Gene expression and alternative splicing analysis in a large-scale Multiple Sclerosis study

Müge Sak[1,2], Julia H. Chariker[1,2], Juw Won Park[1,3,4,5], Eric C. Rouchka[1,6,*]

[1]Kentucky IDeA Network of Biomedical Research Excellence Bioinformatics Core, University of Louisville, Louisville, Kentucky 40202, United States of America.
[2]Department of Neuroscience Training, University of Louisville, Louisville, Kentucky 40202, United States of America.
[3]Department of Medicine, University of Louisville, Louisville, Kentucky 40202, United States of America.
[4]Brown Cancer Center Bioinformatics Core, University of Louisville, Louisville, Kentucky 40202, United States of America.
[5]Center for Integrative Environmental Health Sciences Biostatistics and Informatics Facility Core, University of Louisville, Louisville, Kentucky 40202, United States of America.
[6]Department of Biochemistry and Molecular Genetics, University of Louisville, Louisville, Kentucky 40202, United States of America.

*eric.rouchka@louisville.edu

ABSTRACT

Background: Multiple Sclerosis (MS) is an autoimmune neurodegenerative disease affecting approximately 3 million people globally. Despite rigorous research on MS, aspects of its development and progression remain unclear. Understanding molecular mechanisms underlying MS is crucial to providing insights into disease pathways, identifying potential biomarkers for early diagnosis, and revealing novel therapeutic targets for improved patient outcomes.

Methods: We utilized publicly available RNA-seq data (GSE138614) from post-mortem white matter tissues of five donors without any neurological disorder and ten MS patient donors. This data was interrogated for differential gene expression, alternative splicing and single nucleotide variants as well as for functional enrichments in the resulting datasets.

Results: A comparison of non-MS white matter (WM) to MS samples yielded differentially expressed genes involved in adaptive immune response, cell communication, and developmental processes. Genes with expression changes positively correlated with tissue inflammation were enriched in the immune system and receptor interaction pathways. Negatively correlated genes were enriched in neurogenesis, nervous system development, and metabolic pathways. Alternatively spliced transcripts between WM and MS lesions included genes that play roles in neurogenesis, myelination, and oligodendrocyte differentiation, such as brain enriched myelin associated protein (*BCAS1*), discs large MAGUK scaffold protein 1 (*DLG1*), KH domain containing RNA binding (*QKI*), and myelin basic protein (*MBP*). Our approach to comparing normal appearing WM (NAWM) and active lesion (AL) from one donor and NAWM and chronic active (CA) tissues from two donors, showed that different IgH and IgK gene subfamilies were differentially expressed. We also identified pathways involved in white matter injury repair and remyelination in these tissues. Differentially spliced genes between these lesions were involved in axon and dendrite structure stability. We also identified exon skipping events and spontaneous single nucleotide polymorphisms in membrane associated ring-CH-type finger 1 (*MARCHF1*), UDP glycosyltransferase 8 (*UGT8*), and other genes important in autoimmunity and neurodegeneration.

Conclusion: Overall, we identified unique genes, pathways, and novel splicing events affecting disease progression that can be further investigated as potential novel drug targets for MS treatment.

Keywords: multiple sclerosis, RNA-Seq, differential expression, alternative splicing

INTRODUCTION

Multiple Sclerosis (MS) is an autoimmune neurodegenerative disease of the central nervous system (CNS) that affects about 3 million people worldwide [1]. Magnetic resonance imaging (MRI) of the brain and spinal cord is used to identify characteristic lesions, while evidence of intrathecally-produced immunoglobulin G (IgG) via lumbar puncture is used as a diagnostic test for MS [2]. Vitamin D deficiency and Epstein-Barr virus (EBV) infections have been consistently linked to MS diagnosis [3, 4]. Although MS is not an inherited disease, genetic susceptibility to MS is known [5, 6]. This includes 233 risk loci identified in genome-wide association studies [7]. Thirty of those loci are on the major histocompatibility complex (MHC) suggesting a primary role in autoimmunity [8].

MS is a complex disease with an unclear and multifactorial origin and lesion pathology. Lesions are

characterized by inflammation, demyelination, and neurodegeneration of the CNS. These lesions differ in distribution and composition both among different patients and within individuals [9]. Their specific locations often correlate with clinical symptoms, leading to significant clinical variability [10, 11]. Several biological processes take place during the progression of MS: the autoreactive T-cell and B-cell aggregation in the CNS, secretion of cytokines, damage to myelin sheath, and oligodendrocytes [12-16]. The timing of these processes resulting in neurodegeneration is important for therapeutic strategies [17]. Various imaging strategies have tested variables relevant to timing the disease processes, leading to staging systems for MS [10, 14, 18]. However, there is no standard method for the classification of the lesions. The histological staging system for the post-mortem tissues includes inflammatory cells, glial cells, axonal loss, and myelin staining [18, 19].

The damage to myelin sheaths and oligodendrocytes occurs as a result of inflammation, disruption of the blood-brain barrier (BBB), and immune cell infiltration to CNS [20]. Axons are relatively preserved in the early stages of the disease; however, as the disease progresses, irreversible axonal damage develops [21]. Early studies of MS research focused on the intrathecal immunoglobulin synthesis [22-24]. With the discovery of the risk alleles in the MHC region and the novel findings of T-cells, MS is thought to be T-cell-driven [25, 26]. However, with the recent discovery of the efficacy of B-cell depletion therapies and better technologies to study the adaptive immune system, interest in antibodies in MS has arisen [27] and the profound involvement of the adaptive immune system and B-cells in disease pathology has been shown [28].

Here we utilized publicly available RNA-seq data from postmortem white matter (WM) tissues from 5 non-MS and 10 MS patient donors [29]. The previous analysis with this dataset identified lesion-specific molecular signatures, protein complex networks, and transforming growth factor beta receptor 2 (*TGFβ-R2*) as a central hub [30]. Since each lesion had a high complexity of molecular pathways, the simple molecular categorization of lesion types was not feasible, indicating a dynamic process of lesion evolution [31]. In order to extract additional information from this dataset, we employed differential expression, differential splicing, and variant analysis both among and within donors. This approach identified genes that may be involved in the progression of inflammation. With the non-MS vs MS tissue comparisons, we identified over 6,000 differentially expressed genes (FDR < 0.05). The most differentially expressed genes were involved in inflammation, immune system processes, and cell signaling. In comparing WM to different lesion types, we investigated genes that follow similar expression patterns as the inflammation level in the tissues. The genes that show a positive correlation with inflammation were enriched in immune system processes, immune response and response to stimulus and enriched in KEGG pathways such as cytokine-receptor interaction. The negatively correlated genes were enriched in neurogenesis and nervous system development in biological processes and pathways. Donor-specific differential expression analysis showed that immunoglobulin gene subfamilies were differentially expressed. Finally, differential splicing and variant analyses identified genes that play a role in dysregulation in immune system, metabolism, and axon/dendrite structure stability. Currently, there is no cure for MS. Disease-modifying therapies help recovery from attacks, modify the course of the disease, and manage symptoms. Our comprehensive analysis of differential expression and splicing of genes in MS patients demonstrates the role of previously recognized genes, as well as novel pathways. These results can lead the way to identifying novel genes and drug targets for reversing the damage, preventing MS progression, and helping patients with disease management.

METHODS

Data access and characteristics

The publicly available RNA-seq dataset was retrieved from the GEO repository (accession GSE138614) [32]. The original study that generated the data was approved by the local ethics committee of the original authors. The lesion classification method was explained in El Kajaer et. al [30]. Namely, normal-appearing white matter (NAWM), active (AL), chronic active (CA), inactive (IL), and remyelinating (RL) lesions were characterized by myelin integrity and the inflammatory state [18]. For the comparisons of NAWM tissues with AL or CA lesions within the same donors, the accession numbers are shown in Table 1.

Table 1: Sequence Read Archive accession numbers of the sample-specific comparisons. AL: active lesion; CA: chronic active lesion NAWM: normal-appearing white matter.

Sample ID	Lesion Type	SRA Accession Number
S6	AL	SRR10248474
S6	AL	SRR10248477
S6	AL	SRR10248479
S6	NAWM	SRR10248426
S6	NAWM	SRR10248446
S6	NAWM	SRR10248448
S14	CA	SRR10248408
S14	CA	SRR10248409
S14	CA	SRR10248412
S14	CA	SRR10248440
S14	CA	SRR10248443
S14	NAWM	SRR10248423
S14	NAWM	SRR10248433
S14	NAWM	SRR10248449
S14	NAWM	SRR10248452
S9	CA	SRR10248415
S9	CA	SRR10248420
S9	CA	SRR10248439
S9	NAWM	SRR10248414
S9	NAWM	SRR10248430
S9	NAWM	SRR10248434
S9	NAWM	SRR10248444
S9	NAWM	SRR10248447

Differential expression analysis

Reads from 73 MS samples from 10 MS donors and 25 samples from 5 non-MS donors were mapped to the human genome (hg38) using the Spliced Transcripts Alignment to a Reference (STAR) aligner (version 2.6) [33]. Raw gene counts were determined using HTSeq-count (version 0.10.0) [34]. Raw counts were then normalized using the relative log expression method and filtered to exclude genes with fewer than 10 counts across all samples. Differential expression analysis was performed with DESeq2 [35] using a negative binomial regression model to analyze pairwise comparisons. Statistical significance was determined based on a false discovery rate (FDR) cutoff of 0.05.

Inflammation level analysis

STEM is an algorithmic approach for clustering, comparing, and visualizing short-time series gene expression [36]. STEM selects a set of distinct and representative temporal expression profiles that are independent of the data. Each gene is assigned to the model profile that most closely matches the expression profile, determined by the correlation coefficient. A permutation test determines the assignments of genes to model profiles. The significance of each model profile is calculated by the number of assigned genes under the true ordering of time points compared to the average number assigned to the model profile [36]. STEM was utilized to investigate the correlation between gene expression and tissue inflammation. Instead of using time series samples, the inflammation level of each sample was used to determine the genes that showed positive and negative correlations.

Alternative splicing analysis

Replicate multivariate analysis of transcript splicing (rMATS v3.2.5) [37] was used to identify differentially spliced genes. rMATS employs a modified generalized linear mixed model to identify splicing from RNA-seq data with replicates. Using both splice junction and exon body read counts as input, rMATS computes the *percent spliced in* (PSI) and FDR for five different major types of splicing events: skipped exons (SE), mutually exclusive exons (MXE), retained introns (RI), and 5′ and 3′ alternative splice sites (A5SS and A3SS). Motif enrichment was analyzed in the proximity of alternatively spliced exons using rMAPS2 [38, 39]. rMAPS2 analyzes differential alternative splicing data obtained from rMATS and then graphically displays enriched motif sites.

Functional enrichment analysis

Gene ontology biological processes (GO:BP) enrichments were determined via hypergeometric testing using clusterProfiler [40] for differentially expressed and differentially spliced genes. Gene set enrichment analysis (GSEA 4.3.2) was utilized on pre-ranked differentially expressed genes within sample lesions comparisons [41].

RNA-seq short variant discovery

RNA-seq short variants were identified using GATK's workflow as a guide [42]. Alignment files were preprocessed using Picard's (v_2.25.6) AddOrReplaceGroups and GATK's (v-4.2.0.0) tool MarkDuplicates. SplitNCigarReads formatted the RNA-seq alignments for use in HaplotypeCaller. Base quality and recalibration were performed to correct for systematic errors in base quality scores prior to using HaplotypeCaller for variant calls. SNVs were filtered based on Fisher Strand values (FS > 30.0) and Qual by Depth values (QD < 2.0).

RESULTS

Differentially expressed genes in non-MS versus MS samples

We identified 6,461 differentially expressed genes between MS and non-MS tissue samples (FDR < 0.05) (Figure 1). Among these, the joining chain of multimeric IgA and IgM (*JCHAIN/IGJ*) was identified as the most significant differentially expressed gene in white matter tissue samples from MS donors. This finding aligns with Christensen et al.'s study [43], which reported increased *JCHAIN* expression in cerebrospinal fluid samples from MS patients but not in blood samples, suggesting a specific role in CNS inflammation. Other genes with high expression in MS patient samples were pro-inflammatory, such as immunoglobulin lambda like polypeptide 5 (*IGLL5*), involved in B-cell development [44], collagen type VIII alpha 1 chain (*COL8A1*) [45] and signaling lymphocytic activation molecule family member 7 (*SLAM7*), which regulates B cells and adaptive immunity and affects susceptibility to CNS autoimmunity [46]. We also identified upregulation of paralemmin 3 (*PALM3*) and hepatocyte growth factor (*HGF*) genes that are involved in response to inflammation [47]. Additionally, overexpressed IKAROS family zinc finger 3 (*IKZF3*), MER proto-oncogene, tyrosine kinase (*MERTK*), and Fc receptor-like 5 (*FCRL5*), all of which have risk alleles associated with MS and other autoimmune diseases [48-52] (Figure 1A). The most downregulated genes were primarily involved in various signal transduction and biosynthesis pathways, including tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta (*YWHAQ*), protein tyrosine phosphatase domain containing 1 (*PTPDC1*), RAB guanine nucleotide exchange factor 1 (*RABGEF1*), 1-acylglycerol-3-phosphate O-acyltransferase 4 (*AGPAT4*), and SH3 domain binding protein 5 (*SH3BP5*) (Figure 1A).

Functional enrichment analysis revealed that these differentially expressed genes were significantly involved in stimulus-response, immune system, and cell signaling functions (Figure 1B), as well as proinflammatory cytokine-

Figure 1: Differential expression of WM tissues from MS and non-MS donors. A: Volcano plot showing differentially expressed genes between MS and non-MS donor tissue samples. **B:** Enriched gene ontology biological processes and **C:** KEGG pathways from the differentially expressed genes.

cytokine receptor interaction and TGF-beta signaling pathways (Figure 1C).

Investigation of the genes correlating with the tissue inflammation level

We identified 2,886 genes that showed a positive correlation with the inflammation levels in the tissue (Figure 2A) and 647 genes showed a negative correlation (Figure 2B). Positively correlated genes enriched in biological pathways were similar to the differential expression analysis, such as immune system process, immune response, and response to stimulus (Figure 2C). Additionally, we identified viral protein interaction with cytokine receptors, vitamin digestion and absorption, and natural killer cell-mediated cytotoxicity pathways associated with inflammation levels in MS tissues. Genes showing expression levels that are negatively correlated with the tissue inflammation were enriched in the cellular anatomical entity, morphogenesis, and neurogenesis biological processes and metabolic pathways (Figure 2D).

Differential splicing and differential expression of RNA binding proteins (RBPs)

We identified differentially spliced genes in each tissue compared to non-MS WM with a total of 2,721 genes alternatively spliced in MS tissues. The intersections of alternative splicing events across comparisons are illustrated in Figure 3A. The representative motif bindings for each type of splicing (Supplemental Figure 1) and differentially expressed RNA Binding Proteins (RBPs) that bind to these motifs in MS lesions are shown in Figure 3B. These suggest the mechanism of alternative splicing may result from differentially expressed of RBPs.

Key genes with alternative splicing included myelin basic protein (*MBP*), DEAD-box helicase 5 (*DDX5*), KH domain containing RNA binding (*QKI*), and discs large MAGUK scaffold protein 1 (*DLG1*) (Table 2).

Recent research indicates that deletion of *QKI* isoforms in oligodendrocytes leads to severe CNS hypomyelination accompanied by tremors [53]. Furthermore, *DDX5*, which exhibited A5SS, RI, and SE in MS tissues, influences *MBP* expression [54]. *DLG1* plays a crucial role in lymphocyte activation [55] and has splice variants that regulate p38-dependent and independent effector functions in CD8+ T cells [56]. We also found MS patients had significant SE events in brain enriched myelin associated protein (*BCAS1*) that can result in expression of alternative isoforms (Figure 3C-D). BCAS1 is important for early myelination [57] and carries a known MS risk allele rs2585447 [7].

Comparison of NAWM with AL or CA within donors

We examined differential expression and splicing within the same donors to minimize inter-donor variation and closely analyze the CA and AL lesions by comparing them to the NAWM from the same donor. This approach enabled a detailed and specific examination of the molecular differences in the lesion areas of patient brains. A comparison of NAWM and AL showed increased expression of *IGKV4-1*, *IGHV1-2*, and *IGHV4-59* as well as local inflammation marker *CD163* and interleukin 5 receptor alpha (*IL5RA*) (Figure 4A, Supplemental Table 1). *CD24*, which carries polymorphisms linked to the progression of autoimmune disorders [58] and *JCHAIN* genes were differentially

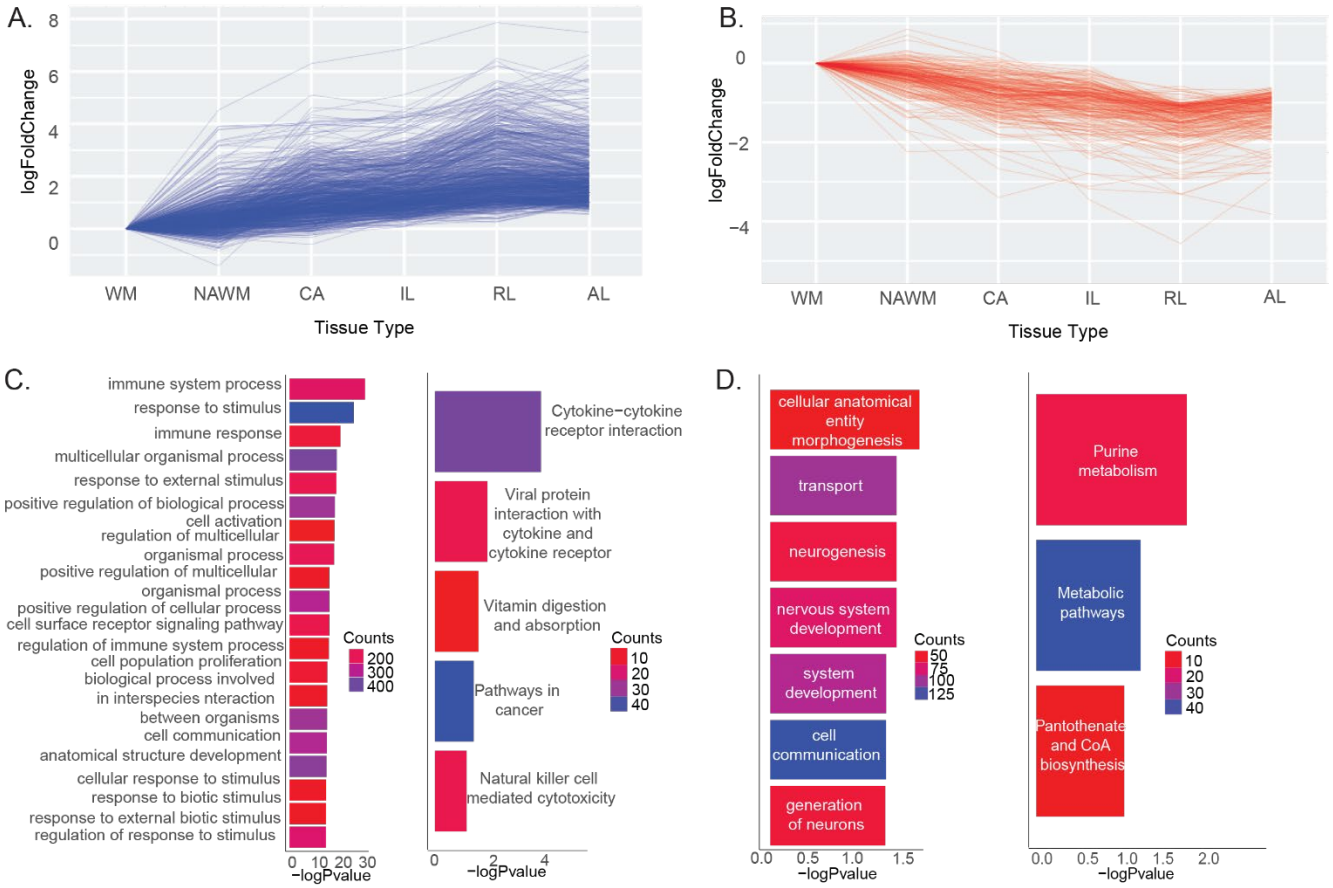

Figure 2: Genes showing A. positive and B. negative correlation in expression changes with tissue inflammation levels. C: Enriched gene ontology biological and KEGG pathways for genes showing positively correlated expression levels with inflammation. **D:** Enriched gene ontology biological and KEGG pathways for genes showing negatively correlated expression levels with inflammation.

expressed in both AL and CA comparisons to NAWM for three donors. The *IGHV3-30* gene was overexpressed in donor S9 for the CA and NAWM comparison. (Figure 4B, Supplemental Table 2). For donor S14, we found the immunoglobulin genes *IGKC*, *IGHV3-7*, *IGHG1*, *IGHV4-39*, and *IGHG3* were overexpressed in CA lesions compared to NAWM (Figure 4C. Supplemental Table 3). In addition to *CD24*, the CA to NAWM comparison in the S9 and S14 donors showed overexpression of hepatocyte growth factor (*HGF*), extracellular matrix protein (*ECM2*), complement 6 (*C6*), ceruloplasmin (*CP*), and insulin growth factor binding protein 7 anti-sense RNA 1 (*IGFBP7-AS1*) (Figure 4B-C, Supplemental Tables 2-3). This suggests these genes may have specific roles in CA lesions.

Differentially expressed AL genes were enriched in cilium movement, mainly through the upregulation of cilia and flagella associated protein 100 (*CFAP100*), cilia and flagella associated protein 45 (*CFA45*), and dynein axonemal heavy chain 5 (*DNAH5*) genes. Negative enrichment of postsynaptic membrane potential regulation indicates a substantial damage to neurons in which synaptic failure eventually leads to brain network alterations and contributes to disabling MS symptoms and disease progression [59] (Supplemental Table 4). Splice variants in AL compared to NAWM lesions were mainly enriched in metabolic processes. We found significant disruption of cellular localizations by alternatively spliced genes (Supplemental Figure 2A) that could be involved in disrupted interleukin receptor localizations in MS tissues [60].

Gene set enrichment analysis for differentially expressed genes in CA and NAWM showed B cell-mediated immunity and axoneme assembly were the most positively enriched pathways. Negatively enriched short-chain fatty acid and acetyl-COA metabolic processes have been shown to have an immunomodulatory potential in MS [61] (Supplemental Tables 5-6). Differentially spliced genes in the CA lesions compared to the NAWM from the same patients were mostly enriched in cellular component organization, nervous system development, and cell morphogenesis pathways, indicating possible pathogenicity by alternative splicing contribution to mitochondrial and metabolic dysfunction of the CA lesions (Supplemental Figure 2B).

Table 2: Differentially spliced genes in the MS patient samples. A3SS: alternative 3' splice site; A5SS: alternative 5' splice site: RI: retained intron; SE: skipped exon.

Gene Symbol	Splicing Event	Gene description
QKI	A3SS	KH Domain-Containing RNA-Binding Protein
DDX5	A5SS	DEAD-Box Helicase 5
NAP1L1	A5SS	Nucleosome Assembly Protein 1 Like 1
DDX5	RI	DEAD-Box Helicase 5
EIF4G2	RI	Eukaryotic Translation Initiation Factor 4 Gamma 2
HMGN1	RI	high mobility group nucleosome binding domain 1
RBM39	RI	RNA Binding Motif Protein 39
RPL7A	RI	Ribosomal Protein L7a
SMARCC2	RI	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin Subfamily C Member 2
ANLN	SE	Anillin, Actin Binding Protein
ARID4B	SE	AT-Rich Interaction Domain 4B
BCAS1	SE	Brain-Enriched Myelin Associated Protein 1
CD44	SE	CD44 Molecule
DDX5	SE	DEAD-Box Helicase 5
DLG1	SE	discs large MAGUK scaffold protein 1
DST	SE	Dystonin
EIF4A2	SE	eukaryotic translation initiation factor 4A2
EPB41L2	SE	Erythrocyte Membrane Protein Band 4.1 Like 2
ERBIN	SE	erbb2 interacting protein
FGF1	SE	fibroblast growth factor 1
FIP1L1	SE	factor interacting with PAPOLA and CPSF1
GLIS3	SE	GLIS family zinc finger 3
HMGN1	SE	high mobility group nucleosome binding domain 1
HNRNPH3	SE	Heterogeneous Nuclear Ribonucleoprotein H3
ITSN1	SE	intersectin 1
LRRC63	SE	leucine-rich repeat containing 63
MACF1	SE	microtubule actin crosslinking factor 1
MAP4	SE	Microtubule Associated Protein 4
MAP4K4	SE	Mitogen-Activated Protein Kinase Kinase Kinase Kinase 4
MBP	SE	Myelin Basic Protein
NCAM1	SE	Neural Cell Adhesion Molecule 1
NCOA6	SE	Nuclear Receptor Coactivator 6
NDRG2	SE	NDRG family member 2
PICALM	SE	Phosphatidylinositol Binding Clathrin Assembly Protein
PRDM2	SE	PR/SET domain 2
PTK2	SE	protein tyrosine kinase 2
PTPN11	SE	protein tyrosine phosphatase non-receptor type 11
PTPRD	SE	protein tyrosine phosphatase receptor type D
PXK	SE	PX Domain Containing Serine/ Threonine Kinase Like
RBM39	SE	RNA Binding Motif Protein 39
SHTN1	SE	Shootin 1
SLTM	SE	SAFB like transcription modulator
SORBS1	SE	Sorbin And SH3 Domain Containing 1
SOX2-OT	SE	SOX2 overlapping transcript
SPECC1	SE	Sperm Antigen with Calponin Homology and Coiled-Coil Domains 1
TLE4	SE	TLE Family Member 4, Transcriptional Corepressor
TMEM165	SE	transmembrane protein 165
WAC	SE	WW Domain Containing Adaptor with Coiled-Coil
ZNF207	SE	Zinc Finger Protein 207
ZNF638	SE	Zinc Finger Protein 638

Table 2 (continued). A3SS: alternative 3' splice site; A5SS: alternative 5' splice site: RI: retained intron; SE: skipped exon.

Spontaneous single nucleotide variations and alternative splicing

A comparison of the single nucleotide variations (SNVs) in MS and non-MS samples did not match any known SNV markers of MS. This is likely due to the small number of donors. We then compared SNVs in AL and CA lesions to NAWM lesions in the S6, S9 and S14 donors and identified a common synonymous variation (rs4947) in the heat shock protein 90 alpha family class A member 1 (*HSP90AA1*) gene across all comparisons. We also identified SNVs on rho GTPase activating protein 21 (*ARHGAP21*), contactin 1 (*CNTN1*), solute carrier family 1 member 2 (*SLC1A2*), UDP glycosyltransferase 8 (*UGT8*), solute carrier organic anion transporter family member 1C1 (*SLCO1C1*), protocadherin 9 (*PCDH9*), and membrane associated ring-CH-type finger 1 (*MARCHF1*). Between CA and NAWM tissues, *ARHGAP21* had significant SE events in both the S9 and S14 donors, as well as in the AL lesions of the donor S6. *SLC1A2* had significant SE events in S14. *CNTN1* also showed an SE variant in S14 CA lesions and S6 AL lesions. The same events for *SLC1A2* and *CNTN1* were found in S9 but were not determined to be significant. *PCDH9* carried spontaneous SNVs in all comparisons and had a significant SE event in both AL and CA lesions from S6 and S9 donors compared to their NAWMs. (Table 3).

MARCHF1 is an E3 ubiquitin ligase and MARCH1-mediated ubiquitination of MHC II impacts the MHC I antigen presentation pathway [62]. *MARCHF1* carries a known MS risk variant rs72989863. We identified different spontaneous SNVs that are upstream variants in CA lesions from both S9 and S14 donors. The S14 donor also had a significant SE variant in CA lesions (Figure 5A). The same region was also skipped in CA lesions from S9 but were not significant. In donor S6, the comparison of AL and NAWM tissues yields the same SE event on *MARCHF1*. This region corresponds to the exon 4 of the ENST00000514618 transcript and this SE event may result in the expression of the ENST0000503008 transcript (Figure 5B). These two transcripts translate into two different MARCHF1 protein isoforms that may alter its E3 ubiquitin ligase activity (uniport IDs Q8TCQ1 and D6RGC4) [63, 64]. However, it is essential to note that our analysis utilized short-read sequencing data which does not allow identification of specific transcripts.

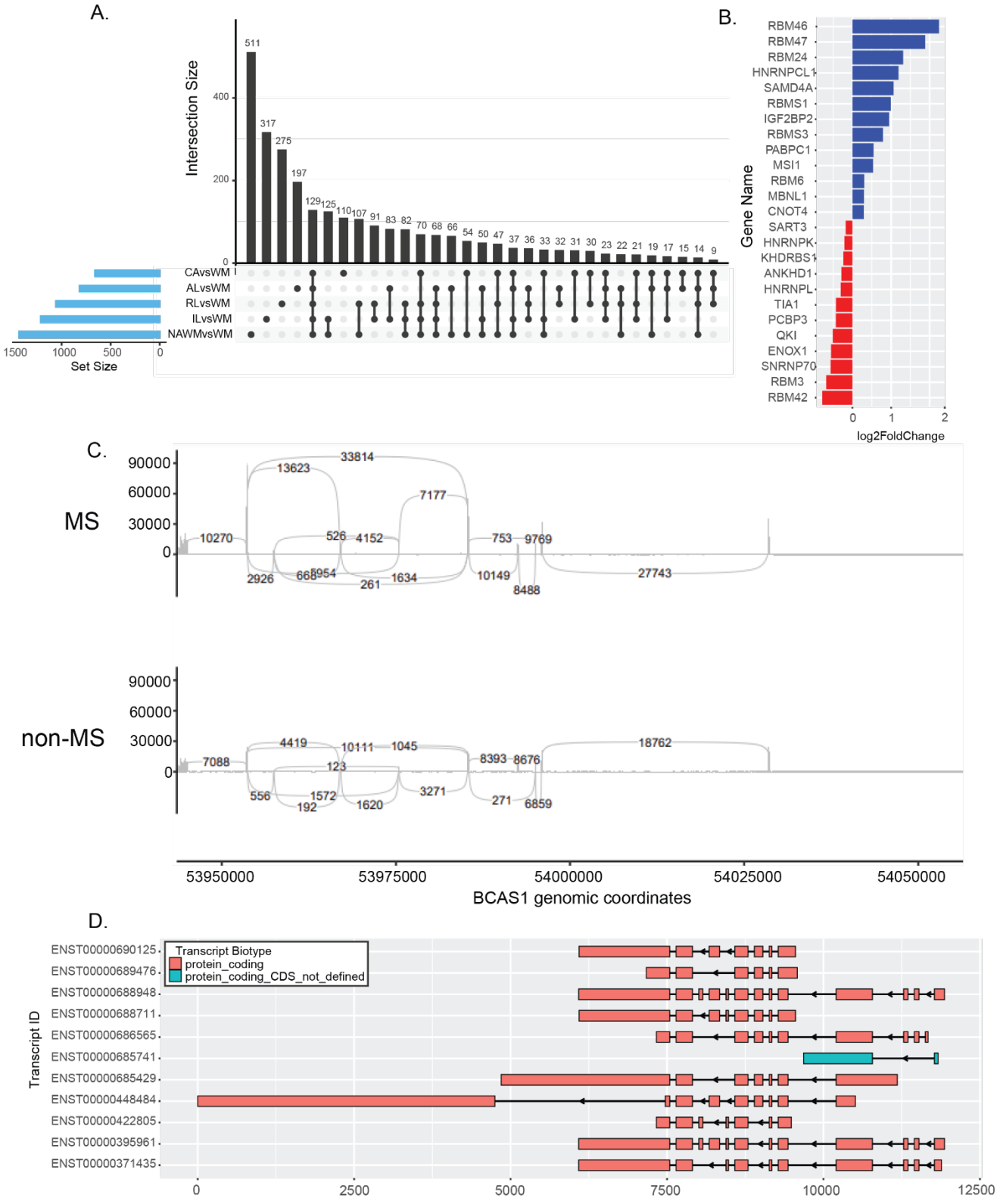

Figure 3: Differentially spliced genes in all tissues from MS donor samples compared to non-MS samples. A: Upset graph showing the number of alternatively spliced genes in each category of tissues and intersections. **B:** Differentially expressed RNA binding proteins that are involved in alternative splicing. **C:** Sashimi plot of *BCAS1* gene that is differentially spliced in MS and non-MS tissues. **D:** Known transcripts of *BCAS1* gene.

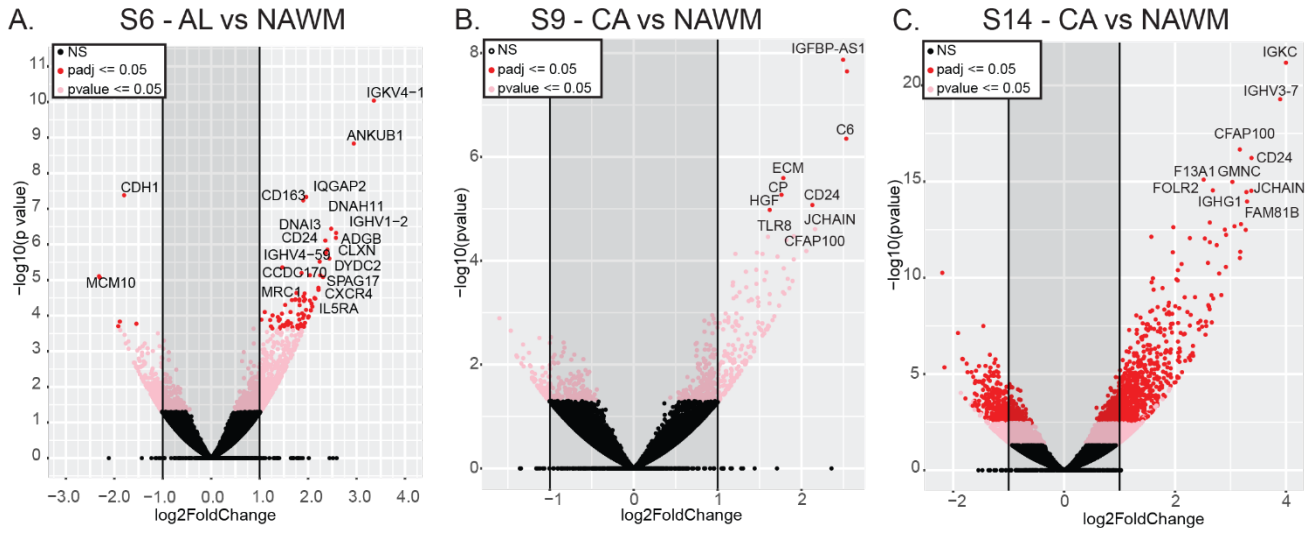

Figure 4: Differentially expressed genes in different tissue types from the same donors. Volcano plots showing differentially expressed genes in **A:** AL compared to NAWM in donor S6, CA compared to NAWM in **B:** donor S9 and **C:** S14.

Table 3: Common SNPs and alternative splicing events in AL and CA tissue comparisons to NAWM. A3SS: alternative 3' splice site; MXE: mutually exclusive exon; RI: retained intron; SE: skipped exon.

CAvsNAWM S9			
Gene ID	**SNP location**	**SNP ID**	**Significant AS Events**
ARHGAP21	chr10:24621826	rs4262623	MXE
SLCO1A2	chr12:21324153	rs10459075	RI
PCDH9	chr13:66988744	rs9540962	SE
SLC1A2	chr11:35259553	rs10742339	
CNTN1	chr12:40733457	rs11178118	
UGT8	chr4:114652891	rs11931776	
UGT8	chr4:139509427	rs1354563728	
MARCHF1	chr4:164241452	rs4337693	
MARCHF1	chr4:163974090	rs6857054	
MARCHF1	chr4:163851505	rs7657544	
HSP90AA1	chr14:102084466	rs4947	
CA vs NAWM S14			
Gene Name	**SNP location**	**SNP ID**	**Significant AS Events**
ARHGAP21	chr10:24620020	not reported	SE
CNTN1	chr12:40943633	rs1056019	SE
HSP90AA1	chr14:102084466	rs4947	
MARCHF1	chr4:163568866	rs1877314	SE
PCDH9	chr13:66953673	rs9540955	
PCDH9	chr13:67216229	rs7489531	
SLC1A2	chr11:35258109	rs10768121	SE/MXE
SLCO1C2	chr12:20704542	rs10770706	
UGT8	chr4:114668146	rs11098262	
HSP90AA1	chr14:102084466	rs4947	
AL vs NAWM S6			
Gene Name	**SNP location**	**SNP ID**	**Significant AS Events**
HSP90AA1	chr14:102084466	rs4947	
PCDH9	chr13:67216229	rs7489531	SE/A3SS
UGT8	chr4:114668146	rs11098262	
ARHGAP21			SE
CNTN1			SE

DISCUSSION

MS is highly complex and variable, both among patients and the stages of the disease [65]. Understanding the molecular processes through disease progression and identifying molecular signatures of the specific pathology of the MS lesions is crucial for understanding the disease development and progression, as well as for patient stratification for prognosis, predicting treatment response, and identifying treatment strategies [66]. The dataset we utilized has a large number of samples from different types of MS lesions which allowed us to investigate differences in these lesions at the level of differential gene expression, alternative splicing, and SNVs. Identifying gene expression patterns correlating with tissue inflammation can be used as a marker for lesion stage determination. However, with a large number of samples, greater variability makes it difficult to cluster samples using their expression profiles (e.g. PCA). The source of this variability may be the result of sample collection from different parts of the brain [67].

Alternative splicing is one of the main mechanisms affecting the expression of genes and gene isotypes. A recent study showed that there are a large number of alternative splicing variants in MS that are not linked to the differential expression [68] and there are splice variants found to be associated with known risk alleles [69]. Many of the alternatively spliced genes we identified play roles in MS-related pathways. Deletion of *QKI* isoforms in oligodendrocytes leads to severe CNS hypomyelination accompanied by tremors [53]. *DLG1* plays a crucial role in lymphocyte activation and has identified spliced variants that regulate p38-dependent and independent effector functions in CD8+ T cells [56]. *BCAS1* is highly expressed in oligodendrocytes and plays role in demyelination in MS [57].

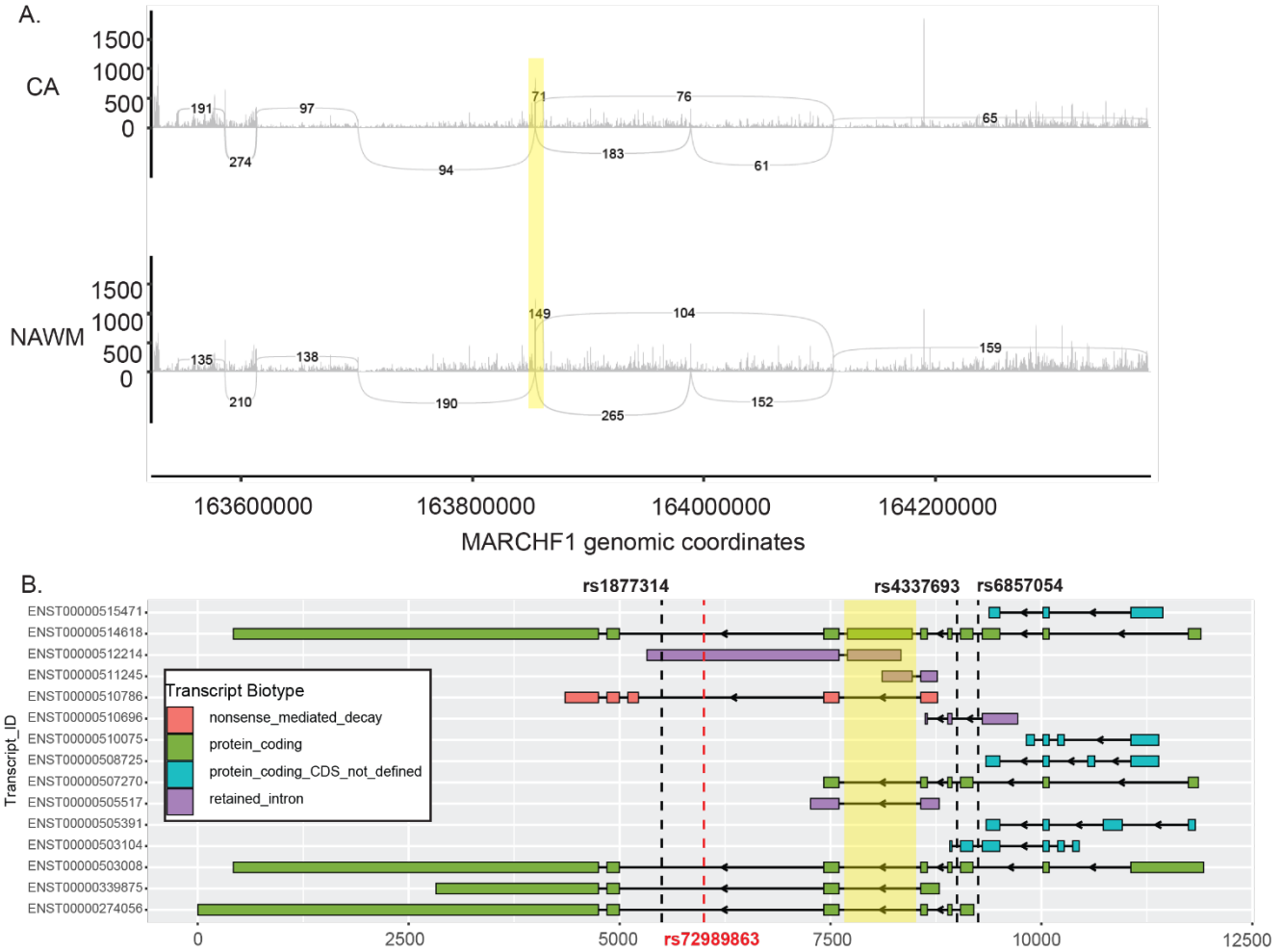

Figure 5: Sashimi plot of the *MARCHF1* gene that is differentially spliced in CA lesions. Skipped exon 4 is highlighted in yellow. **B:** Known transcripts of MARCHF1 gene. Known MS risk allele is shown in red (rs72989863) and the identified spontaneous SNPs are shown in black. Skipped exon 4 is highlighted in yellow.

We discovered *BCAS1,* which carries a known MS risk allele rs2585447 [7], had a splice variant in MS patient samples. Our analysis also identified differential splicing of *MBP* and *MOBP* in different types of MS lesions that show variable expression levels in MS [70, 71]. Studying these alternative isoforms more thoroughly could enhance our comprehension of how MS progresses and how we might be able to halt or reverse it.

Although we had limited samples for the tissue type comparisons within donors, our approach in investigating the differential expression and differential splicing of CA vs NAWM and AL vs NAWM lesions resulted in the identification of specific pathways such as fatty acid and acetyl COA metabolic processes, as well as cellular components and cytoskeleton organization. Lipids are not solely involved in the formation of the myelin sheath but are found to be important components of cell signaling, communication, and transport in CNS and have been shown to be relevant players in neuroinflammation and neurodegeneration [72]. We identified specific pathways that are affected in specific stages of MS lesions, which can be markers for the staging of these lesions, as well as the selection of treatment strategies. They are also the basis for further research in reversing damage and treating disability symptoms of MS. The inflammatory activities in lesions are shown to result in disease progression without any new lesion formation [73]. In AL, we identified markers for neuronal damage by changes in cilia and postsynaptic membrane potential pathways. Targeting these specific proteins and pathways, may block the silent progression and prevent further disability development in the patients.

Comparison of NAWM tissues with CA or AL within patients identified spontaneous SNVs that may be associated with alternative splicing in specific MS lesions. Among the genes we identified, ARHGAP21 is involved in the signaling for synaptic homeostasis and axon/dendritic transport regulation [74]. Serum CNTN1 concentrations are associated with the MS progression in RRMS patients [75]. Solute carrier family 1 member *SLC1A2* is expressed in astrocytes, neurons, and axonal terminals and upregulated in MS. It regulates glutamate concentrations in the CNS preventing excitotoxicity [76, 77]. UGT8 plays an important role in remyelination by mediating the major myelin lipid galactosylceramide [78]. MARCHF1 is an E3 ubiquitin that

mediates the surface turnover of MHC class II (MHCII) and CD86 and plays an essential role in restraining an exhaustion-like program of effector CD4+ T cells [79]. MARCHF1 also regulates type I interferon signaling, T cell activation, and IFN-γ production during infections [80]. Moreover, deletion of the chromosomal positions chr4:164,703,186-165,032,803 encompassing MARCH1 resulted in a previously unreported growth failure, developmental and speech delays and aggressive behavior [81]. It also causes an inflamed tumor microenvironment and suggested as an immune status biomarker for effectiveness of immunotherapy in cancer treatment [82]. We were able to predict the isoforms of *MARCHF1* that are expressed in MS lesions and NAWM; however, utilization of long read sequencing methods are necessary to identify the exact transcripts in each tissue. Although myelin sheath breakdown is the hallmark of MS, the exact targeting and degradation mechanism is unknown. Our results may enable the discovery of the myelin degradation process, as alternative splicing has been shown to affect the selective ubiquitination of different protein isoforms [83]. Identification of the isoforms is important to understand the mechanism, prevent MS progression, and reverse neuronal damage. However, as noted, further studies require more specific methods for the identification of protein isotypes such as IsoSeq and deep proteome sequencing [84, 85].

Both B cell and T cell involvement is known for MS [86]. Our analysis, expectedly, showed the involvement of B-cell mediated immunity in MS lesions. We found brain-derived neurotrophic factor (*BDNF*) as one of the differentially spliced genes that is secreted by B-cells, which prevents axonal loss [87] and highly expressed in the actively demyelinating area [88]. BDNF isoforms have been identified to affect neurogenesis and expression of serotonergic agents [89]. Recent studies show IgG constant region polymorphisms effect antibody stability and dynamics [90]. In MS, a preferential pairing of the IGHV4 gene family with the IGKV1 gene family is shown and IGHV4-39 gene is identified as the most abundant subisotype [91]. Our data showed overexpression of different Ig isotypes and subisotypes in AL and CA lesions. It is important to acknowledge we were able to compare these lesions to NAWM within samples from only 2 donors for CA and 1 donor for AL lesions. However, our results suggest a novel approach in comparison of different categories of lesions could identify specific Ig gene usage and specific biomarkers for molecular classifications via MRI [92].

CONCLUSION

Our analysis explored various methods and approaches to identify novel molecular characteristics of MS lesions in the brain. We discovered the metabolic pathways were highly affected by gene expression changes. We provided additional insight to differential expression with our differential splicing analysis and identified genes with crucial roles in MS related pathways have alternative isotypes in MS lesions. Our findings identified RNA binding motifs that are involved in these alternative splicing events. Additionally, with the comparison of lesions within donors, we found spontaneous SNPs and alternative isoforms in genes that are essential factors in autoimmunity, neuron homeostasis, and myelination that may have pivotal roles in MS lesions. Overall, our results indicate splice variants in specific MS lesions may be used as biomarkers determine the staging of the lesions as well as treatment targets.

ACKNOWLEDGEMENTS

We wish to thank members of the KY INBRE Bioinformatics Core and the SWRM lab members for their helpful insight and feedback.

FUNDING

This work was supported by the National Institutes of Health [P20GM103436 and P30ES030283]. The contents of this work are solely the responsibility of the authors and does not reflect the official views of the National Institutes of Health.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY

The samples utilized in this analysis are publicly available in GEO accession GSE138614.

AUTHOR CONTRIBUTIONS

MS: conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing – original draft, and writing – review and editing. JHC: formal analysis, software, supervision, and writing – review & editing. JWP: formal analysis and writing – review & editing. ECR: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing – review & editing.

REFERENCES

[1] C. Walton *et al.*, "Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition," *Mult Scler,* vol. 26, no. 14, pp. 1816-1821, Dec 2020, doi: 10.1177/1352458520970841.

[2] C. S. Simonsen, H. O. Flemmen, T. Lauritzen, P. Berg-Hansen, S. M. Moen, and E. G. Celius, "The diagnostic value of IgG index versus oligoclonal bands in cerebrospinal fluid of patients with multiple sclerosis," *Mult Scler J Exp Transl Clin,* vol. 6, no. 1, p. 2055217319901291, Jan-Mar 2020, doi: 10.1177/2055217319901291.

[3] J. Nieves, F. Cosman, J. Herbert, V. Shen, and R. Lindsay, "High prevalence of vitamin D deficiency and reduced bone mass in multiple sclerosis," *Neurology,* vol. 44, no. 9, pp. 1687-92, Sep 1994, doi: 10.1212/wnl.44.9.1687.

[4] A. Ascherio and K. L. Munger, "Epstein-barr virus infection and multiple sclerosis: a review," *J Neuroimmune Pharmacol,* vol. 5, no. 3, pp. 271-7, Sep 2010, doi: 10.1007/s11481-010-9201-3.

[5] C. O'Gorman, R. Lin, J. Stankovich, and S. A. Broadley, "Modelling genetic susceptibility to multiple sclerosis with family data," *Neuroepidemiology,* vol. 40, no. 1, pp. 1-12, 2013, doi: 10.1159/000341902.

[6] D. S. Goodin, "The genetic basis of multiple sclerosis: a model for MS susceptibility," *BMC Neurol,* vol. 10, p. 101, Oct 28 2010, doi: 10.1186/1471-2377-10-101.

[7] C. International Multiple Sclerosis Genetics, "Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility," *Science,* vol. 365, no. 6460, Sep 27 2019, doi: 10.1126/science.aav7188.

[8] J. A. Hollenbach and J. R. Oksenberg, "The immunogenetics of multiple sclerosis: A comprehensive review," *J Autoimmun,* vol. 64, pp. 13-25, Nov 2015, doi: 10.1016/j.jaut.2015.06.010.

[9] C. Lucchinetti, W. Bruck, J. Parisi, B. Scheithauer, M. Rodriguez, and H. Lassmann, "Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination," *Ann Neurol,* vol. 47, no. 6, pp. 707-17, Jun 2000, doi: 10.1002/1531-8249(200006)47:6<707::aid-ana3>3.0.co;2-q.

[10] T. Kuhlmann, S. Ludwin, A. Prat, J. Antel, W. Bruck, and H. Lassmann, "An updated histological classification system for multiple sclerosis lesions," *Acta Neuropathol,* vol. 133, no. 1, pp. 13-24, Jan 2017, doi: 10.1007/s00401-016-1653-y.

[11] A. Feinstein, L. D. Kartsounis, D. H. Miller, B. D. Youl, and M. A. Ron, "Clinically isolated lesions of the type seen in multiple sclerosis: a cognitive, psychiatric, and MRI follow up study," *J Neurol Neurosurg Psychiatry,* vol. 55, no. 10, pp. 869-76, Oct 1992, doi: 10.1136/jnnp.55.10.869.

[12] J. M. Frischer *et al.*, "Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque," *Ann Neurol,* vol. 78, no. 5, pp. 710-21, Nov 2015, doi: 10.1002/ana.24497.

[13] B. D. Trapp, L. Bo, S. Mork, and A. Chang, "Pathogenesis of tissue injury in MS lesions," *J Neuroimmunol,* vol. 98, no. 1, pp. 49-56, Jul 1 1999, doi: 10.1016/s0165-5728(99)00081-8.

[14] P. van der Valk and C. J. De Groot, "Staging of multiple sclerosis (MS) lesions: pathology of the time frame of MS," *Neuropathol Appl Neurobiol,* vol. 26, no. 1, pp. 2-10, Feb 2000, doi: 10.1046/j.1365-2990.2000.00217.x.

[15] B. Serafini *et al.*, "Dendritic cells in multiple sclerosis lesions: maturation stage, myelin uptake, and interaction with proliferating T cells," *J Neuropathol Exp Neurol,* vol. 65, no. 2, pp. 124-41, Feb 2006, doi: 10.1097/01.jnen.0000199572.96472.1c.

[16] S. Markovic-Plese and H. F. McFarland, "Immunopathogenesis of the multiple sclerosis lesion," *Curr Neurol Neurosci Rep,* vol. 1, no. 3, pp. 257-62, May 2001, doi: 10.1007/s11910-001-0028-4.

[17] H. Lassmann, "Targets of therapy in progressive MS," *Mult Scler,* vol. 23, no. 12, pp. 1593-1599, Oct 2017, doi: 10.1177/1352458517729455.

[18] R. Reynolds, F. Roncaroli, R. Nicholas, B. Radotra, D. Gveric, and O. Howell, "The neuropathological basis of clinical progression in multiple sclerosis," *Acta Neuropathol,* vol. 122, no. 2, pp. 155-70, Aug 2011, doi: 10.1007/s00401-011-0840-0.

[19] C. J. De Groot *et al.*, "Post-mortem MRI-guided sampling of multiple sclerosis brain lesions: increased yield of active demyelinating and (p)reactive lesions," *Brain,* vol. 124, no. Pt 8, pp. 1635-45, Aug 2001, doi: 10.1093/brain/124.8.1635.

[20] S. Rodriguez Murua, M. F. Farez, and F. J. Quintana, "The Immune Response in Multiple Sclerosis," *Annu Rev Pathol,* vol. 17, pp. 121-139, Jan 24 2022, doi: 10.1146/annurev-pathol-052920-040318.

[21] R. Dutta and B. D. Trapp, "Pathogenesis of axonal and neuronal damage in multiple sclerosis," *Neurology,* vol. 68, no. 22 Suppl 3, pp. S22-31; discussion S43-54, May 29 2007, doi: 10.1212/01.wnl.0000275229.13012.32.

[22] M. K. Sharief, "Intrathecal synthesis of IgM in early multiple sclerosis," *Acta Neurol Scand,* vol. 84, no. 5, pp. 456-7, Nov 1991, doi: 10.1111/j.1600-0404.1991.tb04989.x.

[23] M. K. Sharief and R. Hentges, "Importance of intrathecal synthesis of IgD in multiple sclerosis. A combined clinical, immunologic, and magnetic resonance imaging study," *Arch Neurol,* vol. 48, no. 10, pp. 1076-9, Oct 1991, doi: 10.1001/archneur.1991.00530220098026.

[24] M. K. Sharief and E. J. Thompson, "Intrathecal immunoglobulin M synthesis in multiple sclerosis. Relationship with clinical and cerebrospinal fluid parameters," *Brain,* vol. 114 (Pt 1A), pp. 181-95, Feb 1991. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/1998881.

[25] A. M. Burrell, A. E. Handel, S. V. Ramagopalan, G. C. Ebers, and J. M. Morahan, "Epigenetic mechanisms in multiple sclerosis and the major histocompatibility complex (MHC)," *Discov Med,* vol. 11, no. 58, pp. 187-96, Mar 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21447278.

[26] C. International Multiple Sclerosis Genetics *et al.*, "Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis," *Nature,* vol. 476, no. 7359, pp. 214-9, Aug 10 2011, doi: 10.1038/nature10251.

[27] J. M. Gelfand, B. A. C. Cree, and S. L. Hauser, "Ocrelizumab and Other CD20(+) B-Cell-Depleting Therapies in Multiple Sclerosis," *Neurotherapeutics,* vol. 14, no. 4, pp. 835-841, Oct 2017, doi: 10.1007/s13311-017-0557-4.

[28] A. K. Probstel and S. L. Hauser, "Multiple Sclerosis: B Cells Take Center Stage," *J Neuroophthalmol,* vol. 38, no. 2, pp. 251-258, Jun 2018, doi: 10.1097/WNO.0000000000000642.

[29] T. Frisch *et al.*, "Multiple Sclerosis Atlas: A Molecular Map of Brain Lesion Stages in Progressive Multiple Sclerosis," *Netw Syst Med,* vol. 3, no. 1, pp. 122-129, 2020, doi: 10.1089/nsm.2020.0006.

[30] M. L. Elkjaer *et al.*, "Molecular signature of different lesion types in the brain white matter of patients with progressive multiple sclerosis," *Acta Neuropathol Commun,* vol. 7, no. 1, p. 205, Dec 11 2019, doi: 10.1186/s40478-019-0855-7.

[31] A. Rovira, C. Auger, and J. Alonso, "Magnetic resonance monitoring of lesion evolution in multiple sclerosis," *Ther Adv Neurol Disord,* vol. 6, no. 5, pp. 298-310, Sep 2013, doi: 10.1177/1756285613484079.

[32] T. Barrett *et al.*, "NCBI GEO: archive for high-throughput functional genomic data," *Nucleic Acids Res,* vol. 37, no. Database issue, pp. D885-90, Jan 2009, doi: 10.1093/nar/gkn764.

[33] A. Dobin *et al.*, "STAR: ultrafast universal RNA-seq aligner," *Bioinformatics,* vol. 29, no. 1, pp. 15-21, Jan 1 2013, doi: 10.1093/bioinformatics/bts635.

[34] S. Anders, P. T. Pyl, and W. Huber, "HTSeq—a Python framework to work with high-throughput sequencing data," *bioinformatics,* vol. 31, no. 2, pp. 166-169, 2015.

[35] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome biology,* vol. 15, pp. 1-21, 2014.

[36] J. Ernst and Z. Bar-Joseph, "STEM: a tool for the analysis of short time series gene expression data," *BMC Bioinformatics,* vol. 7, p. 191, Apr 5 2006, doi: 10.1186/1471-2105-7-191.

[37] S. Shen *et al.*, "rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data," *Proc Natl Acad Sci U S A,* vol. 111, no. 51, pp. E5593-601, Dec 23 2014, doi: 10.1073/pnas.1419161111.

[38] J. W. Park, S. Jung, E. C. Rouchka, Y. T. Tseng, and Y. Xing, "rMAPS: RNA map analysis and plotting server for alternative exon regulation," *Nucleic Acids Res,* vol. 44, no. W1, pp. W333-8, Jul 8 2016, doi: 10.1093/nar/gkw410.

[39] J. Y. Hwang, S. Jung, T. L. Kook, E. C. Rouchka, J. Bok, and J. W. Park, "rMAPS2: an update of the RNA map analysis and

plotting server for alternative splicing regulation," *Nucleic Acids Research,* vol. 48, no. W1, pp. W300-W306, 2020.

[40] G. Yu, L.-G. Wang, Y. Han, and Q.-Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *Omics: a journal of integrative biology,* vol. 16, no. 5, pp. 284-287, 2012.

[41] A. Subramanian *et al.*, "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proc Natl Acad Sci U S A,* vol. 102, no. 43, pp. 15545-50, Oct 25 2005, doi: 10.1073/pnas.0506580102.

[42] A. McKenna *et al.*, "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data," *Genome Res,* vol. 20, no. 9, pp. 1297-303, Sep 2010, doi: 10.1101/gr.107524.110.

[43] J. Romme Christensen *et al.*, "Systemic inflammation in progressive multiple sclerosis involves follicular T-helper, Th17- and activated B-cells and correlates with progression," *PLoS One,* vol. 8, no. 3, p. e57820, 2013, doi: 10.1371/journal.pone.0057820.

[44] Z. Liu *et al.*, "Single-cell transcriptional profiling reveals aberrant gene expression patterns and cell states in autoimmune diseases," *Mol Immunol,* vol. 165, pp. 68-81, Jan 2024, doi: 10.1016/j.molimm.2023.12.010.

[45] X. Zhang, Y. Song, X. Chen, X. Zhuang, Z. Wei, and L. Yi, "Integration of Genetic and Immune Infiltration Insights into Data Mining of Multiple Sclerosis Pathogenesis," *Comput Intell Neurosci,* vol. 2022, p. 1661334, 2022, doi: 10.1155/2022/1661334.

[46] P. O'Connell, M. K. Blake, S. Godbehere, A. Amalfitano, and Y. A. Aldhamen, "SLAMF7 modulates B cells and adaptive immunity to regulate susceptibility to CNS autoimmunity," *J Neuroinflammation,* vol. 19, no. 1, p. 241, Oct 3 2022, doi: 10.1186/s12974-022-02594-9.

[47] L. Bai *et al.*, "Hepatocyte growth factor mediates mesenchymal stem cell-induced recovery in multiple sclerosis models," *Nat Neurosci,* vol. 15, no. 6, pp. 862-70, Jun 2012, doi: 10.1038/nn.3109.

[48] M. Comabella, J. Sastre-Garriga, and X. Montalban, "Precision medicine in multiple sclerosis: biomarkers for diagnosis, prognosis, and treatment response," *Curr Opin Neurol,* vol. 29, no. 3, pp. 254-62, Jun 2016, doi: 10.1097/WCO.0000000000000336.

[49] M. Chorazy *et al.*, "Variants of Novel Immunomodulatory Fc Receptor Like 5 Gene Are Associated With Multiple Sclerosis Susceptibility in the Polish Population," *Front Neurol,* vol. 12, p. 631134, 2021, doi: 10.3389/fneur.2021.631134.

[50] C. Ono *et al.*, "Upregulated Fcrl5 disrupts B cell anergy and causes autoimmune disease," *Front Immunol,* vol. 14, p. 1276014, 2023, doi: 10.3389/fimmu.2023.1276014.

[51] M. Stefanovic, L. Stojkovic, I. Zivotic, E. Dincic, A. Stankovic, and M. Zivkovic, "Expression levels of GSDMB and ORMDL3 are associated with relapsing-remitting multiple sclerosis and IKZF3 rs12946510 variant," *Heliyon,* vol. 10, no. 3, p. e25033, Feb 15 2024, doi: 10.1016/j.heliyon.2024.e25033.

[52] K. K. Farh *et al.*, "Genetic and epigenetic fine mapping of causal autoimmune disease variants," *Nature,* vol. 518, no. 7539, pp. 337-43, Feb 19 2015, doi: 10.1038/nature13835.

[53] I. Lavon *et al.*, "QKI-V5 is downregulated in CNS inflammatory demyelinating diseases," *Mult Scler Relat Disord,* vol. 39, p. 101881, Apr 2020, doi: 10.1016/j.msard.2019.101881.

[54] P. Hoch-Kraft, R. White, S. Tenzer, E. M. Kramer-Albers, J. Trotter, and C. Gonsior, "Dual role of the RNA helicase DDX5 in post-transcriptional regulation of myelin basic protein in oligodendrocytes," *J Cell Sci,* vol. 131, no. 9, May 1 2018, doi: 10.1242/jcs.204750.

[55] R. Xavier *et al.*, "Discs large (Dlg1) complexes in lymphocyte activation," *J Cell Biol,* vol. 166, no. 2, pp. 173-8, Jul 19 2004, doi: 10.1083/jcb.200309044.

[56] O. Silva, J. Crocetti, L. A. Humphries, J. K. Burkhardt, and M. C. Miceli, "Discs Large Homolog 1 Splice Variants Regulate p38-Dependent and -Independent Effector Functions in CD8+ T Cells," *PLoS One,* vol. 10, no. 7, p. e0133353, 2015, doi: 10.1371/journal.pone.0133353.

[57] M. K. Fard *et al.*, "BCAS1 expression defines a population of early myelinating oligodendrocytes in multiple sclerosis lesions," *Sci Transl Med,* vol. 9, no. 419, Dec 6 2017, doi: 10.1126/scitranslmed.aam7816.

[58] W. Yang *et al.*, "Association between CD24 Ala/Val polymorphism and multiple sclerosis risk: A meta analysis," *Medicine (Baltimore),* vol. 99, no. 15, p. e19530, Apr 2020, doi: 10.1097/MD.0000000000019530.

[59] L. Bellingacci, A. Mancini, L. Gaetani, A. Tozzi, L. Parnetti, and M. Di Filippo, "Synaptic Dysfunction in Multiple Sclerosis: A Red Thread from Inflammation to Network Disconnection," *Int J Mol Sci,* vol. 22, no. 18, Sep 9 2021, doi: 10.3390/ijms22189753.

[60] S. Hulshof, L. Montagne, C. J. De Groot, and P. Van Der Valk, "Cellular localization and expression patterns of interleukin-10, interleukin-4, and their receptors in multiple sclerosis lesions," *Glia,* vol. 38, no. 1, pp. 24-35, Apr 1 2002, doi: 10.1002/glia.10050.

[61] L. Barcutean, S. Maier, M. Burai-Patrascu, L. Farczadi, and R. Balasa, "The Immunomodulatory Potential of Short-Chain Fatty Acids in Multiple Sclerosis," *Int J Mol Sci,* vol. 25, no. 6, Mar 11 2024, doi: 10.3390/ijms25063198.

[62] K. R. Wilson *et al.*, "MARCH1-mediated ubiquitination of MHC II impacts the MHC I antigen presentation pathway," *PLoS One,* vol. 13, no. 7, p. e0200540, 2018, doi: 10.1371/journal.pone.0200540.

[63] L. W. Hillier *et al.*, "Generation and annotation of the DNA sequences of human chromosomes 2 and 4," *Nature,* vol. 434, no. 7034, pp. 724-31, Apr 7 2005, doi: 10.1038/nature03466.

[64] M. Steinegger and J. Soding, "Clustering huge protein sequence sets in linear time," *Nat Commun,* vol. 9, no. 1, p. 2542, Jun 29 2018, doi: 10.1038/s41467-018-04964-5.

[65] I. Del Negro, S. Pez, G. L. Gigli, and M. Valente, "Disease Activity and Progression in Multiple Sclerosis: New Evidences and Future Perspectives," *J Clin Med,* vol. 11, no. 22, Nov 9 2022, doi: 10.3390/jcm11226643.

[66] A. Gajofatto and M. D. Benedetti, "Treatment strategies for multiple sclerosis: When to start, when to change, when to stop?," *World J Clin Cases,* vol. 3, no. 7, pp. 545-55, Jul 16 2015, doi: 10.12998/wjcc.v3.i7.545.

[67] S. K. Negi and C. Guda, "Global gene expression profiling of healthy human brain and its application in studying neurological disorders," *Sci Rep,* vol. 7, no. 1, p. 897, Apr 18 2017, doi: 10.1038/s41598-017-00952-9.

[68] A. T. Reder, A. Goel, T. Garcia, and X. Feng, "Alternative Splicing of RNA Is Excessive in Multiple Sclerosis and Not Linked to Gene Expression Levels: Dysregulation Is Corrected by IFN-beta," *J Interferon Cytokine Res,* May 2 2024, doi: 10.1089/jir.2024.0032.

[69] E. Putscher *et al.*, "Genetic risk variants for multiple sclerosis are linked to differences in alternative pre-mRNA splicing," *Front Immunol,* vol. 13, p. 931831, 2022, doi: 10.3389/fimmu.2022.931831.

[70] K. Akiyama, S. Ichinose, A. Omori, Y. Sakurai, and H. Asou, "Study of expression of myelin basic proteins (MBPs) in developing rat brain using a novel antibody reacting with four major isoforms of MBP," *J Neurosci Res,* vol. 68, no. 1, pp. 19-28, Apr 1 2002, doi: 10.1002/jnr.10188.

[71] N. Kaushansky, M. Eisenstein, R. Zilkha-Falb, and A. Ben-Nun, "The myelin-associated oligodendrocytic basic protein (MOBP) as a relevant primary target autoantigen in multiple sclerosis," *Autoimmun Rev,* vol. 9, no. 4, pp. 233-6, Feb 2010, doi: 10.1016/j.autrev.2009.08.002.

[72] M. Podbielska, J. O'Keeffe, and A. Pokryszko-Dragan, "New Insights into Multiple Sclerosis Mechanisms: Lipids on the Track to Control Inflammation and Neurodegeneration," *Int J Mol Sci,* vol. 22, no. 14, Jul 7 2021, doi: 10.3390/ijms22147319.

[73] S. F. M. S. E. T. University of California *et al.*, "Silent progression in disease activity-free relapsing multiple sclerosis," *Ann Neurol,* vol. 85, no. 5, pp. 653-666, May 2019, doi: 10.1002/ana.25463.

[74] Z. Kouchi and M. Kojima, "A Structural Network Analysis of Neuronal ArhGAP21/23 Interactors by Computational Modeling," *ACS Omega,* vol. 8, no. 22, pp. 19249-19264, Jun 6 2023, doi: 10.1021/acsomega.2c08054.

[75] Z. Y. van Lierop *et al.*, "Serum contactin-1 as a biomarker of long-term disease progression in natalizumab-treated multiple sclerosis," *Mult Scler,* vol. 28, no. 1, pp. 102-110, Jan 2022, doi: 10.1177/13524585211010097.

[76] A. Vallejo-Illarramendi, M. Domercq, F. Perez-Cerda, R. Ravid, and C. Matute, "Increased expression and function of glutamate transporters in multiple sclerosis," *Neurobiol Dis,* vol. 21, no. 1, pp. 154-64, Jan 2006, doi: 10.1016/j.nbd.2005.06.017.

[77] L. Sheng, Q. Luo, and L. Chen, "Amino Acid Solute Carrier Transporters in Inflammation and Autoimmunity," *Drug Metab Dispos,* Feb 12 2022, doi: 10.1124/dmd.121.000705.

[78] R. Meech *et al.*, "A novel function for UDP glycosyltransferase 8: galactosidation of bile acids," *Mol Pharmacol,* vol. 87, no. 3, pp. 442-50, 2015, doi: 10.1124/mol.114.093823.

[79] C. A. Castellanos *et al.*, "Lymph node-resident dendritic cells drive T(H)2 cell development involving MARCH1," *Sci Immunol,* vol. 6, no. 64, p. eabh0707, Oct 15 2021, doi: 10.1126/sciimmunol.abh0707.

[80] J. Wu *et al.*, "The E3 ubiquitin ligase MARCH1 regulates antimalaria immunity through interferon signaling and T cell activation," *Proc Natl Acad Sci U S A,* vol. 117, no. 28, pp. 16567-16578, Jul 14 2020, doi: 10.1073/pnas.2004332117.

[81] X. G. Pappalardo, M. Ruggieri, R. Falsaperla, S. Savasta, U. Raucci, and P. Pavone, "A Novel 4q32.3 Deletion in a Child: Additional Signs and the Role of MARCH1," *J Pediatr Genet,* vol. 10, no. 4, pp. 259-265, Dec 2021, doi: 10.1055/s-0041-1736458.

[82] Z. Xu, J. Liu, Z. Liu, and H. Zhang, "MARCH1 as a novel immune-related prognostic biomarker that shapes an inflamed tumor microenvironment in lung adenocarcinoma," *Front Oncol,* vol. 12, p. 1008753, 2022, doi: 10.3389/fonc.2022.1008753.

[83] G. Laroia and R. J. Schneider, "Alternate exon insertion controls selective ubiquitination and degradation of different AUF1 protein isoforms," *Nucleic Acids Res,* vol. 30, no. 14, pp. 3052-8, Jul 15 2002, doi: 10.1093/nar/gkf444.

[84] R. M. Miller *et al.*, "Enhanced protein isoform characterization through long-read proteogenomics," *Genome Biol,* vol. 23, no. 1, p. 69, Mar 3 2022, doi: 10.1186/s13059-022-02624-y.

[85] P. Sinitcyn *et al.*, "Global detection of human variants and isoforms by deep proteome sequencing," *Nat Biotechnol,* vol. 41, no. 12, pp. 1776-1786, Dec 2023, doi: 10.1038/s41587-023-01714-x.

[86] J. van Langelaar, L. Rijvers, J. Smolders, and M. M. van Luijn, "B and T Cells Driving Multiple Sclerosis: Identity, Mechanisms and Potential Triggers," *Front Immunol,* vol. 11, p. 760, 2020, doi: 10.3389/fimmu.2020.00760.

[87] V. Nociti and M. Romozzi, "The Role of BDNF in Multiple Sclerosis Neuroinflammation," *Int J Mol Sci,* vol. 24, no. 9, May 8 2023, doi: 10.3390/ijms24098447.

[88] M. Kerschensteiner, C. Stadelmann, G. Dechant, H. Wekerle, and R. Hohlfeld, "Neurotrophic cross-talk between the nervous and immune systems: implications for neurological diseases," *Ann Neurol,* vol. 53, no. 3, pp. 292-304, Mar 2003, doi: 10.1002/ana.10446.

[89] R. B. Foltran and S. L. Diaz, "BDNF isoforms: a round trip ticket between neurogenesis and serotonin?," *J Neurochem,* vol. 138, no. 2, pp. 204-21, Jul 2016, doi: 10.1111/jnc.13658.

[90] A. K. Warrender, J. Pan, C. R. Pudney, V. L. Arcus, and W. Kelton, "Constant domain polymorphisms influence monoclonal antibody stability and dynamics," *Protein Sci,* vol. 32, no. 3, p. e4589, Mar 2023, doi: 10.1002/pro.4589.

[91] I. Lindeman *et al.*, "Stereotyped B-cell responses are linked to IgG constant region polymorphisms in multiple sclerosis," *Eur J Immunol,* vol. 52, no. 4, pp. 550-565, Apr 2022, doi: 10.1002/eji.202149576.

[92] A. Paul, M. Comabella, and R. Gandhi, "Biomarkers in Multiple Sclerosis," *Cold Spring Harb Perspect Med,* vol. 9, no. 3, Mar 1 2019, doi: 10.1101/cshperspect.a029058

Figure 1

Figure 2

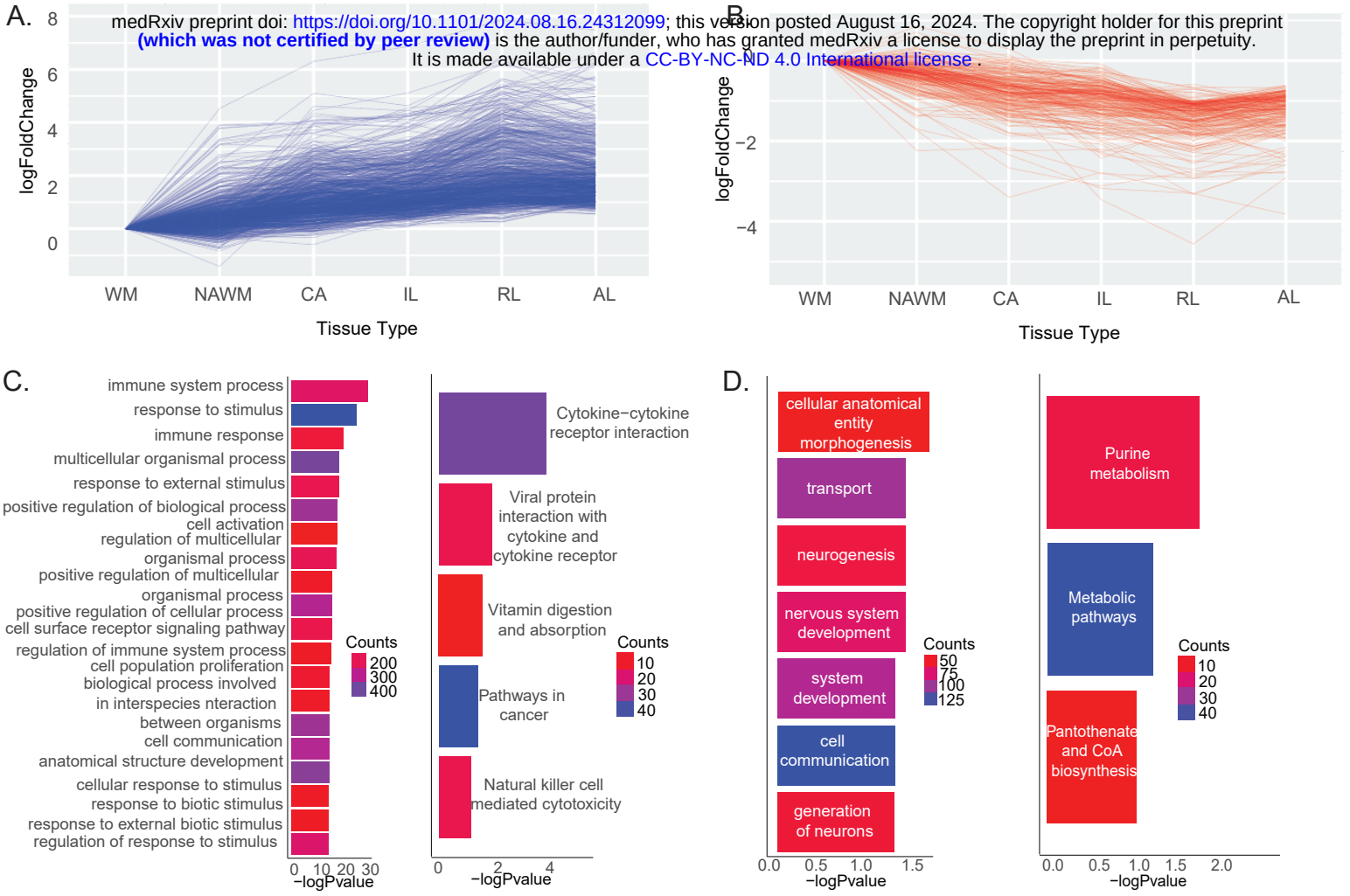

Figure 3

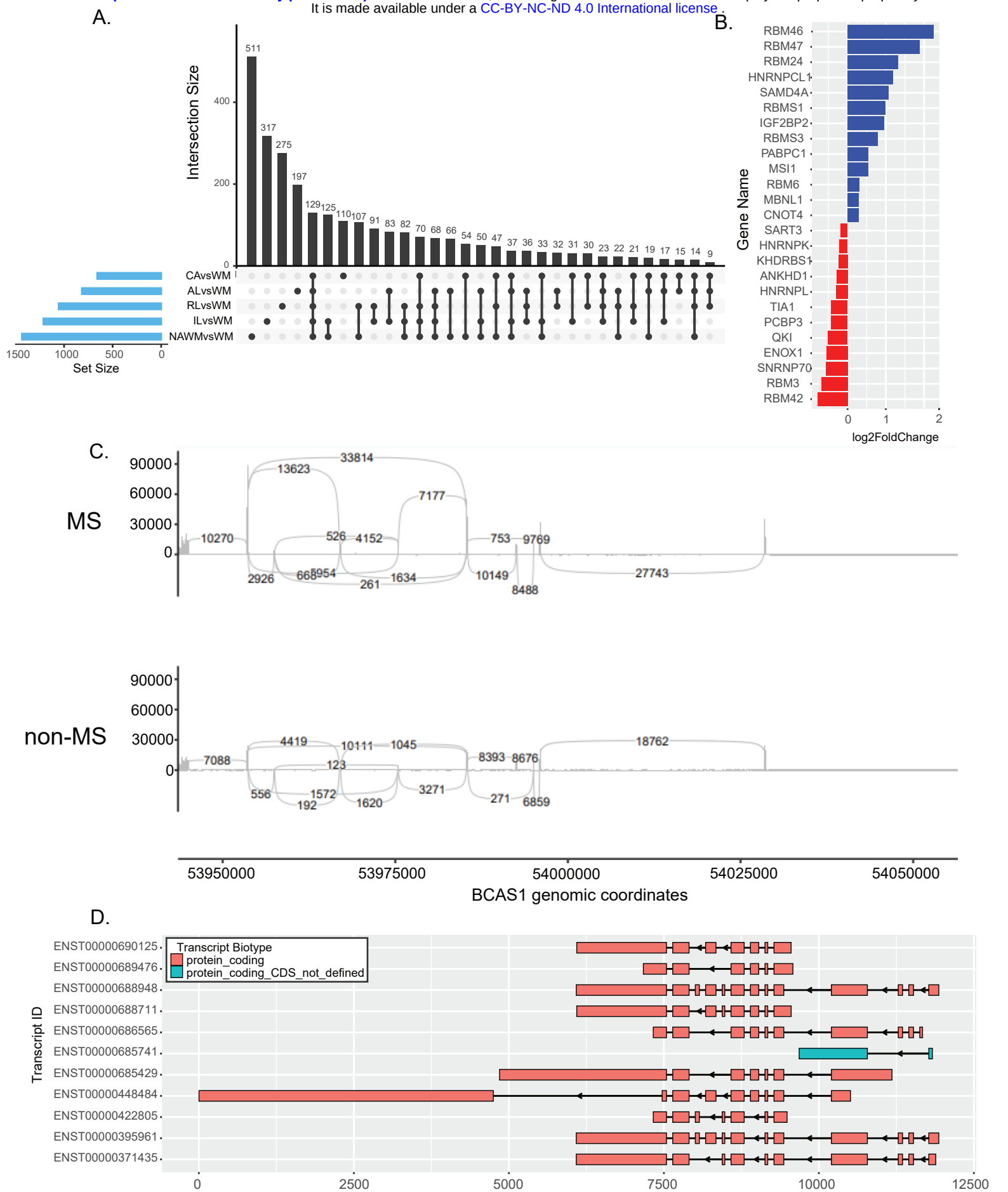

Figure 4

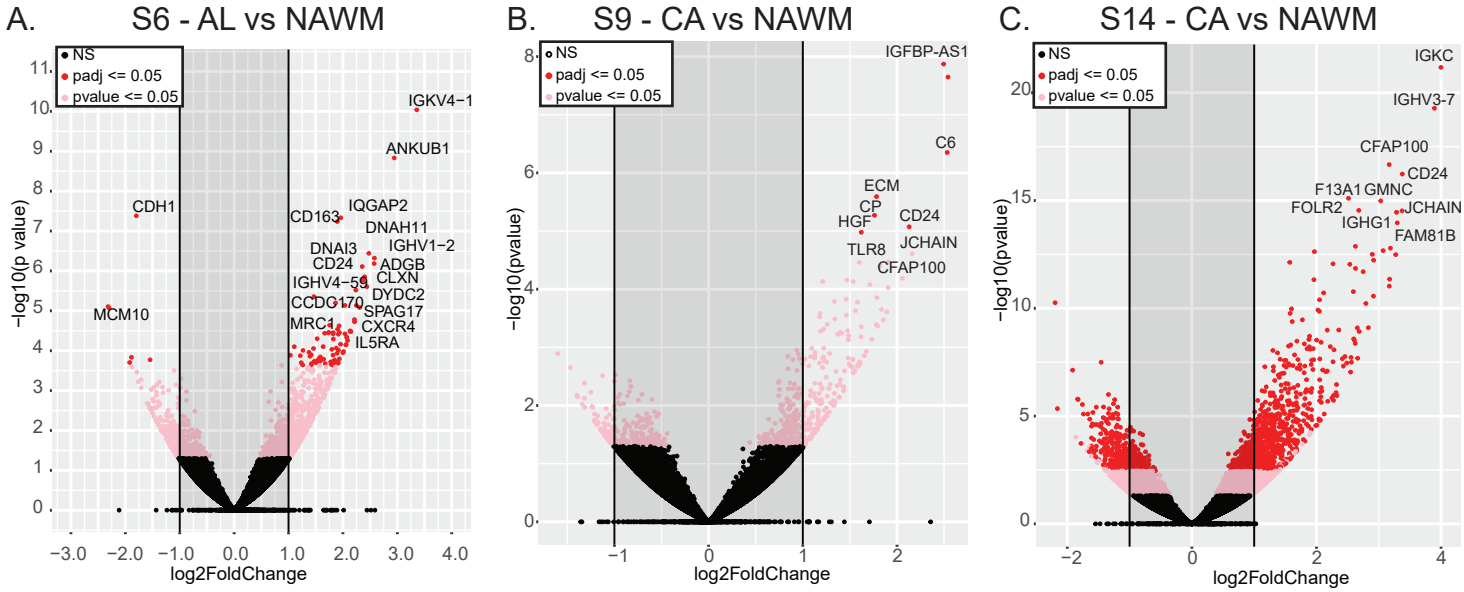

Figure 5