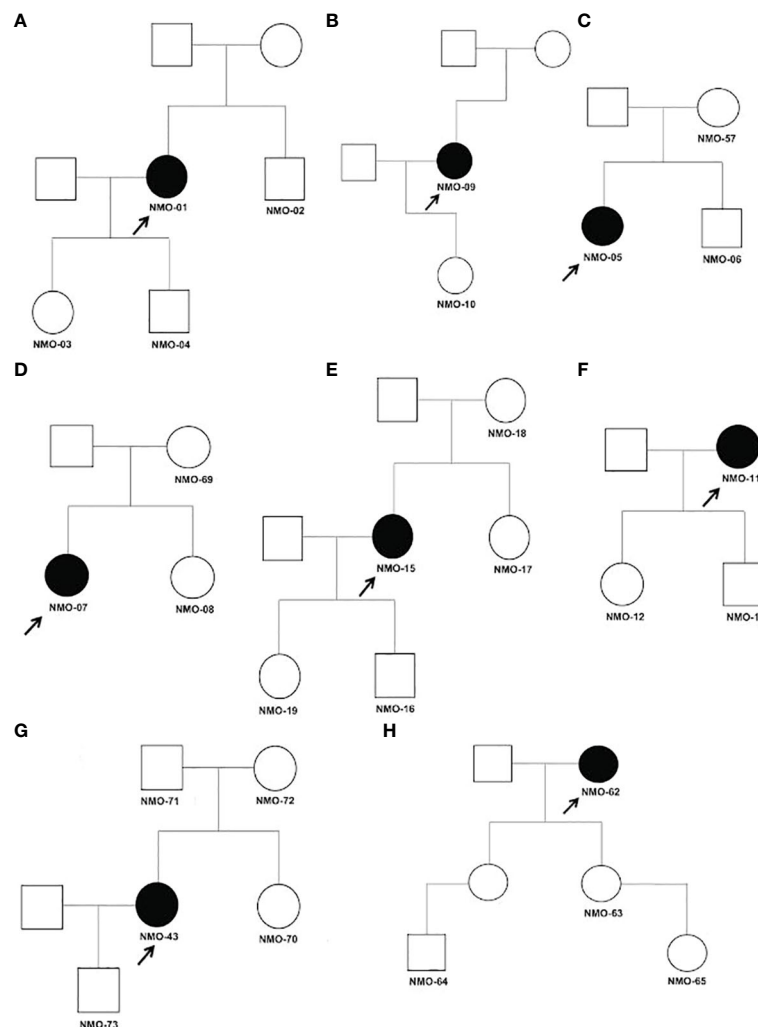
the coding sequence and is therefore not picked up by WES. The second variant identified in the previous study, rs1150757, was also observed in our dataset (MAF = 0.096 in controls and 0.25 in cases, p-value =  $1.52 \times 10^{-5}$ ).

## HLA typing of Serbian Patient Population

Specific HLA alleles are associated with certain autoimmune diseases, and several previous studies were able to observe an association between HLA alleles and NMOSD. However, studies in other populations did not observe this association. One potential explanation for this result is the possibility of linkage disequilibrium with HLA alleles and NMOSD risk loci. To address this possibility, we analyzed HLA haplotypes in NMOSD and compared the prevalence of these haplotypes with the general population. We were able to identify a haplotype association with NMOSD in our Serbian patients. Frequencies of HLA haplotypes, including HLA-A, -B, -DRB1 and HLA-A, -B, -C, -DRB1, -DQB1, have previously been reported in this population (52, 53). We were then able to use these previously available data to calculate the estimated corresponding haplotype frequencies for comparison to the patient population and family members (full comparison result is in Supplementary Table 4). Detailed clinical data for these patients is in Supplementary Table 4. The most significant difference we observed between the Serbian population and the NMOSD patients is in the haplotypes HLA-A\*01, B\*08, DRB1\*03 and HLA-A\*01, B\*08, C\*07, DRB1\*03, DQB1\*02. We observed an increased prevalence of the haplotype HLA-A\*01, B\*08, DRB1\*03, at a frequency of 0.21154 in patients, as opposed to 0.059114 in the Serbian population overall (p-value =  $2e^{-4}$ , Fisher exact test) (Supplementary Table 5, Serbia\_pop3). Likewise, there was an increased prevalence of haplotype HLA-A\*01, B\*08, C\*07, DRB1\*03, DQB1\*02, with a frequency of 0.21154 in NMOSD patients, as opposed to 0.066038 in Serbia\_pop3 (p-value =  $5 \times 10^{-4}$ , Fisher exact test).

Multiple analysis of Serbian haplotypes are available, and a comparison between them indicates that the haplotype frequencies





**FIGURE 2**  
**(A–H)** Pedigrees of tested relatives and probands for trio analysis. Affected individuals are shaded black and probands are indicated by arrows. All individuals tested in our analysis are indicated by their assigned number. Males are represented by squares and females by circles. Detailed clinical data for each family described in [Supplementary Table 4](#).

used for our analysis are accurate. As shown in [Supplementary Table 5](#), the HLA allele frequencies are similar between two datasets we could have used for our analysis, Serbia\_pop3 and Serbia\_Vojvodina. A short list of the HLA alleles with more than 0.03 (3%) difference in frequency between our patients and Serbia\_pop3 is in [Tables 5 and 6](#).

## Molecules encoded by HLA alleles enriched in NMOSD patients can bind disease-relevant antigens

We performed *in silico* modeling of the binding of proteins encoded by these two haplotypes to two NMOSD-relevant antigens, AQP4 and MOG.

The *in silico* modeling generated binding predictions for both HLA class I and class II molecules in the haplotype we identified to AQP4 ([Supplementary Tables 7 and 8](#)) and MOG ([Supplementary Tables 9 and 10](#)). We first examined the haplotype HLA-A\*01, B\*08, C\*07, DRB1\*03, DQB1\*02. We observed that there were two class I molecules that were able to bind peptides within AQP4. The 9-mer peptide at position 104, MVCTRKISI, was predicted to bind B\*0801 and C\*0701 ([Supplementary Table 7](#)). This 9-mer is part of the 15-mer peptide, MVCTRKISIAKSVFY, that is predicted to bind DRB1\*0301 ([Supplementary Table 8](#)).

Our *in silico* modeling of class II binding to AQP4 (74) generated multiple predicted binding peptides, which are summarized in [Supplementary Table 14](#) and [Supplementary Table 15](#). In AQP4, the 20-mer peptide, VWTQAFWKAVTAEFLAMLIF, is predicted

to bind DQA1\*0501-DQB1\*0201 (Supplementary Table 8) and contains a weaker A\*0101 binder, WTQAFWKAV, and a weaker C\*0701 binder, VTAEFLAML (Supplementary Table 7). The 18-mer peptide, GLYEYVFCPDVEFKRRFK, is predicted to bind DRB1\*0301, DQA1\*0501-DQB1\*0201 (Supplementary Table 8) and contains a weaker C\*0701 binder, YEYVFCPDV (Supplementary Table 7). In MOG, the 15-mer peptide, RRLAGQFLEELRNPF, is predicted to bind DRB1\*0301 and DQA1\*0501-DQB1\*0201 (Supplementary Table 7) and contains a promiscuous weak 9-mer binder, FLEELRNPF, predicted to bind A\*0101 and B\*0801 (Supplementary Table 9).

To generate models of which cellular locations and functional domains of the AQP4 molecule may be targeted by the class I and class II molecules, we used the structure of the protein based on X-ray crystallography (74) to predict the location of the binding sites to human AQP4 in PyMol (Figures 3, 4). The binding sites of HLA class I alleles were observed to occur within, on the inside and on the outside of the cell membrane, primarily on transmembrane domains, within the portion of the AQP4 molecule where a crystal structure was available. For HLA class II, the binding sites occurred inside the membrane or inside the cell, for the portion of the AQP4 molecule where a structure was available (Figure 5).

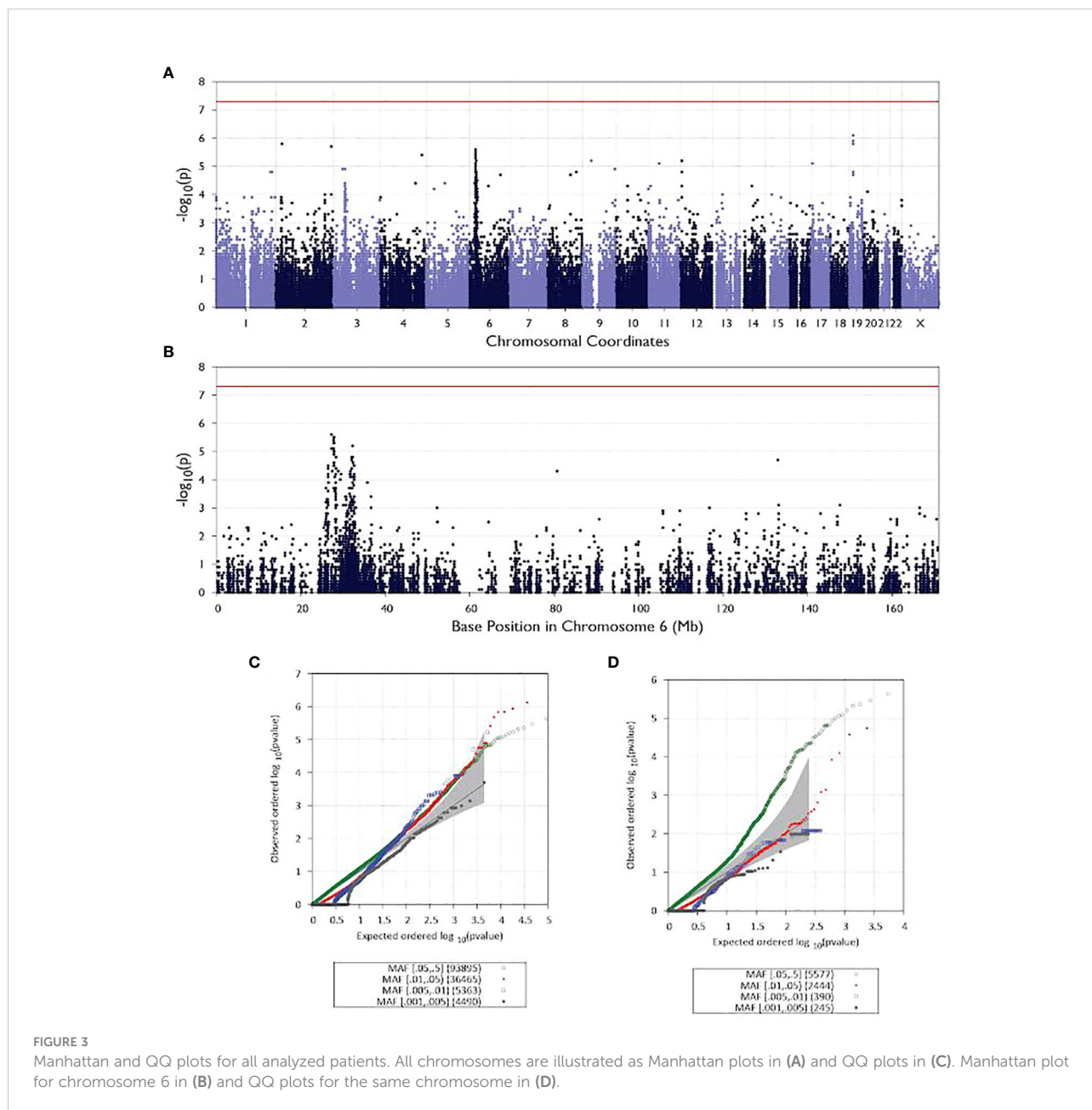


TABLE 5 HLA alleles with more than 0.03 (3%) difference in allele frequencies between the patients and Serbia\_pop3.

Allele	% of patients having the allele	% of Serbians having the allele (Serbia_pop3)	Difference in % of individuals	Allele frequency in patients	Allele frequency in Serbia_pop3	Difference in allele frequency	Pvalue (fisher_exact_test)
DRB1*03	57.69	21	36.69	0.2885	0.1087	0.1798	<b>0.000345819</b>
B*08	42.31	16.47	25.84	0.2115	0.0864	0.1251	<b>0.021468107</b>
DQB1*02	57.69	30.2	27.49	0.2885	0.1698	0.1187	<b>0.038757941</b>
DQB1*05	73.08	54.7	18.38	0.4423	0.3302	0.1121	0.18154181
A*01	46.15	26.25	19.9	0.25	0.1426	0.1074	<b>0.04339775</b>
DRB1*14	23.08	10.7	12.38	0.1154	0.0542	0.0612	0.062586003
DRB1*01	23.08	19.4	3.68	0.1538	0.1029	0.0509	0.486565697
DRB1*08	11.54	6.2	5.34	0.0769	0.0309	0.046	0.220410998
DRB1*16	30.77	20.4	10.37	0.1538	0.1087	0.0451	0.502538467
C*07	46.15	44.02	2.13	0.2885	0.2484	0.0401	0.518981233
C*12	30.77	25.79	4.98	0.1731	0.1352	0.0379	0.414094602
C*03	19.23	11.95	7.28	0.0962	0.0598	0.0364	0.240012755
DQB1*04	11.54	4.4	7.14	0.0577	0.022	0.0357	0.108662136
A*33	11.54	4.92	6.62	0.0577	0.0254	0.0323	0.149070771
B*13	0	6.53	-6.53	0	0.0329	-0.0329	0.415394241
A*24	15.38	21.18	-5.8	0.0769	0.1109	-0.034	0.367454104
A*03	15.38	21.13	-5.75	0.0769	0.1132	-0.0363	0.270804152
B*51	15.38	23.24	-7.86	0.0769	0.1285	-0.0516	0.14625227
DRB1*15	7.69	18.6	-10.91	0.0385	0.0984	-0.0599	0.233720459
C*06	7.69	18.87	-11.18	0.0385	0.1006	-0.0621	0.16514706
DQB1*06	15.38	27	-11.62	0.0769	0.1447	-0.0678	0.107143485
DRB1*07	0	13.9	-13.9	0	0.0713	-0.0713	<b>0.049110681</b>
B*44	3.85	17.22	-13.37	0.0192	0.0916	-0.0724	<b>0.012729771</b>
DRB1*11	19.23	31	-11.77	0.0962	0.1689	-0.0727	0.193298849
DRB1*13	7.69	24.9	-17.21	0.0577	0.1323	-0.0746	0.146018754
DRB1*04	0	17	-17	0	0.0899	-0.0899	<b>0.012695386</b>
DQB1*03	26.92	54.7	-27.78	0.1346	0.3333	-0.1987	<b>0.000522299</b>

The table is sorted by the difference in allele frequency (last column). The full list is in the supplementary file 1 data sheet "Belgrade\_HLA."

Bold values indicate significant alleles.

"\*" means by DNA sequencing.

TABLE 6 Top predicted HLA class II binding peptides in AQP4 and MOG.

Protein	Position	Peptide	Restricted HLA allele
Aquaporin-4	29-48	VWTQAFWKAVTAEFAMLIF	DQA1*0501-DQB1*0201, A*0101, and C*0701
Aquaporin-4	246-263	GLYEYVFCPDVEFKRRFK	DRB1*0301, DQA1*0501-DQB1*020, and C*0701
MOG	38-57	RHPICALVGDVEFLPCRISP	DRB1*0301 and DQA1*0501-DQB1*0201
MOG	233-247	RRLAGQFLELRNPF	DRB1*0301, DQA1*0501-DQB1*0201, A*0101, and B*0801

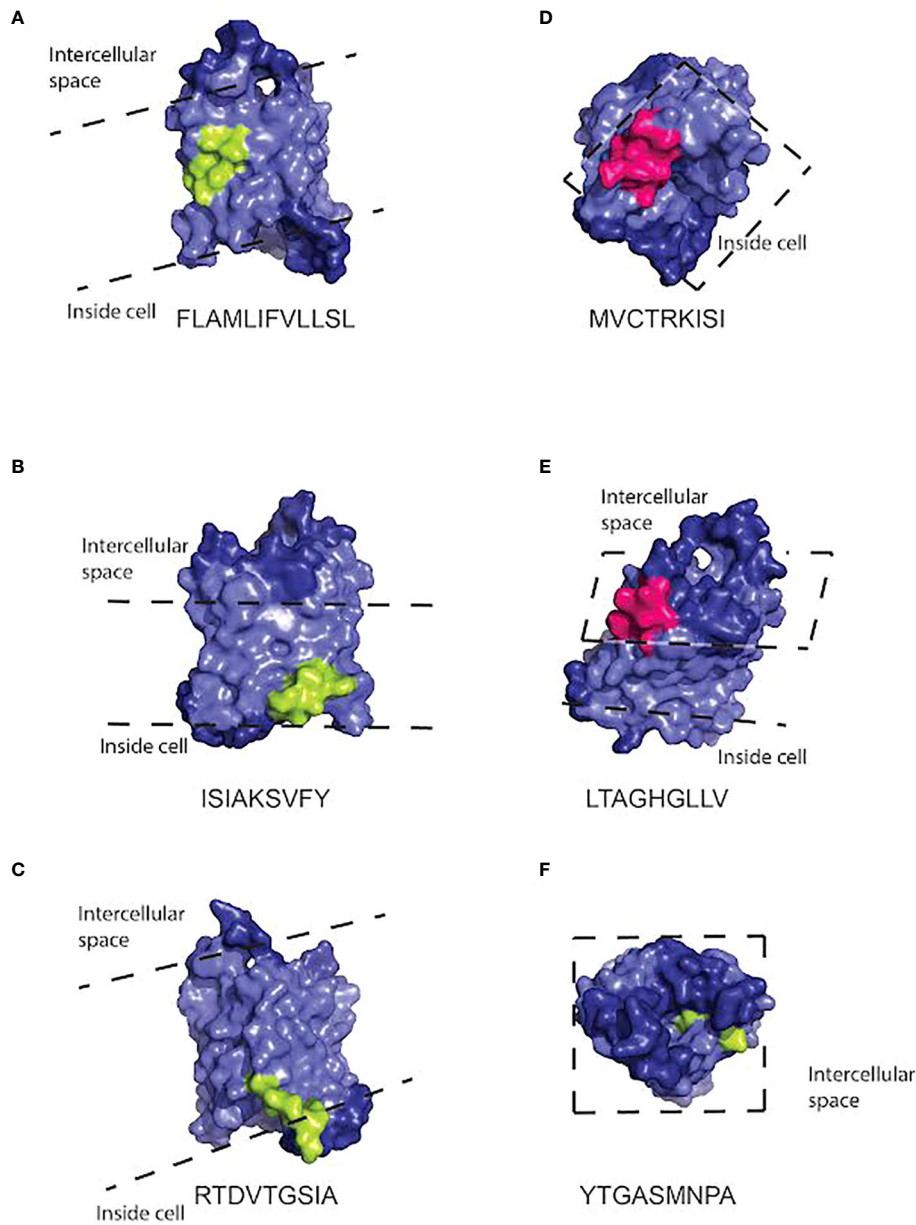
"\*" means by DNA sequencing.

## Discussion

### HLA Typing

Previous studies from multiple populations have shown an association between HLA DRB1\*0301 and NMOSD, but studies in other populations have not (38). Our results are consistent with the recent meta-analysis that indicated that HLA DRB1\*0301 is

associated with NMOSD in populations from Europe and North America, but not populations from Asian countries (38). One potential explanation for different associations of NMOSD and HLA alleles in different populations is that an allele that is inherited in the same haplotype block as DRB1\*03 actually confers NMOSD susceptibility in some populations, but that in other populations, this allele is part of a different haplotype block. Exploring this hypothesis further requires not just looking at



**FIGURE 4**

Three dimensional space-filling cartoon of the AQP4 molecule positioned so that peptides capable of binding HLA class I molecules in the haplotype (either A\*0101, B\*0801, or C\*0701) are visible. The pale blue section of the molecule indicates the intramembrane portion. Dashed lines indicate the approximate position of the extracellular and intracellular surfaces of the cell membrane. Peptides that bind HLA class I molecules strongly are indicated in light green. Promiscuous binders are indicated in pink. Peptide sequences and the HLA molecules binding to them are listed under each model. Peptide sequences: (A) FLAMLIFVLLSL (B) ISIAKSVFY (C) RTDVTGSIA (D) MVCTRKISI (E) LTAGHGLLV (F) YTGASMNPA.

DRB1\*03, but rather analyzing the haplotype block of which it is a part. Our analysis of Serbian NMOSD patients indicated that the haplotypes HLA-A\*01, B\*08, DRB1\*03 and HLA-A\*01, B\*08, C\*07, DRB1\*03, DQB1\*02 are more common in affected individuals than in the Serbian population. Our results indicate that DRB1\*03 is associated with a particular allele configuration at this locus, and therefore, a number of alleles in this haplotype

block are strong candidates for conferring susceptibility to NMOSD in the Caucasian Serbian population. It is possible that only certain alleles in this haplotype increase NMOSD susceptibility. Alternatively, the alleles present within this haplotype block could modify each other, in which case only a particular configuration of alleles would increase NMOSD susceptibility. Future sequencing experiments in multiple

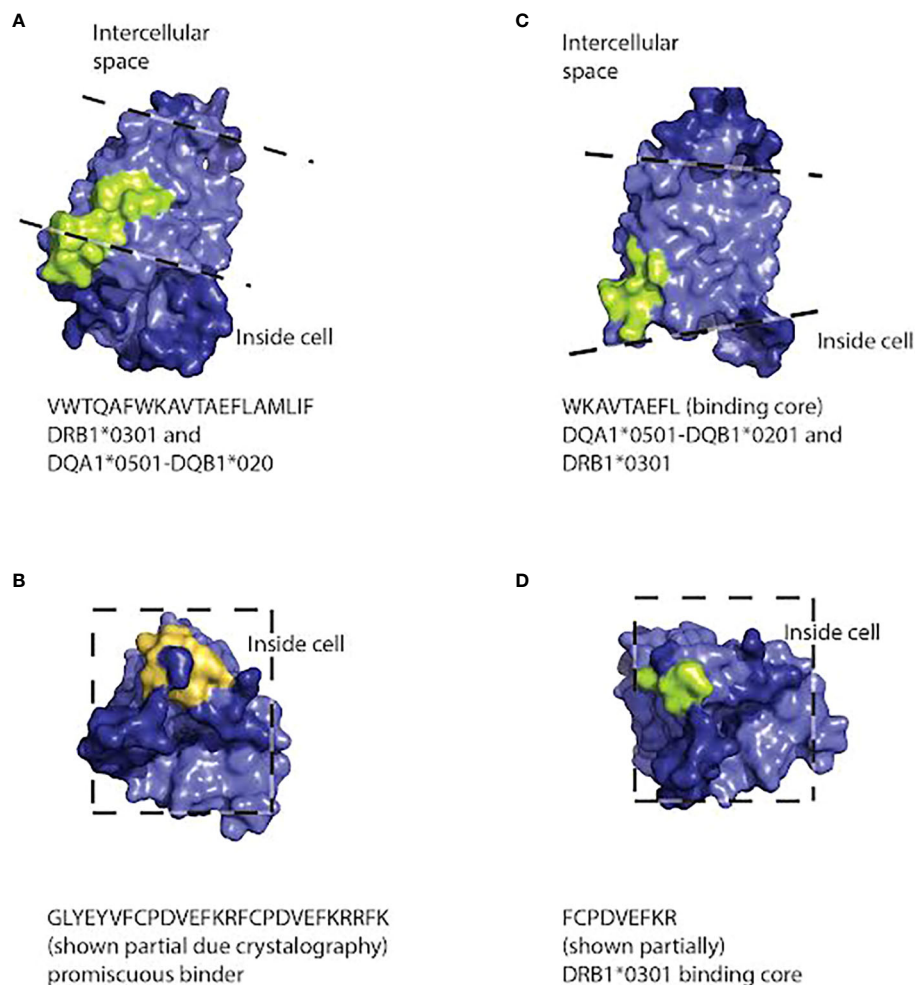


FIGURE 5

Three dimensional space-filling diagram of the AQP4 molecule positioned so that peptides capable of binding HLA class II molecules (DRB1\*0301, DQA1\*0501-DQB1\*0201) are visible. The pale blue section of the molecule indicates the intramembrane portion. Dashed lines indicate the approximate position of the extracellular and intracellular surfaces of the cell membrane. Peptides that bind HLA class II molecules strongly are indicated in light green. Promiscuous binders are indicated in orange. Peptide sequences and the HLA molecules binding to them are listed under each model. Peptide sequences: (A) VWTQAFWKAVTAEFLAMLIF binding to DRB1\*0301 and DQA1\*0501-DQB1\*020. (B) GLYEYVFCPDVEFKRFCPDVEFKRRFK binds promiscuously (C) WKAVTAEFL binding DQA1\*0501-DQB1\*0201 and DRB1\*0301 (D) FCPDVEFKR binds DRB1\*0301.

populations would help distinguish between these possibilities, as would functional and mechanistic studies.

Regarding systems for predicting HLA binding, Gowthaman et al. (2010) evaluated eight prediction systems in predicting peptides binding to seven HLA class I alleles, A\*0101, 0201, 0301, 1101, 2402, B\*0702, and 0801 and found that NetMHC, NetMHCpan, and IEDB showed high overall accuracy (75). Lin et al. evaluated 21 HLA class II prediction methods for their performance in predicting binding peptides of seven common HLA-DR allele, DRB1\*0101, 0301, 0401, 0701, 1101, 1301, and 1501 (76) and found that the HLA class II predictors do not match prediction capabilities of HLA class I predictors. It is more difficult to predict MHC class II binding as, unlike MHC class I molecules, can accommodate the 9-mer binding core of the peptides inside

while peptide termini protrude outside of the grooves (77). NetMHCII is one of the top predictors for DRB1\*0301, and was therefore used in our study. We used different predictors for class I and class II binding, to maximize accuracy.

To test a potential mechanism by which that haplotype could confer susceptibility, we predicted the likely binding sites of the HLA alleles to both AQP4 and MOG because both have been shown to be involved in NMOSD.

The illustrations of our peptide predictions used a published crystal structure of AQP4, 3GD8 (74). This crystal structure has a 1.8 angstrom resolution and it includes all the transmembrane domains, but it doesn't include the N- and C- terminals of the protein, most likely because their structure is less stable. Thus, the peptides on the C-terminus, including the PDZ domain,

were not fully illustrated on the crystal structure. For the incomplete crystal structure available to us, it was notable that most of the peptides predicted to bind HLA alleles are located on the intracellular surface of the molecule or, in one case, inside the pore. Immune cells are unlikely to encounter these peptides unless they are processed and presented on the surface of the cell, but if that does occur, the combination of the peptides and HLA alleles capable of binding them could trigger NMOSD. Since NMOSD are frequently characterized by anti-AQP4 antibodies, it is likely that there is a way that these peptides would be presented on the cell surface. For example, complement-dependent cytotoxicity is thought to be a major mechanism of NMOSD (78). It may also allow intracellular domains of AQP4 to interact with additional immune cells.

Our results are consistent with other epitope mapping studies of AQP4. Based on our current analysis, we have proposed several distinct antigens of AQP4 as predicted based on binding specificities in HLA class I and HLA class II molecules. Future studies to investigate the T cell responses in NMOSD patient PBMCs to these predicted antigens are necessary for discovery and validation of autoimmune NMOSD epitopes. Determining the correct autoimmune antigens may aid in the development of targeted therapies in these patients.

## Whole exome sequencing

Many genes identified by WES are also implicated in other disorders. Mutations in the NOTCH1 gene are associated with aortic valve disease, Adams-Oliver syndrome, T-cell acute lymphoblastic leukemia, chronic lymphocytic leukemia, and head and neck squamous cell carcinoma (79–81). At least some of the mutations thought to be responsible for Adams-Oliver syndrome in patients are heterozygous (82). SPINK5 (serine peptidase inhibitor, Kazal type 5) are associated with Netherton syndrome (OMIM#256500), a disorder involving skin and hair abnormalities and a high risk of allergies, asthma, and an inflammatory skin condition, eczema. Homozygous loss of function in GUSB causes mucopolysaccharidosis type VII (also known as Sly disease) (83, 84), which is characterized by accumulation of glycosaminoglycans within the lysosomes, leading to severe defects in the structure of tissues and organs. Mutations of IL6ST in mice can result in a hyper-responsive T cell phenotype, and development of an arthritis-like autoimmune disease in mice (85). Clinical trials of the anti-CCR4 antibody Mogamulizumab revealed the depletion of Treg cells in a subset of patients, coupled with extreme infiltration of T cells into the skin, resulting in Steven Johnson Syndrome (OMIM#608579) (86, 87). An *in vitro* model of bladder cancer, a partial decrease in the expression of CCR4 has been shown to facilitate metastasis, despite residual protein activity. Certain heterozygous mutations in

FN1 have been proposed to cause or contribute to disease, including glomerulopathy and spondylometaphyseal dysplasia (88, 89).

These genes also have important function in the central-nervous system and immune systems. IL6ST (interleukin 6 signal transducer, also known as GP130) activate another candidate gene NOTCH1 in mouse and human cells (90) and has been implicated in the survival of motor and sensory neuron subtypes in mice (91). CCR4 (C-C Chemokine Receptor 4) is expressed on CD4+ thymocytes, and is potentially involved in negative selection in the thymus (92). CCR4 also expressed on multiple types of T cells, including Th2, Th17, Th22, and Treg (86). FN1 is a component of the blood brain barrier, which is also expressed in astrocytes (93). In rats, it plays a critical role in the axonal regeneration of adult white matter after injury, at least in culture (93). All three genes (SNX4, PDS5A and BIN2) identified by *de novo* mutations are highly specifically expressed in EBV-transformed lymphocytes.

While prior research shows that NMOSD likely has some genetic associations, it is rare to identify multiple cases of the disease within a single family, suggesting that disease susceptibility may be a result of a variety of factors. Recent studies have shown that *de novo* variation in the coding sequences of several genes confer disease susceptibility in probands with neurodevelopmental and psychiatric disorders from otherwise healthy families (50, 51, 94). The results of our trio analysis are consistent with this possibility, as the analysis revealed multiple *de novo* mutations, including predicted pathogenic variants associated with immune function. It is therefore possible that some of these alleles contribute to the development of NMOSD in the probands in this study.

We found that certain alleles from several genes were more prevalent in patients with NMOSD than in controls represented by the gnomAD, indicating that the variants identified in our cohort of patients with NMOSD are presumed to be rare. Certain among these alleles, such as the complement alleles were involved in pathways that have previously been proposed to contribute to the genetic risk of NMOSD (95). While we did not identify the same complement mutation (C4A) as a previous study on NMOSD, we identified mutations in a variety of molecules that are downstream of C4A. The likely explanation for this difference is that our study population of NMOSD patients originated from a different region, and thus, different pathogenic polymorphisms would be expected to be present in this population. The identification of variants in complement genes, including pathogenic mutations, in the NMOSD patient population supports a role for the complement cascade in the pathogenesis of this disease. None of the variants we identified overlapped with the variants detected in the re-analysis of the Estrada et al. data (47, 48), or of a separate association study on a group of Japanese patients (49). However, the latter study identified a variant in the gene encoding a voltage gated potassium channel, KCNMA1, as increasing susceptibility to NMOSD (49). While we did not see a relationship between this particular gene and NMOSD susceptibility in our data, we identified a variety of potentially

deleterious rare variant mutations in NMOSD patients, and potassium channels as a category were detected by Enrichr (Supplementary Tables 1, 2). Given the well-documented role of astrocytes in buffering potassium ions to maintain neuronal membrane potentials, ability to signal and viability, and the involvement of AQP4 in this process, it is reasonable that otherwise benign variations in potassium channels may contribute to the progression of NMOSD. Further analysis of patient populations combined with targeted genetic testing in disease models, including functional studies using patient-specific cells, are necessary to further elucidate the significance of these variants with the ultimate goal to design targeted therapies that reduce progression of NMOSD.

## Coinvestigators

### # Regeneron Genetics Center Banner Author List

All authors/contributors are listed in alphabetical order.

#### RGC Management and Leadership Team

Goncalo Abecasis, Aris Baras, Michael Cantor, Giovanni Coppola, Andrew Deubler, Aris Economides, Luca A. Lotta, John D. Overton, Jeffrey G. Reid, Alan Shuldiner, Katia Karalis and Katherine Siminovitich

#### Sequencing and Lab Operations

Christina Beechert, Caitlin Forsythe, Erin D. Fuller, Zhenhua Gu, Michael Lattari, Alexander Lopez, John D. Overton, Thomas D. Schleicher, Maria Sotiropoulos Padilla, Louis Widom, Sarah E. Wolf, Manasi Pradhan, Kia Manoochehri, Ricardo H. Ulloa.

#### Genome Informatics

Xiaodong Bai, Suganthi Balasubramanian, Boris Boutkov, Gisu Eom, Lukas Habegger, Alicia Hawes, Shareef Khalid, Olga Krasheninina, Rouel Lanche, Adam J. Mansfield, Evan K. Maxwell, Mona Nafde, Sean O'Keeffe, Max Orelus, Razvan Panea, Tommy Polanco, Ayesha Rasool, Jeffrey G. Reid, William Salerno, Jeffrey C. Staples,

#### Research Program Management

Marcus B. Jones and Lyndon J. Mitnaul

## Data availability statement

The datasets presented in this study can be found in online repositories. The name of the repository and accession number can be found below: NCBI database of Genotypes and Phenotypes (dbGaP); accession phs003055.v1.p1.

## Ethics statement

The studies involving human participants were reviewed and approved by Northwell Health Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

IT, AT, and JW designed experiments, interpreted/analyzed data, performed experiments and wrote the paper. GZ performed HLA sequencing analysis and wrote the paper. ID and SM provided expertise on NMOSD, patient samples, and wrote the paper. PJ designed and ran HLA sequencing experiments. DK designed HLA sequencing experiments, interpreted data and wrote the paper. SN provided neurobiology expertise and support for the project, and interpreted data. JNHS designed and supervised experiments, interpreted data, and wrote the paper. VB helped write paper. YS oversaw bioinformatics analysis on exome sequencing. AK, JD, EY, VJ, TdA, MF, AH, MM, MS, VN, WdG, ML, KB, MG provided clinical expertise, recruited patients and processed patient samples.

Regeneron Genetics Center Banner Author List and Contribution Statements

All authors/contributors are listed in alphabetical order.

#### RGC Management and Leadership Team

Goncalo Abecasis, Aris Baras, Michael Cantor, Giovanni Coppola, Andrew Deubler, Aris Economides, Luca A. Lotta, John D. Overton, Jeffrey G. Reid, Alan Shuldiner, Katia Karalis and Katherine Siminovitich

Contribution: All authors contributed to securing funding, study design and oversight. All authors reviewed the final version of the manuscript.

#### Sequencing and Lab Operations

Christina Beechert, Caitlin Forsythe, Erin D. Fuller, Zhenhua Gu, Michael Lattari, Alexander Lopez, John D. Overton, Thomas D. Schleicher, Maria Sotiropoulos Padilla, Louis Widom, Sarah E. Wolf, Manasi Pradhan, Kia Manoochehri, Ricardo H. Ulloa.

Contribution: CB, CF, AL, and JO performed and are responsible for sample genotyping. CB, CF, EF, ML, MP, LW, SW, AL, and JO performed and are responsible for exome sequencing. TS, ZG, AL, and JO conceived and are responsible for laboratory automation. MP, KM, RU, and JO are responsible for sample tracking and the library information management system.

#### Genome Informatics

Xiaodong Bai, Suganthi Balasubramanian, Boris Boutkov, Gisu Eom, Lukas Habegger, Alicia Hawes, Shareef Khalid, Olga Krasheninina, Rouel Lanche, Adam J. Mansfield, Evan K. Maxwell, Mona Nafde, Sean O'Keeffe, Max Orelus, Razvan Panea, Tommy Polanco, Ayesha Rasool, Jeffrey G. Reid, William Salerno, Jeffrey C. Staples,

Contribution: XB, AH, OK, AM, SO, RP, TP, AR, WS, and JR performed and are responsible for the compute logistics, analysis and infrastructure needed to produce exome and genotype data. GE, MO, MN, and JR provided compute infrastructure development and operational support. SB, SK, and JR provide variant and gene annotations and their functional interpretation of variants. EM, JS, RL, BB, AB, LH, and JR conceived and are responsible for creating,

developing, and deploying analysis platforms and computational methods for analyzing genomic data.

#### Research Program Management

Marcus B. Jones and Lyndon J. Mitnaul

Contribution: All authors contributed to the management and coordination of all research activities, planning and execution. All authors contributed to the review process for the final version of the manuscript.

## Funding

IT was supported by a NICHD F32 081835 and a Kavli Institute Pilot Grant. IT, VN, MF, WD, and JS were supported by a grant from the HF Langbert Neuroimmunology Research Fund. MF was funded in part by the Cheryl Manne Research Fellowship.

## Acknowledgments

We are very grateful to Dr Wendy Chung for reading our manuscript and providing essential feedback for our study.

## Conflict of interest

The members of the Regeneron Genetics Center and AK were employed by Regeneron.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Lucchinetti CF, Guo Y, Popescu BF, Fujihara K, Itoyama Y, Misu T. The pathology of an autoimmune astrocytopathy: lessons learned from neuromyelitis optica. *Brain Pathol* (2014) 24(1):83–97. doi: 10.1111/bpa.12099
- Wingerchuk DM, Banwell B, Bennett JL, Cabre P, Carroll W, Chitnis T, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology* (2015) 85(2):177–89. doi: 10.1212/WNL.0000000000001729
- Mealy MA, Wingerchuk DM, Greenberg BM, Levy M. Epidemiology of neuromyelitis optica in the united states: a multicenter analysis. *Arch Neurol* (2012) 69(9):1176–80. doi: 10.1001/archneurol.2012.314
- Cabrera-Gomez JA, Kurtzke JF, Gonzalez-Quevedo A, Lara-Rodriguez R. An epidemiological study of neuromyelitis optica in Cuba. *J Neurol* (2009) 256(1):35–44. doi: 10.1007/s00415-009-0009-0

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.900605/full#supplementary-material>

#### SUPPLEMENTARY TABLE 1

346 variants in 338 genes that identified by PSAP model.

#### SUPPLEMENTARY TABLE 2

Reactome analysis of missense variants identified in the dataset.

#### SUPPLEMENTARY TABLE 3

List of selected rare variants identified by exome sequencing in genes involved in immune function. Full list of detected variants is available in supplementary information.

#### SUPPLEMENTARY TABLE 4

Clinical features and identifiers of probands and family members in trio analysis.

#### SUPPLEMENTARY TABLE 5

Comparison of estimated haplotype frequencies between the Belgrade population (using previously available data) and the patient population and family members included in the present study.

#### SUPPLEMENTARY TABLE 6

Detailed clinical information from Belgrade patients.

#### SUPPLEMENTARY TABLE 7

Summary of the predicted HLA class I binding peptides in AQP4. Predicted weak binders, strong binders, and promiscuous binders (peptides predicted to bind multiple HLA alleles) are highlighted in green, yellow and red, respectively.

#### SUPPLEMENTARY TABLE 8

Predicted HLA class II binding peptides in AQP4. Predicted weak binders, strong binders, and promiscuous binders (peptides predicted to bind multiple HLA alleles) are highlighted in green, yellow and red, respectively.

#### SUPPLEMENTARY TABLE 9

Predicted HLA class I binding peptides in MOG. Predicted weak binders, strong binders, and promiscuous binders (peptides predicted to bind multiple HLA alleles) are highlighted in green, yellow and red, respectively.

#### SUPPLEMENTARY TABLE 10

Predicted HLA class II binding peptides in MOG. Predicted weak binders, strong binders, and promiscuous binders (peptides predicted to bind multiple HLA alleles) are highlighted in green, yellow and red, respectively.

- Uzawa A, Mori M, Kuwabara S. Neuromyelitis optica: concept, immunology and treatment. *J Clin Neurosci* (2014) 21(1):12–21. doi: 10.1016/j.jocn.2012.12.022

- Wildemann B, Jarius S. The expanding range of autoimmune disorders of the nervous system. *Lancet Neurol* (2013) 12(1):22–4. doi: 10.1016/S1474-4422(12)70301-0

- Wingerchuk DM, Weinshenker BG. Neuromyelitis optica spectrum disorder diagnostic criteria: Sensitivity and specificity are both important. *Mult Scler* (2017) 23(2):182–4. doi: 10.1177/1352458516688352

- Crout TM, Parks LP, Majithia V. Neuromyelitis optica (Devic's syndrome): an appraisal. *Curr Rheumatol Rep* (2016) 18(8):54. doi: 10.1007/s11926-016-0599-3



9. Lennon VA, Wingerchuk DM, Kryzer TJ, Pittock SJ, Lucchinetti CF, Fujihara K, et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *Lancet* (2004) 364(9451):2106–12. doi: 10.1016/S0140-6736(04)17551-X
10. Mader S, Lutterotti A, Di Pauli F, Kuenz B, Schanda K, Aboul-Enein F, et al. Patterns of antibody binding to aquaporin-4 isoforms in neuromyelitis optica. *PLoS One* (2010) 5(5):e10455. doi: 10.1371/journal.pone.0010455
11. Jarius S, Wildemann B. AQP4 antibodies in neuromyelitis optica: diagnostic and pathogenetic relevance. *Nat Rev Neurol* (2010) 6(7):383–92. doi: 10.1038/nrneuro.2010.72
12. Waters PJ, McKeon A, Leite MI, Rajasekharan S, Lennon VA, Villalobos A, et al. Serologic diagnosis of NMO: a multicenter comparison of aquaporin-4-IgG assays. *Neurology* (2012) 78(9):663–71. doi: 10.1212/WNL.0b013e318248dec1
13. Weinschenker BG, Wingerchuk DM. Neuromyelitis spectrum disorders. *Mayo Clinic Proc* (2017) 92(4):663–79. doi: 10.1016/j.mayocp.2016.12.014
14. Lalan S, Khan M, Schlakman B, Penman A, Gatlin J, Herndon R. Differentiation of neuromyelitis optica from multiple sclerosis on spinal magnetic resonance imaging. *Int J MS Care* (2012) 14(4):209–14. doi: 10.7224/1537-2073-14.4.209
15. Bradl M, Reindl M, Lassmann H. Mechanisms for lesion localization in neuromyelitis optica spectrum disorders. *Curr Opin Neurol* (2018) 31(3):325–33. doi: 10.1097/WCO.0000000000000551
16. Sabater L, Giral A, Boronat A, Hankiewicz K, Blanco Y, Llufrui S, et al. Cytotoxic effect of neuromyelitis optica antibody (NMO-IgG) to astrocytes: an *in vitro* study. *J Neuroimmunol* (2009) 215(1–2):31–5. doi: 10.1016/j.jneuroim.2009.07.014
17. Roemer SF, Parisi JE, Lennon VA, Benarroch EE, Lassmann H, Bruck W, et al. Pattern-specific loss of aquaporin-4 immunoreactivity distinguishes neuromyelitis optica from multiple sclerosis. *Brain* (2007) 130(Pt 5):1194–205. doi: 10.1093/brain/awl371
18. Hinson SR, McKeon A, Fryer JP, Apiwatanakul M, Lennon VA, Pittock SJ. Prediction of neuromyelitis optica attack severity by quantitation of complement-mediated injury to aquaporin-4-expressing cells. *Arch Neurol* (2009) 66(9):1164–7. doi: 10.1001/archneurol.2009.188
19. Parratt JD, Prineas JW. Neuromyelitis optica: a demyelinating disease characterized by acute destruction and regeneration of perivascular astrocytes. *Mult Scler* (2010) 16(10):1156–72. doi: 10.1177/1352458510382324
20. Marignier R, Nicolle A, Watrin C, Touret M, Cavagna S, Varrin-Doyer M, et al. Oligodendrocytes are damaged by neuromyelitis optica immunoglobulin G via astrocyte injury. *Brain* (2010) 133(9):2578–91. doi: 10.1093/brain/awq177
21. Hillebrand S, Schanda K, Nigritinou M, Tsymala I, Bohm D, Peschl P, et al. Circulating AQP4-specific auto-antibodies alone can induce neuromyelitis optica spectrum disorder in the rat. *Acta Neuropathol* (2019) 137(3):467–85. doi: 10.1007/s00401-018-1950-8
22. Papadopoulos MC, Verkman AS. Aquaporin 4 and neuromyelitis optica. *Lancet Neurol* (2012) 11(6):535–44. doi: 10.1016/S1474-4422(12)70133-3
23. Zeka B, Hastermann M, Hochmeister S, Kogl N, Kaufmann N, Schanda K, et al. Highly encephalitogenic aquaporin 4-specific T cells and NMO-IgG jointly orchestrate lesion location and tissue damage in the CNS. *Acta Neuropathol* (2015) 130(6):783–98. doi: 10.1007/s00401-015-1501-5
24. Zeka B, Hastermann M, Kaufmann N, Schanda K, Pende M, Misi T, et al. Aquaporin 4-specific T cells and NMO-IgG cause primary retinal damage in experimental NMO/SD. *Acta Neuropathol Commun* (2016) 4(1):82. doi: 10.1186/s40478-016-0355-y
25. Pohl M, Kawakami N, Kitic M, Bauer J, Martins R, Fischer MT, et al. T Cell activation in neuromyelitis optica lesions plays a role in their formation. *Acta Neuropathol Commun* (2013) 1:85. doi: 10.1186/2051-5960-1-85
26. Papadopoulos MC, Bennett JL, Verkman AS. Treatment of neuromyelitis optica: state-of-the-art and emerging therapies. *Nat Rev Neurol* (2014) 10(9):493–506. doi: 10.1038/nrneuro.2014.141
27. Jarius S, Ruprecht K, Kleiter I, Borisow N, Asgari N, Pitarokoil K, et al. MOG-IgG in NMO and related disorders: a multicenter study of 50 patients. part 1: Frequency, syndrome specificity, influence of disease activity, long-term course, association with AQP4-IgG, and origin. *J Neuroinflammation* (2016) 13(1):279. doi: 10.1186/s12974-016-0717-1
28. Kitley J, Woodhall M, Waters P, Leite MI, Devenney E, Craig J, et al. Myelin-oligodendrocyte glycoprotein antibodies in adults with a neuromyelitis optica phenotype. *Neurology* (2012) 79(12):1273–7. doi: 10.1212/WNL.0b013e31826aac4e
29. Mader S, Kumpfel T, Meinel E. Novel insights into pathophysiology and therapeutic possibilities reveal further differences between AQP4-IgG- and MOG-IgG-associated diseases. *Curr Opin Neurol* (2020) 33(3):362–71. doi: 10.1097/WCO.0000000000000813
30. Probstel AK, Rudolf G, Dormair K, Collongues N, Chanson JB, Sanderson NS, et al. Anti-MOG antibodies are present in a subgroup of patients with a neuromyelitis optica phenotype. *J Neuroinflammation* (2015) 12:46. doi: 10.1186/s12974-015-0256-1
31. Reindl M, Waters P. Myelin oligodendrocyte glycoprotein antibodies in neurological disease. *Nat Rev Neurol* (2019) 15(2):89–102. doi: 10.1038/s41582-018-0112-x
32. Mader S, Gredler V, Schanda K, Rostasy K, Dujmovic I, Pfaller K, et al. Complement activating antibodies to myelin oligodendrocyte glycoprotein in neuromyelitis optica and related disorders. *J Neuroinflammation* (2011) 8:184. doi: 10.1186/1742-2094-8-184
33. Kim SH, Kim W, Li XF, Jung IJ, Kim HJ. Clinical spectrum of CNS aquaporin-4 autoimmunity. *Neurology* (2012) 78(15):1179–85. doi: 10.1212/WNL.0b013e31824f8069
34. Jarius S, Paul F, Franciotta D, Aktas O, Hohlfeld R, Zipp F, et al. Revised diagnostic criteria for neuromyelitis optica–incorporation of NMO-IgG status. *Nat Clin Pract Neurol* (2007) 3(5):E1. doi: 10.1038/ncpneuro0501
35. Wingerchuk DM. Neuromyelitis optica: effect of gender. *J Neurol Sci* (2009) 286(1–2):18–23. doi: 10.1016/j.jns.2009.08.045
36. Sellner J, Boggild M, Clanet M, Hintzen RQ, Illes Z, Montalban X, et al. EFNS guidelines on diagnosis and management of neuromyelitis optica. *Eur J Neurol* (2010) 17(8):1019–32. doi: 10.1111/j.1468-1331.2010.03066.x
37. Ghezzi A, Bergamaschi R, Martinelli V, Trojano M, Tola MR, Merelli E, et al. Clinical characteristics, course and prognosis of relapsing devic's neuromyelitis optica. *J Neurol* (2004) 251(1):47–52. doi: 10.1007/s00415-004-0271-0
38. Alvarenga MP, do Carmo LF, Vasconcelos CCF, Alvarenga MP, Alvarenga-Filho H, de Melo Bento CA, et al. Neuromyelitis optica is an HLA associated disease different from multiple sclerosis: a systematic review with meta-analysis. *Sci Rep* (2021) 11(1):152. doi: 10.1038/s41598-020-80535-3
39. Zephir H, Fajardy I, Outteryck O, Blanc F, Roger N, Fleury M, et al. Is neuromyelitis optica associated with human leukocyte antigen? *Mult Scler* (2009) 15(5):571–9. doi: 10.1177/1352458508102085
40. Brum DG, Barreira AA, dos Santos AC, Kaimen-Maciel DR, Matiello M, Costa RM, et al. HLA-DRB association in neuromyelitis optica is different from that observed in multiple sclerosis. *Mult Scler* (2010) 16(1):21–9. doi: 10.1177/1352458509350741
41. Alonso VR, de Jesus Flores Rivera J, Garci YR, Granados J, Sanchez T, Mena-Hernandez L, et al. Neuromyelitis optica (NMO IgG+) and genetic susceptibility, potential ethnic influences. *Cent Nerv Syst Agents Med Chem* (2018) 18(1):4–7. doi: 10.2174/1871524916666160229115047
42. Deschamps R, Paturel L, Jeannin S, Chausson N, Olindo S, Bera O, et al. Different HLA class II (DRB1 and DQB1) alleles determine either susceptibility or resistance to NMO and multiple sclerosis among the French afro-Caribbean population. *Mult Scler* (2011) 17(1):24–31. doi: 10.1177/1352458510382810
43. Asgari N, Nielsen C, Stenager E, Kyvik KO, Lillevang ST. HLA, PTPN22 and PD-1 associations as markers of autoimmunity in neuromyelitis optica. *Mult Scler* (2012) 18(1):23–30. doi: 10.1177/1352458511417480
44. Blanco Y, Ercilla-Gonzalez G, Llufrui S, Casanova-Estruch B, Magraner MJ, Ramio-Torrenta L, et al. HLA-DRB1 typing in caucasians patients with neuromyelitis optica. *Rev Neurol* (2011) 53(3):146–52.
45. Brill L, Mandel M, Karussis D, Petrou P, Miller K, Ben-Hur T, et al. Increased occurrence of anti-AQP4 seropositivity and unique HLA class II associations with neuromyelitis optica (NMO), among Muslim arabs in Israel. *J Neuroimmunol* (2016) 293:65–70. doi: 10.1016/j.jneuroim.2016.02.006
46. Yoshimura S, Isobe N, Matsushita T, Yonekawa T, Masaki K, Sato S, et al. Distinct genetic and infectious profiles in Japanese neuromyelitis optica patients according to anti-aquaporin 4 antibody status. *J Neurol Neurosurg Psychiatry* (2013) 84(1):29–34. doi: 10.1136/jnnp-2012-302925
47. Estrada K, Whelan CW, Zhao F, Bronson P, Handsaker RE, Sun C, et al. A whole-genome sequence study identifies genetic risk factors for neuromyelitis optica. *Nat Commun* (2018) 9(1):1929. doi: 10.1038/s41467-018-04332-3
48. Li T, Li H, Li Y, Dong SA, Yi M, Zhang QX, et al. Multi-level analyses of genome-wide association study to reveal significant risk genes and pathways in neuromyelitis optica spectrum disorder. *Front Genet* (2021) 12:690537. doi: 10.3389/fgenet.2021.690537
49. Matsushita T, Masaki K, Isobe N, Sato S, Yamamoto K, Nakamura Y, et al. Genetic factors for susceptibility to and manifestations of neuromyelitis optica. *Ann Clin Transl Neurol* (2020) 7(11):2082–93. doi: 10.1002/acn3.51147
50. Howrigan DP, Rose SA, Samocha KE, Fromer M, Cerrato F, Chen WJ, et al. Exome sequencing in schizophrenia-affected parent-offspring trios reveals risk conferred by protein-coding *de novo* mutations. *Nat Neurosci* (2020) 23(2):185–93. doi: 10.1038/s41593-019-0564-3
51. Nguyen HT, Bryois J, Kim A, Dobbyn A, Huckins LM, Munoz-Manchado AB, et al. Integrated Bayesian analysis of rare exonic variants to identify risk genes for schizophrenia and neurodevelopmental disorders. *Genome Med* (2017) 9(1):114. doi: 10.1186/s13073-017-0497-y
52. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res* (2015) 43(Database issue):D784–788. doi: 10.1093/nar/gku1166
53. Andric Z, Popadic D, Jovanovic B, Jaglicic I, Bojic S, Simonovic R. HLA-a, -b, -c, -DRB1 and -DQB1 allele and haplotype frequencies in the Serbian population. *Hum Immunol* (2014) 75(3):218–26. doi: 10.1016/j.humimm.2013.12.009

54. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* (2016) 32(4):511–7. doi: 10.1093/bioinformatics/btv639
55. Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* (2003) 12(5):1007–17. doi: 10.1110/ps.0239403
56. Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* (2018) 154(3):394–406. doi: 10.1111/imm.12889
57. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* (2010) 38(16):e164. doi: 10.1093/nar/gkq603
58. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* (2011) 43(5):491–8. doi: 10.1038/ng.806
59. Wilfert AB, Chao KR, Kaushal M, Jain S, Zollner S, Adams DR, et al. Genome-wide significance testing of variation from single case exomes. *Nat Genet* (2016) 48(12):1455–61. doi: 10.1038/ng.3697
60. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* (2016) 536(7616):285–91. doi: 10.1038/nature19057
61. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
62. Rentsch P, Schubach M, Shendure J, Kircher M. CADD-splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med* (2021) 13(1):31. doi: 10.1186/s13073-021-00835-9
63. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* (2016) 99(4):877–85. doi: 10.1016/j.ajhg.2016.08.016
64. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinf* (2013) 14:128. doi: 10.1186/1471-2105-14-128
65. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* (2011) 88(1):76–82. doi: 10.1016/j.ajhg.2010.11.011
66. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* (2006) 38(8):904–9. doi: 10.1038/ng1847
67. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* (2012) 7(10):e46688. doi: 10.1371/journal.pone.0046688
68. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. *Nucleic Acids Res* (2020) 48(D1):D498–503. doi: 10.1093/nar/gkz1031
69. Kelley J, de Bono B, Trowsdale J. IRIS: a database surveying known human immune system genes. *Genomics* (2005) 85(4):503–11. doi: 10.1016/j.ygeno.2005.01.009
70. Kaufmann T, Hansch G, Rittner C, Spath P, Tedesco F, Schneider PM. Genetic basis of human complement C8 beta deficiency. *J Immunol* (1993) 150(11):4943–7.
71. Arnold DF, Roberts AG, Thomas A, Ferry B, Morgan BP, Chapel H. A novel mutation in a patient with a deficiency of the eighth component of complement associated with recurrent meningococcal meningitis. *J Clin Immunol* (2009) 29(5):691–5. doi: 10.1007/s10875-009-9295-7
72. Sanchez-Barrera MJ, Vallis Y, Clatworthy MR, Doherty GJ, Veprintsev DB, Evans PR, et al. Bin2 is a membrane sculpting n-BAR protein that influences leucocyte podosomes, motility and phagocytosis. *PLoS One* (2012) 7(12):e52401. doi: 10.1371/journal.pone.0052401
73. Kim NY, Cho MH, Won SH, Kang HJ, Yoon SY, Kim DH. Sorting nexin-4 regulates beta-amyloid production by modulating beta-site-activating cleavage enzyme-1. *Alzheimers Res Ther* (2017) 9(1):4. doi: 10.1371/journal.pone.0052401
74. Ho JD, Yeh R, Sandstrom A, Chorny I, Harries WE, Robbins RA, et al. Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance. *Proc Natl Acad Sci USA* (2009) 106(18):7437–42. doi: 10.1073/pnas.0902725106
75. Gowthaman U, Chodiseti SB, Parihar P, Agrewala JN. Evaluation of different generic *in silico* methods for predicting HLA class I binding peptide vaccine candidates using a reverse approach. *Amino Acids* (2010) 39(5):1333–42. doi: 10.1007/s00726-010-0579-2
76. Lin HH, Zhang GL, Tongchusak S, Reinherz EL, Brusci V. Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. *BMC Bioinf* (2008) 9 Suppl 12:S22. doi: 10.1186/1471-2105-9-S12-S22
77. Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* (1994) 368(6468):215–21. doi: 10.1038/368215a0
78. Phuan PW, Ratelade J, Rossi A, Tradtrantip L, Verkman AS. Complement-dependent cytotoxicity in neuromyelitis optica requires aquaporin-4 protein assembly in orthogonal arrays. *J Biol Chem* (2012) 287(17):13829–39. doi: 10.1074/jbc.M112.344325
79. Agrawal N, Frederick MJ, Pickering CR, Bettgeowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* (2011) 333(6046):1154–7. doi: 10.1126/science.1206923
80. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* (2011) 333(6046):1157–60. doi: 10.1126/science.1208130
81. Das I, Craig C, Funahashi Y, Jung KM, Kim TW, Byers R, et al. Notch oncoproteins depend on gamma-secretase/presenilin activity for processing and function. *J Biol Chem* (2004) 279(29):30771–80. doi: 10.1074/jbc.M309252200
82. Southgate L, Sukalo M, Karountzos ASV, Taylor EJ, Collinson CS, Ruddy D, et al. Haploinsufficiency of the NOTCH1 receptor as a cause of adams-Oliver syndrome with variable cardiac anomalies. *Circ Cardiovasc Genet* (2015) 8(4):572–81. doi: 10.1161/CIRCGENETICS.115.001086
83. Shipley JM, Klinkenberg M, Wu BM, Bachinsky DR, Grubb JH, Sly WS. Mutational analysis of a patient with mucopolysaccharidosis type VII, and identification of pseudogenes. *Am J Hum Genet* (1993) 52(3):517–26.
84. Sly WS, Quinton BA, McAlister WH, Rimoin DL. Beta glucuronidase deficiency: report of clinical, radiologic, and biochemical features of a new mucopolysaccharidosis. *J Pediatr* (1973) 82(2):249–57. doi: 10.1016/S0022-3476(73)80162-3
85. Atsumi T, Ishihara K, Kamimura D, Ikushima H, Ohtani T, Hirota S, et al. A point mutation of tyr-759 in interleukin 6 family cytokine receptor subunit gp130 causes autoimmune arthritis. *J Exp Med* (2002) 196(7):979–90. doi: 10.1084/jem.20020619
86. Yoshie O, Matsushima K, CCR4 and its ligands: from bench to bedside. *Int Immunol* (2015) 27(1):11–20. doi: 10.1093/intimm/ixu079
87. Ishida T, Ito A, Sato F, Kusumoto S, Iida S, Inagaki H, et al. Stevens-Johnson syndrome associated with mogamulizumab treatment of adult T-cell leukemia / lymphoma. *Cancer Sci* (2013) 104(5):647–50. doi: 10.1111/cas.12116
88. Castelletti F, Donadelli R, Banterla F, Hildebrandt F, Zipfel PF, Bresin E, et al. Mutations in FN1 cause glomerulopathy with fibronectin deposits. *Proc Natl Acad Sci USA* (2008) 105(7):2538–43. doi: 10.1073/pnas.0707730105
89. Lee CS, Fu H, Baratang N, Rousseau J, Kumra H, Sutton VR, et al. Mutations in fibronectin cause a subtype of spondylometaphyseal dysplasia with "Corner fractures". *Am J Hum Genet* (2017) 101(5):815–23. doi: 10.1016/j.ajhg.2017.09.019
90. Taniguchi K, Wu LW, Grivennikov SI, de Jong PR, Lian I, Yu FX, et al. A gp130-Src-YAP module links inflammation to epithelial regeneration. *Nature* (2015) 519(7541):57–62. doi: 10.1038/nature14228
91. Nakashima K, Wiese S, Yanagisawa M, Arakawa H, Kimura N, Hisatsune T, et al. Developmental requirement of gp130 signaling in neuronal survival and astrocyte differentiation. *J Neurosci* (1999) 19(13):5429–34. doi: 10.1523/JNEUROSCI.19-13-05429.1999
92. Annunziato F, Romagnani P, Cosmi L, Beltrame C, Steiner BH, Lazzari E, et al. Macrophage-derived chemokine and EB1-ligand chemokine attract human thymocytes in different stage of development and are produced by distinct subsets of medullary epithelial cells: possible implications for negative selection. *J Immunol* (2000) 165(1):238–46. doi: 10.4049/jimmunol.165.1.238
93. Andrews RN, Caudell DL, Metheny-Barlow LJ, Peiffer AM, Tootz JA, Bourland JD, et al. Fibronectin produced by cerebral endothelial and vascular smooth muscle cells contributes to perivascular extracellular matrix in late-delayed radiation-induced brain injury. *Radiat Res* (2018) 190(4):361–73. doi: 10.1667/RR14961.1
94. Jin ZB, Li Z, Liu Z, Jiang Y, Cai XB, Wu J. Identification of *de novo* germline mutations and causal genes for sporadic diseases using trio-based whole-exome/genome sequencing. *Biol Rev Camb Philos Soc* (2018) 93(2):1014–31. doi: 10.1111/brv.12383
95. Chen Y, Li R, Wu AM, Shu YQ, Lu ZQ, Hu XQ. The complement and immunoglobulin levels in NMO patients. *Neuro Sci* (2014) 35(2):215–20. doi: 10.1007/s10072-013-1481-y

## COPYRIGHT

© 2022 Tabansky, Tanaka, Wang, Zhang, Dujmovic, Mader, Jegathanan, DeAngelis, Funaro, Harel, Messina, Shabbir, Nursey, DeGouvia, Laurent, Blitz, Jindra, Gudesblatt, Regeneron Genetics Center, King, Drulovic, Yunis, Brusci, Shen, Keskin, Najjar and Stern. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.