




ORIGINAL ARTICLE

Identification and characterization of a novel intronic splicing mutation in *CSF1R*-related leukoencephalopathy

Yilai Han¹ | Jinming Han¹  | Zhen Li¹ | Siqi Chen¹ | Ju Liu² | Ruxing Zhou¹ | Shufang Zhao¹ | Dawei Li¹ | Zheng Liu¹ | Yinan Zhao¹ | Junwei Hao^{1,3,4}  | Guoliang Chai^{1,3,5} 

¹Department of Neurology, Xuanwu Hospital Capital Medical University, National Center for Neurological Disorders, Beijing, China

²Department of Neurology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

³Beijing Municipal Geriatric Medical Research Center, Beijing, China

⁴Key Laboratory for Neurodegenerative Diseases of Ministry of Education, Beijing, China

⁵Chinese Institutes for Medical Research, Beijing, China

Correspondence

Guoliang Chai and Junwei Hao,
Department of Neurology, Xuanwu
Hospital Capital Medical University,
National Center for Neurological
Disorders, Beijing 100053, China.
Email: guoliang.chai@xwh.ccmu.edu.cn
and haojunwei@vip.163.com

Funding information

National Key Research and Development
Program of China, Grant/Award
Number: 2021YFA1101403; National
Natural Science Foundation of China,
Grant/Award Number: 82301526 and
82090043; Beijing Municipal Public
Welfare Development and Reform Pilot
Project for Medical Research Institutes,
Grant/Award Number: JYY2023-7;
Project for Innovation and Development
of Beijing Municipal Geriatric Medical
Research Center, Grant/Award Number:
11000023T000002041657; Youth Beijing
Scholar, Grant/Award Number: NO.020;
Project of Construction and Support for
high-level Innovative Teams of Beijing
Municipal Institutions, Grant/Award
Number: BPHR20220112

Abstract

Aims: Colony stimulating factor 1 receptor (*CSF1R*)-related leukoencephalopathy is a rapidly progressing neurodegenerative disease caused by *CSF1R* gene mutations. This study aimed to identify and investigate the effect of a novel intronic mutation (c.1754-3C>G) of *CSF1R* on splicing.

Methods: A novel intronic mutation was identified using whole-exome sequencing. To investigate the impact of this mutation, we employed various bioinformatics tools to analyze the transcription of the *CSF1R* gene and the three-dimensional structure of its encoded protein. Furthermore, reverse transcription polymerase chain reaction (RT-PCR) was performed to validate the findings.

Results: A novel mutation (c.1754-3C>G) in *CSF1R* was identified, which results in exon 13 skipping due to the disruption of the 3' splice site consensus sequence NYAG/G. This exon skipping event was further validated in the peripheral blood of the mutation carrier through RT-PCR and Sanger sequencing. Protein structure prediction indicated a disruption in the tyrosine kinase domain, with the truncated protein showing significant structural alterations.

Conclusions: Our findings underscore the importance of intronic mis-splicing mutations in the diagnosis and management of *CSF1R*-related leukoencephalopathy.

KEYWORDS

CSF1R-related leukoencephalopathy, exon skipping, intronic mutation, RNA splicing

The first two authors contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *CNS Neuroscience & Therapeutics* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Colony-stimulating factor 1 receptor (*CSF1R*)-related leukoencephalopathy is a rare progressive neurological disease caused by heterozygous mutations in the *CSF1R* gene.¹ Approximately 300 cases have been reported related to *CSF1R* gene mutations until the year 2022,^{2,3} while its genetic and phenotypic spectrum remains to be defined. The disease manifests with adult-onset cognitive decline, psychiatric symptoms, gait disturbance, and bradykinesia. Both familial and sporadic cases have been reported, exhibiting an autosomal dominant inheritance pattern with a high penetrance.⁴ Female patients tend to exhibit clinical symptoms earlier than males.⁵ Magnetic resonance imaging (MRI) features of *CSF1R*-related leukoencephalopathy include ventricular enlargement, cerebral atrophy, periventricular calcifications, and thinning of the corpus callosum.⁶ Typical neuropathologic findings include widespread white matter degeneration with a loss of myelin and axons, abundant neuroaxonal spheroids, and lipid-laden and pigmented macrophages.⁷

Microglia, the resident immune cells of the central nervous system (CNS), rely on *CSF1R* for their survival, maintenance, and proliferation. Thus, *CSF1R*-related leukoencephalopathy is now considered a primary CNS microgliopathy.⁸ The *CSF1R* gene encodes a ligand-dependent tyrosine kinase receptor that belongs to the platelet-derived growth factor (PDGF) receptor family, with *CSF1* and Interleukin-34 (IL-34) serving as its two ligands. *CSF1R* contains one intracellular tyrosine kinase domain (TKD), one transmembrane domain, and five immunoglobulin-like domains within its extracellular domain.⁸ Upon ligand binding, *CSF1R* undergoes dimerization and autophosphorylation, generating phosphotyrosine motifs that serve as docking sites for downstream effector pathways.⁹ Most previously reported variants lead to the inactivation of the tyrosine kinase and then reduced the *CSF1R* signaling.¹⁰ This molecular mechanism is in line with observations in patients with *CSF1R*-related leukoencephalopathy, where microglial cell numbers and dendritic arborizations are significantly reduced.

Pathogenic *CSF1R* gene mutations may cause disease through either haploinsufficiency or dominant-negative effects.^{11,12} Most previously reported *CSF1R* mutations were located in the TKD. Although *CSF1R* missense mutations lose their kinase activity and show autosomal dominant traits, some sporadic cases can be explained by incomplete penetrance and genetic mosaicism.¹³ The haploinsufficiency or dominant negative effect cannot fully explain penetrance. When a mutated allele coexists with a highly expressed wild-type allele, the latter is sufficient to ensure normal cell function. In both *CSF1R*-related leukoencephalopathy patients and the knock-in mouse model (*Csf1r*^{E631K/+}), there is a notable reduction of microglial cell numbers and dendritic arborizations.¹⁴ Microglia, but not neuronal reduction of *CSF1R* (*Cx3cr1-Cre*^{+/-}; *Csf1r*^{Flox/+}), is sufficient to induce adult-onset leukoencephalopathy in mice, further supporting a primary microgliopathy.

Pre-mRNA splicing, the process of removing non-coding introns, is critical for the expression of human genes and represents

a common post-transcriptional regulatory process in eukaryotes. Approximately 10%–62% of pathogenic single nucleotide variants affect alternative splicing processes by disrupting splicing sites.¹⁵ Accurate splicing requires the precise recognition of consensual donor and acceptor splice sites at the 5' and 3' ends of intron-exon junctions, respectively.¹⁶ Maintaining the structural integrity of splicing sites is essential for ensuring proper splicing. Specifically, the canonical 5' splice site contains the "GU" dinucleotide, which is recognized by U1 snRNP. On the other hand, the 3' splice site, which comprises the "AG" dinucleotide and the preceding polypyrimidine tract, can also be referred to as the NYAG/G rule.¹⁷ It is worth noting that a minority of splicing sites may not follow the NYAG splicing rule, such as NAGNAG.¹⁸ These elements are crucial for splicing and are associated with U2AF1 and U2AF65, respectively.^{19,20} Notably, the vast majority of pathogenic splicing mutations identified thus far are found at the 5' 'GU' or 3' 'AG' sites, within the ± 1 or ± 2 nucleotide range. This predominance may be attributed to the relatively straightforward predictability of the major impact of mutations on splicing. Apart from splicing site mutations, mutations occurring within intronic regions present a great challenge in predicting their effects on splicing.

Due to the rarity of *CSF1R*-related leukoencephalopathy, its pathogenic mechanisms remain unclear. In this study, we identified a novel intronic mutation, c.1754-3C>G, in the *CSF1R* gene within a family affected by *CSF1R*-related leukoencephalopathy. This mutation disrupts the 3' splice site NYAG/G consensus sequence, causing exon 13 skipping. Notably, while the skipping of exon 13 did not introduce a premature termination codon, it resulted in the truncation of intracellular TKD, thereby revealing a new pathogenic mechanism for splicing mutations in this disease. This discovery facilitates the monitoring of disease progression during the early stages.

2 | METHODS

2.1 | Subjects

A Chinese family, including one woman affected with *CSF1R*-related leukoencephalopathy and three family members, was comprehensively evaluated by two experienced neurologists. All individuals provided informed written consent for the genetic analysis before enrollment. The Ethics Committee of Xuanwu Hospital Capital Medical University approved this study.

2.2 | Whole-exome sequencing

Genomic DNA was isolated from the peripheral blood of the proband and three family members. All processes were conducted according to the manufacturer's protocols. DNA purity and integrity were assessed using spectrophotometry (Thermo Fisher Scientific, Waltham, MA, USA) and agarose gel electrophoresis, respectively. Subsequently, 1 μ g of DNA was broken into 250–300bp fragments

for terminal repair and joint connection using the Covaris ultrasonic instrument. Agilent SureSelect Human All Exon V6 was used to capture exons of amplified samples to prepare DNA nanospheres. Captured libraries were constructed according to the manufacturer's protocols. Sequencing was conducted using the Illumina HiSeq2000 sequencer (Illumina, San Diego, CA, USA) with a 100× sequencing coverage depth.

2.3 | Sequencing data processing and nucleotide variation calling

The comprehensive quality control software fastp was used to filter the low-quality reads and trim the sequencing adapter.²¹ Clean reads were aligned to reference genomes (hg19) using the BWA-MEM software (V0.7.17) and converted into bam format files.²² The high credibility short variants, including single nucleotide variations (SNVs) and indels, were conducted using the GATK4 Best Practices of Germline Short Variant Discovery. The argument MarkDuplicated was used to label repeated sequences generated by sequencing. The BaseRecalibrator and ApplyBQSR commands in GATK4 were used to recalibrate the base quality score. Subsequently, the mutation detection causing both mutation and non-mutation locations was shown in the genome variation call format (gVCF) files. Finally, the gVCF files from all samples were combined and preliminarily filtered with the screening criteria of min-mean DP >3 and minGQ >3.

2.4 | Variations annotation, filtering, and pathogenicity prediction

Based on the ACMG (American College of Medical Genetics) guidelines for variation classification, a variety of databases and tools were used for filtering and predicting potential pathogenicity. First, functional effects were predicted using the SnpEff toolbox. Combined annotation-dependent depletion (CADD) is a tool for scoring the deleteriousness of SNVs and insertion/deletion variants in the human genome. A CADD score (PHREAD >10), deleterious prediction by SIFT (sorting intolerant from tolerant), or probably/possibly damaging prediction by PolyPhen was the screening criteria for missense mutations. Second, we retained the variations that lead to loss-of-function (LoF) variation, termination mutation (stop-gain), termination loss (stop-loss), frame-shift insertions/deletions, canonical splice sites and initiation loss (start-loss). The residual variation intolerance score (RVIS) sequenced genes according to the level of intolerance to functional genetic variants. The top 50% was performed to further filter genetic variation. Third, reported SNPs were annotated according to the dbSNP. All variants identified in the proband were screened against the 1000 Genomes Project, gnomAD database, and OMIM database. The ClinVar database was used to filter allele frequency (AF), disease citations, and other in silico attributes. The variants with minor allele frequencies (MAF ≥0.1%) were excluded.

2.5 | Splicing site analysis

The UCSC Genome Browser (<https://genome.ucsc.edu/>) was used to obtain sequencing information of target fragments and transform mutant loci coordinates from hg19 to GRCh38. The splice site scores were evaluated using the Splice Prediction by Neural Network Site software (https://www.fruitfly.org/seq_tools/splice.html) to identify canonical, cryptic splicing sites between exons 12 and 14 in the CSF1R gene. SpliceAid2 (http://193.206.120.249/splicing_tissue.html) and Human Splicing Finder (HSF) version 2.4 (<http://www.umd.be/HSF/>) were utilized as web-based bioinformatics tools to investigate the potential impact of the mutation and identified the profile changes of the exon splicing enhancer (ESE), exon splicing silencer (ESS), canonical, and cryptic splicing sites in the variable region. EX-SKIP (<https://ex-skip.img.cas.cz/>) was used to predict the effect of the mutation on the transcription process and determine if the mutation caused an exon-skipping event. The Jalview 2.11.2.5 (<https://www.jalview.org/download/>) software was used for homologous nucleic acid/amino acid sequence alignment and visualization. The Open Reading Frame Finder (ORFfinder) in NCBI (<https://www.ncbi.nlm.nih.gov/orffinder/>) was used to identify the premature termination codons in mutated transcripts and amino acid sequences. AlphaFold2 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) was used to predict a 3D structure of the mutated CSF1R protein. The protein data bank (PDB) file of the wild-type CSF1R protein was obtained from the AlphaFold Structure Database (<https://alphafold.ebi.ac.uk/>). The 3D structures of mutant/wild-type CSF1R were visualized using PyMOL 2.2.3 (<https://pymol.org>).

2.6 | Mutation validation by Sanger sequencing

PCR and Sanger sequencing were performed to validate identified gene mutations in peripheral genome DNA from the proband and other family members. CSF1R-F: AACCCAAGTCCTACCCCT and CSF1R-R: TACCAGGTCCGCTGGAAGA were designed according to identified pathogenic mutation loci in the hg19.

2.7 | RNA isolation, cDNA synthesis, cDNA amplification, and sequencing

To immediately stabilize intracellular RNA, fresh blood samples from the proband's sister (carrying CSF1R gene mutation) and healthy control were collected and preserved in the PAXgene Blood RNA Tubes (BD Bioscience). The total RNA was isolated using PAXgene Blood RNA Kits and evaluated with the Qubit 2100. Genomic DNA was removed from total RNA and then reversed transcription of RNA into cDNA using the HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper, Vazyme). The cDNA was amplified with primers (F: CGAGAGCTATGAGGGCAACAG, R: GTCATGATCTTCAGCTCGGAC) spanning from exons 12–14. The

exon 12–14 cDNA region was amplified by PCR, and the levels of different transcripts were compared in RT-PCR products through Sanger sequencing.

3 | RESULTS

3.1 | Clinical presentation and exome analysis

A 44-year-old Chinese woman was admitted to a local hospital due to cognitive decline and right upper extremity weakness for 7 months. She exhibited impaired social communication skills, characterized by blunted responses and unclear expressions. Approximately 2 months later, right lower extremity weakness was noted in this patient. Antiplatelet drugs and statins were administered while her clinical symptoms did not show significant improvement. Upon admission, neurological examinations revealed tremors (right-sided obviousness) and abnormal cerebellar signs without disability. The patient's father and aunt had experienced progressive walking difficulties and had passed away (Figure 1A). No remarkable clinical symptoms were reported in the patient's child, brother, and sister.

Routine laboratory investigations indicated the presence of hyperlipidemia, while renal and liver functions appeared to be within the normal level. Tests for arylsulphatase, metachromatic granules, and very long-chain fatty acids yielded negative results. Cerebrospinal fluid (CSF) analysis showed normal cell counts and protein levels. Brain axial T2-weighted and T2-weighted fluid-attenuated inversion recovery (FLAIR) images revealed periventricular white

matter hyperintensities and abnormal signals on bilateral frontoparietal subcortex and semi-oval centers (Figure 1B). These lesions showed low signals on T1-weighted image, without enhancement. Additionally, some brain lesions showed diffusion restriction on diffusion-weighted imaging (DWI).

We performed whole-exome sequencing on the proband and identified a novel variant (c.1754-3C>G) in the *CSF1R* gene. To analyze the significance of this mutation, we conducted comprehensive in silico analyses, including CADD, PolyPhen, and SIFT. To further validate our findings and investigate the presence of this SNV in the family, we employed Sanger sequencing (Figure 1C). Genetic testing could not be performed on her father, aunt, and grandfather since they had passed away. Among tested individuals, one sibling (III-4) was found to carry the same *CSF1R* gene mutation, while remaining asymptomatic at the age of 49. Conversely, another sibling (III-1) and her cousin (III-10) were clinically unaffected and did not carry *CSF1R* gene mutations.

3.2 | Molecular mechanisms of the splicing procedure

The consensus sequence for the acceptor site adheres to the pattern NYAG/G (N, any nucleotide; Y, pyrimidine; “/,” inline-exon boundary). Across various species, the mutated site within *CSF1R* follows this rule, featuring either cytosine or thymine (Figure 2). Therefore, we predicted that the c.1754-3C>G variant may disrupt the consensus acceptor site and subsequently impact the splicing

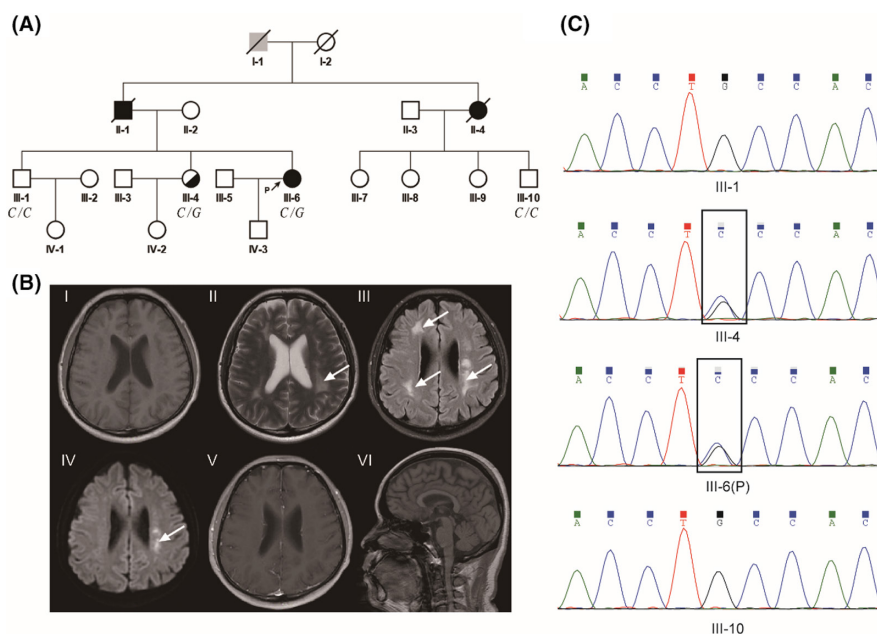


FIGURE 1 Genetic and MRI findings in a family with *CSF1R*-related leukoencephalopathy. (A) The family pedigree spans four generations and includes 19 individuals. The proband (III-6) is highlighted with an arrow. Solid black symbols denote affected members, while half-filled symbols indicate carriers. (B) I: T1-weighted axial image; II: T2-weighted axial image; III: Fluid-attenuated inversion recovery (FLAIR) image; IV: Diffusion-weighted imaging (DWI) image; V: Post-contrast T1-weighted image; VI: T1-weighted sagittal image. (C) Sanger sequencing chromatograms are shown for the proband (III-6), proband's sister (III-4), another sibling (III-1), and a cousin (III-10). Sequences framed in black highlight the carriers of the c.1754-3C>G mutation in the *CSF1R* gene.

of *CSF1R*. To assess the potential impact of this mutation on RNA splicing, we conducted a series of in silico analyses. We first employed a neural network algorithm, which revealed that the nearest acceptor site of this mutation was not present in the allele (Table S1). Furthermore, we utilized HSF ver. 3.1 to evaluate the consequences of the c.1754-3C>G mutation, including its potential effects on the distribution of canonical and cryptic splicing sites, as well as exonic splicing enhancers (ESE) and exonic splicing silencers (ESS) within the region spanning exons 12 and 14. According to the MaxEnt model, the mutation likely affects splicing, as indicated by a notable low score (6.61) at the nearest acceptor site of the mutation (Table S2). In addition to the disruption of the acceptor site, we identified two compromised ESE sites and one ESS site, which may collectively lead to a detrimental splicing event between exons 12 and 13 (Table S2). This mutation transforms the RNA binding sequence from "GCAGG/CAGGUAAGAC" to "UGGGA," which may impact its interaction with specific RNA-binding proteins (Figure S1). Consequently, these changes may affect the binding site of the RNA-binding proteins (RBP), shifting from a splicing-enhancing RBP (SF2/ASF) to a splicing-suppressing RBP (hnRNP H1/F; Tables S3 and S4, Figure S1). In support of this, we further showed that the mutation

allele had a higher chance of exon skipping than the wild-type allele through the analysis of EX-SKIP.

3.3 | Identification of splicing alteration

To further validate the splicing alteration, we aimed to conduct RT-PCR from the peripheral blood cells expressing *CSF1R*. The patient had previously undergone an allogeneic hematopoietic stem cell transplant (HSCT) after diagnosis. Therefore, we decided to assess the impact of the mutation on her sister, who is an asymptomatic carrier of this mutation. We performed RT-PCR and Sanger sequencing to analyze the splicing pattern between exons 12 and 14 in the patient's sister and a healthy control (Figures 3B and S2). Our results revealed one amplification product in the healthy control and two amplification products in the mutation carrier (Figure 3A). Upon sequencing these amplification products, we observed that the longer product (256bp) from the carrier matched the sequence from the healthy control containing the product with exons 12, 13, and 14 (Figure 3B). In contrast, the shorter product (151bp) contained an exon skipping event involving exon 13.

	Intron12	Junction	Exon13
Affected	a a g g g a c c t g t g t g c g t g g	g a g	G T A A G A C C C T C G G A G C
Human	a a g g g a c c t g t g t g c g t g g	c a g	G T A A G A C C C T C G G A G C
Rhesus	a a g g g a c c t g t g t g c g t g g	c a g	G T A A G A C C C T C G G A G C
Mouse	a a g g a g c c c a t g t g t g t g a t	a g	G T A A G A C T C T A G G A G C
Dog	a a g t a g g c t g t g t g t g t c a	c a g	G G A A G A C T C T C G G A G C
Elephant	a a g g a g t c t g t g t a t g t g g t	a g	G T A A G A C C C T T G G A G C
Chicken	- - - - -	c a g	G A A A A A C T C T T G G A G C
Zebrafish	- - - - - g t g t g t g t g -	t a g	G G A A G A T C C T G G G T G C

c.1754-3C>G

Purine Pyrimidine

FIGURE 2 A conservative analysis is conducted on the c.1754-3C>G mutation carrier in humans and other species. The gray frame indicates the disrupted NYAG/G conserved sequence in the mutation carrier. The red box represents Purine, and the blue signifies Pyrimidine.

3.4 | Impact of the patient variant on *CSF1R* protein

We further evaluated the impact of exon 13 skipping on *CSF1R* protein. Exon 13 has a length of 105bp, and its absence may lead to a premature stop or truncation of the protein. We performed an ORF finder analysis and our results showed that pretermination codons are not present in the mutant cDNA sequence, indicating that this mutation is likely to cause a truncated *CSF1R* protein (Figure 4A). Upon comparing the protein sequences of the wild-type and truncated *CSF1R*, we noted that the mutant isoform lacked 35 amino acids and introduced a nonsynonymous mutation, changing lysine to

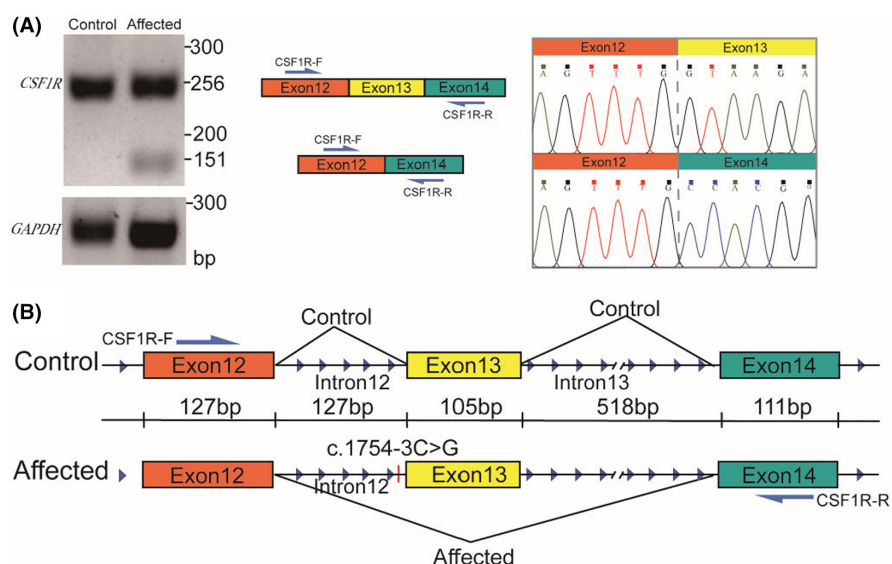


FIGURE 3 The nucleotide variation c.1754-3C>G leads to the skipping of exon 13. (A) The RT-PCR and Sanger results are presented for the c.1754-3C>G mutation carrier (Affected) and a healthy control (Control). (B) The splicing patterns are depicted for Control and Affected samples between exon 12 and exon 14.

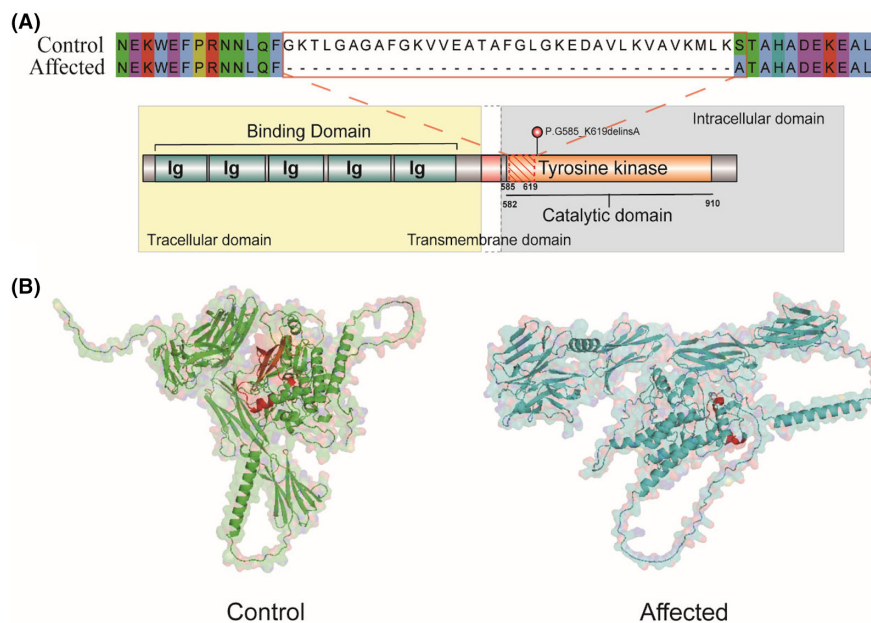


FIGURE 4 The exon skipping disrupts the TKD, resulting in a truncated protein. (A) A schematic illustration of the CSF1R protein is provided. The section highlighted with a red shadow depicts the truncated region of the TKD as a result of the mutation. The red frame showcases a comparison of amino acid sequences between the affected and control protein. (B) The 3D structure of CSF1R is displayed. In the overview of the protein structure, the red region marks the truncated section of the mutation-affected protein.

alanine (P.G585_K619delinsA). A critical region situated within the conserved tyrosine kinase domain is missing, markedly impairing the enzyme's catalytic activity (Figure 4A). Subsequent analyses demonstrated that this region, inclusive of exon 13 and adjacent sequences, displays significant conservation across a variety of species. These findings highlight the indispensable role of this region in sustaining the normal function of CSF1R protein (Figure S3). We further analyzed the predicted CSF1R structure using the AlphaFold model, suggesting that the affected peptide was located within a segment of the alpha helix and β -turn angle (Figure 4B). Additionally, the structure of the truncated CSF1R isoform exhibited significant alterations. These findings collectively emphasize the potential functional consequences of this mutation on CSF1R protein.

4 | DISCUSSION

In this study, we reported a novel intronic mis-splicing mutation in the *CSF1R* gene associated with *CSF1R*-related leukoencephalopathy. This mutation has not been previously documented in the literature. Interestingly, this mutation represents a non-canonical splicing variant that does not directly affect the AG dinucleotide splicing acceptor site, and as such, might be overlooked by ACMG guidelines.^{15,23} However, the c.1754-3C>G mutation induces the skipping of exon 13 by perturbing the NYAG/G sequence rule, leading to the production of a truncated CSF1R protein within the TKD. Identifying the consequences of non-canonical splicing variants at the RNA level is beneficial for the diagnosis of genetic diseases.

The *CSF1R* gene plays an essential role in microglial survival. Loss of the integrity of TKD within the *CSF1R* can cause dysfunctional CSF1R activity and microglial functions.²⁴ *CSF1R*-related leukoencephalopathy is characterized by a predominant involvement of microglia, classifying as microgliopathy.^{10,25} Approximately 95% of previously reported *CSF1R* variants linked to leukoencephalopathy are located

within the TKD, leading to the impairment of autophosphorylation. In vitro experiments have shown that cells expressing mutant *CSF1R* fail to exhibit ligand-induced autophosphorylation.^{20,26-28} Among identified mutations, over 100 missense mutations have been reported in the *CSF1R* gene.^{4,27} Compared with missense mutations, splicing mutations in *CSF1R* gene were relatively rare. To date, 17 splicing site mutations have been identified related to *CSF1R*-related leukoencephalopathy (summarized in Table 1). Among them, four mutations occur in intron 12, six in intron 18, and three in introns 13, 20, and one in 17, respectively. Similar to c.1754-3C>G, previously reported splicing mutations affected the regions located in the kinase domain. Within the identified variation, 10 mutations adjoin the donor site, with more than six in the acceptor site and a ratio of 1.5:1, which is similar to the ratio of reported splicing mutations in HGMD (Human Gene Mutation Database).^{29,30}

Our study identified a novel pathogenic mutation in *CSF1R* (c.1754-3C>G) that damages the acceptor site of intron 12 and causes exon 13 skips. The absence of exon 13 disrupts the function of *CSF1R* by deleting up to 40 consecutive amino acids within the TKD. This could result in the generation of enzymatically inactive receptors, potentially causing aberrant microglial function via either haploinsufficiency or a dominant-negative mechanism. Some previously reported *CSF1R* mutations were noted in different loci near exon 13.^{8,12,31} For example, Rosa et al. identified a splicing acceptor site mutation in *CSF1R* (c.1754-2A>G) in monozygotic twins, who died of *CSF1R*-related leukoencephalopathy at 40 and 41 years of age, respectively.⁸ The skipping of exon 13 was identified in the transcript from *CSF1R* in the mutation carriers.⁸ Nynke and colleagues reported a homozygous *CSF1R* splicing site mutation (c.1754-1G>C) in a patient with pediatric-onset leukoencephalopathy.³¹ This mutation can also cause exon 13 skipping. Interestingly, her parents are heterozygous carriers and did not display any clinical symptoms. Another *CSF1R* gene mutation (c.1858+1G>T) in a family with *CSF1R*-related leukoencephalopathy was located in a splicing

TABLE 1 Summary of the CSF1R splicing site mutations.

ID	Variant	Location	Exon	Familial (F)/Sporadic (S) Disease	Trait	Reported by
1	c.1754-10T>A	Intron 12	12_13	S	NA	Wu X et al. (2022)
2	c.1754-3C>G	Intron 12	12_13	F (Father, Aunt)	Incomplete penetrance	Novel Mutation
3	c.1754-2A>G	Intron 12	12_13	F (Brother)	Complete penetrance	Rademakers et al. (2011)
4	c.1754-1G>C	Intron 12	12_13	F	NA	Oosterhof et al. (2019)
5	c.1858+1G>T	Intron 13	13_14	F (Mother, Brother, sister)	Complete penetrance	Yang et al. (2019)
6	c.1858+5G>A	Intron 13	13_14	F (grandma, father, uncle, brother, sister, son)	Complete penetrance	Jiang et al. (2022)
7	c. 1859-119G>A	Intron 13	13_14	F (grandma, father, uncle, brother, sister, son)	Incomplete penetrance	Guo et al. (2019)
8	c.2320-2A>G	Intron 17	17_18	F (Mother)	Complete penetrance	Rademakers et al. (2011)
9	c.2442+1G>T	Intron 18	18_19	S	NA	Konno et al. (2014)
10	c.2442+1G>A	Intron 18	18_19	S	NA	Lakshmanan et al. (2017)
11	c.2442+2T>C	Intron 18	18_19	F (Mother, Brother, sister)	Incomplete penetrance	YaYang et al. (2019)
12	c. 2442+2_2442+3 dupT	Intron 18	18_19	Unknown	NA	Chu et al. (2021)
13	c.2442+5G>C	Intron 18	18_19	S	NA	Konno et al. (2017)
14	c.2442+5G>A	Intron 18	18_19	F (Father)	Complete penetrance	Konno et al. (2017)
15	c.2655-2A>G	Intron 20	20_21	Unknown	NA	Guerreiro et al. (2013)
16	c.2654+1G>A	Intron 20	20_21	F (Grandma, Mother)	Complete penetrance	Jiang et al. (2023)
17	c.2654_2654+3del	Intron 20	20_21	F (Father, Brother)	Complete penetrance	Xie et al. (2019)

Note: The bold text indicates the novel mutation identified in this study.

donor site leading to exon 13 skipping from *CSF1R* mRNA.¹² Notably, all these previously reported mutations that cause exon 13 skipping disrupt the canonical GU-AG splicing site. The c.1754-1G>C and c.1754-2A>G mutants of *CSF1R* disrupted the splicing acceptor site AG of intron 12, while the c.1858+1G>T destroyed the splicing donor site GU of intron 13.^{8,12,31} According to the ACMG/AMP variant classification guidelines, the disruption of highly conserved GT-AG splicing motif (± 1 or 2 splice sites) is considered as strong evidence for variants' pathogenicity.²³ However, some SNVs introduce missense/nonsense codons, while others affect auxiliary splicing cis-elements or generate cryptic GT-AG dinucleotides.¹⁹ Therefore, some mutations may cause one part of the intron to be erroneously spliced into mature transcripts as if it were an exon. Mutation (C.1858+5G>A) disrupts the consensus sequence of the 5' ss (from CAG/GUAAGUAAU to CAG/GUAAAUAAU) and breaks the complementary strand to the AUACUUACCUG sequence of U1 snRNA.^{28,32} It also influences the construction of the E1 complex and causes an exon 13-skipping event and retention of intron 12 in the *CSF1R* gene. In our study, the mutation (c.1754-3C>G) affected the splicing procedure by breaking the NYAG/G rule near the splicing acceptor site, with the (c.1754-3C>G) being G at intron-3 (G at Int-3). "G at Int-3" is frequently observed in exons that alternatively skip in the human genome.³³ The effect of "G at Int-3" was further identified through cryo-electron microscopy and isothermal titration calorimetry,

showing that "G at Int-3" decreases a binding affinity for U2AF1.³⁴ The changes of NYAG/G due to the conversion of conserved pyrimidines into purines are the main causes for exon 13 skipping. The RBP binding site changing from SRSF1 to hnRNPH may cause exon 13 to skip in the *CSF1R* gene. SRSF1 binding prevents the skipping exon according to the uniprot, and the hnRNP induces the exon skipping event. The antagonistic regulation of SRSF1 and hnRNP for *COLQ* exon 16 was part of pathological processes of congenital myasthenic syndrome.³⁵

Homozygous *Csf1r* knockout (*Csf1r*^{-/-}) mice exhibited global defects of brain development and then died perinatally.^{36,37} Heterozygous (*Csf1r*^{+/-}) mice mimicking *CSF1R*-related leukoencephalopathy symptoms due to *CSF1R* mutations showed cognitive and sensorimotor deficits.³⁸ MRI revealed enlarged lateral ventricles and a thinner corpus callosum. *CSF1R* is mainly expressed in microglia and cortical layer V neurons, with their expression being declined in *Csf1r*^{+/-} mice with age. Behavioral and pathological traits, along with elevated *Csf2* expression in *Csf1r*^{+/-} mice, were replicated by microglial *Csf1r* heterozygous deletion (MCsf1r^{het}) rather than neural deletion.³⁹ Increased *Csf2* expression was observed in various cells in *Csf1r*^{+/-} mice, with microglial activation evident in MCsf1r^{het} mice.³⁸ The consistent rise in microglial density indicates that *CSF1R*-related leukoencephalopathy primarily arises from microglial abnormalities. Lesions in patients with *CSF1R*-related leukoencephalopathy are

located in the frontal and parietal lobes, periventricular, and deep white matter.⁶ In the diagnostic criteria, the core imaging features are bilateral cerebral white matter lesions and corpus callosum thinning,⁴⁰ which is associated with morphological changes in microglia.⁴¹ Persistent diffusion restriction can frequently be observed, which may distinguish CSF1R-related leukoencephalopathy from other demyelinating diseases.

The dominant negative effect and haploinsufficiency are genetic dominance forms in some CSF1R-related leukoencephalopathy families.^{11,27} The former reduces the formation of active CSF-CSF1R dimer on the cell surface through binding between the wild type and mutant CSF1R.¹³ Haploinsufficiency, where a single functional copy of a gene is insufficient for normal cellular function, has been implicated in dominantly inherited diseases.⁴² Several reports have confirmed that the severity of phenotypes is negatively correlated with the dosage of CSF1R.⁴³ In patients, nonsense or frameshift mutations, as well as some splice site mutations are subjected to nonsense-mediated mRNA decay.⁴⁴ The haploinsufficiency can also be noted in the mouse model, with a heterozygous genotype developing CSF1R-related leukoencephalopathy-like symptoms.^{38,39} In our study, the absence of exon 13 may fail to induce nonsense-mediated decay because premature termination codons could not be identified. The production of truncated CSF1R due to c.1754-3C>G mutation may reduce the number of active cell surface CSF1R dimer and then lead to the disease.

In our pedigree, CSF1R-related leukoencephalopathy was consistent with the dominant inheritance pattern, while one carrier (the elder sister, 5 years older than the proband) displayed no clinical symptoms yet. The family declined to perform whole-exome sequencing for their children. Incomplete penetrance has also been observed in other reported pedigrees, including those with a splicing mutation causing exon 13 skipping.^{8,12,31} Factors such as genetic modifiers, epigenetic modifications, or environmental exposures may potentially explain incomplete penetrance observed in this disease. Further research is required to figure out the intricate interplay of these factors in determining the clinical manifestation of CSF1R-related leukoencephalopathy.

AUTHOR CONTRIBUTIONS

YH and JH have made equal contributions to this study. They were responsible for bioinformatics analyses, and drafting and revising the manuscript. ZL verified the nucleotide mutations and exon skipping using PCR and Sanger sequencing. SC revised the figures. JL collected samples and gathered clinical data. RZ and SZ analyzed the protein structure. YZ collected the published mutation. DL and ZL gathered clinical data and provided recommendations. JH and GC spearheaded the design and coordination of the study, offered critical revisions to the manuscript, and ensured the accuracy of the data analysis.

ACKNOWLEDGMENTS

The authors express gratitude to the patients and their families for their invaluable contribution to this study. The authors thank for the

support of the Chinese Institute for Brain Research, Beijing. We are grateful to the Neuroimmune Center team at Xuanwu Hospital for their dedicated laboratory support.

FUNDING INFORMATION

This study was supported by the National Key Research and Development Program of China (2021YFA1101403), the National Natural Science Foundation of China (82301526 and 82090043), Beijing Municipal Public Welfare Development and Reform Pilot Project for Medical Research Institutes (JYY2023-7), the Project for Innovation and Development of Beijing Municipal Geriatric Medical Research Center (11000023T000002041657), the Youth Beijing Scholar (NO. 020), and the Project of Construction and Support for high-level Innovative Teams of Beijing Municipal Institutions (BPHR20220112).

CONFLICT OF INTEREST STATEMENT

The authors affirm that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CONSENT FOR PUBLICATION

Not applicable.

ORCID

Jinming Han  <https://orcid.org/0000-0002-6084-3275>

Junwei Hao  <https://orcid.org/0000-0002-3075-5434>

Guoliang Chai  <https://orcid.org/0000-0001-9658-6381>

REFERENCES

1. Rosenstein I, Andersen O, Victor D, et al. Four Swedish cases of CSF1R-related leukoencephalopathy: visualization of clinical phenotypes. *Acta Neurol Scand.* 2022;145(5):599-609.
2. Wu J, Cheng X, Ji D, et al. The phenotypic and genotypic spectrum of CSF1R-related disorder in China. *Mov Disord.* 2024;39:798-813.
3. Papapetropoulos S, Gelfand JM, Konno T, et al. Clinical presentation and diagnosis of adult-onset leukoencephalopathy with axonal spheroids and pigmented glia: a literature analysis of case studies. *Front Neurol.* 2024;15:1320663.
4. Tian WT, Zhan FX, Liu Q, et al. Clinicopathologic characterization and abnormal autophagy of CSF1R-related leukoencephalopathy. *Transl Neurodegener.* 2019;8(1):32.
5. Konno T, Yoshida K, Mizuno T, et al. Clinical and genetic characterization of adult-onset leukoencephalopathy with axonal spheroids and pigmented glia associated with CSF1R mutation. *Eur J Neurol.* 2017;24(1):37-45.
6. Mickeviciute GC, Valiuskyte M, Plattén M, et al. Neuroimaging phenotypes of CSF1R-related leukoencephalopathy: systematic review, meta-analysis, and imaging recommendations. *J Intern Med.* 2022;291(3):269-282.
7. Axelsson R, Røyttä M, Sourander P, Akesson HO, Andersen O. Hereditary diffuse leukoencephalopathy with spheroids. *Acta Psychiatr Scand Suppl.* 1984;314:1-65.
8. Rademakers R, Baker M, Nicholson AM, et al. Mutations in the colony stimulating factor 1 receptor (CSF1R) gene cause hereditary diffuse leukoencephalopathy with spheroids. *Nat Genet.* 2012;44(2):200-205.

9. Han J, Chitu V, Stanley ER, Wszolek ZK, Karrenbauer VD, Harris RA. Inhibition of colony stimulating factor-1 receptor (CSF-1R) as a potential therapeutic strategy for neurodegenerative diseases: opportunities and challenges. *Cell Mol Life Sci*. 2022;79(4):219.
10. Konno T, Kasanuki K, Ikeuchi T, Dickson DW, Wszolek ZK. CSF1R-related leukoencephalopathy: a major player in primary microgliopathies. *Neurology*. 2018;91(24):1092-1104.
11. Guo L, Ikegawa S. From HDLS to BANDDOS: fast-expanding phenotypic spectrum of disorders caused by mutations in CSF1R. *J Hum Genet*. 2021;66(12):1139-1144.
12. Yang X, Huang P, Tan Y, Xiao Q. A novel splicing mutation in the CSF1R gene in a family with hereditary diffuse leukoencephalopathy with axonal spheroids. *Front Genet*. 2019;10:491.
13. Pridans C, Sauter KA, Baer K, Kissel H, Hume DA. CSF1R mutations in hereditary diffuse leukoencephalopathy with spheroids are loss of function. *Sci Rep*. 2013;3(1):3013.
14. Stables J, Green EK, Sehgal A, et al. A kinase-dead Csf1r mutation associated with adult-onset leukoencephalopathy has a dominant inhibitory impact on CSF1R signalling. *Development*. 2022;149(8):dev200237.
15. Li K, Xiao J, Ling Z, et al. Prioritizing de novo potential non-canonical splicing variants in neurodevelopmental disorders. *EBioMedicine*. 2024;99:104928.
16. Hujová P, Grodecká L, Souček P, Freiberger T. Impact of acceptor splice site NAGTAG motif on exon recognition. *Mol Biol Rep*. 2019;46(3):2877-2884.
17. Zhang MQ. Statistical features of human exons and their flanking regions. *Hum Mol Genet*. 1998;7(5):919-932.
18. Hiller M, Huse K, Szafranski K, et al. Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. *Nat Genet*. 2004;36(12):1255-1257.
19. Ohno K, Takeda JI, Masuda A. Rules and tools to predict the splicing effects of exonic and intronic mutations. *Wiley Interdiscip Rev RNA*. 2018;9(1):e1451.
20. Wu S, Romfo CM, Nilsen TW, Green MR. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*. 1999;402(6763):832-835.
21. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890.
22. Jung Y, Han D. BWA-MEME: BWA-MEM emulated with a machine learning approach. *Bioinformatics*. 2022;38(9):2404-2413.
23. Richards S, Aziz N, Bale S, 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):405-424.
24. Green KN, Crapser JD, Hohsfield LA. To kill a microglia: a case for CSF1R inhibitors. *Trends Immunol*. 2020;41(9):771-784.
25. Wang YL, Wang FZ, Li R, Jiang J, Liu X, Xu J. Recent advances in basic research for CSF1R-microglial encephalopathy. *Front Aging Neurosci*. 2021;13:792840.
26. Tsai PC, Fuh JL, Yang CC, et al. Clinical and genetic characterization of adult-onset leukoencephalopathy caused by CSF1R mutations. *Ann Clin Transl Neurol*. 2021;8(11):2121-2131.
27. Hu B, Duan S, Wang Z, et al. Insights into the role of CSF1R in the central nervous system and neurological disorders. *Front Aging Neurosci*. 2021;13:789834.
28. Roca X, Krainer AR. Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. *Nat Struct Mol Biol*. 2009;16(2):176-182.
29. Krawczak M, Thomas NST, Hundrieser B, et al. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat*. 2007;28(2):150-158.
30. Pros E, Gómez C, Martín T, Fábregas P, Serra E, Lázaro C. Nature and mRNA effect of 282 different NF1 point mutations: focus on splicing alterations. *Hum Mutat*. 2008;29(9):E173-E193.
31. Oosterhof N, Chang IJ, Karimiani EG, et al. Homozygous mutations in CSF1R cause a pediatric-onset leukoencephalopathy and can result in congenital absence of microglia. *Am J Hum Genet*. 2019;104(5):936-947.
32. Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA. Are snRNPs involved in splicing? *Nature*. 1980;283(5743):220-224.
33. Shibata A, Okuno T, Rahman MA, et al. IntSplice: prediction of the splicing consequences of intronic single-nucleotide variations in the human genome. *J Hum Genet*. 2016;61(7):633-640.
34. Yoshida H, Park SY, Sakashita G, et al. Elucidation of the aberrant 3' splice site selection by cancer-associated mutations on the U2AF1. *Nat Commun*. 2020;11(1):4744.
35. Rahman MA, Azuma Y, Nasrin F, et al. SRSF1 and hnRNP H antagonistically regulate splicing of COLQ exon 16 in a congenital myasthenic syndrome. *Sci Rep*. 2015;5:13208.
36. Chitu V, Stanley ER. Regulation of embryonic and postnatal development by the CSF-1 receptor. *Curr Top Dev Biol*. 2017;123:229-275.
37. Dai XM, Ryan GR, Hapel AJ, et al. Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects. *Blood*. 2002;99(1):111-120.
38. Biundo F, Chitu V, Shlager GGL, et al. Microglial reduction of colony stimulating factor-1 receptor expression is sufficient to confer adult onset leukodystrophy. *Glia*. 2021;69(3):779-791.
39. Chitu V, Gokhan S, Gulinello M, et al. Phenotypic characterization of a Csf1r haploinsufficient mouse model of adult-onset leukodystrophy with axonal spheroids and pigmented glia (ALSP). *Neurobiol Dis*. 2015;74:219-228.
40. Konno T, Yoshida K, Mizuta I, et al. Diagnostic criteria for adult-onset leukoencephalopathy with axonal spheroids and pigmented glia due to CSF1R mutation. *Eur J Neurol*. 2018;25(1):142-147.
41. Kinoshita M, Oyanagi K, Kondo Y, et al. Pathologic basis of the preferential thinning of the corpus callosum in adult-onset leukoencephalopathy with axonal spheroids and pigmented glia (ALSP). *eNeurologicalSci*. 2021;22:100310.
42. Johnson AF, Nguyen HT, Veitia RA. Causes and effects of haploinsufficiency. *Biol Rev*. 2019;94(5):1774-1785.
43. Arreola MA, Soni N, Crapser JD, et al. Microglial dyshomeostasis drives perineuronal net and synaptic loss in a CSF1R^{+/-} mouse model of ALSP, which can be rescued via CSF1R inhibitors. *Sci Adv*. 2021;7(35):eabg1601.
44. Konno T, Tada M, Tada M, et al. Haploinsufficiency of CSF-1R and clinicopathologic characterization in patients with HDLS. *Neurology*. 2014;82(2):139-148.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Han Y, Han J, Li Z, et al.

Identification and characterization of a novel intronic splicing mutation in CSF1R-related leukoencephalopathy. *CNS Neurosci Ther*. 2024;00:e14815. doi:[10.1111/cns.14815](https://doi.org/10.1111/cns.14815)