



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/235293/>

Version: Published Version

Article:

Fasham, J., Rankin, J., Schot, R. et al. (2026) Palindrome-mediated 16p13.3 triplications cause a recognizable neurodegenerative disorder with ataxia. *The American Journal of Human Genetics*, 113 (1). pp. 221-233. ISSN: 0002-9297

<https://doi.org/10.1016/j.ajhg.2025.11.011>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

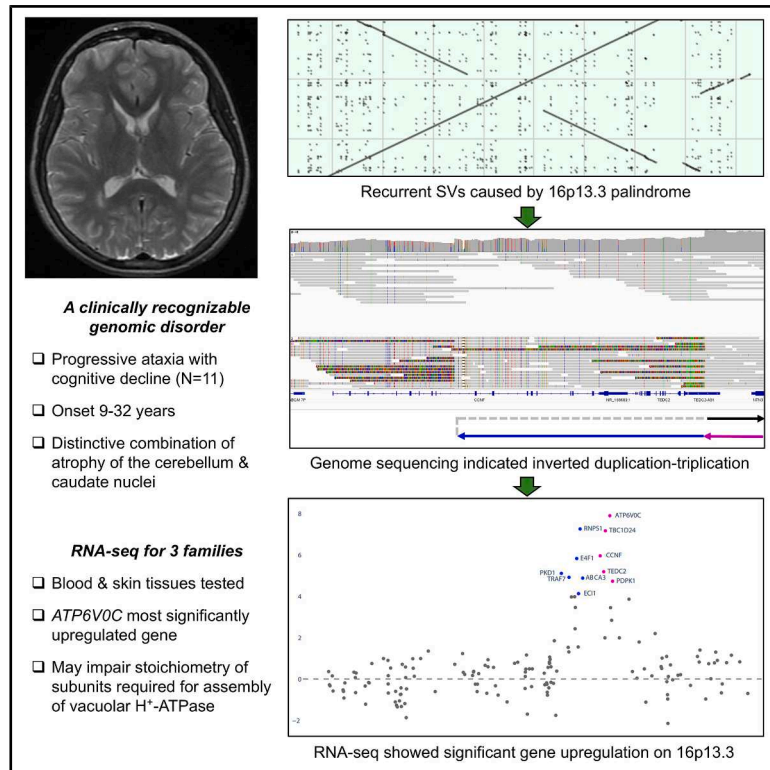
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Palindrome-mediated 16p13.3 triplications cause a recognizable neurodegenerative disorder with ataxia

Graphical abstract



Authors

James Fasham, Julia Rankin, Rachel Schot, ..., Andrew H. Crosby, Emma L. Baple, Alistair T. Pagnamenta

Correspondence

e.baple@exeter.ac.uk (E.L.B.),
a.pagnamenta@exeter.ac.uk (A.T.P.)

Fasham and colleagues report a palindrome-mediated genomic disorder causing a recognizable, severe phenotype marked by early-onset progressive ataxia, cognitive decline, and cerebellar atrophy. Microarray and short/long-read genome sequencing uncovered overlapping inverted duplication-triplications on 16p13.3. RNA-seq showed upregulation of multiple genes within the variant, with *ATP6V0C* showing the most significant increase.

Palindrome-mediated 16p13.3 triplications cause a recognizable neurodegenerative disorder with ataxia

James Fasham,^{1,2,31} Julia Rankin,^{2,31} Rachel Schot,³ Susan M. White,^{4,5} Katrina M. Bell,^{4,6} Matthew N. Wakeling,¹ Lucy J. Mallin,⁷ Alex Shah,⁸ Michelle G. de Silva,^{4,6} David I. Francis,⁴ Maie Walsh,⁹ Emily E. Jones,¹⁰ Kayal Vijayakumar,¹¹ Katie Johnson,¹² Francis H. Sansbury,^{13,14} Johann te Water Naudé,¹⁵ Paola Giunti,¹⁶ Marios Hadjivassiliou,¹⁷ Andrea H. Nemeth,^{18,19} George K. Tofaris,^{18,20} Carlo Rinaldi,^{20,21} Benito Banos-Pinero,²² Marianna Selikhva,²³ Nishanka Ubeyratna,¹ Anneke Kievit,³ Frank Sleutels,³ Joey van Giessen,³ Tahsin Stefan Barakat,³ Timothy S. Hall,¹ Alan Whone,²⁴ Eleanor Thomas,²⁵ Joseph S. Leslie,¹ Rosemary A. Bamford,¹ Aaron R. Jeffries,²⁶ Jenny Lord,²⁷ Susan Walker,²⁸ Tjakko J. van Ham,³ Sue L. Hill,²⁹ Lucy McGavin,^{8,30} Andrew Parrish,⁷ Andrew H. Crosby,¹ Emma L. Baple,^{1,2,32,*} and Alistair T. Pagnamenta^{1,32,*}

Summary

Complex neurodegenerative conditions have occasionally been associated with copy-number gains. Using microarray and genome sequencing on DNA samples from eleven individuals from nine unrelated families, we show that copy-number gains at 16p13.3 cause a severe, recognizable disorder characterized by early-onset progressive ataxia and cognitive decline (9–32 years). Most affected individuals also displayed peripheral neuropathy and scoliosis. Optic atrophy, nystagmus, and dystonia were more variable features. The neuroradiological phenotype comprises a distinctive combination of atrophy of the cerebellum and caudate nuclei. Co-segregation data showed that the structural variant (SV) had occurred *de novo* in 5 individuals and, for one other individual, had been inherited from a mosaic, unaffected parent. Triplicated segments of 16p13.3 were identified within the duplications. Although these varied in size (30–811 kb), the minimal region of overlap included a single gene (*ATP6VOC*) that is highly expressed in the cerebellum. RNA sequencing (RNA-seq) using whole-blood and fibroblast/lymphoblast cultures indicated increased expression of several genes within the SV, with *ATP6VOC* showing the most significant increase (up to 4-fold). In most cases, the central segment of the SV was proven to be inverted and lay immediately distal to a 144 kb palindrome. Across 500,000 individuals from the UK Biobank, we identified 19 duplications but no triplications at this locus. Further analysis of the consequences of *ATP6VOC* overexpression on the stoichiometry within the vacuolar H⁺-ATPase heteromer and on neurological function will provide valuable pathomechanistic insights. Together, our findings define palindrome-mediated triplication on 16p13.3 as the cause of a clinically distinct childhood-onset neurodegenerative disorder.

Childhood-onset neurodegenerative disorders comprise a heterogeneous group of rare conditions that present with progressive neurological and/or cognitive impairment.¹ Features can include regression of motor or cognitive

skills, seizures, behavioral changes, ataxia, spasticity, sensory deficits, and intellectual disability. Cerebellar atrophy is a frequent finding with a wide differential diagnosis, including mitochondrial disorders, such as Leigh

¹Department of Clinical and Biomedical Sciences, Faculty of Health and Life Sciences, University of Exeter, EX2 5DW Exeter, UK; ²Department of Clinical Genetics, Royal Devon University Hospital, EX1 2ED Exeter, UK; ³Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; ⁴Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, Australia; ⁵Department of Paediatrics, University of Melbourne, Parkville, VIC, Australia; ⁶Murdoch Children's Research Institute, Melbourne, VIC, Australia; ⁷Exeter Genomics Laboratory, Royal Devon University Healthcare NHS Foundation Trust, EX2 5DW Exeter, UK; ⁸University Hospitals Plymouth NHS Trust, Derriford Road, Crownhill, Plymouth, PL6 8DH Devon, UK; ⁹Genomic Medicine, the Royal Melbourne Hospital, Parkville, VIC, Australia; ¹⁰Bristol Genetics Laboratory, North Bristol NHS Trust, BS10 5NB Bristol, UK; ¹¹Department of Paediatric Neurology, University Hospitals Bristol NHS Foundation Trust, BS1 3NU Bristol, UK; ¹²Nottingham Regional Genetics Service, Nottingham City Hospital Campus, The Gables, NG5 1PB Nottingham, UK; ¹³All Wales Medical Genomics Service, NHS Wales Cardiff and Vale University Health Board, Wales Genomic Health Centre, Cardiff Edge Business Park, Longwood Drive, Whitchurch, CF14 7YU Cardiff, UK; ¹⁴Division of Cancer & Genetics, School of Medicine, Cardiff University, CF14 4XN Cardiff, UK; ¹⁵Department of Paediatric Neurology, University Hospital of Wales, CF14 4XW Cardiff, UK; ¹⁶Queen Square Institute of Neurology, WC1N 3BG London, UK; ¹⁷Academic Department of Neurosciences, Sheffield Teaching Hospitals NHS Trust and University of Sheffield, S10 2JF Sheffield, UK; ¹⁸Nuffield Department of Clinical Neurosciences, University of Oxford, OX3 9DU Oxford, UK; ¹⁹Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, OX3 7LD Oxford, UK; ²⁰Department of Clinical Neurology, Oxford University Hospitals NHS Foundation Trust, OX3 9DU Oxford, UK; ²¹Institute of Developmental and Regenerative Medicine, University of Oxford, OX3 7TY Oxford, UK; ²²Oxford Genetics Laboratories, Oxford University Hospitals NHS Foundation Trust, OX3 9DU Oxford, UK; ²³Neurology Department, Southmead Hospital, BS10 5NB Bristol, UK; ²⁴Southmead Hospital, BS10 5NB Bristol, UK; ²⁵Department of Child Health, Royal Devon University Hospital, EX1 2ED Exeter, UK; ²⁶Biosciences, University of Exeter, EX1 2LU Exeter, UK; ²⁷Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, S10 2HQ Sheffield, UK; ²⁸Genomics England, EC1M 6BQ London, UK; ²⁹NHS England, London, UK; ³⁰University of Plymouth, PL4 8AA Plymouth, UK

³¹These authors contributed equally

³²These authors contributed equally

*Correspondence: e.baple@exeter.ac.uk (E.L.B.), a.pagnamenta@exeter.ac.uk (A.T.P.)

<https://doi.org/10.1016/j.ajhg.2025.11.011>

© 2025 The Authors. Published by Elsevier Inc. on behalf of American Society of Human Genetics.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

syndrome (MIM: 256000 and 500017),² ataxia telangiectasia (MIM: 208900),³ disorders of neurodegeneration with brain iron accumulation (NBIA),⁴ and others. These conditions are commonly autosomal-recessive or X-linked disorders, caused by loss of function genetic variants. Chromosomal gains involving triplosensitive genes are a rare cause of progressive childhood-onset neurological disorders. Notable examples include whole-gene duplication or triplication of *PLP1* (MIM: 300401) in Pelizaeus-Merzbacher disease (MIM: 312080)⁵ and duplication or, rarely, triplication of *PMP22* (MIM: 601097) in Charcot-Marie-Tooth 1A (MIM: 118220).^{6,7}

Technologies such as microarray and short-read genome sequencing (GS) are well suited to detecting microdeletions and microduplications but often fail to identify complex structures. In some cases, this can lead to diagnostic structural variants (SVs) being overlooked. Advances in computational algorithms for copy-number variant (CNV) and SV detection from GS and long-read sequencing now enable improved characterization of more complex chromosomal rearrangements. One notable example of such a rearrangement is the duplication-triplication. These configurations typically have an inverted central segment and are often mediated by oppositely oriented segmental duplications (SDs).⁸ Complex SVs can also recur at palindromic DNA sequences, as these can promote double-strand breaks. Here, we used a combination of microarray, exome, GS, and RNA sequencing (RNA-seq) technologies to define palindrome-mediated duplication-triplications on 16p13.3 as the cause of a clinically and radiologically distinct early-onset neurodegenerative disorder with ataxia.

We initially identified individual 1, who presented at 10 years of age with childhood-onset, severe, and progressive ataxia, associated with axonal neuropathy and neuroradiological evidence of reduced cerebellar volume (Note S1). Array-based comparative genome hybridization (aCGH) analysis (8x60k constitutive v.3.0, Oxford Gene Technology) followed by parental segregation studies identified a rare *de novo* CN gain of 448 kb involving 16p13.3. The *de novo* nature of this variant prompted us to seek further individuals with similar chromosomal gains through international collaboration, using DECIPHER, GeneMatcher,⁹ and the Genomics England National Genomic Research Library (NGRL; ethical approval from Cambridge South/Central RECs, 14/EE/1112, 20/EE/0035, and 25/EE/0125). Our analysis used Manta and Canvas calls for 100,000 Genomes Project (100kGP) participants, which were analyzed using SVRare.^{10–12} For NHS Genomic Medicine Service (GMS) data, Dragen CNV calls were assessed with custom scripts. Genomic coordinates are reported using GRCh38.

In total, after resolving duplicate individuals identified across studies (Note S2), we ascertained 11 affected individuals from 9 distinct families (Figure 1A) with overlapping CN gains involving 16p13.3, progressive ataxia, and cerebellar atrophy. Phenotype information was ob-

tained with informed consent by clinical care providers using a standardized proforma. Clinical findings for the 11 affected individuals are summarized in Table 1 and Note S1.

Affected individuals exhibited a shared phenotype characterized by progressive ataxia and cognitive decline. Ataxia, sometimes manifesting with dysarthria (10/11) and nystagmus (4/11), was usually noted in the second decade (range: 9–32 years). Mobility was typically affected, with a walking aid or wheelchair required in the third or fourth decade. Cognitive impairment was also universally present, with the majority (8/10) demonstrating progressive decline. This cognitive deterioration was often accompanied by behavioral challenges and sometimes required supportive care. In individuals 4, 7, and 8, either cognitive delay or ataxia predominated initially, with the other symptoms emerging later. Peripheral neuropathy, typically axonal and sensory predominant, was also common (9/11), as were absent deep tendon reflexes (9/10), *pes cavus* (5/6), and scoliosis/kyphosis (7/10). Additional, more variable neurological features, such as dystonia (2/10) and spasticity (1/9), were also observed. There was no craniofacial dysmorphism and no confirmed clinical or electrophysiological seizures reported. MRI neuroimaging findings were striking and consistent across all affected individuals, with cerebellar volume loss, caudate atrophy, and a high T2/FLAIR signal in the basal ganglia (Figures 1B–1I and S1). Progressive volume loss was also observed where multiple images were available for review (Figure 1E).

The pathogenicity of the overlapping gains identified in these 9 families is supported by inheritance studies (Figure 1A), with 5 confirmed *de novo* (inheritance not determined for 2 families). Family 1 comprised 3 affected individuals. A gain on 16p13.3 was identified through GS in the proband (III-3), and co-segregation studies showed the presence of the SV in her similarly affected brother (III-2) but not in the eldest unaffected sibling (III-1). The SV had been inherited from the affected mother (II-2), but as DNA was unavailable for the maternal grandfather, the inheritance for the mother is unknown.

Across the cohort, SV sizes ranged from 157 to 918 kb (Tables 1 and S1; Figure 2A). In all 7 families where short-read GS data were available (individuals 2 and 4–8 and family 1), the SV comprised a triplicated segment embedded within a duplication. As a representative example, IGV screenshots showing GS read alignments for individual 7 are shown in Figure 2B. Read alignments for other affected individuals are shown in Figures S2–S7. In all but one case (individual 5; Figure S4), split read pairs at the distal end (both mapping to the negative strand) suggest that the central portion of the rearrangement has been inverted. For individual 5, the duplication-triplication was also present in the unaffected father in a mosaic state (Figure S4), estimated to be heterozygous in 34%–37% of blood cells (Note S3).

Our systematic analysis of SVs in the 100kGP data strongly suggested that the presence of a 4-copy region is

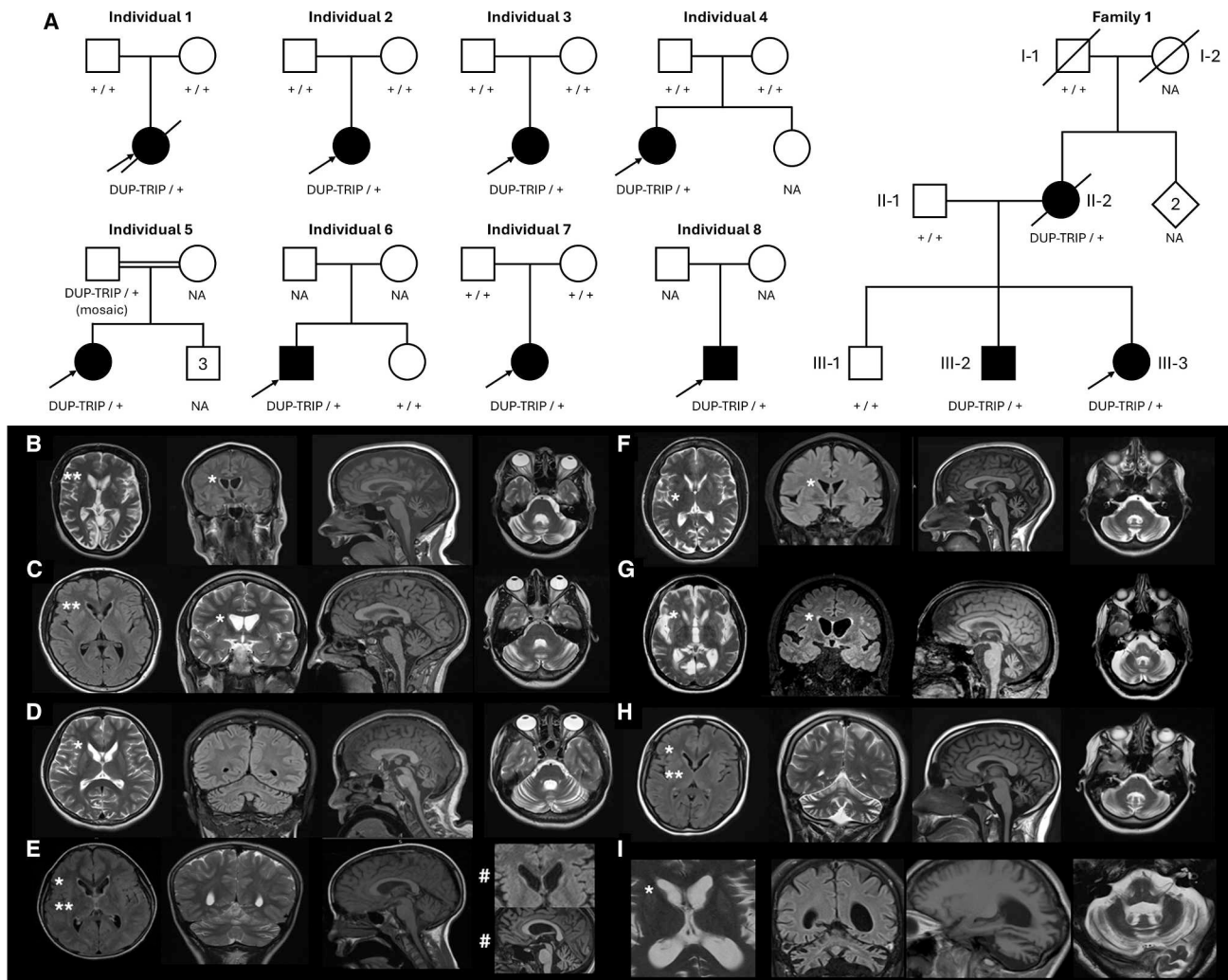


Figure 1. Family structures and MRI data for 11 individuals with overlapping gains on 16p13.3

(A) Pedigrees for 8 unrelated individuals and one multiplex family with overlapping SVs on 16p13.3. Individuals 1–3 were ascertained locally and via the DECIPHER database. Individuals 4–8 were identified from 100kGP and GMS data available in the National Genomic Research Library. Family 1 was ascertained via GeneMatcher, where *ATP6VOC* and *PDPK1* were entered as the candidate genes of interest. + represents the normal copy of 16p13.3.

(B–I) MRI scans from individuals 1–8 acquired at ages 36, 18, 9, 10 and 20 (#), 31, 53, 20, and 29. Axial and coronal T2 or FLAIR images demonstrate caudate atrophy of varying degrees (*) in all individuals and a high T2/FLAIR signal in the basal ganglia (**) in individuals 1, 2, 4, and 7. Sagittal T1 and axial T2 images (columns 3 and 4) show the cerebellar hemisphere and vermal atrophy in all individuals. There is progressive caudate and cerebellar atrophy in individual 4 (row E). For further MRI images, see [Figure S1](#).

critical for the early-onset progressive ataxia phenotype ([Note S4](#)), and this was supported by the results for individuals 4–8 and family 1. This prompted closer scrutiny of the existing data for individuals 1–3. The datasets available for individuals 1 and 3 were limited to exome sequencing and aCGH. Lower-resolution aCGH data could not distinguish between duplicated and hypothetically triplicated regions, and data from a ~60,000-resolution array for individual 3 are shown as a representative example in [Figure S8](#). In contrast, retrospective coverage analysis of the previously uninformative exome data was consistent with segments of triplication embedded within the respective duplications ([Figure S9](#)). For individual 2, repeat analysis using a higher-resolution array with ~1.8 M probes identified a nested region of triplication ([Figure S10](#)), undetected by the original

SNP microarray. This was subsequently confirmed by short-read GS ([Figure S2](#)). In individual 7, multiplex ligation-dependent probe amplification testing prior to 100kGP recruitment yielded fold change (FC) results intermediate between CNs 3 and 4 ([Figure S11](#)). These data highlight the difficulties of using non-GS methodologies to distinguish between duplicated and triplicated segments of under 500 kb.

The triplicated segments within the 16p13.3 SVs varied in size from 30 to 811 kb across the 11 individuals reported here ([Table 1](#)). The minimal region of triplication overlap was shown to be 26.1 kb and defined by the distal breakpoint in individual 5 and the proximal breakpoint in individual 1 (g.2503226–2529370 [GenBank: NC_000016.10]; [Figure 2B](#); [Table S1](#)). A single protein-coding gene, *ATP6VOC*

Table 1. Clinical and neuroradiological findings in individuals with overlapping duplication/triplications involving 16p13.3

Individual ID (DECIPHER ID)	Individual 1 (381641)	Individual 2 (339803)	Individual 3 (359492)	Individual 4 (341728)	Individual 5	Individual 6	Individual 7	Individual 8	Family 1, II-2	Family 1, III-2	Family 1, III-3
Age at last assessment	died at 41 y	27 y	17 y	20 y	40 y	65 y	36 y	40 y	died at 56 y	36 y	27 y
Ancestry ^a	British	Australian ^b	Sri Lankan	British	Pakistani	British	Irish	British	Dutch	Dutch	Dutch
Sex	female	female	female	female	female	male	female	male	female	male	female
SV detection method	array + ES	array + GS	array, ES, LR	array + GS	GS	GS	GS + LR	GS	array + GS	array	array + GS
Duplication size	448 kb	704 kb	640 kb	918 kb	230 kb	157 kb	187 kb	378 kb	514 kb	514 kb	514 kb
Triplication size	30 kb	179 kb	317 kb	811 kb	64 kb	85 kb	105 kb	296 kb	250 kb	250 kb	250 kb
Inheritance	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	<i>de novo</i>	father mosaic	NK	<i>de novo</i>	NK	NK ^c	maternal	maternal
Clinical features											
Cognitive impairment	yes (<10 y)	mild (5 y)	yes, special education	mild/Mod. with ADHD	yes	significant	moderate	yes, initially normal	yes	yes, special education	yes, special education
Regression (8/10)	yes (in 20s)	no	yes	no	NK	yes	yes	yes (16 y)	yes (9 y)	yes	yes
Progressive ataxia ^d	yes (10 y)	yes (13 y)	yes (9 y)	yes (15 y)	yes	yes (32 y)	yes	yes (23 y)	yes (9 y)	yes (12 y)	yes (12 y)
Dysarthria (10/11)	yes	yes	yes	no	yes	yes	yes	yes	yes	yes (15 y)	yes
Nystagmus (4/11)	no	no	no	yes (15 y)	no	yes	yes	probable	no	no	no
Axonal neuropathy (9/11)	sensory (17 y)	no	yes	no	sensory	sensory	yes	sensory	sensory (46 y)	sensory (15 y)	sensory (9 y)
<i>Pes cavus</i> (5/6)	yes	NK	yes	no	NK	yes	NK	yes	NK	NK	yes
Dystonia (2/10)	no	no	no	no	yes	no	yes	no	no	NK	no
Scoliosis/kyphosis (7/10)	no	scoliosis (10 y)	scoliosis	scoliosis (14 y)	scoliosis	no	NK	scoliosis	kyphosis	no	scoliosis
Optic atrophy (2)	NK	NK	NK	NK	NK	yes	NK	no	yes	NK	NK
Other features	seizures suspected	hearing loss	hypertonia	macrocephaly clinodactyly	cataract, CVI, dysphagia	seizures suspected	chorea	–	spasticity	–	hypertonia, strabismus, hearing loss
Neuroradiology											
Cerebellar and caudate atrophy (11/11)	yes (35 y, 36 y)	yes (18 y, 23 y)	yes (9 y)	progressive (10 y, 20 y)	yes (31 y)	yes (53 y)	yes (20 y)	yes (29 y)	yes	yes (15 y)	yes (9 y)
Other findings	–	–	–	syrinx T4–T9	–	global atrophy	–	cerebral volume loss	–	–	–

The age at which the phenotype was first recorded is shown in parentheses. CNV sizes are approximate, as in most cases, there remains some uncertainty about the precise coordinates, particularly at the proximal breakpoints. Full genomic coordinates are shown in Table S1. Abs, absent; ADHD, attention-deficit hyperactivity disorder; CVI, cortical visual impairment; ES, exome sequencing; GS, genome sequencing; LR, long-read genome sequencing; Mod., moderate; y, years; NK, not known.

^aAncestry is as per self-report.

^bEuropean-Australian.

^cPresumed *de novo*; not paternally inherited and no DNA available from the deceased mother.

^dAtaxia was diagnosed during childhood in most patients. In others, the diagnosis was later, although it is possible that childhood symptoms preceded any formal diagnosis.

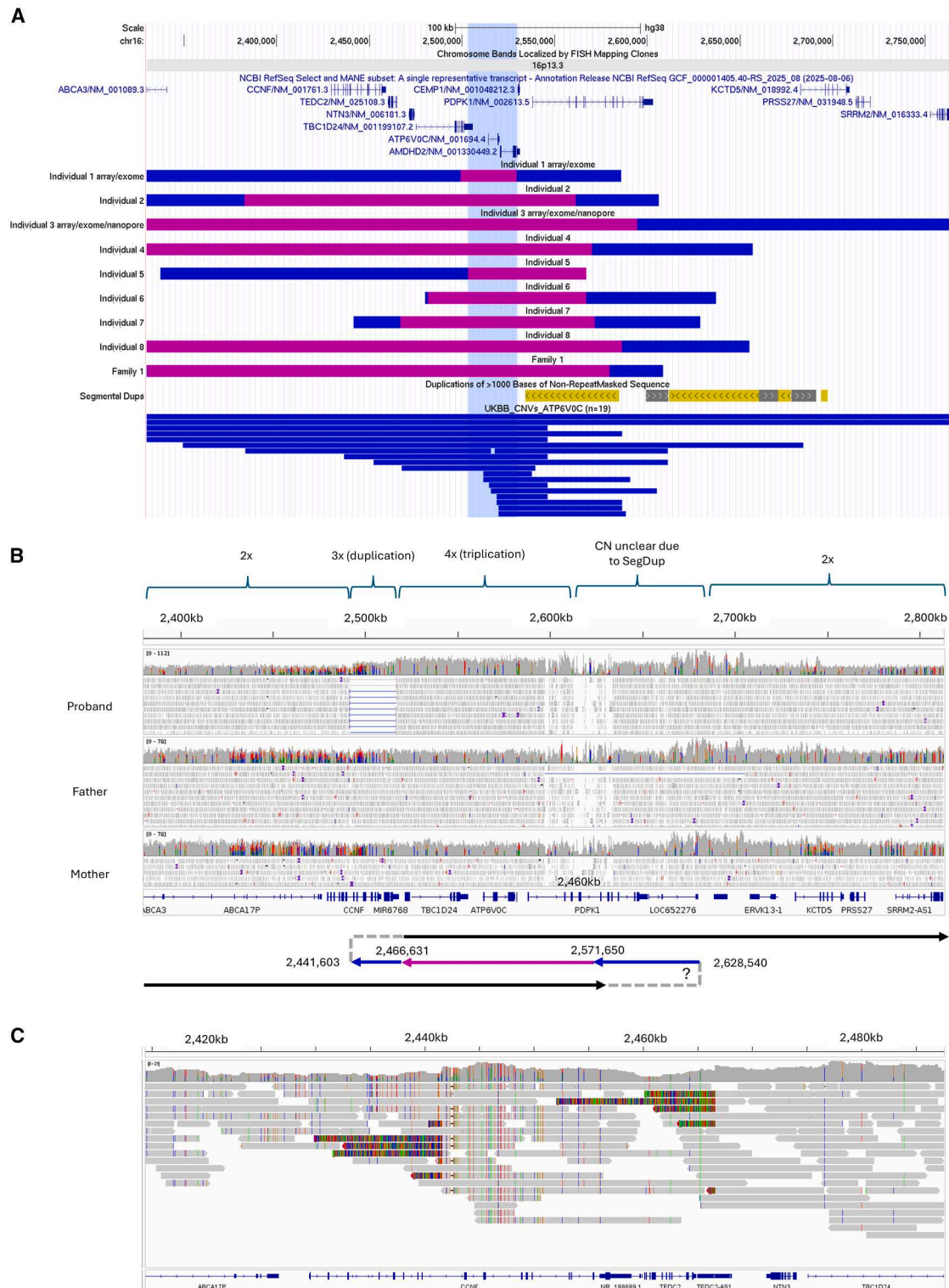


Figure 2. Genetic findings across 11 individuals with overlapping SVs on 16p13.3

(A) Custom UCSC browser session for the region chr16:2329790–2762907 showing relative positions of SVs across individuals 1–8 and family 1. Blue shading denotes duplication, and purple shading represents triplicated segments. For individuals 1 and 3, the array datasets available are of low resolution, so the presence of triplication is confirmed only by exome data and, hence, is approximate. The minimal region of triplication overlap is highlighted in light blue and overlaps a single protein-coding gene in its entirety, *ATP6VOC*. Segmental duplications are shown below the custom SV track and where yellow shading denotes similarity of 98%–99%. The bottom track shows 19 duplications identified in the UK Biobank that overlapped *ATP6VOC*. An interactive version of this image is available at <https://genome.ucsc.edu/s/ExeterGenetics/16p13.3-TRIPv9>.

(B) Short-read alignments shown in IGV that support the presence of the SV in individual 7, chosen as a representative example. The x axis coordinates correspond to the same region shown in (A) (coordinates lifted over to hg19). Regions of 3× and 4× are labeled, and

(legend continued on next page)

(MIM: 108745), lies fully within this region. *ATP6VOC* is highly expressed in the cerebellum, with a median “transcripts per million” (TPM) of 643.3 (GTEx data, accessed April 24, 2025). Other genes partially within or nearby the shared triplicated region include *TBC1D24* (MIM: 613577), *AMDHD2* (MIM: 620864), and *CEMP1* (MIM: 611113), and the evidence supporting the relative strength of candidacy for each of these is shown in [Table S2](#).

Due to the presence of the 144 kb palindrome-like repeat, precise proximal breakpoint locations remain uncertain in most cases. The exception to this is the multiplex kindred (family 1), where, despite low mapping quality, split read pairs mapping to the positive strand helped resolve the proximal breakpoint junction ([Figure S12](#)). We next undertook PacBio HiFi GS for individual 7 using the Revio instrument, seeking to resolve more precisely the structure of these complex palindrome-mediated SVs. Analysis used the HiFi-human-whole-genome sequencing (WGS)-Workflow Description Language (WDL) pipeline ([web resources](#)). We obtained a mean coverage of 36 \times and validated the distal breakpoints that had previously been identified using short reads ([Figure 2C](#)). Although SV calling with HiFiCNV (v.1.0.0) and pbsv (v.2.9.0) identified a 120 kb duplication (g.2460001–2580000 [GenBank: NC_000016.10], imprecise) and a 25 kb inversion (g.2441607–2466627 [GenBank: NC_000016.10]) at the distal end, the proximal breakpoints remained algorithmically unresolved. Assigning read alignments into phase groups with HiPhase¹³ and *de novo* assembly with hifiasm¹⁴ was also unable to resolve the proximal end. However, by analyzing the phase of two *cis*-morphisms that are found 4.2 kb apart in the middle of the highly homologous repeat ([Figure S13](#)), we inferred approximate proximal breakpoint positions ([Table S1](#)). Similarly, we undertook nanopore sequencing for individual 3 according to the SQK-LSK110 protocol ([Note S5](#)) but obtained a mean genome-wide coverage of only $\sim 5\times$. Although the distal breakpoint of the SV was identified ([Figure S14](#)), there were only four supporting reads, ranging from 849 bp to 35.9 kb in size. Due to higher per-base error rates with nanopore sequencing, *cis*-morphism analysis was not attempted, so the precise variant structure remains unclear.

To determine the likely transcriptional effect of these SVs, RNA-seq was undertaken for two affected individuals (III-3 in family 1 and individual 2), with consistent findings across these and a third already available RNA-seq dataset (described below). Fibroblast-derived RNA was utilized for family 1, III-3. Data analysis employed an adapted version of the OUTRIDER method described previously¹⁵ and a transcript set based on RefSeq MANE v.1.3 (GRCh38). *Z* scores, FCs, and *p* values were calcu-

lated by comparing results with 133 control individuals, all unrelated patients for whom RNA-seq-based diagnostic testing had been requested at the Erasmus MC Clinical Genetics department. *ATP6VOC* was the most significantly dysregulated gene ([Figures 3A](#) and [3B](#)), with a *Z* score of 7.90, an adjusted *p* value of 1.42×10^{-13} , and a FC of 1.83. Of the 15 upregulated genes with *Z* scores > 4 and *p* < 0.0025 , 11 were localized within the 16p13.3 duplication/triplication ([Table S3](#); [Figure S15](#)). RNA-seq data also identified a fusion transcript involving *PKD1* (MIM: 601313) and *ABCA3* (MIM: 601615) ([Figures S16](#) and [S17](#)). However, given the variable distal breakpoints of overlapping SVs in this cohort, the fusion transcript’s significance is unclear.

RNA-seq analysis for individual 2 used cultured lymphoblast-derived RNA. OUTRIDER analysis was performed across 114 samples sent in for RNA-seq-based diagnostics to identify genes with significant expression differences compared with the rest of the cohort. While the initial analysis used basic Gencode annotations and yielded non-significant results for *ATP6VOC* (FC = 2.10, *Z* score = 2.34), the results were confounded by Gencode annotations for fusion genes (Gencode: ENSG00000260272 and ENSG00000259784) that overlap the primary transcript for *ATP6VOC* (Gencode: ENST00000330398). Repeating this analysis using just the “MANE select” annotations, *ATP6VOC* became the most significantly upregulated gene (FC = 1.95, *Z* score = 6.86; [Figures S18](#) and [S19](#)). Overall, 15/16 of the upregulated genes are from the inverted duplication-triplication region ([Table S4](#)). A split read pair with similar genomic coordinates to those seen in the GS data confirmed the distal breakpoint ([Figure S2](#)).

RNA-seq data were also available for individual 7 as part of the Genomics England transcriptomics study.¹⁶ In brief, blood-derived RNA for 5,546 probands was collected using archived PaxGene tubes taken at recruitment to the 100kGP. RNA samples were depleted for rRNA/globin and sequenced using 100 bp paired-end reads. Alignment and transcript quantification were performed using Illumina’s DRAGEN pipeline (v.3.8.4). Compared with the other 5,545 RNA-seq datasets and normalized to *ACTB* or *GAPDH*, the relative expression in this individual was 4.16–4.95 \times above the mean for *ATP6VOC* and 3.48–4.48 \times above the mean for *AMDHD2* ([Figures 3C](#), [3D](#), and [S20](#)). Accurate measurement of *CEMP1* expression was not possible with short-read RNA-seq, as the gene is embedded within the 3’ UTR of *AMDHD2*.

To investigate the incidence of similar 16q13.3 rearrangements in the general population, we interrogated UK Biobank (UKB) data, which comprises $\sim 500,000$ UK

split read pairs (blue) are seen on the distal end of the SV, where both reads map to the negative strand and coincide with the stepped increases in coverage. The patchiness seen at the proximal end of the SV is due to reads with low mapping quality, shown in white. The purple/blue arrows in the subway plot denote that the central segment of the SV has been inverted. Other hypothetical conformations are shown in [Figure 4](#).

(C) PacBio HiFi data for the same individual confirm the distal ends of the SV. Instead of split read pairs, the junction is visible due to soft-clipped bases present in a subset of reads. Data are shown according to the genome build GRCh38.

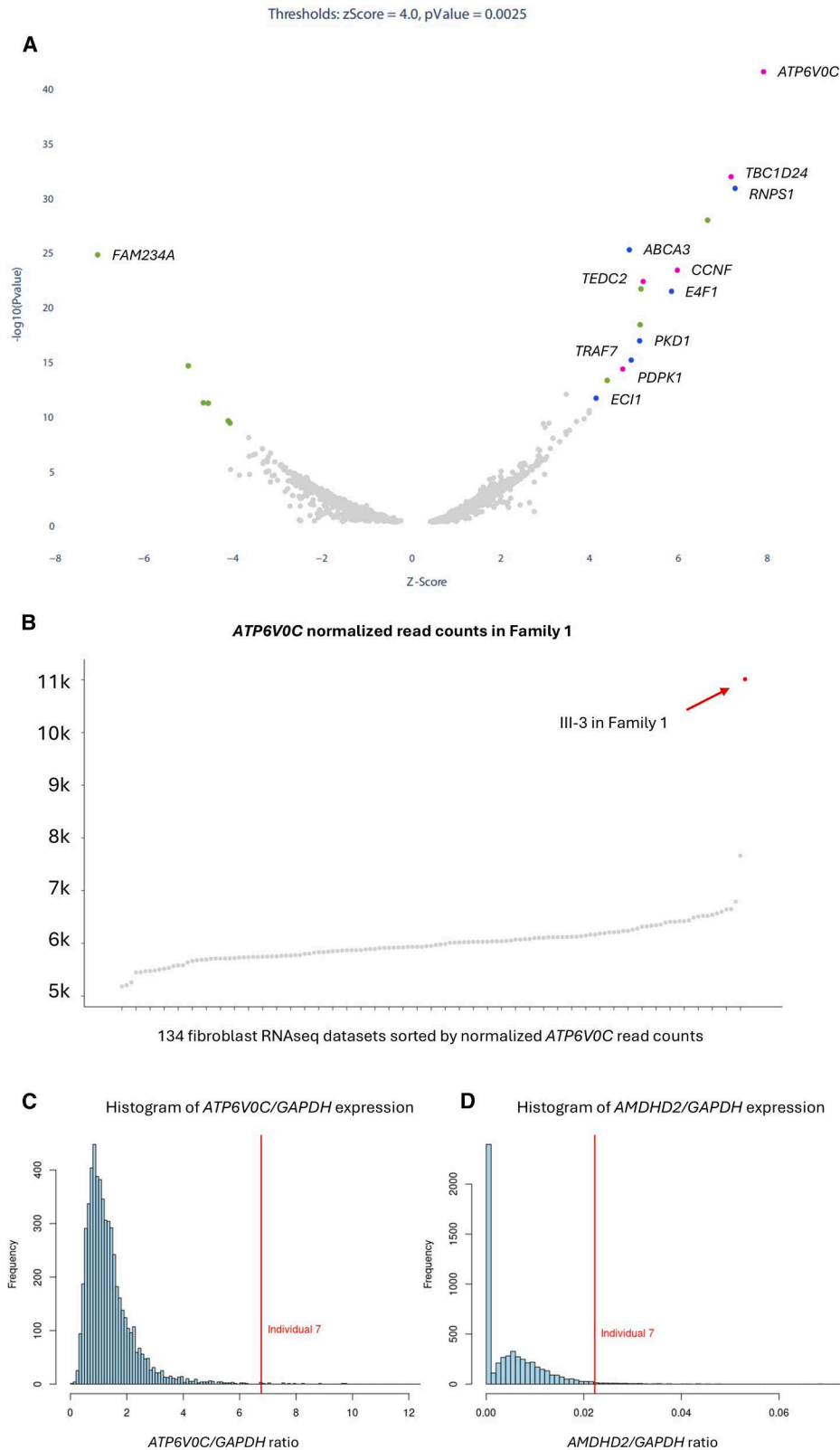


Figure 3. RNA-seq data for two individuals with 16p13.3 rearrangements confirm aberrant gene expression

(A) Genome-wide volcano plot showing genes that had significantly aberrant expression (Z score < -4 or > 4 , $p < 0.0025$) in individual III-3 from family 1. Significantly upregulated genes that lie inside or outside the 16p13.3 duplication/triplication are highlighted in blue/purple or green, respectively. The most significant result was *ATP6V0C*, which lies inside the triplicated segment and has a Z score of 7.90. Although *FAM234A* also lies on 16p13.3, it is not in the duplicated region, and the aberrant expression detected was more

(legend continued on next page)

individuals between the ages of 40 and 69 years at recruitment. Of these, 490,640 participants underwent GS, as previously described.¹⁷ CNVs were called using the DRAGEN v.3.7.8 CNV module, and data were searched using a custom RStudio script and a 10 kb size threshold. Nineteen individuals were identified harboring duplications of between 25.4 kb and 2.94 Mb that overlap *ATP6VOC* (Figure 2A). No triplications (i.e., a CN of 4) were identified, and segment CN ratios were 1.27–1.56. All of these individuals were healthy enough to be recruited to the UKB as adults. By reviewing hospital episode statistics, general practitioner (GP) records, and other self-reported clinical information, only one individual was identified who could plausibly be consistent with a milder form of the 16p13.3-triplication-linked condition, presenting with idiopathic familial dystonia and mild cognitive disorder (ICD-10 codes G24.1 and F06.7).

The clinical and genomic data described here define palindrome-mediated 16p13.3 triplication as the cause of a recognizable neurodegenerative disorder characterized by progressive ataxia and intellectual disability, with additional neurological features in many individuals. Cognitive impairment from childhood was universal and progressive in 80% of affected individuals. Notably, individual 1 developed severe, progressive dementia, raising the possibility of similar trajectories in other adult patients. Ataxia typically emerged during childhood or adolescence and progressed gradually. Peripheral neuropathy, spinal anomalies, and cerebellar signs, such as dysarthria and nystagmus, were frequent, though variably present. While some individuals had dystonia or spasticity, seizures were notably absent. Importantly, all patients shared a strikingly consistent MRI pattern, with cerebellar and caudate atrophy and basal ganglia signal abnormalities, supporting a shared, progressive neurodegenerative trajectory. This constellation of clinical and radiological features should prompt consideration of this disorder in undiagnosed individuals.

This condition is distinct among neurodegenerative ataxic disorders. While several conditions combine progressive ataxia with predominantly sensory axonal peripheral neuropathy (e.g., Friedreich ataxia [MIM: 229300],¹⁸ ataxia with vitamin E deficiency [MIM: 277460], and CANVAS [MIM: 614575]^{19,20}), none are typically accompanied by progressive cognitive decline. Progressive intellectual impairment does occur in multisystemic conditions such as mitochondrial diseases and Refsum disease (MIM: 266500), but these typically also manifest outside the nervous system (e.g., as retinitis pigmentosa or cardiac impairment).

Although recent advances have improved our understanding of the genetic basis of rare hereditary ataxias and neurodegenerative disorders, current diagnostic yields remain below 50%,²¹ suggesting that additional causative mechanisms remain to be identified. SVs, including copy gains, represent a class of genomic alterations that is technically challenging to detect and fully characterize and may account for a subset of undiagnosed individuals. Although CN gains have been found to cause complex neurodegenerative conditions, pathogenic triplications are relatively sparse in the literature. In many cases where triplication is a disease mechanism (*PLP1*, Pelizaeus-Merzbacher disease²²; *SNCA*, Parkinson disease [MIM: 605543]; 17q22, retinitis pigmentosa [MIM: 600852]²³; and *APP*, Alzheimer disease [MIM: 104300]²⁴), the phenotype is more severe than the equivalent condition caused by duplication of the same locus. The neurodegenerative disorder described here contrasts with these, as there does not appear to be a consistent, milder ataxia/neurological phenotype in duplication carriers, supporting a higher dosage threshold for phenotypic expression. Based on this model, one would hypothesize that a relative CN of 4 due to a homozygous duplication would result in the same phenotype as a single triplicated allele. Alternatively, the mechanism may specifically depend on triplication.

Genomic palindromes can form hairpin or cruciform-type structures that promote double-strand breaks and lead to the formation of complex SVs. For instance, several inter/intra-chromosomal insertions into a 180 bp palindrome have been described ~80 kb downstream of *SOX3* (MIM: 313430). Depending on which sequences have been inserted, these can result in a wide range of phenotypes.^{25–28} It is estimated that 718 palindromes > 200 bp are spread across the human genome.²⁹ Future studies should investigate these genomic loci for complex SVs that may also represent novel genomic disorders. There is likely considerable background of benign structural variation in the general population at the 16p13.3 locus due to the palindrome, and we note that GRCh38 and the CHM13 v.2.0 reference genomes differ substantially (Figure S21). Furthermore, the overall structure of inverted duplication-triplications is often ambiguous unless much longer sequencing reads or optical mapping data are available that span the central inverted segment (Figure 4). Resolving inverted duplication-triplication structures may become a key use case for long-read sequencing technologies as effective read lengths increase, and this locus, with its 33 kb segments of 99.8% sequence identity (Figure S22), is particularly challenging.

likely due to a rare intron 1 variant in the 5' UTR (c.–140+2T>G [GenBank: NM_032039.4]). Z scores and p values were calculated by comparing results with 133 controls.

(B) Normalized read counts for *ATP6VOC* for all 134 fibroblast datasets confirm individual III-3 from family 1 to be the most significant outlier, with a relative fold change of 1.83. Similar RNA-seq results for individual 2 are presented in Figures S18 and S19.

(C and D) RNA-seq data from the 100kGP shows individual 7 to be an outlier for (C) *ATP6VOC* expression (4.95× above the mean) and (D) *AMDHD2* expression (4.48× above the mean), following normalization to *GAPDH*, when compared to 5,545 other RNA-seq datasets. Both genes lie within or partially within the triplicated segment that is shared across all 9 families. Data normalized to a second housekeeping gene are shown in Figure S20.

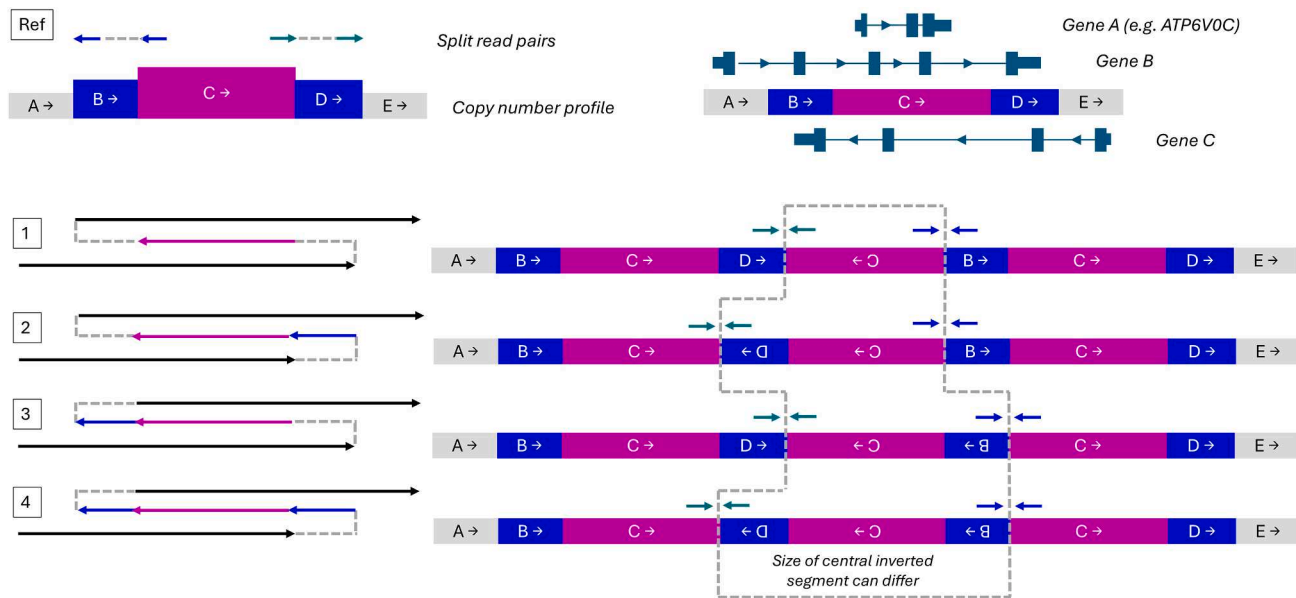


Figure 4. Schematic diagram showing the split-read pairs and copy number profile changes that are associated with inverted duplication-triplications

Although these types of rearrangements appear complex and can be called as 2 inversions, 2 duplications, and a central triplication, they involve just two breakpoints. The subway diagram highlights that there are typically 4 possible configurations that can all explain the short-read data. The size of the central inverted segment depends on which of these structures is true, and this can impact the interpretation. For instance, in this hypothetical example, the rearranged chromosome contains 3 copies of gene A (e.g., *ATP6VOC* in the present series of structural variants), irrespective of which precise structure is involved. In contrast, gene B is disrupted in solutions 2 and 4, while gene C would be disrupted in solutions 3 and 4. Ultra-long-read sequencing data or optical genome-mapping technologies where reads/labeled molecules span both breakpoint junctions can help distinguish which of the 4 possible solutions is right.

Using 100kGP data, we previously identified 3 individuals with intellectual disability and genetic haploinsufficiency of serine/arginine repetitive matrix 2 (*SRRM2*; MIM: 606032) resulting from complex deletion-inversions, with distal breakpoints clustered within the same 16p13.3 palindrome.³⁰ An additional *SRRM2* deletion-inversion identified from the GMS further supports the recurrent nature of complex SVs at this locus (Figure S23). Heterozygous loss-of-function variants in *SRRM2* (HGNC: 16639), a gene encoding a splicing factor, had previously been shown to result in autosomal-dominant intellectual developmental disorder type 72 (MIM: 620439).^{31,32} The findings described here support our hypothesis that the symmetry of this 144 kb palindromic repeat would make the formation of complex SVs equally likely in both directions. In contrast to the proximal deletion-inversion events, where the molecular mechanism is *SRRM2* haploinsufficiency, RNA-seq has demonstrated a robust increase in expression for several genes within these more distal inverted duplication-triplications. Interpretation is limited by the fact that RNA data were available for only 3 of 11 individuals, from three different laboratories, and using non-disease-relevant tissues. Future studies assessing RNA and protein levels in disease-relevant cell types derived from postmortem tissues or differentiated induced pluripotent stem cells (iPSCs) would be valuable. It remains to be seen whether further instances of 16p13.3 triplication could be uncovered in genetically unsolved individuals with ataxia using an RNA-seq-first approach.

The identification of critical dosage-sensitive genes within disease-associated CNVs is challenging. Although several mechanistic questions remain, the most compelling candidate gene within the 16p13.3 SV interval is *ATP6VOC*. Deleterious variants in this gene have previously been linked to neurodevelopmental abnormalities, with epilepsy reported in several cases.^{33,34} Microdeletions including this gene have also been implicated in neurodevelopmental disorders, with some sharing proximal breakpoints near the palindromic repeat.³⁵ *ATP6VOC* encodes the C subunit of the membrane-bound integral domain of a vacuolar proton pump enzyme complex, which acidifies organelles and establishes a proton gradient critical for several cellular processes. Pathogenic variants found in individuals with neurodevelopmental abnormalities were shown to impair V-ATPase function.³⁶ In family 1 and individual 2, RNA-seq indicated that the most significant gene expression changes occurred in *ATP6VOC*. We therefore speculate that the triplications described here may disturb the correct stoichiometry of subunits required for efficient assembly of the vacuolar H⁺-ATPase. While a peripheral domain is involved with ATP hydrolysis, the integral membrane V0 domain mediates proton pumping via a rotary mechanism. This V0 domain contains a c-ring structure that rotates and is made up of 9 subunits encoded by *ATP6VOC* and just one encoded by *ATP6VOB*.³⁶ Future studies assessing how enzyme assembly is altered in cell lines from affected individuals and whether this leads to abnormal proton

gradients being established would provide valuable pathomechanistic insights.

Genes encoding other subunits of this large enzyme complex have also been linked to neurological disorders.³⁶ Notably, *de novo* variants in *ATP6V1A*, which encodes the A subunit, cause developmental encephalopathy with epilepsy.³⁷ More recently, variants in *ATP6VOA1* (which encodes the brain-enriched isoform of the a subunit in the VO domain) have been shown to cause progressive myoclonus epilepsy and developmental and epileptic encephalopathy. Of relevance to the present study, 8 of 13 patients with *ATP6VOA1* variants presented with ataxia, and 25% had cerebellar atrophy.^{38,39}

Another gene that lies within the region of triplication in all but one case is *CEMP1*, a single-exon gene embedded within the final exon of *AMDHD2*. This gene regulates the deposition of cementum that, together with the periodontal ligament, helps anchor teeth to the jawbone⁴⁰ and thus is not a strong functional candidate. Another candidate gene that lies proximal to *ATP6VOC/CEMP1* on 16p13.3 and is likely impacted by this SV is *PDPK1* (MIM: 605213). Based on an analysis of large rare CNVs in almost 1 M individuals,⁴¹ *PDPK1* is predicted to be triplosensitive (pTripto = 0.99). This gene encodes 3-phosphoinositide-dependent protein kinase 1 and is crucial for mammalian brain development. Mice with conditional knockout of the orthologous gene exhibit decreased cerebellar size and ataxia-like behavior, suggesting that *pdpk1* may be critical for motor balance and coordination.⁴² GTEx data indicate that *PDPK1* has high relative expression in the cerebellum (a median TPM of 33.73). Although the gene does not fall within the minimal region of overlap, it must be noted that proximal breakpoints were imprecise due to the 144 kb palindromic repeat, and so the regions of triplication could potentially extend to include this gene in some cases. Even with the ~20 kb PacBio reads available for individual 7, proximal breakpoints of the rearrangement were unclear, and we hypothesized that they are situated between two *cis*-morphisms. Imprecise SV characterization is also pertinent to individuals 1 and 3 due to the low resolution of aCGH data, the low coverage of nanopore data, and the inability to obtain fresh samples, and the presence of a triplication was inferred from exome sequencing data.

While a simple dosage threshold model is an attractive option, we cannot rule out the occurrence of aberrant topologically associating domains (TADs) leading to ectopic gene expression. For other genomic disorders linked to specific SVs, 3D chromatin conformation using Hi-C and related methods has demonstrated how the altered distribution of TADs can impact gene expression.^{43–45} One example is split-hand/foot malformation type 3 (MIM: 246560), where ~0.5 Mb duplications on 10q24 show a high degree of clustering of both start and endpoints.⁴⁶ This similarity is due to the precise enhancer repositioning, which mouse studies show is required to lead to the ectopic expression of *Lbx1* and *Btrc*.⁴³ The heterogeneity

of sizes for the 16p13.3 SVs described here argues against a similar mechanism. However, future experimental studies using Hi-C should be complemented with deep learning computational methods⁴⁷ to confirm whether these complex genomic rearrangements affect chromatin structure and contribute to pathology.

We note that a complex rearrangement comprising 5 copies of *ATP6VOC* was recently described in a single individual presenting with ataxia.⁴⁸ The age of onset in this patient appears to have been earlier than in the 11 cases described in the present study (Table 1), with frequent falls and poor motor skills already noted by the age of 4. Additional features included amblyopia, hyperopia, dysarthric speech, coughing with liquid intake, excessive drooling, a progressive decrease in IQ, and cerebellar atrophy, which was first noted at 11 years (K. Hamanaka, personal communication). Future studies showing a correlation between higher CN states and disease severity would be of benefit for conclusively elucidating the pathomechanistic basis and clinical spectrum of this neurodegenerative disorder.

Together, our studies define a childhood-onset complex neurological disorder caused by overlapping 16p13.3 duplication-triplications and characterized by neurodegenerative features, including progressive ataxia, cognitive decline, and neuropathy, and a distinctive neuroradiological phenotype. Proximal breakpoints clustered within the 16p13.3 palindrome, making this the second complex genomic disorder linked to this mutational hotspot. Refinement of the spectrum of complex SVs that can lead to this disorder, combined with analyses of the consequences of increased *ATP6VOC* dosage on vacuolar H⁺-ATPase formation and function, will be crucial to determining the underlying cause of disease and thus potential therapeutic interventions for this devastating disorder.

Data and code availability

Data for individuals 1–4 are available via the DECIPHER database (www.deciphergenomics.org). For individuals 4–8, GS data (including structural VCFs and gene-level SV reports) and transcriptomic data (for individual 7 only) are available in the NGRL v.5.1, Genomics England. Information about the NGRL can be found at <https://doi.org/10.6084/m9.figshare.4530893.v7>, and details of how researchers can apply for access can be found at www.genomicsengland.co.uk/join-us. For family 1, genome and RNA-seq data are stored in a repository at the Clinical Genetics Department of the Erasmus MC. The long-read GS data for individuals 3 and 7 and the exome sequencing data for individuals 1 and 3 are stored by the Exeter Genomics Laboratory. The consent agreements for these datasets only permit sharing anonymized data for affected individuals, and therefore, we are unable to share BAM or VCF files. The code to query SVs in the NGRL is available in the Genomics England Research Environment User Guide, found at https://re-docs.genomicsengland.co.uk/structural_variant. The other code used in this study has already been published.

Acknowledgments

First and foremost, we are grateful to the families in this study. We thank Jing Yu for sharing SVRare reports and Mark Nellist for useful discussions. This research was funded by the NHS Rare and Inherited Disease Genomic Network of Excellence and the National Institute for Health and Care Research (NIHR) Exeter Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. This research was made possible through access to data in the NGRL, which is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The NGRL is funded by the NIHR and NHS England. The Wellcome Trust, Cancer Research UK, and the Medical Research Council funded research infrastructure. This study makes use of data generated by the DECIPHER community. A full list of centers that contributed to the generation of the data is available from <https://deciphergenomics.org/about/stats>. DECIPHER is hosted by EMBL-EBI. DECIPHER project funding was provided by the Wellcome Trust (WT223718/Z/21/Z). This research was conducted using the UKB Resource under application number 103356 and used data provided by patients collected by the NHS as part of their care and support. UKB protocols were approved by the National Research Ethics Service Committee. The Barakat lab was supported by the Netherlands Organisation for Scientific Research (ZonMw Vidi, grant 09150172110002). The Rare Disease Flagship acknowledges financial support from the Royal Children's Hospital Foundation.

Author contributions

E.L.B., J.R., and A.H.C. conceived and supervised the study. L.M. analyzed the radiological images. E.L.B., A.H.C., R.S., T.S.B., S.M.W., and S.L.H. coordinated the collaborations. A.H.N., A.K., A.S., A.W., B.B.-P., C.R., E.L.B., E.T., F.H.S., G.K.T., J.F., J.R., J.t.W.N., K.J., K.V., M.G.d.S., M.H., M.S., M.W., P.G., and S.M.W. recruited the patients and/or assessed the clinical information. K.M.B., L.J.M., and T.S.H. provided bioinformatic support. D.I.F. and E.E.J. analyzed the array data. A.P. and A.R.J. supervised the GS analysis. R.A.B. and N.U. performed the genetic studies. A.T.P., F.S., J.F., J.L., J.S.L., J.v.G., M.N.W., R.S., S.W., and T.J.v.H. analyzed the genomic and/or transcriptomic datasets. A.T.P. and J.F. wrote the initial draft of the manuscript.

Declaration of interests

S.W. is an employee at Genomics England.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist with improving the grammar and clarity of some sentences. The authors subsequently reviewed and edited all content to ensure accuracy and take full responsibility for the final manuscript.

Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.ajhg.2025.11.011>.

Web resources

DECIPHER database, www.deciphergenomics.org
Gencode, <https://www.gencodegenes.org>
GeneMatcher, <https://genematcher.org>
GTEx, <https://gtexportal.org>
hifiasm, <https://github.com/chhypl23/hifiasm>
HiFi-human-WGS-WDL pipeline, <https://github.com/PacificBiosciences/HiFi-human-WGS-WDL>
HGNC, <https://www.genenames.org>
NCBI BLAST tool, <https://blast.ncbi.nlm.nih.gov>
OMIM, <https://www.omim.org>

Received: July 28, 2025

Accepted: November 12, 2025

References

1. Wilson, D.M., 3rd, Cookson, M.R., Van Den Bosch, L., Zetterberg, H., Holtzman, D.M., and Dewachter, I. (2023). Hallmarks of neurodegenerative diseases. *Cell* *186*, 693–714. <https://doi.org/10.1016/j.cell.2022.12.032>.
2. Schubert Baldo, M., and Vilarinho, L. (2020). Molecular basis of Leigh syndrome: a current look. *Orphanet J. Rare Dis.* *15*, 31. <https://doi.org/10.1186/s13023-020-1297-9>.
3. Rothblum-Oviatt, C., Wright, J., Lefton-Greif, M.A., McGrath-Morrow, S.A., Crawford, T.O., and Lederman, H.M. (2016). Ataxia telangiectasia: a review. *Orphanet J. Rare Dis.* *11*, 159. <https://doi.org/10.1186/s13023-016-0543-7>.
4. Kurian, M.A., McNeill, A., Lin, J.P., and Maher, E.R. (2011). Childhood disorders of neurodegeneration with brain iron accumulation (NBIA). *Dev. Med. Child Neurol.* *53*, 394–404. <https://doi.org/10.1111/j.1469-8749.2011.03955.x>.
5. Wolf, N.I., Sistermans, E.A., Cundall, M., Hobson, G.M., Davis-Williams, A.P., Palmer, R., Stubbs, P., Davies, S., Endziniene, M., Wu, Y., et al. (2005). Three or more copies of the proteolipid protein gene PLP1 cause severe Pelizaeus-Merzbacher disease. *Brain* *128*, 743–751. <https://doi.org/10.1093/brain/awh409>.
6. Liu, P., Gelowani, V., Zhang, F., Drory, V.E., Ben-Shachar, S., Roney, E., Medeiros, A.C., Moore, R.J., DiVincenzo, C., Burnette, W.B., et al. (2014). Mechanism, prevalence, and more severe neuropathy phenotype of the Charcot-Marie-Tooth type 1A triplication. *Am. J. Hum. Genet.* *94*, 462–469. <https://doi.org/10.1016/j.ajhg.2014.01.017>.
7. Veitia, R.A., Zschocke, J., and Birchler, J.A. (2025). Gene Dosage Sensitivity and Human Genetic Diseases. *J. Inherit. Metab. Dis.* *48*, e70058. <https://doi.org/10.1002/jimd.70058>.
8. Carvalho, C.M.B., Ramocki, M.B., Pehlivan, D., Franco, L.M., Gonzaga-Jauregui, C., Fang, P., McCall, A., Pivnick, E.K., Hines-Dowell, S., Seaver, L.H., et al. (2011). Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat. Genet.* *43*, 1074–1081. <https://doi.org/10.1038/ng.944>.
9. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* *36*, 928–930. <https://doi.org/10.1002/humu.22844>.
10. Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A.J., Kruglyak, S., and Saunders, C.T. (2016). Manta: rapid detection of structural variants and

- indels for germline and cancer sequencing applications. *Bioinformatics* 32, 1220–1222. <https://doi.org/10.1093/bioinformatics/btv710>.
11. Roller, E., Ivakhno, S., Lee, S., Royce, T., and Tanner, S. (2016). Canvas: versatile and scalable detection of copy number variants. *Bioinformatics* 32, 2375–2377. <https://doi.org/10.1093/bioinformatics/btw163>.
 12. Yu, J., Szabo, A., Pagnamenta, A.T., Shalaby, A., Giacopuzzi, E., Taylor, J., Shears, D., Pontikos, N., Wright, G., Michaelides, M., et al. (2022). SVRare: discovering disease-causing structural variants in the 100K Genomes Project. *medRxiv* 2021, 2010.2015–21265069. <https://doi.org/10.1101/2021.10.15.21265069>.
 13. Holt, J.M., Saunders, C.T., Rowell, W.J., Kronenberg, Z., Wenger, A.M., and Eberle, M. (2024). HiPhase: jointly phasing small, structural, and tandem repeat variants from HiFi sequencing. *Bioinformatics* 40, btac042. <https://doi.org/10.1093/bioinformatics/btac042>.
 14. Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* 18, 170–175. <https://doi.org/10.1038/s41592-020-01056-5>.
 15. Dekker, J., Schot, R., Bongaerts, M., de Valk, W.G., van Veghel-Plandsoen, M.M., Monfils, K., Douben, H., Elfferich, P., Kasteleijn, E., van Unen, L.M.A., et al. (2023). Web-accessible application for identifying pathogenic transcripts with RNA-seq: Increased sensitivity in diagnosis of neurodevelopmental disorders. *Am. J. Hum. Genet.* 110, 251–272. <https://doi.org/10.1016/j.ajhg.2022.12.015>.
 16. Pagnamenta, A.T., Yu, J., Walker, S., Noble, A.J., Lord, J., Dutta, P., Hashim, M., Camps, C., Green, H., Devaiah, S., et al. (2024). The impact of inversions across 33,924 families with rare disease from a national genome sequencing project. *Am. J. Hum. Genet.* 111, 1140–1164. <https://doi.org/10.1016/j.ajhg.2024.04.018>.
 17. UK Biobank Whole-Genome Sequencing Consortium (2025). Whole-genome sequencing of 490,640 UK Biobank participants. *Nature* 645, 692–701. <https://doi.org/10.1038/s41586-025-09272-9>.
 18. Delatycki, M.B., and Corben, L.A. (2012). Clinical features of Friedreich ataxia. *J. Child Neurol.* 27, 1133–1137. <https://doi.org/10.1177/0883073812448230>.
 19. Sullivan, R., Kaiyrzhanov, R., and Houlden, H. (2021). Cerebellar ataxia, neuropathy, vestibular areflexia syndrome: genetic and clinical insights. *Curr. Opin. Neurol.* 34, 556–564. <https://doi.org/10.1097/WCO.0000000000000961>.
 20. Szmulewicz, D.J., Roberts, L., McLean, C.A., MacDougall, H.G., Halmagyi, G.M., and Storey, E. (2016). Proposed diagnostic criteria for cerebellar ataxia with neuropathy and vestibular areflexia syndrome (CANVAS). *Neurol. Clin. Pract.* 6, 61–68. <https://doi.org/10.1212/CPJ.0000000000000215>.
 21. Meli, A., Montano, V., Palermo, G., Fogli, A., Rocchi, A., Gerfo, A.L., Maltomini, R., Cori, L., Siniscalchi, A., Bernardini, C., et al. (2025). Diagnosis of hereditary ataxias: a real-world single center experience. *J. Neurol.* 272, 111. <https://doi.org/10.1007/s00415-024-12772-9>.
 22. Bahrambeigi, V., Song, X., Sperle, K., Beck, C.R., Hijazi, H., Grochowski, C.M., Gu, S., Seeman, P., Woodward, K.J., Carvalho, C.M.B., et al. (2019). Distinct patterns of complex rearrangements and a mutational signature of microhomeology are frequently observed in PLP1 copy number gain structural variants. *Genome Med.* 11, 80. <https://doi.org/10.1186/s13073-019-0676-0>.
 23. de Bruijn, S.E., Fiorentino, A., Ottaviani, D., Fanucchi, S., Melo, U.S., Corral-Serrano, J.C., Mulders, T., Georgiou, M., Rivolta, C., Pontikos, N., et al. (2020). Structural Variants Create New Topological-Associated Domains and Ectopic Retinal Enhancer-Gene Contact in Dominant Retinitis Pigmentosa. *Am. J. Hum. Genet.* 107, 802–814. <https://doi.org/10.1016/j.ajhg.2020.09.002>.
 24. Grangeon, L., Cassinari, K., Rousseau, S., Croisile, B., Formaglio, M., Moreaud, O., Boutonnat, J., Le Meur, N., Miné, M., Coste, T., et al. (2021). Early-Onset Cerebral Amyloid Angiopathy and Alzheimer Disease Related to an APP Locus Triplication. *Neurol. Genet.* 7, e609. <https://doi.org/10.1212/NXG.0000000000000609>.
 25. Boschann, F., Moreno, D.A., Mensah, M.A., Sczakiel, H.L., Skipalova, K., Holtgrewe, M., Mundlos, S., and Fischer-Zirnsak, B. (2022). Xq27.1 palindrome mediated interchromosomal insertion likely causes familial congenital bilateral laryngeal abductor paralysis (Plott syndrome). *J. Hum. Genet.* 67, 405–410. <https://doi.org/10.1038/s10038-022-01018-z>.
 26. Bowl, M.R., Nesbit, M.A., Harding, B., Levy, E., Jefferson, A., Volpi, E., Rizzoti, K., Lovell-Badge, R., Schlessinger, D., Whyte, M.P., and Thakker, R.V. (2005). An interstitial deletion-insertion involving chromosomes 2p25.3 and Xq27.1, near SOX3, causes X-linked recessive hypoparathyroidism. *J. Clin. Investig.* 115, 2822–2831. <https://doi.org/10.1172/JCI24156>.
 27. Gardner, J.C., Jovanovic, K., Ottaviani, D., Melo, U.S., Jackson, J., Guarascio, R., Ziaka, K., Hau, K.L., Lane, A., Taylor, R.L., et al. (2025). Inter-chromosomal insertions at Xq27.1 associated with retinal dystrophy induce dysregulation of LINC00632 and CDR1as/ciRS-7. *Am. J. Hum. Genet.* 112, 523–536. <https://doi.org/10.1016/j.ajhg.2025.01.007>.
 28. Rahikkala, E., Komulainen-Ebrahim, J., Tolonen, J.P., Vorimo, S., Suo-Palosaari, M., Vieira, P., Piispala, J., Uusimaa, J., Pylkäs, K., and Mantere, T. (2024). Optical Genome Mapping Identifies a Second Xq27.1 Rearrangement Associated With Charcot-Marie-Tooth Neuropathy CMTX3. *Mol. Genet. Genomic Med.* 12, e70014. <https://doi.org/10.1002/mgg3.70014>.
 29. Ganapathiraju, M.K., Subramanian, S., Chaparala, S., and Karunakaran, K.B. (2020). A reference catalog of DNA palindromes in the human genome and their variations in 1000 Genomes. *Hum. Genome Var.* 7, 40. <https://doi.org/10.1038/s41439-020-00127-5>.
 30. Pagnamenta, A.T., Yu, J., Willis, T.A., Hashim, M., Seaby, E.G., Walker, S., Xian, J., Cheng, E.W.Y., Tavares, A.L.T., Forzano, F., et al. (2023). A Palindrome-Like Structure on 16p13.3 Is Associated with the Formation of Complex Structural Variations and SRRM2 Haploinsufficiency. *Hum. Mutat.* 2023, 6633248.
 31. Cuinat, S., Nizon, M., Isidor, B., Stegmann, A., van Jaarsveld, R.H., van Gassen, K.L., van der Smagt, J.J., Volker-Touw, C.M.L., Holwerda, S.J.B., Terhal, P.A., et al. (2022). Loss-of-function variants in SRRM2 cause a neurodevelopmental disorder. *Genet. Med.* 24, 1774–1780. <https://doi.org/10.1016/j.gim.2022.04.011>.
 32. Kaplanis, J., Samocho, K.E., Wiel, L., Zhang, Z., Arvai, K.J., Eberhardt, R.Y., Gallone, G., Lelieveld, S.H., Martin, H.C., McRae, J.F., et al. (2020). Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature* 586, 757–762. <https://doi.org/10.1038/s41586-020-2832-5>.

33. Rong, M., Marques, P.T., Ali, Q.Z., Morcos, R., Chandran, I., Qaiser, F., Møller, R.S., Bayat, A., Rubboli, G., Gardella, E., et al. (2025). Variants in ATP6V0C are associated with Dravet-like developmental and epileptic encephalopathy. *Epilepsia* *66*, 2046–2052. <https://doi.org/10.1111/epi.18346>.
34. Tinker, R.J., Burghel, G.J., Garg, S., Steggall, M., Cuvertino, S., and Banka, S. (2021). Haploinsufficiency of ATP6V0C possibly underlies 16p13.3 deletions that cause microcephaly, seizures, and neurodevelopmental disorder. *Am. J. Med. Genet.* *185*, 196–202. <https://doi.org/10.1002/ajmg.a.61905>.
35. Mucha, B.E., Banka, S., Ajeawung, N.F., Molidpere, S., Chen, G.G., Koenig, M.K., Adejumo, R.B., Till, M., Harbord, M., Perrier, R., et al. (2019). A new microdeletion syndrome involving TBC1D24, ATP6V0C, and PDPK1 causes epilepsy, microcephaly, and developmental delay. *Genet. Med.* *21*, 1058–1064. <https://doi.org/10.1038/s41436-018-0290-3>.
36. Mattison, K.A., Tossing, G., Mulroe, F., Simmons, C., Butler, K.M., Schreiber, A., Alsadah, A., Neilson, D.E., Naess, K., Wedell, A., et al. (2023). ATP6V0C variants impair V-ATPase function causing a neurodevelopmental disorder often associated with epilepsy. *Brain* *146*, 1357–1372. <https://doi.org/10.1093/brain/awac330>.
37. Fassio, A., Esposito, A., Kato, M., Saito, H., Mei, D., Marini, C., Conti, V., Nakashima, M., Okamoto, N., Olmez Turker, A., et al. (2018). De novo mutations of the ATP6V1A gene cause developmental encephalopathy with epilepsy. *Brain* *141*, 1703–1718. <https://doi.org/10.1093/brain/awy092>.
38. Aoto, K., Kato, M., Akita, T., Nakashima, M., Mutoh, H., Akasaka, N., Tohyama, J., Nomura, Y., Hoshino, K., Ago, Y., et al. (2021). ATP6V0A1 encoding the $\alpha 1$ -subunit of the V0 domain of vacuolar H(+)-ATPases is essential for brain development in humans and mice. *Nat. Commun.* *12*, 2107. <https://doi.org/10.1038/s41467-021-22389-5>.
39. Bott, L.C., Forouhan, M., Lieto, M., Sala, A.J., Ellerington, R., Johnson, J.O., Speciale, A.A., Criscuolo, C., Filla, A., Chitayat, D., et al. (2021). Variants in ATP6V0A1 cause progressive myoclonus epilepsy and developmental and epileptic encephalopathy. *Brain Commun.* *3*, fcab245. <https://doi.org/10.1093/braincomms/fcab245>.
40. Campos, M.N., Giraldo, E.L., Del Rio Portilla, F., Fernández-Velasco, D.A., Arzate, H., and Romo-Arévalo, E. (2023). Solution NMR structure of cementum protein 1 derived peptide (CEMP1-p1) and its role in the mineralization process. *J. Pept. Sci.* *29*, e3494. <https://doi.org/10.1002/psc.3494>.
41. Collins, R.L., Glessner, J.T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L.-M., Ulirsch, J., et al. (2022). A cross-disorder dosage sensitivity map of the human genome. *Cell* *185*, 3041–3055.e25. <https://doi.org/10.1016/j.cell.2022.06.036>.
42. Liu, R., Xu, M., Zhang, X.Y., Zhou, M.J., Zhou, B.Y., Qi, C., Song, B., Fan, Q., You, W.Y., Zhu, J.N., et al. (2020). PDK1 Regulates the Maintenance of Cell Body and the Development of Dendrites of Purkinje Cells by pS6 and PKC γ . *J. Neurosci.* *40*, 5531–5548. <https://doi.org/10.1523/JNEUROSCI.2496-19.2020>.
43. Cova, G., Glaser, J., Schöpflin, R., Prada-Medina, C.A., Ali, S., Franke, M., Falcone, R., Federer, M., Ponzi, E., Ficarella, R., et al. (2023). Combinatorial effects on gene expression at the Lbx1/Fgf8 locus resolve split-hand/foot malformation type 3. *Nat. Commun.* *14*, 1475. <https://doi.org/10.1038/s41467-023-37057-z>.
44. Dimartino, P., Zadorozhna, M., Yumiceba, V., Basile, A., Cani, I., Melo, U.S., Henck, J., Breur, M., Tonon, C., Lodi, R., et al. (2024). Structural Variants at the LMNB1 Locus: Deciphering Pathomechanisms in Autosomal Dominant Adult-Onset Demyelinating Leukodystrophy. *Ann. Neurol.* *96*, 855–870. <https://doi.org/10.1002/ana.27038>.
45. Nmezi, B., Rodriguez Bey, G., Oranburg, T.D., Dudnyk, K., Lardo, S.M., Herdman, N., Jacko, A., Rubio, S., Loeza-Alcocer, E., Kofler, J., et al. (2025). An oligodendrocyte silencer element underlies the pathogenic impact of lamin B1 structural variants. *Nat. Commun.* *16*, 1373. <https://doi.org/10.1038/s41467-025-56378-9>.
46. Holder-Espinasse, M., Jamsheer, A., Escande, F., Andrieux, J., Petit, F., Sowinska-Seidler, A., Socha, M., Jakubiuk-Tomaszuk, A., Gerard, M., Mathieu-Dramard, M., et al. (2019). Duplication of 10q24 locus: broadening the clinical and radiological spectrum. *Eur. J. Hum. Genet.* *27*, 525–534. <https://doi.org/10.1038/s41431-018-0326-9>.
47. Maroofian, R., Pagnamenta, A.T., Navabazam, A., Schwesinger, R., Roberts, H.E., Lopopolo, M., Dehghani, M., Vahidi Mehrjardi, M.Y., Haerian, A., Soltanianzadeh, M., et al. (2024). Familial severe skeletal Class II malocclusion with gingival hyperplasia caused by a complex structural rearrangement at the KCNJ2-KCNJ16 locus. *HGG Adv.* *5*, 100352. <https://doi.org/10.1016/j.xhgg.2024.100352>.
48. Hamanaka, K., Fujita, A., Miyatake, S., Misawa, K., Koshimizu, E., Uchiyama, Y., Tsuchida, N., Seyama, R., Sakamoto, M., Iwama, K., et al. (2025). Genome sequencing provides high diagnostic yield and new etiological insights for intellectual disability and developmental delay. *NPJ Genom. Med.* *10*, 60. <https://doi.org/10.1038/s41525-025-00521-4>.