21. Hair J, Anderson R, Black B, Babin B. Multivariate Data Analysis. New York, NY: Pearson Education; 2016.

22. Chalmers RP. Mirt: a multidimensional item response theory package for the R environment. J Stat Softw 2012;48:1–29.

## Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

# Detection and Characterization of a De Novo *Alu* Retrotransposition Event Causing *NKX2-1*-Related Disorder


Francesca Magrinelli, MD, PhD,[1*] Clarissa Rocca, MSc,[2,3] Roberto Simone, PhD,[2] Riccardo Zenezini Chiozzi, PhD,[4] Zane Jaunmuktane, MD, FRCPath,[1] Niccolò E. Mencacci, MD, PhD,[5] Michele Tinazzi, MD, PhD,[6] Sandeep Jayawant, MD,[7] Andrea H. Nemeth, MD, PhD,[8,9] German Demidov, PhD,[10] Henry Houlden, MD, PhD,[2†] and Kailash P. Bhatia, MD, DM, FRCP[1†]

[1]*Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, London, United Kingdom* [2]*Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, United Kingdom* [3]*William Harvey Research Institute, School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom* [4]*Mass-Spectrometry, Science Technology Platforms, University College London, London, United Kingdom* [5]*Ken and Ruth Davee Department of Neurology and Simpson Querrey Center for Neurogenetics, Northwestern University, Feinberg School of Medicine, Chicago, Illinois, USA* [6]*Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy* [7]*Paediatric Neurology, Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom* [8]*Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom* [9]*Oxford Centre for Genomic Medicine, Oxford University Hospitals National Health Service Foundation Trust, Oxford, United Kingdom* [10]*Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany*



*__Correspondence to:__ Dr. Francesca Magrinelli, Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK; E-mail: f.magrinelli@ucl.ac.uk

†Henry Houlden and Kailash P. Bhatia contributed equally to this work.

**Relevant conflicts of interest/financial disclosures:** The authors received no specific funding for this study and have no conflict of interest to declare concerning this work.

**Funding agency:** This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 779257 (Solve-RD).

**Received:** 5 September 2022; **Revised:** 17 October 2022; **Accepted:** 6 November 2022

**Published online 23 November 2022 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29280**

**ABSTRACT: Background:** Heterozygous *NKX2-1* loss-of-function variants cause combinations of hyperkinetic movement disorders (MDs, particularly childhood-onset chorea), pulmonary dysfunction, and hypothyroidism. Mobile element insertions (MEIs) are potential disease-causing structural variants whose detection in routine diagnostics remains challenging.
**Objective:** To establish the molecular diagnosis of two first-degree relatives with clinically suspected *NKX2-1*-related disorder who had negative *NKX2-1* Sanger (SS), whole-exome (WES), and whole-genome (WGS) sequencing.
**Methods:** The proband's WES was analyzed for MEIs. A candidate MEI in *NKX2-1* underwent optimized SS after plasmid cloning. Functional studies exploring *NKX2-1* haploinsufficiency at RNA and protein levels were performed.
**Results:** A 347-bp *Alu*Ya5 insertion with a 65-bp poly-A tail followed by a 16-bp duplication of the pre-insertion wild-type sequence in exon 3 of *NKX2-1* (ENST00000354822.7:c.556_557ins*Alu*541_556dup) segregated with the disease phenotype.
**Conclusions:** We identified a de novo exonic *Alu*Ya5 insertion causing *NKX2-1*-related disorder in SS/WES/WGS-negative cases, suggesting that MEI analysis of short-read sequencing data or targeted long-read sequencing could unmask the molecular diagnosis of unsolved MD cases. © 2022 The Authors. *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

**Key Words:** brain-lung-thyroid syndrome; chorea; dystonia; mobile element insertion; thyroid transcription factor-1

## Introduction

*NKX2-1* is a homeobox gene encoding thyroid transcription factor-1 (TTF1), a key regulator of tissue-specific gene expression mainly involved in thyroid, lung, and ventral forebrain morphogenesis.[1] Heterozygous *NKX2-1* loss-of-function variants account for a clinical spectrum, including childhood-onset chorea, hypothyroidism, and pulmonary dysfunction, in isolation or any combination, with brain-lung-thyroid (BLT) syndrome being the most severe phenotype.[2,3] Other possible manifestations include motor delay, dystonia,

ataxia, urinary tract abnormalities, and malignancies.[4,5] Magnetic resonance imaging (MRI) pituitary abnormalities have been recognized in 13% of *NKX2-1*-variant carriers and 26% of BLT syndrome patients reported thus far.[6] *NKX2-1* point and copy number variants were detected in only 26.7% of patients in a large series of benign hereditary chorea/BLT syndrome.[2]

Mobile elements are discrete DNA segments constituting about two-thirds of the human genome.[7] Evolutionarily, they have been able to mobilize between genomic regions via a direct cut-and-paste mechanism (*transposons*) or target-primed reverse transcription of an RNA intermediate (*retrotransposons*).[8] Among retrotransposons, a small fraction of long (LINEs) and short interspersed nuclear elements (SINEs, which include *Alu* elements) as well as SINE-variable number of tandem repeat (VNTR)-*Alu* (SVA) elements retain intragenomic transpositional competence and can cause human genetic diseases because of their activity.[8-15] Detection of mobile element insertions (MEIs) by routine diagnostics is ignored as they are ultra-rare, remains challenging due to their intrinsic characteristics (eg, size, highly repetitive sequence), and requires to run dedicated variant-calling algorithms on short-read sequencing (SRS) data.[16-20]

We identified and characterized the first *Alu*-SINE retrotransposition event in *NKX2-1* associated with infancy-onset levodopa-responsive choreo-dystonia in two first-degree relatives (Fig. 1A) for whom no molecular diagnosis had previously been established by Sanger sequencing (SS) of full-length *NKX2-1* and canonical pipelines for whole-exome (WES) and short-read whole-genome sequencing (SR-WGS) data analysis.

## Patients and Methods

### Patients

#### *Proband (II:2)*

This 46-year-old white British woman was born full-term from extended traumatic breech delivery. She was diagnosed with respiratory distress syndrome and required neonatal intensive care. She experienced frequent respiratory infections during infancy. Her motor milestones were delayed (age sitting: 12 months, standing with support: 2 years, walking with rollator: 2.5–3 years). She manifested with limb and truncal twitching movements from age 1. She was diagnosed with athetoid cerebral palsy at age 2. During adolescence, a trial of levodopa (100 mg thrice per day) resulted in gait improvement and allowed her to walk using crutches. Her past medical history included kyphoscoliosis, urinary tract infections, mixed urinary incontinence and mild urinary hesitancy since childhood, dyspnea on exertion requiring inhalers since her 30s, and one episode of acute urinary retention at about age 35. Subclinical hypothyroidism was diagnosed at age 45, which led to initiation of levothyroxine. She was born to non-consanguineous parents. Her father (I:1), mother (I:2), and two brothers were healthy, whereas her only daughter was similarly affected. On examination (Video S1), she had generalized chorea with dystonic posturing of her fingers and legs, particularly when walking, hyperreflexia, and impaired postural reflexes. Brain MRI revealed a cystic lesion of the neurohypophysis (Fig. 1B). Peripheral smear for acanthocytes, phenylalanine loading test, and cerebrospinal fluid analysis of pterins and monoamine metabolites were normal. Electroencephalogram and nerve conduction study/electromyogram were unremarkable. Muscle biopsy revealed slight type I fiber predominance, one ragged red fiber with increased lipid staining, one cytochrome oxidase–deficient fiber, and occasional endomysial T cells. Neuropsychometry was within normal limits. SS of *NKX2-1* was negative.

#### *Proband's Daughter (III:1)*

This 15-year-old white British girl was born full-term from uncomplicated elective caesarean section and was diagnosed with congenital talipes equinovarus at birth. She was noted to be fidgety and have poor control of her posture during early childhood. She manifested delayed motor milestones (age sitting: 16 months, walking with Kaye walker: 2 years). Her gait improved after levodopa initiation at age 2. She walked with crutches at age 6, started walking unsupported at age 10, and became fully independent with walking at age 13. Brain MRI showed prominent pituitary fossa and gland, with an isointense lesion located behind the anterior pituitary gland (Fig. 1B). Her past medical history included scoliosis, recurrent bronchiolitis during childhood, stress incontinence requiring toilet training at age 3, and oxybutynin since early childhood. Her thyroid function was normal at age 14. On examination (Video S1), she had generalized chorea with dystonic posturing of her neck, fingers, and legs, particularly when walking, hyperreflexia, and impaired postural reflexes.

### Methods

All subjects provided written informed consent. Blood for genomic DNA and total RNA extraction was obtained from subjects I:1, I:2, II:2, and III:1. Skin biopsies were collected from subjects I:2, II:2, and III:1. WES was performed in subject II:2 and SR-WGS in all subjects (Illumina Technologies, San Diego, California, USA). Canonical pipelines, including de novo analysis, were applied to WES and SR-WGS data. The proband's biological parenthood was ascertained as described elsewhere.[21] The germline MEI detection tool SCRAMble (https://github.com/GeneDx/scramble) was run on the proband's WES bam files with default settings.[18,19]

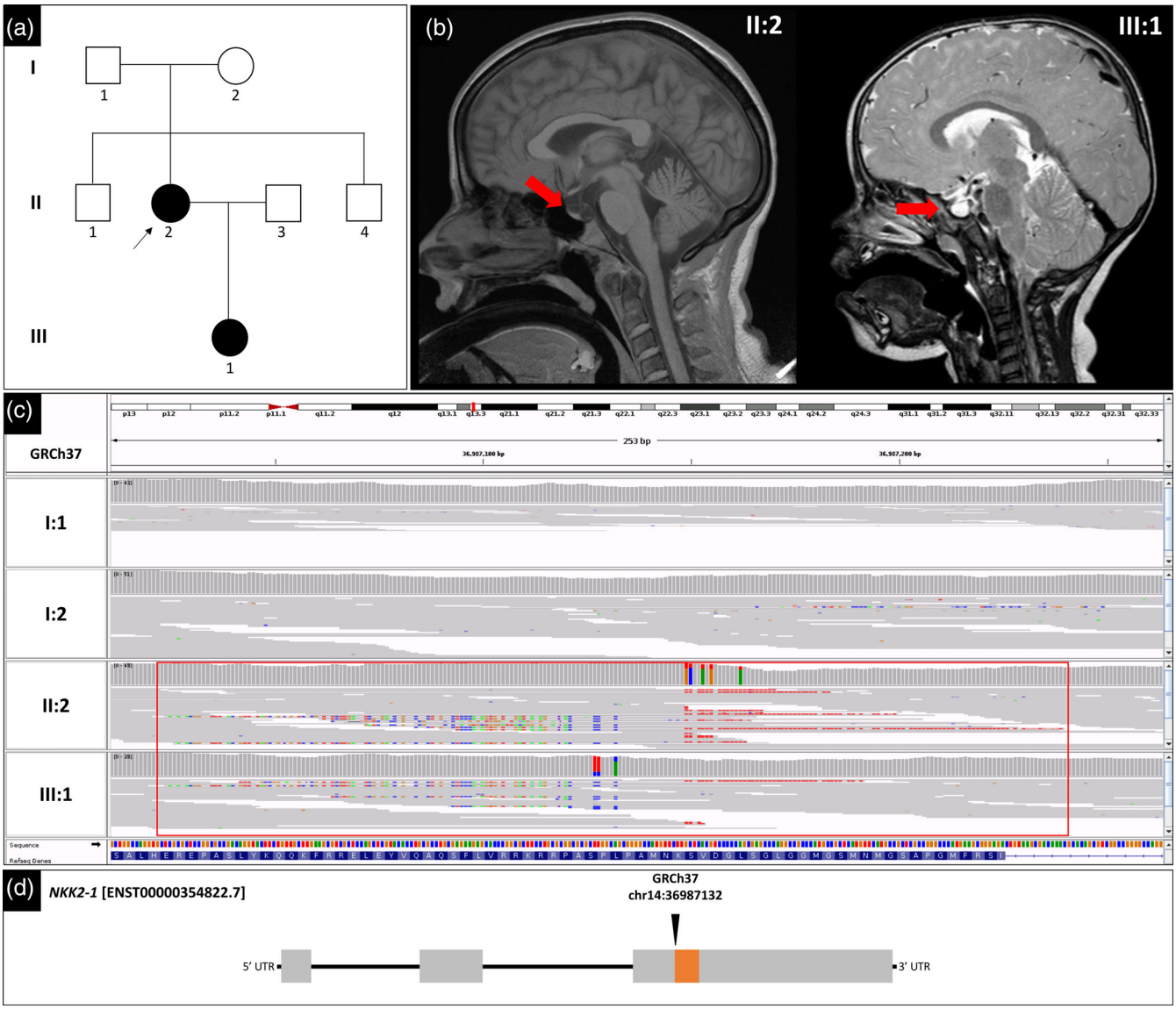


**FIG. 1.** (**A**) Family tree of the pedigree with an *Alu* element exonic insertion in *NKX2-1*. The proband is indicated by a black arrow. Filled black symbols denote clinically affected individuals, and open symbols indicate unaffected family members. (**B**) Brain MRI (magnetic resonance imaging) showed a cystic lesion of the posterior pituitary gland in the proband (II:2, age 42; sagittal T1 midline) and prominent pituitary fossa and pituitary gland with a CSF (cerebrospinal fluid)-isointense lesion situated behind the anterior pituitary gland in her affected daughter (III:1, age 2; sagittal T2 midline). (**C**) Integrative Genomics Viewer screen capture showing the alignment of the *NKX2-1* alleles. The paired-end reads obtained from the two affected family members (II:2, III:1) and the proband's unaffected parents (I:1, I:2) were aligned to the human genome GRCh37 in the genomic range chr14:36,987,000–36,987,300. Chimeric reads containing the soft-clipped poly-T reflecting the 3′-end poly-A tail (right red stretch) and the soft-clipped 5′ start (left multicolor stretch) of the *Alu* element in individuals II:2 and III:1 are shown in the lower boxes. The large *NKX2-1* insertion is absent in the proband's unaffected parents (I:1, I:2). (**D**) Schematic of the *NKX2-1* gene with insertion of the *Alu* element (orange box) in exon 3. [Color figure can be viewed at wileyonlinelibrary.com]

Solve-RD WES cohort was used to estimate allele frequency, which shows that this event was ultra-rare, occurring in 1 of 21,780 alleles.[22] Neither this nor other structural variants (SV) in *NKX2-1* were detected in gnomAD SV 2.1 (accessed on October 10, 2022). All subjects' SR-WGS bam files were inspected for the presence of the candidate MEI using Integrative Genomics Viewer (IGV).[23] Polymerase chain reaction (PCR) primers were designed to amplify the *NKX2-1* region carrying the identified MEI (Fig. 2A). PCR and agarose gel electrophoresis of PCR products were performed. DNA fragments were excised from agarose gel and purified using the Monarch DNA Gel Extraction Kit (New England Biolabs, Ipswich, Massachusetts, USA). Purified DNA was next Sanger sequenced. DNA purified from gel bands predicted to carry the MEI was cloned in ccompetent *Escherichia coli* using a TOPO TA Cloning Kit (Invitrogen, Waltham, Massachusetts, USA), and SS of recombinant plasmids was performed using the same primers. Annotation was carried out by

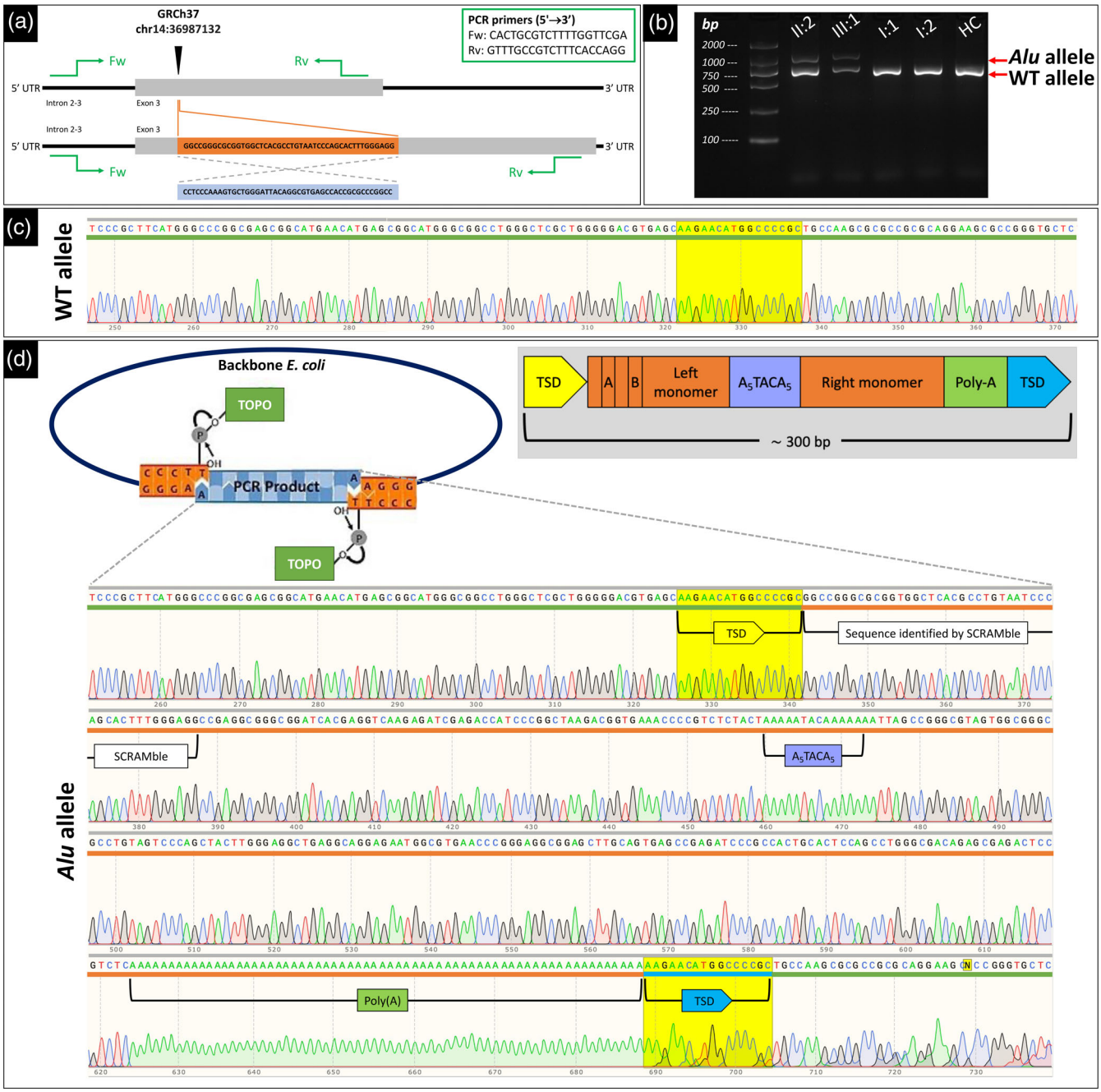


**FIG. 2.** (**A**) Schematic of biallelic *NKX2-1* exon 3. The 46-bp fragment of the *Alu* element as identified by the tool SCRAMble onto a negative strand of *NKX2-1* is reported in the light blue box, whereas its reverse complement counterpart as expected by Sanger sequencing (SS) is reported in the orange box. Target-specific primers for PCR (polymerase chain reaction) are indicated in the green-rimmed box, with a predicted amplicon size of 706 bp for the wild-type allele and ~1000 bp for the allele carrying the *Alu* element insertion. (**B**) Agarose (2%) gel electrophoresis of the PCR products obtained for a fragment of *NKX2-1* exon 3 from the proband (II:2), her affected daughter (III:1), her unaffected father (I:1) and mother (I:2), and one healthy control (HC) showed an additional DNA band of ~1000 bp in affected individuals (II-1, III-1), who are heterozygotes for the presence of the *Alu* insertion (*Alu* allele), whereas healthy subjects (I:1, I:2, HC) are homozygotes for the wild-type (WT) allele. (**C**) Chromatogram showing the result of SS of DNA purified from the ~700-bp gel band using the aforementioned primers. The sequence corresponds to the wild-type sequence of *NKX2-1* exon 3. The sequence highlighted in yellow represents the target site duplication (TSD) used by the *Alu* element to insert in the genome. (**D**) Chromatogram showing the result of SS of DNA purified from the ~1000-bp gel band and cloned in chemically competent *Escherichia coli* using a TOPO TA Cloning Kit (Invitrogen; upper left corner). Recombinant plasmids were sequenced using the aforementioned primers. The schematic of a full-length *Alu* element is shown in the gray box, including a left monomer (with boxes A and B containing RNA polymerase III promoter binding sites), an internal element $A_5TACA_5$, a right monomer, a long poly-A tail on the 3′ end, and TSD of the flanking region. These canonical *Alu* components are highlighted in the *Alu* element characterized here with labels of the same colors used in the *Alu* schematic. In the chromatogram, bases underlined in green represent the wild-type *NKX2-1* sequence, bases underlined in orange the *Alu* element, and bases underlined in blue the duplicated TSD. The sequence highlighted in yellow represents the TSD immediately preceding the *Alu* insertion site, which is duplicated after the poly-A tail and used by the *Alu* element to insert in the genome. The sequence whose reverse-complement counterpart was detected by SCRAMble is indicated by a white label. [Color figure can be viewed at wileyonlinelibrary.com]

performing a Blast comparison between the rebuilt sequence and RepeatMasker data.[24] Expasy Translate (https://web.expasy.org/translate/) allowed prediction of the NKX2-1 amino acid sequence generated by *Alu* insertion.[25] Formalin-fixed paraffin-embedded skin biopsies underwent p62 and NKX2-1 immunostaining. Real-time PCR (RT-PCR) of *NKX2-1* transcripts from whole blood, agarose gel electrophoresis of RT-PCR products, and immunoblotting and mass spectrometry (MS/MS) of proteins from lysate of whole-blood mononuclear cells isolated using Ficoll gradient centrifugation[26] (Ficoll-Paque PLUS, Cytiva, Marlborough, Massachusetts, USA) were performed.

Detailed clinical descriptions and methods are provided in the Supplementary Data.

## Results

Standard analysis of WES/SR-WGS data did not reveal candidate variants in genes associated with movement disorders (MDs).[22] Genetic kinship analysis inferred parent–offspring relationship between the proband and the known parents (kinship coefficient proband father: 0.246, proband mother: 0.247).[21] SCRAMble identified a 46-bp *Alu* sequence inserted onto the antisense strand of *NKX2-1* at genomic position chr14:36987132 (GRCh37). Twenty-seven reads covering the insertion site contained soft-clipped parts, forming a cluster with 100% consensus alignment quality to the *Alu* sequence.[19] The candidate MEI was visually identified through IGV in *NKX2-1* exon 3 in subjects II:2 and III:1 only (Fig. 1C,D).[23] Gel electrophoresis of PCR products revealed two bands in affected family members and one band in unaffected family members and in one healthy control (Fig. 2B). SS of the ~700-bp DNA fragment shared by all subjects corresponded to the wild-type *NKX2-1* sequence (Fig. 2C). Due to failure of SS, the ~1000-bp DNA fragment detected in II:2 and III:1 was sequenced after plasmid cloning, which revealed a 347-bp *Alu* element insertion with a 65-bp poly-A tail and 16-bp direct repeats on both sides of the element, corresponding to target site duplications (Fig. 2D).[27,28] This *Alu* sequence corresponds to an *Alu*Ya5 element with 0.3% divergence to the respective consensus sequence based on RepeatMasker.[24] Predicted mutant amino acid sequence due to *Alu* insertion revealed a premature stop codon 31 amino acids downstream of the insertion.[25]

Functional analyses are detailed in the Supplementary Data.

## Discussion

We identified a de novo *Alu* insertion in exon 3 of *NKX2-1* segregating with infancy-onset levodopa-responsive choreo-dystonia, respiratory and urinary dysfunctions, and MRI abnormalities of the pituitary in two first-degree relatives and with subclinical hypothyroidism in one of them. Deep phenotyping fueled firm clinical suspicion toward *NKX2-1*-related disorder and prompted additional genetic analyses despite previous extensive candidate-gene and hypothesis-free testing being unrevealing. The novel variant ENST00000354822.7 (*NKX2-1*):c.556_557ins*Alu*541_556dup p.(Leu186Argfs*32) is pathogenic per the ACMG/AMP guidelines (PVS1; PS2;PM2;PP1),[29] with the *Alu* sequence identifying an *Alu*Ya5 element.[24]

SINE-*Alu* elements are primate-specific retrotransposons originating from a 5′-to-3′ fusion of the *7SL* RNA gene and amplified throughout the human genome up to >1 million copies over ~65 million years.[27,30,31] Being non-autonomous, they require the enzymatic machinery of the autonomous LINE-1 retroelements for their intragenomic mobilization.[32] Although most *Alu* elements map to non-coding regions, mounting evidence that they can induce insertional mutagenesis and modulate gene expression posttranscriptionally exists, thus representing a wide source of potential disease-causing SVs and regulatory functions.[8,13-15,33] In particular, *Alu*Y retrotransposons are among the evolutionarily youngest *Alu* subfamily, including some mobilization-competent elements,[11] as proven by the new retrotransposition event characterized here.

Our study confirmed the clinical relevance of integrating dedicated algorithms for MEI detection in routine pipelines for SRS analysis.[17-20] Interestingly, SS of full-length *NKX2-1* had previously failed to detect MEI. First, the *Alu* insertion caused preferential amplification of the shorter wild-type allele resulting in mutant allele dropout, as proven by our multiple assays to optimize PCR conditions by increasing the amount of starting DNA, minimizing the cycle number, and extending the elongation time.[34] Second, SS of the *Alu* allele PCR product succeeded only after it was cloned in plasmids, suggesting that the *Alu* sequence generated secondary structures.

The mechanism by which this *Alu* insertion caused *NKX2-1* haploinsufficiency remains undetermined even after functional tests, most likely due to low-to-none *NKX2-1* expression in whole-blood cells and skin fibroblasts, as documented by public repositories (GTEx, The Human Protein Atlas).[35,36] *NKX2-1* is in fact highly and selectively expressed in the thyroid, lung, and pituitary gland, and, to a lesser extent, restricted brain regions, including the hypothalamus and basal ganglia,[35,36] which is in keeping with the pleiotropic functions of TTF1 and reflects the phenotypic spectrum of *NKX2-1*-related disorder.[37] In rodent and human thyroid and lung, expression of *NKX2-1* is consistent throughout life stages from embryonic development to adult tissues. On the contrary, *NKX2-1* expression is found in both diencephalic and telencephalic domains

during brain development but not in adult neurons of the basal ganglia.[38] Based on its rebuilt sequence, *Alu* insertion is predicted to introduce a premature stop codon 31 nucleotide triplet after the first codon change.[25] Because the *Alu* exonic insertion is ultimately analogous to a frameshift variant and located in the last coding exon of *NKX2-1*, the mutant mRNA is likely to evade nonsense-mediated decay and result in a non-functional truncated protein.[39] NKX2-1 was identified on immunoblotting only using Femto chemiluminescence using both NKX2-1 antibodies, which confirms extremely low protein concentrations in mononuclear cell lysate. This could explain why semiquantitative analysis of immunoblotting did not show statistically significant differences between affected and unaffected individuals and MS/MS failed to detect NKX2-1, which is expected to localize in the same spectrum as highly expressed housekeeping proteins (eg, actins) due to its molecular weight (42 kDa) and could therefore be masqueraded by their peak.

In conclusion, our study highlights the importance of deep phenotyping and clinicogenetic correlation to red-flag false negatives due to intrinsic limitations of genetic testing.[37] It supports the inclusion of dedicated MEI detection pipelines in routine SRS data analysis or the reanalysis of selected MD cases with targeted long-read sequencing to improve molecular diagnostic yield. ■

**Acknowledgments:** We are most grateful to the family discussed here for participating in this study. We would like to acknowledge the entire staff of Solve-RD as well as the entire staff of the Human Protein Atlas project and GTEx project. We thank Dr. David Pellerin (Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, University College London, United Kingdom) for his kind review of the final manuscript.

## Data Availability Statement

The data that supports the findings of this study are available in the supplementary material of this article and from the corresponding author upon reasonable request.

# References

1. Kimura S, Hara Y, Pineau T, et al. The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. Genes Dev 1996;10(1):60–69.
2. Thorwarth A, Schnittert-Hubener S, Schrumpf P, et al. Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. J Med Genet 2014;51(6):375–387.
3. Patel NJ, Jankovic J. NKX2-1-related disorders. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. GeneReviews®. University of Washington, Seattle, WA; 1993-2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK185066/.
4. Inzelberg R, Weinberger M, Gak E. Benign hereditary chorea: an update. Parkinsonism Relat Disord 2011;17(5):301–307.
5. Peall KJ, Kurian MA. Benign hereditary chorea: an update. Tremor Other Hyperkinet Mov 2015;5:314.
6. Thust S, Veneziano L, Parkinson MH, et al. Altered pituitary morphology as a sign of benign hereditary chorea caused by TITF1/NKX2.1 mutations. Neurogenetics 2022;23(2):91–102.
7. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet 2011;7(12):e1002384.
8. Solyom S, Kazazian HH Jr. Mobile elements in the human genome: implications for disease. Genome Med 2012;4(2):12.
9. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet 2009;10(10):691–703.
10. Brouha B, Schustak J, Badge RM, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci U S A 2003;100(9):5280–5285.
11. Bennett EA, Keller H, Mills RE, et al. Active Alu retrotransposons in the human genome. Genome Res 2008;18(12):1875–1883.
12. Wang H, Xing J, Grover D, et al. SVA elements: a hominid-specific retroposon family. J Mol Biol 2005;354(4):994–1007.
13. Kaer K, Speek M. Retroelements in human disease. Gene 2013;518(2):231–241.
14. Hancks DC, Kazazian HH Jr. Roles for retrotransposon insertions in human disease. Mobile DNA 2016;7:9.
15. Larsen PA, Hunnicutt KE, Larsen RJ, Yoder AD, Saunders AM. Warning SINEs: Alu elements, evolution of the human brain, and the spectrum of neurological disease. Chromosome Res 2018;26(1–2):93–111.
16. Stewart C, Kural D, Stromberg MP, et al. A comprehensive map of mobile element insertion polymorphisms in humans. PLoS Genet 2011;7(8):e1002236.
17. Tattini L, D'Aurizio R, Magi A. Detection of genomic structural variants from next-generation sequencing data. Front Bioeng Biotechnol 2015;3:92.
18. Torene RI, Galens K, Liu S, et al. Mobile element insertion detection in 89,874 clinical exomes. Genet Med 2020;22(5):974–978.
19. Demidov G, Park J, Armeanu-Ebinger S, et al. Detection of mobile elements insertions for routine clinical diagnostics in targeted sequencing data. Mol Genet Genomic Med 2021;9(12):e1807.
20. Niu Y, Teng X, Zhou H, et al. Characterizing mobile element insertions in 5675 genomes. Nucleic Acids Res 2022;50(5):2493–2508.
21. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26(22):2867–2873.
22. Zurek B, Ellwanger K, Vissers L, et al. Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. Eur J Hum Genet 2021;29(9):1325–1331.
23. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013;14(2):178–192.
24. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013-2015.
25. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 2003;31(13):3784–3788.
26. Fuss IJ, Kanof ME, Smith PD, Zola H. Isolation of whole mononuclear cells from peripheral blood and cord blood. Curr Protoc Immunol 2009;85(1):7.1.1–7.1.8.
27. Hasler J, Strub K. Alu elements as regulators of gene expression. Nucleic Acids Res 2006;34(19):5491–5497.
28. Kim S, Cho CS, Han K, Lee J. Structural variation of Alu element and human disease. Genomics Inform 2016;14(3):70–77.
29. 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.
30. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. Nat Rev Genet 2002;3(5):370–379.

31. Kriegs JO, Churakov G, Jurka J, Brosius J, Schmitz J. Evolutionary history of 7SL RNA-derived SINEs in Supraprimates. Trends Genet 2007;23(4):158–161.

32. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet 2003;35(1): 41–48.

33. Deininger P. Alu elements: know the SINEs. Genome Biol 2011; 12(12):236.

34. Walsh PS, Erlich HA, Higuchi R. Preferential PCR amplification of alleles: mechanisms and solutions. PCR Methods Appl 1992;1(4): 241–250.

35. Consortium GT. The genotype-tissue expression (GTEx) project. Nat Genet 2013;45(6):580–585.

36. Thul PJ, Lindskog C. The human protein atlas: a spatial map of the human proteome. Protein Sci 2018;27(1):233–244.

37. Magrinelli F, Balint B, Bhatia KP. Challenges in Clinicogenetic correlations: one gene - many phenotypes. Mov Disord Clin Pract 2021;8(3):299–310.

38. Krude H, Schutz B, Biebermann H, et al. Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. J Clin Invest 2002;109(4):475–480.

39. Brogna S, Wen J. Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol 2009;16(2):107–113.

## Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.