

## ARTICLE OPEN



# Germline mosaicism in *TCF20*-associated neurodevelopmental disorders: a case study and literature review

Jessie Poquérusse<sup>1,2</sup>, Whitney Whitford<sup>1,2</sup>, Juliet Taylor<sup>3</sup>, Nerine Gregersen<sup>3</sup>, Donald R. Love<sup>4,7</sup>, Bobby Tsang<sup>5</sup>, Kylie M. Drake<sup>6</sup>, Russell G. Snell<sup>1,2</sup>, Klaus Lehnert<sup>1,2</sup> and Jessie C. Jacobsen<sup>1,2</sup>✉

© The Author(s) 2025

Autosomal dominant variants in transcription factor 20 (*TCF20*) can result in *TCF20*-associated neurodevelopmental disorder (TAND), a condition characterized by developmental delay and intellectual disability, autism, dysmorphisms, dystonia, and variable other neurological features. To date, a total of 91 individuals with TAND have been reported; ~67% of cases arose de novo, while ~10% were inherited, and, intriguingly, ~8% were either confirmed or suspected to have arisen via germline mosaicism. Here, we describe two siblings with a developmental condition characterized by intellectual disability, autism, a circadian rhythm sleep disorder, and attention deficit hyperactivity disorder (ADHD) caused by a novel heterozygous single nucleotide deletion in the *TCF20* gene, NM\_001378418.1:c.4737del; NP\_001365347.1:p.Lys1579Asnfs\*36 (GRCh38/hg38). The variant was not detected in DNA extracted from peripheral blood in either parent by Sanger sequencing of PCR-generated amplicons, or by deep sequencing of PCR amplicons using MiSeq and MinION. However, droplet digital PCR (ddPCR) of DNA derived from early morning urine detected the variation in 3.2% of the father's urothelial cells, confirming germline mosaicism. This report is only the second to confirm with physical evidence *TCF20* germline mosaicism and discusses germline mosaicism as a likely under-detected mode of inheritance in neurodevelopmental conditions.

*Journal of Human Genetics* (2025) 70:215–222; <https://doi.org/10.1038/s10038-025-01323-3>

## INTRODUCTION

Transcription factor 20 (*TCF20*, MIM 603107) has emerged as a syndromic neurodevelopmental disorder (NDD)-associated gene [1–4]. *TCF20* is expressed in human brain pre- and early postnatally, reaching peak developmental expression between 7 weeks post-conception and 2 years of age [5]. *TCF20* is located in the distal region of chromosome 22 (22q13.2) which is involved in human neurodevelopmental processes, and is predicted to be highly intolerant to loss of function (LOF) variants (pLI = 1) [6]. It is ubiquitously expressed through adulthood, and highly expressed in the cerebellum [7].

Animal models have shown that *TCF20* is highly expressed in the pre-migratory neural crest cells of chicken embryos [8], and in the hippocampus and cerebellum of mice during brain development [9]. The *TCF20* protein, also named AR1, SPBP, or SPRE-binding protein, acts as a transcriptional activator which localizes to the nucleus and binds and controls the regulatory region of the gene encoding the extracellular matrix breakdown protein metalloproteinase 3 (MPP3)/stromelysin [10]. It has been found to also upregulate expression of a number of transcription factors, such as JUN, SP1, PAX6 and ETS1 [11, 12]. Recent functional work in mice has shown that *TCF20* is essential for neurogenesis, promoting the expression of *TDG*, which controls DNA methylation at the T-cell factor 4 (TCF-4) promoter, thereby regulating its

expression and modulating neural differentiation, deficits in which are associated with autism [13].

In humans, variants in the *TCF20* gene can result in developmental and intellectual disability, autism, dysmorphisms, and neurological features, alongside various other impairments spanning gastrointestinal issues, hepatic issues, and skeletomuscular systems. Pathogenic variants in *TCF20* act in an autosomal dominant mode (MIM 618430; Orphanet 35099) and result in *TCF20*-associated neurodevelopmental disorders (TAND) [14].

The *TCF20* and *RAI* (retinoic acid-induced 1) genes have likely evolved from a common ancestor by genome duplication [15]. Deletions or loss of function variants in *RAI* result in Smith-Magenis syndrome (MIM 182290), while duplications in *RAI* result in Potocki-Lupski syndrome (MIM 610883), both characterized by a similar constellation of mental disability and various congenital craniofacial and skeletal dysmorphisms.

To date, this case report and nine other publications [3, 16–23] including four large cohort studies [14, 24–26], have identified 76 unique pathogenic variants in or affecting (e.g. by chromoanagenesis) the *TCF20* gene in 91 individuals from 83 different families (Fig. 3 and Supplementary Table 1). Among these, Babbs et al. and Torti et al. described cases of a likely germline mosaic nature [16, 26], while only Schneeweiss et al. confirmed germline mosaic inheritance, of maternal origin, of a *TCF20* variant in an affected sibling [21].

<sup>1</sup>School of Biological Sciences, The University of Auckland, Auckland, New Zealand. <sup>2</sup>Centre for Brain Research, The University of Auckland, Auckland, New Zealand. <sup>3</sup>Genetic Health Service New Zealand, Auckland City Hospital, Auckland, New Zealand. <sup>4</sup>Diagnostic Genetics, LabPLUS, Auckland City Hospital, Auckland, New Zealand. <sup>5</sup>Pediatrics and Newborn Services, Waitakere Hospital, Auckland, New Zealand. <sup>6</sup>Canterbury Health Laboratories, Christchurch, New Zealand. <sup>7</sup>Present address: Genetic Pathology, Sidra Medicine, Doha, Qatar. ✉email: [jjacobsen@auckland.ac.nz](mailto:jjacobsen@auckland.ac.nz)

Received: 11 December 2024 Revised: 5 February 2025 Accepted: 6 February 2025

Published online: 26 February 2025

Here, we report two affected siblings with developmental and intellectual disability, autism, dysmorphic features, and neurological variances. Whole exome sequencing (WES) of peripheral blood-derived DNA from the eldest sibling (II:1) identified a previously unreported *TCF20* frameshift variant, NC\_000022.11:g.42210575del; NM\_001378418.1:c.4737del; NP\_001365347.1:p.Lys1579Asnfs\*36 (GRCh38/hg38; MANE Select transcripts). The variant was also identified in peripheral blood-derived DNA in the younger sibling (II:2), but not identified in the parents' peripheral blood by Sanger sequencing of PCR-generated amplicons. It was, however, detected by droplet digital PCR (ddPCR) in 3.2% of the father's urothelial cells from an early morning urine sample.

This case report is only the second confirmed case with physical evidence of germline mosaicism in *TCF20*-associated neurodevelopmental disorders and the first to confirm the inheritance of the same *TCF20* variant in siblings. Our findings suggest germline mosaicism could be an under-detected mode of transmission for neurodevelopmental conditions with important implications for genetic counseling.

## MATERIALS AND METHODS

### Clinical report

**II:1.** The proband was the first child of a healthy Cook Island mother and Samoan father. She was born in New Zealand after a normal pregnancy at 40 weeks gestation. She experienced early post-natal global developmental delay. She was seen in the clinic for mild global development delay at 2 years and 4 months of age and met formal criteria for autism with intellectual disability. On examination at age 15, she was found to display immature speech, using short phrases and pointing to attract attention, while making reasonable eye contact. She has no obvious stereotypies or sensory processing issues. She has a circadian rhythm sleep disorder which is successfully controlled with melatonin and suffers from attention hyperactivity disorder (ADHD) which is successfully treated with methylphenidate. She experiences gastrointestinal distress in the form of intermittent constipation, and has mild left-hand side weakness and joint hypermobility. She has dysmorphisms in the form of posteriorly rotated small ears with small ear lobes. She has single palmar creases in the right hand and three café-au-lait macules. A brain MRI scan at age 3 years revealed no abnormal findings.

**II:2.** The proband's younger brother was also New Zealand-born at 40 weeks gestation. He experienced early post-natal global developmental delay, intellectual disability, and autism. At 4.5 years of age, he was found to have left facial weakness with generalized hypotonia. He also suffers from a developmental coordination disorder. At 5.5 years of age, he was found to have symptoms of ADHD and dyspraxia, leading to a formal diagnosis of ADHD. He also has obesity and right exotropia. He is non-dysmorphic.

### DNA extraction and whole exome sequencing

High quality genomic DNA for II:1 was extracted from whole blood cells using the Qiagen Gentra Puregene Blood Kit according to manufacturer's instructions, after which the DNA was quantified by Nanodrop, Qubit, and agarose gel electrophoresis.

Whole exome sequencing (WES) was carried out using the SureSelect XT Human All Exon v5 (Agilent, Santa Clara, CA, USA) and Illumina HiSeq 2000/2500 (Illumina Corporation, San Diego, CA, USA) to generate 100 bp paired-end reads at an average read depth of 48x. Sequence alignment and variant calling for single nucleotide variations (SNVs) and insertions and deletions (indels) of WES data was carried out as previously described, but converted to GRCh38/hg38 coordinates using UCSC's LiftOver tool [27]. WES alignments were viewed using the Integrated Genomics Viewer (IGV, v2.3.55) [28].

### WES analysis

A combined total of 93,403 high-confidence variants were identified in exonic and splice-junction regions across individual II:1's whole exome sequence. Variants were first removed if observed at a minor allele frequency (MAF) > 0.01 in healthy populations from the ExAC database [29], or at a MAF > 0.05 in unrelated cases from our in-house database. This

yielded a total of 4548 variants. Variants were filtered such as to only keep homozygous, compound heterozygous, X-linked and de novo variants, and retain splice site, stop gain/loss, start gain/loss, frameshift, inframe insertion/deletion, and missense variants, as well as variants in 5' and 3' UTRs, as annotated by Ensembl's Variant Effect Predictor tool [30]. This retained a total of 980 variants. These remaining variants were prioritized based on their likelihood of having a highly deleterious functional consequence based on the gene's pLI score [31] and in silico predictions by SIFT [32], PolyPhen2 [33] and CAROL [34] functional pathogenicity prediction algorithms, previously established associations with phenotypes or disorders, relevant biochemical/biological function, and fetal or early postnatal brain-enriched spatiotemporal expression [7].

### Sanger sequencing

Variants identified in *TCF20* by WES were validated by PCR followed by Sanger sequencing.

Primers were designed using Primer-BLAST [35]. For II:1, the forward primer 5' TGTACAAACGGCTCCAAGTTC 3' and the reverse primer 5' AGTGACGATTTCACCGAAGC 3' were used to generate a 391 bp product encompassing the variant NC\_000022.11:g.42210575del.

All PCR products were generated using the Expand High Fidelity PCR System (Roche) with the following thermocycler conditions: 5 min initial denaturation at 94 °C, with 30 cycles of 30 s denaturation at 94 °C, 30 s of annealing at 55 °C, and 45 s of extension at 72 °C, followed by 7 min of final extension at 72 °C.

Sanger sequencing was performed by Auckland Genomics, The University of Auckland, New Zealand. Electropherograms were viewed using Geneious (v8.1.5) (<http://www.geneious.com>).

Confirmation of the variant in II:2 was performed by All Wales Medical Genetics Service (United Kingdom) using PCR and Sanger sequencing. Both parents were also tested for the variant.

### Barcoded Illumina MiSeq sequencing

In an attempt to detect the presence of the variant in low level mosaicism form in the parents, the genomic *TCF20* region described above was amplified from peripheral whole blood of each parent.

Targeted amplification was performed using PCR with the PCR primers previously described, adapted to include the Illumina overhang adapter sequences (underlined); forward primer (5' TCGTCGGCAGCGTCAGATGTGTA-TAAGAGACAGTGTACAACGGCTCCAAGTTC 3') and reverse primer (5' GTCTCTGTGGGCTCGGAGATGTGTATAAGAGACAGAGTGACGATTTCACCGAAGC 3').

Each PCR was performed using the Expand High Fidelity PCR System (Roche), with the following conditions: 5 min initial denaturation at 94 °C, with 30 cycles of 30 s denaturation at 94 °C, 30 s of annealing at 55 °C, and 45 s of extension at 72 °C, followed by 7 min of final extension at 72 °C. Amplified product was purified using AMPure XP beads (Beckman Coulter). The purified DNA was quantified using Qubit 2.0 fluorometer (Invitrogen) via the Qubit dsDNA Broad Range Assay Kit. MiSeq sequencing was performed by Auckland Genomics, The University of Auckland, New Zealand. Sequence reads were aligned to the target reference (GRCh38/hg38) using Burrow-Wheeler Aligner (BWA-MEM) [36] and visualized using the Integrated Genomics Viewer (IGV, v2.3.55) [28].

### MinION deep sequencing

To further probe the presence of the variant in rare form (e.g. less than 10% allele fraction) in the parents' blood, we performed MinION sequencing (Oxford Nanopore Technologies) of PCR amplicons from parental peripheral whole blood DNA.

Targeted *TCF20* amplification was performed using the primers and conditions described for Sanger sequencing. Amplified product was purified using AMPure XP beads (Beckman Coulter). The purified DNA was quantified using Qubit 2.0 fluorometer (Invitrogen) via the Qubit dsDNA Broad Range Assay Kit.

The library was prepared for Oxford Nanopore Technologies' (ONT) MinION sequencing as in the 1D Native barcoding DNA (with EXP-NBD103 and SQK-LSK108) protocol (ONT). The MinION sequencer was run using a single new R9.4 flowcell. Base calling was performed using guppy basecaller v2.3.7 (ONT) from the raw fast5 sequencing files from the MinION sequencer, followed by barcode demultiplexing also using guppy barcoder v2.3.7 (ONT), generating a fastq file for each barcode. Alignment of barcoded reads to the *TCF20* target region was performed using minimap2 v2.16 (using the GRCh37/hg19 reference genome) and visualized using the Integrated Genomics Viewer (IGV, v2.3.55) [28].

### Droplet digital PCR and fluorescence

To assess parental mosaicism in cells derived from a different germ layer, droplet digital PCR was performed on DNA from urine (Canterbury Health Labs, Christchurch, New Zealand). The *TCF20* familial variant was amplified by PCR in the presence of a fluorescent intercalating dye, EvaGreen (BioRad), using the manufacturer's recommended conditions. The wildtype forward primer 5' GAGAGCCAAAGCCAAAAACAGAGGCAAAGG 3', variant forward primer 5' GAGAGCCAAAGCCAAAAACAGAGGCAAAGG 3' and the reverse primer 5' TTTGATCTCAGTTCTTGGGGTTCACA 3' were used to generate a 122 bp amplicon. Droplet fluorescence was then measured on a QX200 Droplet Digital PCR system and data was analyzed using QuantaSoft software. Quantitation was carried out by normalizing the variant allele to the wildtype allele. In a known heterozygous control (II:1) this ratio was 50%.

## RESULTS

### Variant identification

The sole variant in II:1 (case 81 in Musgrave et al. [37]) retained after filtering was a heterozygous single nucleotide deletion of a thymidine in exon 1 (of 6) of *TCF20*, NC\_000022.11:g.42210575del; NM\_001378418.1:c.4737del; NP\_001365347.1:p.Lys1579Asnfs\*36 (GRCh38/hg38; MANE Select transcript), Class 5 Pathogenic (PVS1, PM2\_Supp, PS2\_mod) [38] – Fig. 1. The variation lies within a 7-nucleotide thymidine homopolymer tract. This variant was subsequently confirmed to be present in sibling II:2, but was not detectable in electropherograms obtained from Sanger sequencing of the parents.

### Assessment of parental mosaicism

To investigate potential parental mosaicism, deep sequencing of amplicons generated from DNA extracted from peripheral blood encompassing the variant in both parents and the proband was performed by Illumina MiSeq short-read sequencing. Read depth at the variant locus exceeded 105,000 in every sample. The deletion was present in 50,825/105,340 (48.2%) of reads for II:1, confirming heterozygosity at this locus. The single thymidine deletion was present in 2036/150,247 (1.36%) and 1851/233,680 (0.79%) of the reads for the mother I:2 and father I:1, respectively.

Next, potential parental mosaicism was also investigated using ONT MinION sequencing for the amplicon encompassing the variant in both parents and proband also in DNA extracted from peripheral blood. The read depth for each individual exceeded 15,000 reads. Alignments for all three individuals showed over 50% of reads with a 6-nucleotide long thymidine homopolymer. Additionally, >20% of aligned reads contained a 5-nucleotide long thymidine homopolymer.

Finally, parental mosaicism was investigated using ddPCR from DNA extracted from urothelial cells from early morning urine from

the parents. First morning urine was chosen as a DNA source for ddPCR, to have a population of cells that more closely resembles the original gonadal tissue (compared to white blood cells). Interestingly, the variant was detected by ddPCR in 3.2% of paternal DNA (Fig. 2) and was absent in the maternal sample.

## DISCUSSION

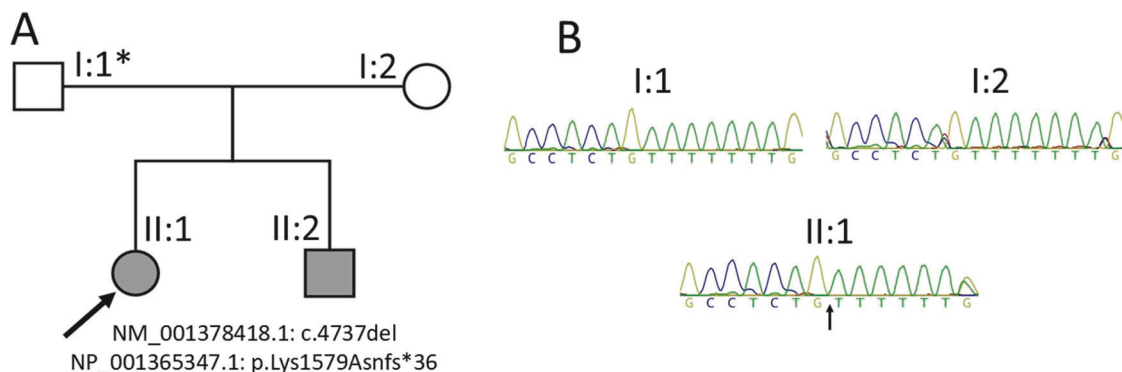
### *TCF20* variant and predicted consequence

WES and Sanger sequencing analysis revealed a previously unreported single base pair heterozygous deletion, NM\_001378418.1:c.4737del; NP\_001365347.1:p.Lys1579Asnfs\*36 (GRCh38/hg38), in the proband (II:1, Fig. 1) and her brother (II:2) located within the region encoding the N2 domain of the *TCF20* protein amidst a stretch of four other previously reported neurodevelopmental disorder-associated variants (spanning c.4737 to c.4786; Fig. 3), reported as pathogenic in ClinVar (NCBI) (accession number VCV002571818.3). The genetic variant deletes a thymidine in a homopolymer tract of seven thymidine residues (22:g.42210568-42210575), inducing a frameshift resulting in the replacement of the C-terminal 381 amino acid residues with 34 residues not found in the wild-type protein. The frameshift replaces 21 residues of the 25 amino-acids long second nuclear localization signal (N2) [39], potentially affecting the ability of the truncated *TCF20* protein to localize to the nucleus to exert its transcription factor binding activity. As the variant is located in the second of six exons, degradation of the *TCF20* mRNA harboring the variant via nonsense mediated decay is another potential consequence.

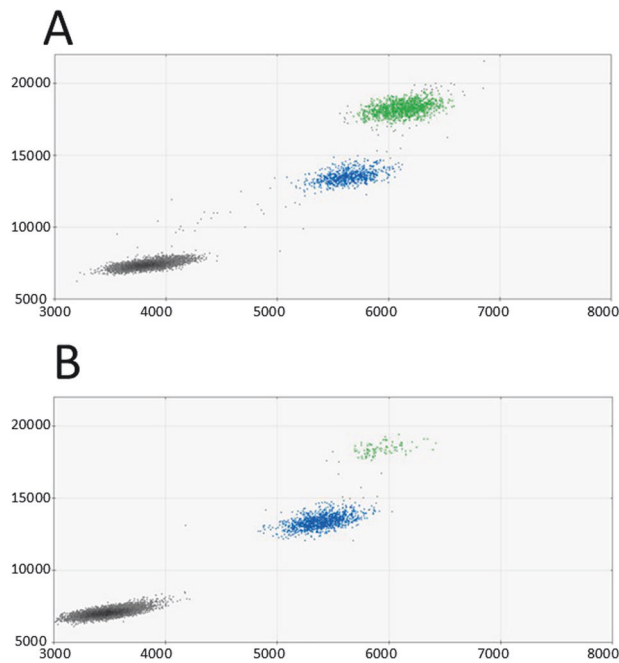
Given that *TCF20* is highly intolerant to loss of function variants, with a pLI of 1.0 [31], its ranking in the top 23.45% of genes most intolerant to haploinsufficiency [40], and the overlap of the symptoms observed in both siblings with previous *TCF20*-associated disorders (see Supplementary Table 1 for details), the variant was considered causative of the siblings' phenotype.

### *TCF20*-associated neurodevelopmental disorders

Including the siblings reported here, a total of 76 unique disorder-causing variants in *TCF20* have now been reported across 91 individuals from 83 different families to date (see Supplementary Table 1 for details) [3, 14, 16–24, 26, 41]. Causative variants in *TCF20* include SNVs (36.3%), indels (50.5%), copy number variants (CNVs) (9.9%), and structural variants (SVs) (3.3%). Of the 87 individuals in whom gender was reported, these *TCF20* variants affected 32 females (36.8%) and 55 males (63.2%). Most individuals have developmental/intellectual disability (present in 85 out of 91 cases, i.e. 93.4% of cases), autism (present in 48 out of 86 reported cases, i.e. 55.8% of cases), and dystonia [22]. However,



**Fig. 1 Summary of the *TCF20* variant.** **A** Pedigree showing the immediate family of the proband (II:1, indicated by the arrow) and affected sibling (II:2). The germline mosaic genotype of the father of the affected siblings is marked with an asterisk. **B** Sanger sequencing chromatograms of peripheral blood-derived DNA of the proband and her parents, with the arrow pointing to the thymidine deletion in the proband



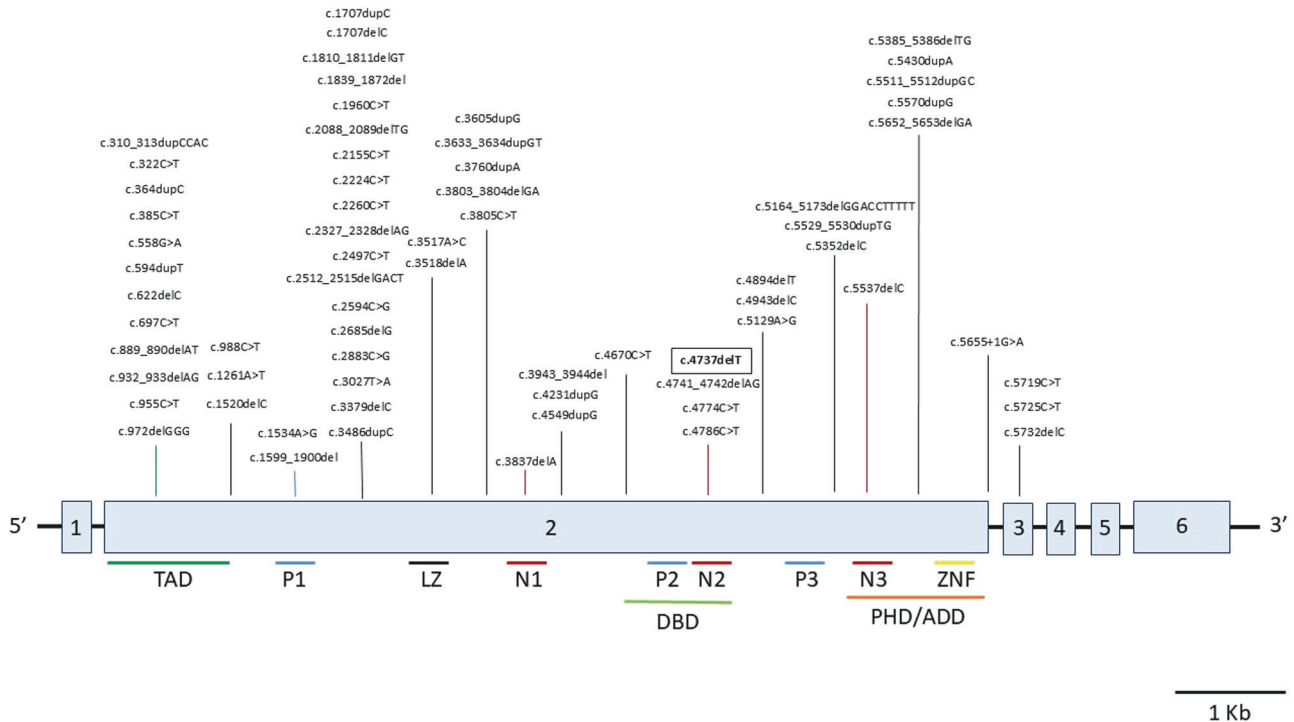
**Fig. 2** Two-dimensional ddPCR scatter plots for the analysis of the **TCF20** variant. **A** Proband's ddPCR scatter plot from peripheral blood. **B** Proband's father's ddPCR scatter plot from urothelial cells. Colored dots represent single droplets of emulsion carrying amplified DNA harboring the variant (green) or wildtype DNA (blue). Black dots represent droplets with no amplification of the target DNA

dysmorphisms, neurological findings, hepatic issues [23], and other organ system-specific impairments appear to be variable among cases (see Supplementary Table 1 for details), highlighting the morphological heterogeneity of *TCF20*-associated neurodevelopmental disorder.

### Germline mosaicism in *TCF20*

As the MiSeq error rates for homopolymer runs are equivalent to the proportion of reads with a single T deletion in the parents [42], we initially concluded there was no definitive indication of mosaicism in either parent using this approach in DNA extracted from blood. On MinION sequencing, the alignments were indistinguishable between the proband and two parents: Alignments from all three individuals showed over 50% of reads with a 6-nucleotide long thymine homopolymer, and 20% of aligned reads contained a 5-nucleotide long thymine homopolymer, highlighting the difficulty of accurately determining the count of residues in homopolymer sequences. The variant was, however, confidently detected by droplet digital PCR in 3.2% of paternal DNA extracted from urine (urothelial cells), confirming germline inheritance.

There are now four independent reports of either suspected (affected siblings carrying the same variant not detectable in parents) or confirmed germline mosaic variants in seven individuals with *TCF20*-associated disorders from four different families. The cases with germline mosaicism consist of 1) a pericentric chromosome 22 inversion affecting both the *TCF20* and *ACTN6B* genes resulting in intellectual disability and autism in brothers not detected in parental blood/lymphoblastoid cell-derived DNA [16], 2) a substitution (NM\_005650.1:c.2224 C > T) in siblings with autism not identified in parental blood-derived DNA [26], 3) a substitution (NM\_005650.1:c.558 G > A) in a male



**Fig. 3** Previously reported *TCF20* variants. All SNVs and indels (76 total) in *TCF20* are represented. Structural and copy number variants are not shown (see Supplementary Table 1 for a full list of previously reported copy number and structural variants). The variant from this report is framed in bold. Exons are indicated in light blue and numbered. Protein domains are detailed under the exon diagram. Abbreviations: TAD = transactivation domain; P1/2/3 = PEST domains 1/2/3; LZ = leucine zipper; N1/2/3 = nuclear localization signals 1/2/3; DBD = DNA-binding domain; PHD/ADD = Plant Homeodomain/ADD; ZNF = zinc finger domain



with intellectual disability and autism, found in low-level germline mosaic form (6.98%) in his mother's peripheral whole blood [21], and 4) in this report, a thymidine deletion (NM\_001378418.1:c.4737GT>G) in siblings with intellectual disability and autism identified in low-level germline mosaic form in the father's urine. These are in addition to 67.0% (61 out of 91) of cases in which *TCF20* variants have arisen apparently de novo (see Table 1), which could be cases of undetected germline mosaic inheritance from a parent.

Interestingly, Amiel et al. [43] found that 12 of 20 cases with *PAX2*-associated renal-coloboma syndrome were caused by either an insertion or a deletion of a single residue from a stretch of homopolymer of seven guanines. This may indicate the presence of a variant hotspot linked to strand slippage during replication, including in one familial case of suspected paternal germline mosaicism. A similar mechanism could be at play in the *TCF20* thymidine tract.

### Germline mosaicism proposed as a recurring mode of inheritance in neurodevelopmental disorders

Since germline mosaicism was first identified in studies of the mutagenic effects of mustard gas [44], it has been invoked to explain the presence of recurrent rare dominant disorders in children of unaffected parents [45–47]. More recently, attention has focused increasingly on the genetic architecture of neurodevelopmental disorders which appear to be highly enriched in de novo variants in key brain developmental genes [48–53], including in germline mosaic form [54]. Conditions with germline mosaicism as a recurrent mode of inheritance include Rett syndrome [55], Fragile X syndrome [56], and tuberous sclerosis complex (TSC) [57]. A recent study using a single-molecule Molecular Inversion Probe (smMIP) enrichment method followed by parental blood DNA sequencing at a depth of ~7000 found that nearly 4% of all cases of reported de novo variants in individuals affected by a range of NDDs were transmitted from mosaic parents, a proportion which is even higher in epileptic individuals with NDDs [58].

In addition, a number of case reports for a range of conditions have shown that variants which were initially thought to have arisen de novo were inherited from a parent with low-level (<10% in allele fractions, as defined in [59, 60] germline mosaicism [61–63]. In a study by Xu et al., approximately 10% of variants in the sodium channel gene *SCN1A* (which cause Dravet syndrome), initially considered de novo, were identified by deep amplicon resequencing to have resulted from either maternal or paternal mosaicism at allele frequencies ranging from 1.1 to 32.6% [62]. Consistently, Zillhardt et al. found that both maternal and paternal germline mosaicism accounts for up to 15% of recurrent malformations of cortical development [64], while Breuss et al. on analyzing sperm DNA from fathers of children with autism, revealed that many harbored germline mosaic autism-associated variants, with allele fractions ranging from 2 to 15% [63].

Furthermore, the inheritance of low-level germline mosaic variants is not limited to pathogenic variants, as many germline mosaic variants have been identified in healthy families with no associated clinical phenotypes. Rahbari et al. assessed rates of de novo SNVs in unaffected multi-sibling families to find that ~3.8% of these were in fact result of parental germline mosaicism, at alternate allele fractions ranging from 0.57% to 10.24%, with roughly half of maternal and half of paternal origin. In parallel, 1.3% of de novo variants were found to be shared among siblings, reaching 24% for de novo variants which were mosaic in >1% of parental blood-derived cells, and 50% for de novo variants that were present in >6% of parental blood-derived cells [65]. Another large study conducted by Jónsson et al. found that 57.2% of de novo variants (including SNVs and indels) shared by siblings were confirmed to be found in mosaic form in parental blood [47].

The selfish spermatogonial selection hypothesis suggests a mechanism to explain the prevalence of such pathogenic variants in the male germline. It highlights that a sperm cell that has acquired a variant that confers a selective advantage, in quantity (by clonal expansion) and/or quality (by increased motility and survivability), is more likely to survive and fertilize, even if this variant eventually leads to a disorder over the course of development [66–70]. This is of particular salience to the context of neurodevelopmental disorders since variants which enhance processes of cellular hyperproliferation and motility will, when brain-expressed, amplify the very mechanisms by which certain disorders of brain development, such as autism, macrocephaly and RASopathies, arise or are exacerbated [71]. Interestingly, this spermatogonial selection hypothesis appears to provide a parsimonious explanation for the correlation between de novo variants and advanced paternal age [68, 72].

### Limitations in the detection of germline mosaicism

The likely under-detection of germline mosaicism could be a consequence of both limitations with biological sampling and DNA sequencing technology.

First, if only one sibling inherits a variant, germline mosaicism in a parent will not be conspicuously suspected. Therefore, cases of germline mosaicism in which only one sibling is affected are likely to go unnoticed [73].

Second, either paternal or maternal germline mosaicism is difficult to identify if the variant exclusively affects cell types not included in the sequenced sample. In our study for example, the identification of the *TCF20* variant in a proportion of urothelial cells raises the possibility that the variant arose in the father early on in embryonic development, presumably affecting cells giving rise to sperm and urothelial cells, but not blood cells, a process described by Rhabari et al. [65]. This is consistent with a study showing that only about one third of sperm mosaic variants are detectable in blood cells [74], and data demonstrating that the analysis of the fathers' blood alone may underestimate the

**Table 1.** Variant types and modes of inheritance of disorder-associated *TCF20* variants, including those previously reported and those from our case report

	De novo	Inherited (maternally)	Inherited (paternally)	Germline mosaic (suspected or confirmed)	Undetermined	TOTAL	Fraction of all variants (%)
SNV	20	2	3	3	5	33	36,3
Indel	32	2	2	2	8	46	50,5
CNV	8	0	0	0	1	9	9,9
Other SV	1	0	0	2	0	3	3,3
Total number of variants	61	4	5	7	14	91	
Fraction of all variants (%)	67,0	4,4	5,5	7,7	15,4		

presence of germline mosaicism in fathers of children with intellectual disability syndromes [75].

Finally, even at very high read depths (e.g. 100,000x), limitations inherent to sequencing target preparation and sequencing technologies limit their ability to detect certain types of variants, irrespective of the allele frequency. First, Illumina MiSeq sequencing has a published error rate of 1.85–2.5% when sequencing homopolymer runs, specifically, 1.85% for guanine or cytosine 6- to 7 bp homopolymer runs [76], and ~2.5% for thymidine or adenine 7-bp homopolymer runs [42]. These previously reported error rates for homopolymer runs are equivalent to the proportion of reads with a single T deletion observed in the MiSeq results for both parents, which could, in our case, have been misinterpreted as the presence of the *TCF20* variant in the parents. Second, similar to Illumina MiSeq sequencing, the difficulty of basecalling ONT sequence-surrounding stretches of low complexity is a well-reported limitation of the technology [77–80]. For the flowcells used in this experiment (R9.4), the base composition in the pore is primarily determined from the three central nucleotides located within the pore. Thus, accurately determining the length of a homopolymer tract greater than four nucleotides (including the 7 nucleotide thymidine homopolymer seen in the proband) is technically challenging [81]. In our case this meant that neither MiSeq sequencing at 105,000 read depth nor MinION Oxford Nanopore sequencing at >15,000 read depth of targeted PCR products were able to reliably detect the proband's deletion [82, 83].

Considering the limitations linked to the complexity of hierarchical multicellular organism development—and the need for tissue source and sequencing technology to match a variant's distribution, nature, and allele fraction—germline mosaic variants may remain under-detected in *TCF20*-associated neurodevelopmental disorders and neurodevelopmental conditions more generally. In the future, it will be interesting to profile the distribution of different types of genetic variants that arise in germline mosaic form, investigate any mutagenic mechanistic links to environmental exposures [84], and understand how they might differentially affect coding versus noncoding genomic regions.

## CONCLUSION

In conclusion, we here broaden the spectrum of genetic variation, provide the second known case of confirmed germline mosaic *TCF20*-associated disorders and first known case of confirmed germline mosaic inheritance of the same *TCF20* variant in siblings. In light of a clearly growing subset of patients with a positive family history of *TCF20*-associated neurodevelopmental disorders [21], our report demonstrates that parental testing of individuals with *TCF20*-associated neurodevelopmental disorders is critical, and that special attention should be paid to capturing low-level mosaicism in order to provide the most accurate genetic counselling and recurrence risk assessments possible. Germline mosaic variants may represent a clinically salient yet under-detected mode of disorder-associated inheritance not only in *TCF20*-associated disorder, but other neurodevelopmental disorders as well [85–87].

## REFERENCES

- Wilfert AB, Sulovari A, Turner TN, Coe BP, Eichler EE. Recurrent de novo mutations in neurodevelopmental disorders: properties and clinical implications. *Genome Med.* 2017;9:101.
- Lin Y, Afshar S, Rajadhyaksha AM, Potash JB, Han S. A Machine Learning Approach to Predicting Autism Risk Genes: Validation of Known Genes and Discovery of New Candidates. *Front Genet.* 2020;11:500064.
- Huang S, Xu J, Li Y, Mo W, Lin X, Wang Y, et al. A syndrome featuring developmental disorder of the nervous system induced by a novel mutation in the *TCF20* gene, rarely concurrent immune disorders: a case report. *Front Genet.* 2023;14:1192668.
- Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell.* 2020;180:568–584.e23.
- Shen EH, Overly CC, Jones AR. The Allen Human Brain Atlas. *Trends in Neurosciences.* 2012;35:1–92. <https://doi.org/10.1016/j.tins.2012.09.005>
- Mitz AR, Philyaw TJ, Boccuto L, Shcheglovitov A, Sarasua SM, Kaufmann WE, et al. Identification of 22q13 genes most likely to contribute to Phelan McDermid syndrome. *European Journal of Human Genetics.* 2018;26:293–302.
- Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreservation and Biobanking.* 2015;13:307–8.
- Adams MS, Gammill LS, Bronner-Fraser M. Discovery of transcription factors and other candidate regulators of neural crest development. *Developmental Dynamics.* 2008;237:1021–33.
- Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature.* 2007;445:168–76.
- Sanz L, Moscat J, Diaz-Meco MT. Molecular characterization of a novel transcription factor that controls stromelysin expression. *Molecular and cellular biology.* 1995;15:3164–70.
- Rekdal C, Sjøttem E, Johansen T. The Nuclear Factor SPBP Contains Different Functional Domains and Stimulates the Activity of Various Transcriptional Activators. *Journal of Biological Chemistry.* 2000;275:40288–300.
- Gburcik V, Bot N, Maggolini M, Picard D. SPBP is a phosphoserine-specific repressor of estrogen receptor alpha. *Molecular and cellular biology.* 2005;25:3421–30.
- Feng C, Zhao J, Ji F, Su L, Chen Y, Jiao J. *TCF20* dysfunction leads to cortical neurogenesis defects and autistic-like behaviors in mice. *EMBO Rep.* 2020;21:e49239.
- Vetrini F, McKee S, Rosenfeld JA, Suri M, Lewis AM, Nugent KM, et al. De novo and inherited *TCF20* pathogenic variants are associated with intellectual disability, dysmorphic features, hypotonia, and neurological impairments with similarities to Smith-Magenis syndrome. *Genome Medicine.* 2019;11:12.
- Darvekar S, Rekda C, Johansen T, Sjøttem E. A phylogenetic study of SPBP and RAI1: evolutionary conservation of chromatin binding modules. *PLoS One.* 2013;8:e78907.
- Babbs C, Lloyd D, Pagnamenta AT, Twigg SR, Green J, McGowan SJ, et al. De novo and rare inherited mutations implicate the transcriptional coregulator *TCF20*/SPBP in autism spectrum disorder. *Journal of Medical Genetics.* 2014;51:737–47.
- Schäffgen J, Cremer K, Becker J, Wieland T, Zink AM, Kim S, et al. De novo nonsense and frameshift variants of *TCF20* in individuals with intellectual disability and postnatal overgrowth. *European Journal of Human Genetics.* 2016;24:1739–45.
- Kurtas N, Arrigoni F, Errichiello E, Zucca C, Maghini C, D'Angelo MG, et al. Chromothripsis and ring chromosome 22: A paradigm of genomic complexity in the Phelan-McDermid syndrome (22q13 deletion syndrome). *Journal of Medical Genetics.* 2018;55:269–77.
- Upadia J, Gonzales PR, Atkinson TP, Schroeder HW, Robin NH, Rudy NL, et al. A previously unrecognized 22q13.2 microdeletion syndrome that encompasses *TCF20* and *TNFRSF13C*. *American Journal of Medical Genetics, Part A.* 2018;176:2791–7.
- Lévy J, Cogan G, Maruani A, Maillard A, Dupont C, Drunat S, et al. Rare and de novo duplications containing *TCF20* are associated with a neurodevelopmental disorder. *Clinical Genetics.* 2022;101:364–70. <https://doi.org/10.1111/cge.14099>
- Schneeweiss, MR, Dale, B & Ejaz, R. Diagnosis and clinical presentation of two individuals with a rare *TCF20* pathogenic variant. *BMJ Case Reports* (2022), 15 <https://doi.org/10.1136/bcr-2022-248995>.
- Svorenova T, Romito LM, Colangelo I, Han V, Jech R, Prokisch H, et al. Dystonia as a prominent feature of *TCF20*-associated neurodevelopmental disorder: Expanding the phenotype. *Parkinsonism Relat Disord.* 2022;102:89–91.
- Córdoba-Jover B, Ribera J, Portolés I, Lecue E, Rodríguez-Vita J, Pérez-Sisqués L, et al. *Tcf20* deficiency is associated with increased liver fibrogenesis and alterations in mitochondrial metabolism in mice and humans. *Liver Int.* 2023;43:1822–36.
- Lelieveld SH, Reijnders MR, Pfundt R, Yntema HG, Kamsteeg EJ, de Vries P, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nature Neuroscience.* 2016;19:1194–6.
- McRae JF, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature.* 2017;542:433–8. <https://doi.org/10.1038/nature21062>
- Torti E, Keren B, Palmer EE, Zhu Z, Afenjar A, Anderson IJ, et al. Variants in *TCF20* in neurodevelopmental disability: description of 27 new patients and review of literature. *Genetics in Medicine.* 2019;21:2036–42. <https://doi.org/10.1038/s41436-019-0454-9>
- Lift Genome Annotations. <https://genome.ucsc.edu/cgi-bin/hgLiftOver>.

28. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*. 2013;14:178–92.
29. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91.
30. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biology*. 2016;17:122.
31. Samocha KE, Robinson EB, Sanders SJ, Stevens C, Sabo A, McGrath LM, et al. A framework for the interpretation of de novo mutation in human disease. *Nature Genetics*. 2014;46:944–50.
32. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research*. 2003;31:3812–4.
33. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*. 2013;Chapter 7:7 <https://doi.org/10.1002/0471142905.hg0720s76>
34. Lopes MC, Joyce C, Ritchie GR, John SL, Cunningham F, Asimit J, et al. A combined functional annotation score for non-synonymous variants. *Human heredity*. 2012;73:47–51.
35. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*. 2012;13:134.
36. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–60.
37. Musgrave, SM et al. Genetic diagnostic outcomes from a 10-year research programme in autism in Aotearoa New Zealand. *Journal of the Royal Society of New Zealand* 1–17 (2024) <https://doi.org/10.1080/03036758.2024.2394128>.
38. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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;17:405–24.
39. Uniprotkb P, Helpy, URE UniProtKB. *Urease Subunit alpha* vol. 14916 1–7 <https://www.uniprot.org/uniprot/P14916> (2018).
40. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *American Journal of Human Genetics*. 2009;84:524–33.
41. McRae, JF et al. Prevalence, phenotype and architecture of developmental disorders caused by de novo mutation: The Deciphering Developmental Disorders Study. *bioRxiv* (2016) <https://doi.org/10.1101/049056>.
42. Stoler N, Nekrutenko A. Sequencing error profiles of Illumina sequencing instruments. *NAR Genomics and Bioinformatics*. 2021;3:019 <https://doi.org/10.1093/nargab/lqab019>
43. Amiel J, Audollent S, Joly D, Dureau P, Salomon R, Tellier AL, et al. PAX2 mutations in renal-coloboma syndrome: mutational hotspot and germline mosaicism. *Eur J Hum Genet*. 2000;8:820–6.
44. AUERBACH C. Chemically induced mosaicism in *Drosophila melanogaster*. *Proceedings of the Royal Society of Edinburgh. Section B: Biology*. 1946;62:211–22.
45. Reed TE, Falls HF. A pedigree of aniridia with a discussion of germinal mosaicism in man. *American journal of human genetics*. 1955;7:28–38.
46. Bakker E, Van Broeckhoven C, Bonten EJ, van de Vooren MJ, Veenema H, Van Hul W, et al. Germline mosaicism and Duchenne muscular dystrophy mutations. *Nature*. 1987;329:554–6.
47. Jónsson H, Sulem P, Arnadóttir GA, Pálsson G, Eggertsson HP, Kristmundsdóttir S, et al. Multiple transmissions of de novo mutations in families. *Nature Genetics*. 2018;50:1674–80.
48. Iossifov I, O’Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515:216–21.
49. Krumm N, O’Roak BJ, Shendure J, Eichler EE. A de novo convergence of autism genetics and molecular neuroscience. *Trends Neurosci*. 2014;37:95–105.
50. Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. *Genome Biology*. 2016;17:241.
51. Turner TN, Coe BP, Dickel DE, Hoekzema K, Nelson BJ, Zody MC, et al. Genomic Patterns of De Novo Mutation in Simplex Autism. *Cell*. 2017;171:710–722.e12.
52. Alonso-Gonzalez A, Rodríguez-Fontenla C, Carracedo A. De novo Mutations (DNMs) in Autism Spectrum Disorder (ASD): Pathway and Network Analysis. *Front. Genet*. 2018;9:406.
53. Veltman JA, Brunner HG. De novo mutations in human genetic disease. *Nat Rev Genet*. 2012;13:565–75.
54. The Deciphering Developmental Disorders Study. Prevalence, phenotype and architecture of developmental disorders. *bioRxiv* (2016).
55. Wen Y, Wang J, Zhang Q, Yang X, Wei L, Bao X. MECP2 germline mosaicism plays an important part in the inheritance of Rett syndrome: a study of MECP2 germline mosaicism in males. *BMC medicine*. 2023;21:155.
56. Chiurazzi P, Kozak L, Neri G. Unstable triplets and their mutational mechanism: Size reduction of the CGG repeat vs. germline mosaicism in the fragile X syndrome. *American Journal of Medical Genetics*. 1994;51:517–21.
57. Rose VM, Au KS, Pollom G, Roach ES, Prashner HR, Northrup H. Germ-line mosaicism in tuberous sclerosis: how common? *American journal of human genetics*. 1999;64:986–92.
58. Lecoquiere F, Cassinari K, Drouot N, May A, Fournieux S, Charbonnier F, et al. Assessment of parental mosaicism rates in neurodevelopmental disorders caused by apparent de novo pathogenic variants using deep sequencing. *Scientific Reports*. 2024;14:5289.
59. Jin P, Gao X, Wang M, Qian Y, Yang J, Yang Y, et al. Case Report: Identification of Maternal Low-Level Mosaicism in the Dystrophin Gene by Droplet Digital Polymerase Chain Reaction. *Frontiers in Genetics*. 2021;12:686993 <https://doi.org/10.3389/fgene.2021.686993>
60. Karolák JA, Liu Q, Xie NG, Wu LR, Rocha G, Fernandes S, et al. Highly Sensitive Blocker Displacement Amplification and Droplet Digital PCR Reveal Low-Level Parental FOXF1 Somatic Mosaicism in Families with Alveolar Capillary Dysplasia with Misalignment of Pulmonary Veins. *Journal of Molecular Diagnostics*. 2020;22:447–56. <https://doi.org/10.1016/j.jmoldx.2019.12.007>
61. Gajecka M. Unrevealed mosaicism in the next-generation sequencing era. *Molecular Genetics and Genomics*. 2016;291:513–30.
62. Xu X, Yang X, Wu Q, Liu A, Yang X, Ye AY, et al. Amplicon Resequencing Identified Parental Mosaicism for Approximately 10% of “de novo” *SCN1A* Mutations in Children with Dravet Syndrome. *Human Mutation*. 2015;36:861–72.
63. Breuss, MW et al. Quantification of autism recurrence risk by direct assessment of paternal sperm mosaicism. *bioRxiv* 208165 (2017) <https://doi.org/10.1101/208165>.
64. Zillhardt JL, Poirier K, Broix L, Lebrun N, Elmorjani A, Martinovic J, et al. Mosaic parental germline mutations causing recurrent forms of malformations of cortical development. *European Journal of Human Genetics*. 2016;24:611–4.
65. Rahbari R, Wuster A, Lindsay SJ, Hardwick RJ, Alexandrov LB, Al Turki S, et al. Timing, rates and spectra of human germline mutation. *Nature Genetics*. 2016;48:126–33. <https://doi.org/10.1038/ng.3469>
66. Choi S-K, Yoon S-R, Calabrese P, Arnheim N. A germ-line-selective advantage rather than an increased mutation rate can explain some unexpectedly common human disease mutations. *Proceedings of the National Academy of Sciences*. 2008;105:10143–8.
67. Goriely A, McVean GAT, Röjmyr M, Ingemarsson B, Wilkie AOM. Evidence for Selective Advantage of Pathogenic FGFR2 Mutations in the Male Germ Line. *Science*. 2003;301:643–6.
68. Yoon S-R, Choi SK, Eboreime J, Gelb BD, Calabrese P, Arnheim N. Age-Dependent Germline Mosaicism of the Most Common Noonan Syndrome Mutation Shows the Signature of Germline Selection. *The American Journal of Human Genetics*. 2013;92:917–26.
69. Qin J, Calabrese P, Tiemann-Boege I, Shinde DN, Yoon SR, Gelfand D, et al. The Molecular Anatomy of Spontaneous Germline Mutations in Human Testes. *PLoS Biology*. 2007;5:e224.
70. Lim J, Maher GJ, Turner GD, Dudka-Ruszkowska W, Taylor S, Rajpert-De Meyts E, et al. Selfish spermatogonial selection: Evidence from an immunohistochemical screen in testes of elderly men. *PLoS ONE*. 2012;7:42382 <https://doi.org/10.1371/journal.pone.0042382>
71. Courchesne, E, T Pramparo, VH Gazestani, MV Lombardo, K Pierce, NE Lewis et al. The ASD Living Biology: from cell proliferation to clinical phenotype. *Mol Psychiatry*. 2018;24:88–107 <https://doi.org/10.1038/s41380-018-0056-y>.
72. Goriely A, McGrath JJ, Hultman CM, Wilkie AOM, Malaspina D. “Selfish Spermatogonial Selection”: A Novel Mechanism for the Association Between Advanced Paternal Age and Neurodevelopmental Disorders. *American Journal of Psychiatry*. 2013;170:599–608.
73. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nat Rev Genet*. 2013;14:307–20.
74. Yang, X et al. Temporal stability of human sperm mosaic mutations results in life-long threat of transmission to offspring. *bioRxiv* (2020).
75. Frisk S, Wachtmeister A, Laurell T, Lindstrand A, Jäntti N, Malmgren H, et al. Detection of germline mosaicism in fathers of children with intellectual disability syndromes caused by de novo variants. *Molecular Genetics and Genomic Medicine*. 2022;10:1880 <https://doi.org/10.1002/mggg.3.1880>
76. Shin S, Park J. Characterization of sequence-specific errors in various next-generation sequencing systems. *Molecular BioSystems*. 2016;12:914–22. <https://doi.org/10.1039/c5mb00750j>
77. Delahaye C, Nicolas J. Sequencing DNA with nanopores: Troubles and biases. *PLoS ONE*. 2021;16:0257521 <https://doi.org/10.1371/journal.pone.0257521>
78. Cornelis S, Gansemans Y, Deleye L, Deforce D, Van Nieuwerburgh F. Forensic SNP Genotyping using Nanopore MinION Sequencing. *Scientific Reports*. 2017;7:41759 <https://doi.org/10.1038/srep41759>

79. Srivathsan A, Baloğlu B, Wang W, Tan WX, Bertrand D, Ng A, et al. A MinION™-based pipeline for fast and cost-effective DNA barcoding. *Molecular Ecology Resources*. 2018. <https://doi.org/10.1111/1755-0998.12890>
80. Oikonomopoulos S, Wang YC, Djambazian H, Badescu D, Ragoussis J. Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Scientific Reports*. 2016;6:31602 <https://doi.org/10.1038/srep31602>
81. Clive Brown | Owl Stretching with Examples - YouTube. <https://www.youtube.com/watch?v=JmncdnQgaIE>.
82. Laehnmann D, Borkhardt A, McHardy AC. Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. *Briefings in Bioinformatics*. 2016;17:154–79. <https://doi.org/10.1093/bib/bbv029>
83. Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol*. 2018;19:90.
84. Beal MA, Meier MJ, Williams A, Rowan-Carroll A, Gagné R, Lindsay SJ, et al. Paternal exposure to benzo(a)pyrene induces genome-wide mutations in mouse offspring. *Communications Biology*. 2019;2:228 <https://doi.org/10.1038/s42003-019-0476-5>
85. Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, et al. Parental Somatic Mosaicism Is Underrecognized and Influences Recurrence Risk of Genomic Disorders. *The American Journal of Human Genetics*. 2014;95:173–82.
86. Zhang Q, Yang X, Wang J, Li J, Wu Q, Wen Y, et al. Genomic mosaicism in the pathogenesis and inheritance of a Rett syndrome cohort. *Genetics in Medicine*. 2019;21:1330–8. <https://doi.org/10.1038/s41436-018-0348-2>
87. Krupp DR, Barnard RA, Duffourd Y, Evans SA, Mulqueen RM, Bernier R, et al. Exonic Mosaic Mutations Contribute Risk for Autism Spectrum Disorder. *The American Journal of Human Genetics*. 2017;101:369–90.

## ACKNOWLEDGEMENTS

The family is gratefully acknowledged for their participation in the study. We thank Kristine Boxen from Auckland Genomics at The University of Auckland for assistance with Sanger sequencing. The study was supported by the New Zealand eScience Infrastructure ([www.nesi.org.nz](http://www.nesi.org.nz)). JCJ was supported by the Neurological Foundation of New Zealand and a government-funded Rutherford Discovery Fellowship administered by the Royal Society of New Zealand.

## AUTHOR CONTRIBUTIONS

JCJ, RGS, and KL designed the experiments. JP, WW performed DNA-based laboratory experiments. JT, NG and BT performed the clinical evaluation, and JT, NG and DRL clinically evaluated the proposed variants. KL and JCJ performed data and bioinformatic analysis. JP and JCJ performed the *TCF20* variant analyses. KMD performed ddPCR analyses. JP and JCJ wrote the manuscript with contribution from all authors.

## FUNDING

Open Access funding enabled and organized by CAUL and its Member Institutions.

## COMPETING INTERESTS

The authors declare no competing interests.

## ETHICS APPROVAL

Ethical approval was obtained by the Northern B Health and Disability Ethics Committee (12/NTB/59).

## PATIENT CONSENT

All procedures were performed in accordance with the ethical standards of the institutional and national responsible committees. Written informed consent to be included in the study was obtained for all participants.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s10038-025-01323-3>.

**Correspondence** and requests for materials should be addressed to Jessie C. Jacobsen.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025