

## OPEN ACCESS

### \*CORRESPONDENCE

Abigail L. Pfaff,  
✉ [abigail.pfaff@murdoch.edu.au](mailto:abigail.pfaff@murdoch.edu.au)

RECEIVED 12 March 2025

ACCEPTED 06 May 2025

PUBLISHED 03 June 2025

### CITATION

Pfaff AL and Köks S (2025)  
Retrotransposition-competent L1s are  
increased in the genomes of individuals  
with amyotrophic lateral sclerosis.  
*Exp. Biol. Med.* 250:10575.  
doi: 10.3389/ebm.2025.10575

### COPYRIGHT

© 2025 Pfaff and Köks. This is an open-  
access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Retrotransposition-competent L1s are increased in the genomes of individuals with amyotrophic lateral sclerosis

Abigail L. Pfaff<sup>1,2\*</sup> and Sulev Köks<sup>1,2</sup>

<sup>1</sup>Perron Institute for Neurological and Translational Science, Perth, WA, Australia, <sup>2</sup>Personalised  
Medicine Centre, Health Futures Institute, Murdoch University, Murdoch, WA, Australia

## Abstract

An individual's genetics contributes to their risk of developing amyotrophic lateral sclerosis (ALS); however, there is still a large proportion of the heritability of ALS to be understood. Part of this missing heritability may lie in complex variants, such as the long interspersed element 1 (L1) retrotransposon, which have yet to be evaluated. The majority of L1 insertions in the human genome are no longer able to retrotranspose, but to date 279 retrotransposition-competent (RC) L1s have been reported. Many RC-L1s are polymorphic for their presence/absence; therefore, each individual will have a different number and complement of RC-L1s. These elements have been hypothesized to be involved in disease processes by multiple mechanisms such as somatic mutation by retrotransposition, the triggering of neuroinflammation and DNA damage. We hypothesize that L1s may influence disease development either through their effects on endogenous genes or through the properties that enable them to retrotranspose. Whole genome sequencing data from the New York Genome Center ALS consortium were used to characterize L1 variation identifying 2,803 polymorphic L1 elements and association analysis was performed in European individuals (ALS/ALS with other neurological disorder (ALSND)  $n = 2,653$ , controls  $n = 320$ ). There were no individual L1 elements associated with disease, but we did identify a significant increase in the number of RC-L1s in ALS/ALSND genomes ( $p = 0.01$ ) and the presence of  $\geq 46$  RC-L1s showed the most significant association (OR = 1.09 (1.02–1.16),  $p = 0.01$ ) with disease. Analysis of individual L1s and their association with age at onset and survival identified one L1 whose presence was significantly associated with a lower age at onset (52.7 years) compared to homozygous absent individuals (59.2 years) ( $p_{adj} = 0.009$ ). Our study has identified novel genetic factors for both disease risk and age at onset in ALS providing further evidence for the role of L1 retrotransposons in neurodegenerative diseases.

### KEYWORDS

retrotransposons, L1, amyotrophic lateral sclerosis, neurodegeneration, genetics

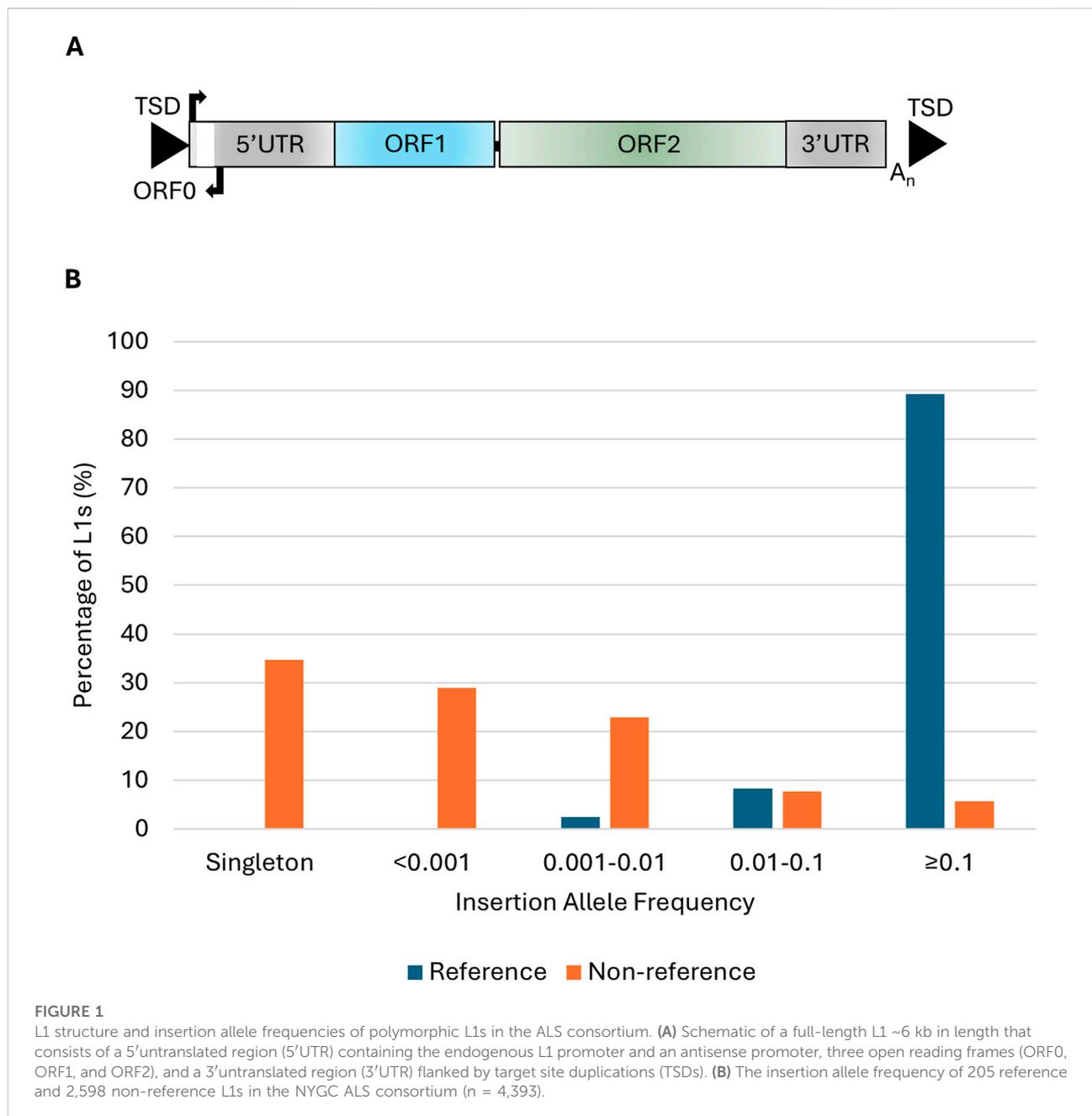
## Impact statement

This is the first study to assess genome-wide L1 variation in a large number of ALS and control genomes and to identify novel genetic factors involved in ALS risk and age at onset. ALS is a genetically heterogeneous disease with a large proportion of its heritability yet to be identified. Our study addressed this missing heritability by evaluating the presence/absence of the L1 retrotransposon and focusing on the subset of these elements that are still able to mobilize in the human genome. By understanding the mechanisms and risk factors that lead to

ALS, new therapeutic targets can be identified for a disease with very limited treatment options.

## Introduction

The long interspersed element-1 (L1) is the only autonomous family of retrotransposons in the human genome that contains elements currently able to mobilize and has played an important role in shaping the structure and function of the human genome [1]. L1s propagate through a “copy and paste” mechanism,



termed target primed reverse transcription [2, 3], and mobilize the non-autonomous *Alu* and SINE-VNTR-Alu retrotransposons. The ongoing mobilization of L1s has led to genetic variation between individuals and can affect the function and expression of endogenous genes. There are at least 29 instances of disease caused by an L1 insertion, generally through loss-of-function mutations or aberrant splicing [4]. A full-length L1 element is ~6 kb in size, contains both a 5' and 3'-untranslated region (UTR), three open reading frames (ORF0, ORF1, ORF2), a poly A tail at its 3' end, and is flanked by variable target site duplications (Figure 1A). The proteins encoded by ORF1, a ~40 kDa protein with RNA binding and chaperone activities, and ORF2, a ~150 kDa protein with endonuclease and reverse transcriptase activities, are required for retrotransposition [5–8]. Despite L1s contributing to 17% of the human genome and more than one million L1s annotated in hg38 only 146 L1s are full-length with intact ORFs [9]. The majority are unable to mobilize due to internal deletions or rearrangements, 5' truncations and mutations in the ORFs encoding the proteins required for retrotransposition [10, 11].

The dysregulation and elevated expression of L1s, along with other families of retrotransposons, have been detected in several neurological conditions including the neurodegenerative disease amyotrophic lateral sclerosis (ALS) [12, 13]. ALS, the most common form of motor neuron disease (MND), is a progressive disease that is characterized by the loss of motor neurons in the brain and spinal cord and the life expectancy after diagnosis is 2–5 years [14]. It is a clinically and genetically heterogeneous condition, and the neurodegenerative mechanisms involved are not fully understood. Several processes have been shown to be altered or dysregulated in the disease, such as protein aggregation, axonal transport, oxidative stress, mitochondrial dysfunction, RNA processing and the expression of retrotransposable elements [12, 15]. Changes in the expression of multiple classes and families of retrotransposons have been identified in subsets of individuals with ALS [16, 17]. An increase in expression of several families of retrotransposons was identified in the frontal cortex of individuals with ALS who carried the C9ORF72 expansion when compared to those without the expansion and healthy controls [16]. Another study using machine learning identified three subtypes of ALS based on transcriptomic data from the frontal and motor cortices, one of which was characterized by the activation of retrotransposons and TAR DNA-binding protein 43 (TDP-43) dysfunction [17]. Cytoplasmic accumulation of TDP-43, encoded by the gene TARDBP, in which mutations can cause ALS, is a hallmark of the majority of ALS cases. There is evidence in human tissues and cell lines that TDP-43 binds to retrotransposon transcripts of different classes, which is thought to aid in the repression of these elements [18]. Heterochromatin formation over repetitive elements, including L1s, acts to suppress their expression, but loss of

nuclear TDP-43 from neurons results in chromatin decondensation over L1 elements [19].

We hypothesize that polymorphic L1s may play a role in ALS development either through their effects on endogenous genes, such as loss-of-function mutations, or through properties that are inherent in their ability to mobilize. These include somatic mutation through retrotransposition, the triggering of neuroinflammation via the interferon pathway and DNA damage through the endonuclease activity of the ORF2 protein [13, 20]. The ability of specific L1s to retrotranspose can be detected using cellular retrotransposition assays and by tracing the source elements of both germline and somatic insertions using 3' transductions [21–23]. Retrotransposition-competent L1s (RC-L1s) annotated in the human reference genome and non-reference RC-L1s can be polymorphic for their presence or absence. Therefore, each individual will have a different number and complement of RC-L1s in their genome, which could potentially lead to differing levels of functional L1 mRNA.

To investigate the role of L1s in ALS we utilized whole genome sequencing from the New York Genome Center (NYGC) ALS consortium to characterize the landscape of polymorphic L1s in ALS and control genomes to evaluate their effects on ALS risk, age at onset, and survival. These analyses were performed for both individual L1 insertions and the total number of RC-L1s present. We identified an increased burden of RC-L1s, those still able to mobilize, in the genomes of individuals with ALS and a single L1 associated with a reduction in the age at onset of the disease.

## Materials and methods

### Genotyping of L1s using whole genome sequencing from the ALS consortium

Whole genome sequencing data in cram file format aligned to hg38 were obtained from the NYGC ALS consortium. The ALS consortium includes individuals with a range of diagnoses such as ALS spectrum MND (ALS), other MND, other neurological disorders (including Parkinson's disease and dementias) and ALS with other neurological disorders (ALSND) along with non-neurological controls (NNCs). The structural variant caller Delly<sup>1</sup>, with default settings, was used to identify structural variants in a subset of individuals (n = 244) [24]. The structural variants from each individual were merged and those deletions that overlapped with reference L1s were extracted. This generated a list of reference L1s that were absent in at least one of the 244 individuals. This panel of

1 <https://github.com/dellytools/delly>

TABLE 1 Demographics of the NYGC ALS consortium cohort in which association analysis was performed.

Demographic		NNC (n = 320)	ALS/ALSND (n = 2,653)
Gender	Male	157 (49.1%)	1,594 (60.1%)
	Female	163 (50.9%)	1,059 (39.9%)
Age <sup>a</sup> Mean (min-max)		57.4 (17–90)	59.1 (12–90)

<sup>a</sup>For NNC age at collection (43 unknown) and ALS/ALSND age at symptom onset (159 unknown).

reference L1s was used in the second call step of Delly to generate genotypes for the entire ALS consortium cohort available (n = 4,393). Non-reference L1s were genotyped using Mobile Element Locator Tools (MELT version 2.2.2 in MELT-split mode) [22]. The L1 insertions detected were filtered to keep those supported by  $\geq 2$  split reads and an assessment score  $\geq 3$  and that had passed the filtering criteria performed by MELT. There were 210 reference and 2,649 non-reference polymorphic L1s identified in the ALS consortium cohort and after filtering for Hardy-Weinberg equilibrium ( $p < 1 \times 10^{-6}$  in NNCs) 205 reference and 2,598 non-reference L1s remained.

## Identification of retrotransposition competent L1s

RC-L1s were identified using published data on those L1s that displayed activity in a cellular retrotransposition assay or were identified as source elements for germline or somatic insertions using 3' transduction events and details can be found in Pfaff et al [25]. The list of 198 RC-L1s from Pfaff et al has been extended using data from a more recent publication [26] and consists of 279 RC-L1s, 102 reference and 177 non-reference L1s.

## Association, age at onset and survival analysis

Association analysis was performed on those individuals who were >90% European according to the ALS consortium metadata and those diagnosed with ALS-spectrum MND or ALS-spectrum MND with another neurological condition (n = 2,653) and compared to non-neurological controls (n = 320) (see Table 1 for demographics). Association analysis of 501 polymorphic L1s (minor allele frequency >0.01) with ALS was performed using logistic regression adjusted for age, sex and sequencing preparation in PLINK (v1.07) and p values were adjusted for multiple testing (Bonferroni correction). Of the polymorphic L1s genotyped in the NYGC ALS consortium 106 were retrotransposition competent and 93 of these were

polymorphic in the European subset. The 3 RC-L1s located on the X chromosome were removed from further analysis so that men and women could be compared and 1 RC-L1s was filtered out because >5% of genotypes were missing. Only individuals with genotypes for all 89 RC-L1s were retained for analysis. Linear regression adjusted for sex, age and sequencing preparation was used to analyze the association between the total number of RC-L1s present and disease status. In addition, NNCs and individuals with ALS/ALSND were categorized based on the number of alleles with an RC-L1 present to analyze the association between the likelihood of having ALS/ALSND and a certain number of RC-L1s. Logistic regression adjusted for sex, age and sequencing preparation was performed on these RC-L1 groupings to determine if having more than a certain number of these defined elements was associated with ALS/ALSND. P values were adjusted for multiple testing using Bonferroni correction.

Age at onset analysis was performed using linear regression of age at onset on L1 genotype or RC-L1 number with sex, sequencing platform and site of onset as covariates and p values were adjusted for multiple testing (Bonferroni correction). Survival analysis was completed using the Cox proportional hazards model from the 'coxme' package in R with sex, sequencing platform, age at onset and site of onset as covariates and p values were adjusted for multiple testing (Bonferroni correction). Individuals in the ALS consortium dataset who were still alive were censored at their last follow-up.

## L1s in ALS-associated loci

To identify polymorphic L1s from the ALS consortium located in ALS-associated loci the L1 coordinates were intersected with the coordinates of ALS-associated genes. The list of ALS associated genes was generated from the amyotrophic lateral sclerosis online database (ALSoD)<sup>2</sup> using those defined as definitive ALS genes or with strong evidence for their association.

<sup>2</sup> <https://alsod.ac.uk/>

TABLE 2 Top 5 L1s significantly associated with ALS before correction.

L1 ID	Minor allele	MAF		OR (95% CI)	Unadj p value	Bonferroni p value
		NNCs	ALS/ALSND			
NRL1_20_43323466	P	0.431	0.371	0.76 (0.63–0.91)	0.003	1
NRL1_13_90973977	P	0.155	0.124	0.72 (0.57–0.92)	0.007	1
NRL1_20_29125080	P	0.025	0.014	0.46 (0.25–0.83)	0.011	1
NRL1_7_17055090	P	0.339	0.290	0.77 (0.64–0.94)	0.011	1
NRL1_10_89957889	P	0.081	0.105	1.54 (1.09–2.16)	0.013	1

## Expression quantitative trait loci analysis in medial and lateral motor cortex

Matrix eQTL was used to calculate the L1 loci regulating the expression of transcript variants [27]. We used an additive linear model with covariates, age and sex, and FDR was used to correct multiple testing with only the results that remained significant after FDR correction (<0.05) reported here. Matrix eQTL also reports effect-size estimates as beta values or slope coefficients.

## Results

### The landscape of polymorphic L1s in the NYGC ALS consortium

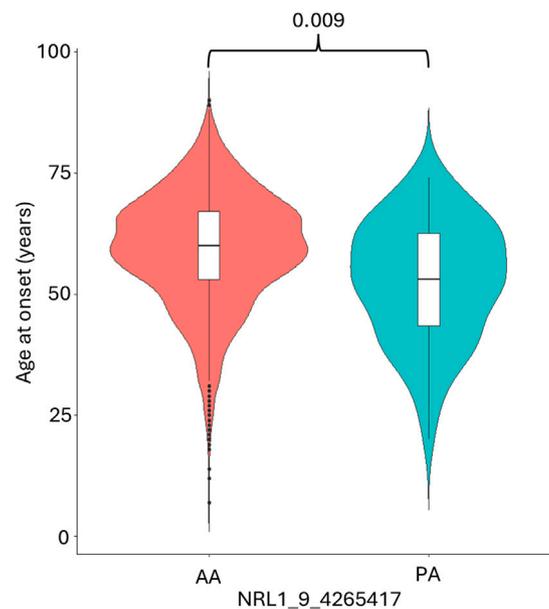
In 4,393 whole genomes analyzed from the NYGC ALS consortium 205 reference and 2,598 non-reference L1s were identified as polymorphic for their presence/absence. The insertion allele frequencies (IAF) of the reference L1s were higher than those of the non-reference elements with 89.3% of the reference L1s having an IAF of  $\geq 0.1$  compared to 5.7% of the non-reference L1s (Figure 1B). The majority of L1 insertions were found in intergenic regions (59.5%) followed by 39.9% of the L1s found in introns. There were 3 (0.1%) L1 insertions located in exons, 2 (0.07%) in the 5'UTR and 12 (0.4%) in the 3'UTR of genes. The three exonic insertions were located in coding exons of the following genes *AASDH*, *HLA-DRB1* and *FSTL4*. The non-reference L1s located in the *AASDH* and *HLA-DRB1* genes were each found in a single individual diagnosed with ALS. The third exonic L1 is a common insertion (IAF = 0.77) found in the reference genome and the *FSTL4* transcript initiates from the antisense promoter of this L1 element. There were two L1 insertions located in the introns of ALS-associated genes *ERBB4* and *SCFD1* with IAF of 0.13 and 0.15 respectively in the ALS consortium, neither of which was associated with disease.

Association analysis was conducted on 501 L1 insertions (MAF >0.01) in the individuals of European descent in the ALS consortium as this formed the largest proportion of the cohort

(Table 1 for demographics of this subset). There were 25 L1s associated with ALS/ALSND, but these did not survive correction for multiple testing (Supplementary Table S1). The top 5 L1s significantly associated with ALS/ALSND before correction are shown in Table 2. Age at onset analysis was performed identifying 31 associated L1s before correction with one L1 (NRL1\_9\_4265417) surviving correction for multiple testing (Supplementary Table S2). The presence of NRL1\_9\_4265417 was associated with a lower age at onset (est = -6.55, SE = 1.51, padj = 0.009). The average age at onset for individuals with the PA genotype was 52.7 years compared to 59.2 years for those with the AA genotype (Figure 2). NRL1\_9\_4265417 is a rare insertion with an IAF of 0.017 and located in intron 2 of the *GLIS3* gene. Survival analysis was performed identifying 26 associated L1s before correction for multiple testing (Supplementary Table S3), the top 5 of which are shown in Figure 3.

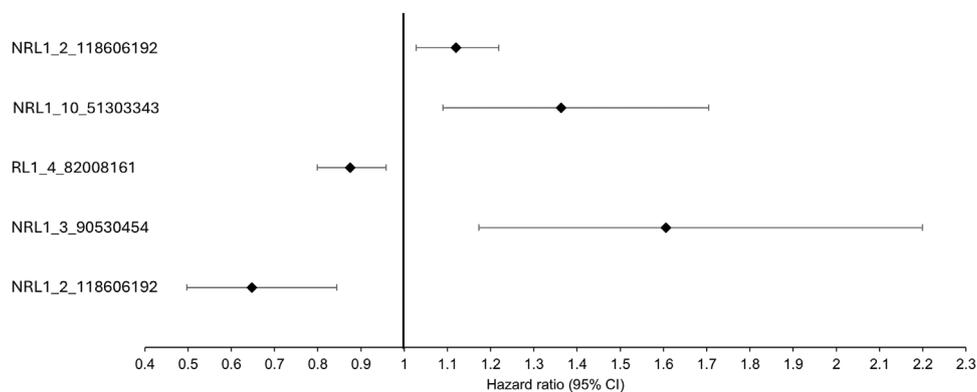
### The number of RC-L1s is significantly increased in ALS/ALSND genomes

In the NYGC ALS consortium 106 RC-L1s, 34 reference and 72 non-reference, were identified as polymorphic for the presence or absence from the compiled list of 279 RC-L1s. In the European subset of the cohort 93 RC-L1s were polymorphic. Not all the RC-L1s were detected as variable in our cohort as some were fixed in the population or were very rare and may only be present in specific population groups. The number of present alleles of 89 of the polymorphic RC-L1s in the European individuals, those on the sex chromosomes and with missing genotypes >5% were removed, was significantly higher in the genomes of individuals with ALS/ALSND compared to NNCs ( $\beta = 0.40$ ,  $p = 0.01$ ). The number of alleles present in NNCs genomes ranged from 32 to 56 and in ALS/ALSND genomes from 30 to 63 and their distribution is shown in Figure 4A. The ALS/ALSND and NNC individuals were grouped into 6 different categories to determine whether having more than a certain number of RC-L1s present at the 89 loci was associated with disease. The presence of  $\geq 45$ ,  $\geq 46$  and  $\geq 47$  polymorphic RC-L1s



**FIGURE 2**

The presence of *NRL1\_9\_4265417* is associated with a lower age at onset. The average age at onset of individuals diagnosed with ALS/ALSND was significantly lower in individuals with the PA genotype (52.7 years) compared to the AA genotype (59.2 years). PA  $n = 66$  and AA  $n = 2,429$ .



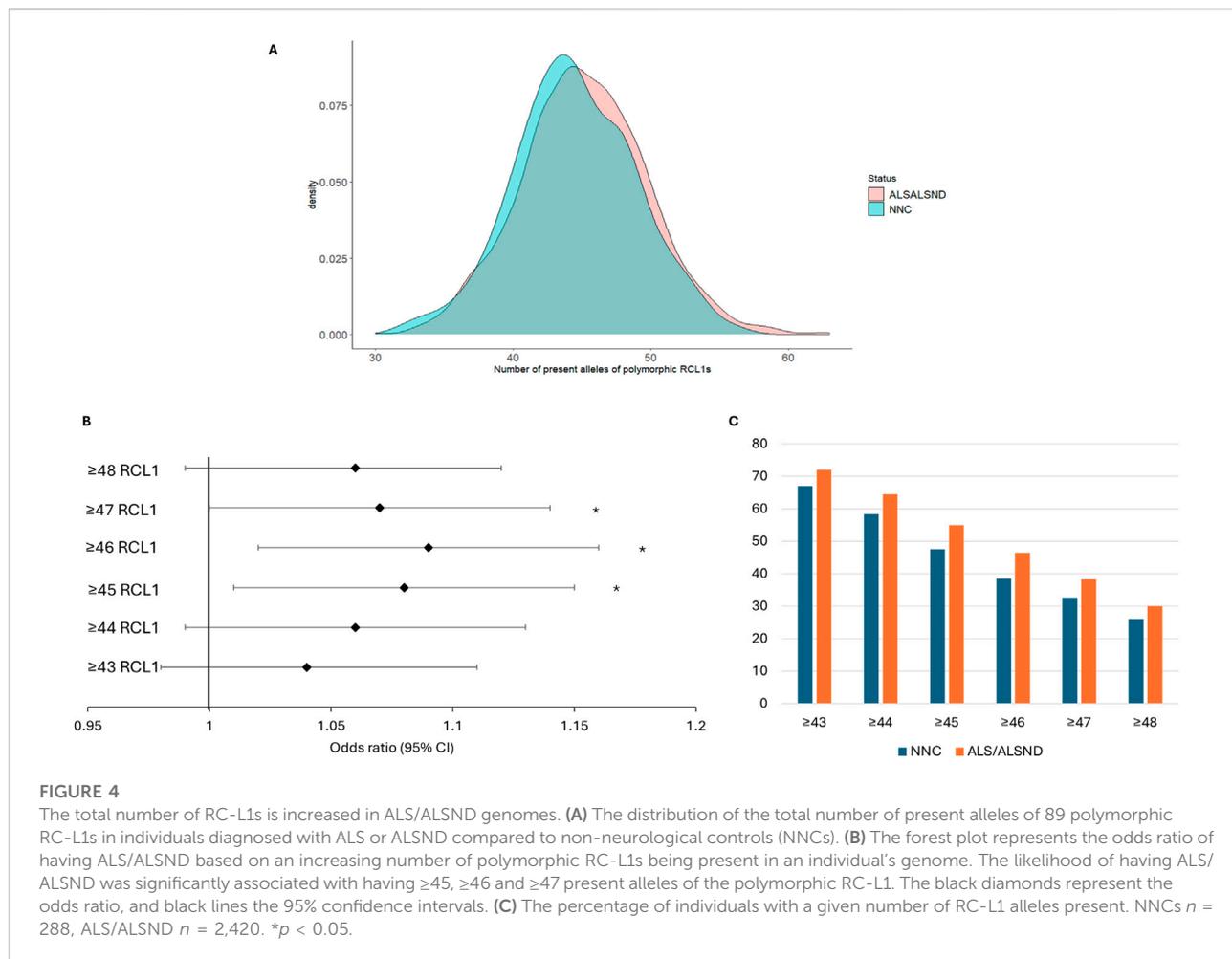
**FIGURE 3**

The top 5 most significant L1s associated with survival before correction. The black diamonds represent the hazard ratio and the black lines the 95% confidence intervals.

was significantly associated with ALS/ALSND (Figure 4B) and the percentage of individuals with a certain number of RC-L1s is shown in Figure 4C. The presence of  $\geq 46$  RC-L1s showed the most significant association (OR = 1.09 (1.02–1.16),  $p = 0.01$ ) with 38.5% of NNCs having  $\geq 46$  compared to 46.5% of individuals with ALS/ALSND. The number of RC-L1s present in the genomes was not associated with age at onset or survival.

To determine whether the RC-L1s might influence ALS pathogenesis through effects on gene expression, an analysis

was performed using genotypes of the 2,803 polymorphic L1s in this study and transcriptomic data from the medial and lateral motor cortex. This eQTL analysis identified 252 L1s that were significantly associated with expression changes of 395 different transcripts, but only 11 of these L1s were retrotransposition competent (Supplementary Table S4). The proportion of RC-L1s acting as eQTLs (4.4%) was not significantly different from the genome-wide proportion (3.8%) (Fisher's exact test  $p = 0.60$ ) suggesting that they are not enriched as eQTLs over those L1s



that are not able to retrotranspose. There were 25 transcripts whose expression was associated with the 11 RC-L1s and the majority (72%) were either pseudogenes or novel transcripts with limited functional information available. Genes whose expression was associated with the presence/absence of RC-L1s included zinc finger protein 367 (ZNF367), olfactory receptor family 5 subfamily K member 1 (OR5K1) and maturin (MTURN).

## Discussion

Here we present data from a genome-wide analysis of polymorphic L1s using 4,393 whole genomes from the NYGC ALS consortium to identify novel genetic factors involved in ALS development and age at onset. ALS is a genetically heterogeneous disease with different patterns of inheritance and levels of gene penetrance with more than 30 genes associated with its development. The majority of large-scale studies have focused on single-nucleotide variants and suggest a heritability of 10–20%, which leaves 30–40% to be explained [28]. This

missing heritability is likely to be found for other types of genetic variants and structural variants have already been associated with ALS [29]. Our study focused on polymorphic reference and non-reference L1s, evaluating their association with ALS as individual elements and in combination for those that can still mobilize within the human genome.

In the NYGC ALS consortium data 2,803 polymorphic L1s (205 reference and 2,598 non-reference) were identified. Association analysis performed on 501 L1s (MAF >0.01) did not identify individual L1s associated with disease after correction for multiple testing; however, when evaluating RC-L1s we identified an increased number of these elements in the genomes of individuals with ALS compared to controls. When analyzing the number of alleles present in 89 RC-L1s the total number of RC-L1s was higher in ALS/ALSND genomes (Figure 4A) and having  $\geq 46$  RC-L1s was most significantly associated with disease (Figures 4B,C). RC-L1s encode functional proteins that enable them to mobilize in the human genome and have the potential to generate somatic insertions, cause DNA damage and generate cytosolic RNA: DNA hybrids that activate the innate immune response and

affect normal cellular function [13, 20]. These features of RC-L1s that are inherent in their ability to retrotranspose are important factors in how they may contribute to neurodegeneration and neuroinflammation. An increased number of these elements present in a genome could lead to higher levels of functional mRNA in a cell when these elements are expressed and potentially greater effects. Elevated expression of retrotransposons, including L1s, in ALS has been reported by multiple studies and was often restricted to a subset of individuals [12, 16, 17]. Analysis of locus specific expression of reference L1 elements encoding intact proteins showed an overall decrease in expression in the brains of individuals with ALS, but in a small number of individuals with ALS this expression was massively increased, up to 25 times the average [30]. This study was limited to reference L1s and included all L1s encoding intact proteins (not specifically those with evidence of their ability to retrotranspose); therefore, evaluation of both reference and non-reference RC-L1s would provide important data regarding their expression profile in ALS.

Methylation of a CpG island located in the L1 5'UTR is involved in regulating L1 expression and a reduction in methylation could indicate which elements are most likely to be expressed. For example, the expression of a specific L1 located on chromosome 13 occurred alongside its hypomethylation during hESC neurodifferentiation and in adult tissues such as the hippocampus [31]. Changes in methylation of L1s may also be associated with disease states, for example L1s were differentially methylated in the prefrontal cortices of individuals with psychiatric disorders [32]. When analyzing a small number of highly active RC-L1s we identified a reduction in the methylation of selected RC-L1s in the motor cortex of individuals with ALS compared to healthy controls [33]. A global evaluation of RC-L1 methylation and expression in ALS could help identify those elements that are altered in the disease to refine the list of RC-L1s used in our analysis. For example, when analyzing the burden of RC-L1s in Parkinson's disease (PD) we identified an increased number of highly active RC-L1s in PD genomes rather than an increase in the presence of all RC-L1s [25].

Another potential mechanism by which RC-L1s may contribute to ALS pathogenesis is through their effects on endogenous gene expression. Transcriptomic analysis identified 252 L1s acting as eQTLs in either the medial or lateral motor cortex, of which 11 were RC-L1s affecting 25 different transcripts. The global association of RC-L1s as opposed to individual elements being linked to disease and the small number affecting gene expression suggest that it is their retrotransposition properties that are more likely to be involved in disease processes.

L1s, along with other retrotransposons, are therapeutic targets in a number of neurological conditions [34]. Reverse transcriptase inhibitors that target the L1 protein have been

tested in a clinical trial for the neurodevelopmental disorder Aicardi-Goutieres syndrome [35, 36] and trials are currently underway in C9orf72 expansion positive individuals with ALS or frontotemporal dementia and in individuals with progressive supranuclear palsy. By analyzing RC-L1 variation and activity the aim would be to identify those individuals with ALS in whom L1s are contributing to disease development and who may benefit the most from these therapies targeting L1 activity.

We performed age at onset and survival analysis for both individual L1s and the burden of RC-L1s, identifying a single L1 (NRL1\_9\_4265417) whose presence was associated with a reduced age at onset (Figure 2). Age at onset of ALS is influenced by sex, family history and genetics [37]. Carriers of a Mendelian genetic variant have a lower age at onset and two recent age at onset GWAS have identified only two associated variants [28, 38, 39]. NRL1\_9\_4265417 is not in a known ALS-associated region and is located in intron 2 of the *GLIS3* gene, which is a transcriptional regulator and has been shown to modulate genes involved in autophagy and neuronal function [40]. Variants in *GLIS3* have also been associated with levels of tau and ptau in Alzheimer's disease [41]. Analysis of the functional consequences of this L1 insertion will be required to determine its effects and how it may act as a disease modifier.

This study has identified novel genetic factors for both disease risk and age at onset in ALS providing further evidence for the role of L1 retrotransposons in this neurodegenerative disease and their potential as a therapeutic target.

## Author contributions

Conceptualization, AP and SK; methodology, AP and SK; formal analysis, AP; data interpretation, AP and SK; writing – original draft preparation, AP; writing – review and editing, SK All authors contributed to the article and approved the submitted version.

## Data availability

The WGS sequencing analysed in this study from the ALS consortium were obtained upon application to the New York Genome Center and data requests can be made by completing a genetic data request form at ALSData@nygenome.org. <https://www.nygenome.org/science-technology/collaborative-research>.

## Ethics statement

The studies involving humans were approved by Research ethics and integrity, Murdoch University (2020/040). The studies

were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. AP and SK are funded by the MSWA and the Perron Institute for Neurological and Translational Science.

## Acknowledgments

We would like to acknowledge the Target ALS Human Postmortem Tissue Core, the New York Genome Center for Genomics of Neurodegenerative Disease, the Amyotrophic Lateral Sclerosis Association and the TOW Foundation for the whole genome sequencing data used in this publication. This work was supported by resources provided by the Pawsey

## References

- Kazazian HH, Jr, Moran JV. Mobile DNA in health and disease. *N Engl J Med* (2017) **377**:361–70. doi:10.1056/nejmra1510092
- Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription *in vitro*. *EMBO J* (2002) **21**:5899–910. doi:10.1093/emboj/cdf592
- Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* (1993) **72**:595–605. doi:10.1016/0092-8674(93)90078-5
- Hancks DC, Kazazian HH, Jr. Roles for retrotransposon insertions in human disease. *Mobile DNA* (2016) **7**:9. doi:10.1186/s13100-016-0065-9
- Feng Q, Moran JV, Kazazian HH, Jr, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* (1996) **87**:905–16. doi:10.1016/s0092-8674(00)81997-2
- Khazina E, Weichenrieder O. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci U S A* (2009) **106**:731–6. doi:10.1073/pnas.0809964106
- Mathias SL, Scott AF, Kazazian HH, Jr, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. *Science* (1991) **254**:1808–10. doi:10.1126/science.1722352
- Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. High frequency retrotransposition in cultured mammalian cells. *Cell* (1996) **87**:917–27. doi:10.1016/s0092-8674(00)81998-4
- Penzkofer T, Jager M, Figlerowicz M, Badge R, Mundlos S, Robinson PN, et al. L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes. *Nucleic Acids Res* (2017) **45**:D68–D73. doi:10.1093/nar/gkw925
- Lander ES, Linton LM, Birren B, Nussbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature* (2001) **409**:860–921. doi:10.1038/35057062
- Myers JS, Vincent BJ, Udall H, Watkins WS, Morrish TA, Kilroy GE, et al. A comprehensive analysis of recently integrated human Ta L1 elements. *The Am J Hum Genet* (2002) **71**:312–26. doi:10.1086/341718
- Savage AL, Schumann GG, Breen G, Bubbs VJ, Al-Chalabi A, Quinn JP. Retrotransposons in the development and progression of amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* (2019) **90**:284–93. doi:10.1136/jnnp-2018-319210
- Tam OH, Ostrow LW, Gale Hammell M. Diseases of the nervous system: retrotransposon activity in neurodegenerative disease. *Mobile DNA* (2019) **10**:32. doi:10.1186/s13100-019-0176-1

Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

## Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.ebm-journal.org/articles/10.3389/ebm.2025.10575/full#supplementary-material>

- Masrori P, Van Damme P. Amyotrophic lateral sclerosis: a clinical review. *Eur J Neurol* (2020) **27**:1918–29. doi:10.1111/ene.14393
- van Es MA, Hardiman O, Chio A, Al-Chalabi A, Pasterkamp RJ, Veldink JH, et al. Amyotrophic lateral sclerosis. *The Lancet* (2017) **390**:2084–98. doi:10.1016/s0140-6736(17)31287-4
- Prudencio M, Gonzales PK, Cook CN, Gendron TF, Daugherty LM, Song Y, et al. Repetitive element transcripts are elevated in the brain of C9orf72 ALS/FTLD patients. *Hum Mol Genet* (2017) **26**:3421–31. doi:10.1093/hmg/ddx233
- Tam OH, Rozhkov NV, Shaw R, Kim D, Hubbard I, Fennessey S, et al. Postmortem cortex samples identify distinct molecular subtypes of ALS: retrotransposon activation, oxidative stress, and activated glia. *Cell Rep* (2019) **29**:1164–77.e5. doi:10.1016/j.celrep.2019.09.066
- Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PLoS One* (2012) **7**:e44099. doi:10.1371/journal.pone.0044099
- Liu EY, Russ J, Cali CP, Phan JM, Amlie-Wolf A, Lee EB. Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. *Cell Rep* (2019) **27**:1409–21.e6. doi:10.1016/j.celrep.2019.04.003
- Saleh A, Macia A, Muotri AR. Transposable elements, inflammation, and neurological disease. *Front Neurol* (2019) **10**:894. doi:10.3389/fneur.2019.00894
- Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* (2003) **100**:5280–5. doi:10.1073/pnas.0831042100
- Gardner EJ, Lam VK, Harris DN, Chuang NT, Scott EC, Pittard WS, et al. The Mobile Element Locator Tool (MELT): population-scale mobile element discovery and biology. *Genome Res* (2017) **27**:1916–29. doi:10.1101/gr.218032.116
- Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, Zamora J, Supek F, Demeulemeester J, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nat Genet* (2020) **52**:306–19. doi:10.1038/s41588-019-0562-0
- Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* (2012) **28**:i333–i339. doi:10.1093/bioinformatics/bts378
- Pfaff AL, Bubbs VJ, Quinn JP, Koks S. An increased burden of highly active retrotransposition competent L1s is associated with Parkinson's disease risk and

progression in the PPMI cohort. *Int J Mol Sci* (2020) **21**:6562. doi:10.3390/ijms21186562

26. Nam CH, Youk J, Kim JY, Lim J, Park JW, Oh SA, et al. Widespread somatic L1 retrotransposition in normal colorectal epithelium. *Nature* (2023) **617**:540–7. doi:10.1038/s41586-023-06046-z

27. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* (2012) **28**:1353–8. doi:10.1093/bioinformatics/bts163

28. Al-Chalabi A, Andrews J, Farhan S. Recent advances in the genetics of familial and sporadic ALS. *Int Rev Neurobiol* (2024) **176**:49–74. doi:10.1016/bs.irm.2024.04.007

29. Al Khleifat A, Iacoangeli A, van Vugt J, Bowles H, Moisse M, Zwamborn RAJ, et al. Structural variation analysis of 6,500 whole genome sequences in amyotrophic lateral sclerosis. *NPJ Genom Med* (2022) **7**:8. doi:10.1038/s41525-021-00267-9

30. Pfaff AL, Bubb VJ, Quinn JP, Koks S. Locus specific reduction of L1 expression in the cortices of individuals with amyotrophic lateral sclerosis. *Mol Brain* (2022) **15**:25. doi:10.1186/s13041-022-00914-x

31. Sanchez-Luque FJ, Kempen MJHC, Gerdes P, Vargas-Landin DB, Richardson SR, Troskie RL, et al. LINE-1 evasion of epigenetic repression in humans. *Mol Cell* (2019) **75**:590–604 e12. doi:10.1016/j.molcel.2019.05.024

32. Watanabe R, Nakachi Y, Matsubara H, Ueda J, Ishii T, Ukai W, et al. Identification of epigenetically active L1 promoters in the human brain and their relationship with psychiatric disorders. *Neurosci Res* (2023) **195**:37–51. doi:10.1016/j.neures.2023.05.001

33. Savage AL, Lopez AI, Iacoangeli A, Bubb VJ, Smith B, Troakes C, et al. Frequency and methylation status of selected retrotransposition competent L1 loci

in amyotrophic lateral sclerosis. *Mol Brain* (2020) **13**:154. doi:10.1186/s13041-020-00694-2

34. Eisenstein M. Startups probe hidden viruses in the 'dark genome' to treat disease. *Nat Biotechnol* (2024) **42**:539–41. doi:10.1038/s41587-024-02215-1

35. Crow YJ, Briggs TA, Eleftheriou D, Parida A, Battison C, Giddings A, et al. Reverse transcriptase inhibitors in Aicardi-Goutières syndrome: a crossover clinical trial. *Dev Med Child Neurol* (2025) **67**:750–7. doi:10.1111/dmcn.16199

36. Rice GI, Meyzer C, Bouazza N, Hully M, Boddaert N, Semeraro M, et al. Reverse-transcriptase inhibitors in the aicardi-goutieres syndrome. *N Engl J Med* (2018) **379**:2275–7. doi:10.1056/nejmc1810983

37. Chiò A, Moglia C, Canosa A, Manera U, D'Ovidio F, Vasta R, et al. ALS phenotype is influenced by age, sex, and genetics: a population-based study. *Neurology* (2020) **94**:e802–e810. doi:10.1212/wnl.00000000000008869

38. Li C, Lin J, Jiang Q, Yang T, Xiao Y, Huang J, et al. Genetic modifiers of age at onset for amyotrophic lateral sclerosis: a genome-wide association study. *Ann Neurol* (2023) **94**:933–41. doi:10.1002/ana.26752

39. Li C, Wei Q, Hou Y, Lin J, Ou R, Zhang L, et al. Genome-wide analyses identify NEAT1 as genetic modifier of age at onset of amyotrophic lateral sclerosis. *Mol Neurodegeneration* (2023) **18**:77. doi:10.1186/s13024-023-00669-6

40. Calderari S, Ria M, Gerard C, Nogueira TC, Villate O, Collins SC, et al. Molecular genetics of the transcription factor GLIS3 identifies its dual function in beta cells and neurons. *Genomics* (2018) **110**:98–111. doi:10.1016/j.ygeno.2017.09.001

41. Cruchaga C, Kauwe JS, Harari O, Jin SC, Cai Y, Karch CM, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. *Neuron* (2013) **78**:256–68. doi:10.1016/j.neuron.2013.02.026