Received: 27 December 2021 | Revised: 4 February 2022 | Accepted: 18 February 2022

DOI: 10.1002/ajmg.a.62743

ORIGINAL ARTICLE

AMERICAN JOURNAL OF medical genetics PART A WILEY

Germline mosaicism of a missense variant in *KCNC2* in a multiplex family with autism and epilepsy characterized by long-read sequencing

Elvisa Mehinovic[1] | Teddi Gray[2] | Meghan Campbell[2] | Jenny Ekholm[3] | Aaron Wenger[3] | William Rowell[3] | Ari Grudo[3] | Jane Grimwood[4] | Jonas Korlach[3] | Christina Gurnett[5] | John N. Constantino[2] | Tychele N. Turner[1]

[1]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA

[2]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA

[3]Pacific Biosciences, Menlo Park, California, USA

[4]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA

[5]Department of Neurology, Washington University School of Medicine, St. Louis, Missouri, USA

Correspondence
Tychele N. Turner, Department of Genetics, Washington University School of Medicine, 4523 Clayton Avenue, Campus Box 8232, St. Louis, MO 63110, USA.
Email: tychele@wustl.edu

Funding information
National Institutes of Health, Grant/Award Numbers: R00MH117165, P50HD103525; Pacific Biosciences

Abstract

Currently, protein-coding de novo variants and large copy number variants have been identified as important for ~30% of individuals with autism. One approach to identify relevant variation in individuals who lack these types of events is by utilizing newer genomic technologies. In this study, highly accurate PacBio HiFi long-read sequencing was applied to a family with autism, epileptic encephalopathy, cognitive impairment, and mild dysmorphic features (two affected female siblings, unaffected parents, and one unaffected male sibling) with no known clinical variant. From our long-read sequencing data, a de novo missense variant in the *KCNC2* gene (encodes Kv3.2) was identified in both affected children. This variant was phased to the paternal chromosome of origin and is likely a germline mosaic. In silico assessment revealed the variant was not in controls, highly conserved, and predicted damaging. This specific missense variant (Val473Ala) has been shown in both an ortholog and paralog of Kv3.2 to accelerate current decay, shift the voltage dependence of activation, and prevent the channel from entering a long-lasting open state. Seven additional missense variants have been identified in other individuals with neurodevelopmental disorders ($p = 1.03 \times 10^{-5}$). *KCNC2* is most highly expressed in the brain; in particular, in the thalamus and is enriched in GABAergic neurons. Long-read sequencing was useful in discovering the relevant variant in this family with autism that had remained a mystery for several years and will potentially have great benefits in the clinic once it is widely available.

KEYWORDS
autism, channel, epilepsy, genetics, genomics, long-read sequencing

1 | INTRODUCTION

Autism is a complex neurodevelopmental disorder with a genetic component from both common and rare variation. Common variation contributes to ~50% of autism (Gaugler et al., 2014) and individuals with autism have excess polygenic risk (Weiner et al., 2017). In terms of rare variation, ~30% of all cases can be explained by de novo copy number variants and de novo single-nucleotide variants or small insertions/deletions that are loss-of-function or severe missense changes (Iossifov et al., 2014). Several other types of rare genetic variants have been implicated in autism including rare inherited variants (Iossifov et al., 2015; Krumm et al., 2015; Wilfert et al., 2021), recessive variants (Doan et al., 2019), and de novo noncoding putative regulatory variants (An et al., 2018; Padhi et al., 2021; Turner et al., 2016; Turner et al., 2017; Zhou et al., 2019). Another important feature of autism is its sex ratio of 4:1 with 80% of all cases being male. There has been a documented excess of rare de novo variants (DNVs) and recessive variants in females with autism (Doan et al., 2019; Iossifov et al., 2014; Jacquemont et al., 2014; Levy et al., 2011; Neale et al., 2012; Sanders et al., 2015; Turner et al., 2019) and families with multiple affected females with autism have been prioritized for gene discovery (Turner et al., 2015).

Recently, long-read sequencing technologies have emerged and enabled access to variation previously intractable with short-read sequencing (Wenger et al., 2019). Long-read sequencing is promising for the objective of precision genomics. We previously defined precision genomics as "determining all possible relevant genomic variation within an individual to the precise nucleotide" (Sams et al., 2022) and emphasized the importance of precision genomics in the framework of precision medicine (Sams et al., 2022). Furthermore, family-based studies have been highly successful in identifying de novo variants and long-read sequencing coupled with family-based analyses should prove especially fruitful for gene discovery in families with no known genetic cause.

This study was focused on a family with two females with autism and epilepsy, their unaffected male sibling, and their unaffected parents. Both affected siblings had previously undergone karyotype, chromosome microarray analyses in 2010 (Affymetrix 6.0 arrays), and exome sequencing in 2013 (Baylor Medical Genetics Lab) in the clinic with no identification of relevant genomic variation. Since there were two affected females in the family and they also shared additional phenotypes including epilepsy, we hypothesized that long-read sequencing technologies would uncover the relevant genomic variation in this family. Pacific Biosciences long-read sequencing, 10× Genomics sequencing, and Bionano optical mapping were applied for each family member and from this analysis a missense de novo variant in the *KCNC2* gene (encodes the Kv3.2 potassium channel) was identified in both females with autism but not in their unaffected brother. This variant arose on the paternal chromosome in both individuals and is likely to be a germline mosaic event. There are multiple lines of evidence that support the importance of this variant for the family in this study and provides an example of the use of precision genomics, via long-read sequencing, in autism.

2 | MATERIALS AND METHODS

2.1 | Clinical features of affected sisters

The probands are sisters who were 25 and 23 at the time of publication. The oldest sister was diagnosed with epileptic encephalopathy, hypotonic cerebral palsy, and developmental delay in the first year of life. She was diagnosed with autism at age 2. She has cognitive impairment, microcephaly, kyphoscoliosis, femoral anteversion and external tibial torsion, hip dysplasia, and a pilonidal cyst. Her seizures consisted of atonic and atypical absence and were accompanied by generalized epileptiform discharges on EEG. The younger sister was also diagnosed with epileptic encephalopathy, hypotonic cerebral palsy, and developmental delay in the first year of life. She has cognitive impairment, autism, atypical absence epilepsy with bifrontally predominant or generalized 3–4 Hz discharges on EEG. They both had clinical chromosome microarray analysis in 2010 (Affymetrix 6.0 arrays) and exome sequencing in 2013 (Baylor Medical Genetics Lab).

2.2 | Family enrollment in the study

The Washington University in St. Louis Institutional Review Board approved this research under IRB ID #202002147. The Washington University Federal Wide Assurance (FWA) number is FWA00002284, the Washington University IRB is IRB00009237, and the Washington University Protocol Adherence Review Committee (PARC) IRB is IRB00005594. The family (PB.100) was consented using approved Consent and Assent forms (approved in IRB ID #202002147) during a visit to the Washington University in St. Louis Child Psychiatry Clinic. During this visit a total of 2 × 25 ml tubes of blood were collected from each of the five family members (parents, two female children with autism and epilepsy, and one unaffected male sibling). One 25 ml tube of blood was taken to the McDonnell Genome Institute for high molecular weight DNA extraction through their core services and the other was taken to the Washington University in St. Louis Genome Engineering and iPSC center for PBMC storage through their core services.

2.3 | Genomic technologies

Both the 10× Genomics and the Bionano technologies were applied to DNA from all family members at the McDonnell Genome Institute through their core services following standard protocols. The PacBio HiFi sequencing from DNA in all family members was performed at the HudsonAlpha Genome Sequencing Center through a PacBio SMRT grant.

2.4 | Analysis of 10× Genomics genome data

Longranger 2.2.2 (https://support.10xgenomics.com/genome-exome/software/downloads/latest) was run on everyone using the longranger wgs command, fastq as input, FreeBayes (Garrison, 2012) as the variant

caller, and the 10x GRCh38-2.1.0 reference genome data. The output loupe files were visualized in the Loupe 2.1.2 browser (https://support.10xgenomics.com/genome-exome/software/downloads/latest). Summary statistics and variant calls were assessed in the browser.

2.5 | Analysis of PacBio HiFi genome data in family PB.100

Each CCS fastq file was aligned to build 38 of the human genome (GRCh38_full_analysis_set_plus_decoy_hla.fa) using pbmm2 (https://github.com/PacificBiosciences/pbmm2) version 1.3.0 align. Structural variants were called using pbsv (https://github.com/PacificBiosciences/pbsv) version 2.3.0. SNVs/indel GVCFs were called using DeepVariant (Poplin et al., 2018) version 1.0.0 and the GVCFs were joint genotyped using GLNexus version 1.2.7 (Yun et al., 2020). Post-calling, PLINK (Purcell et al., 2007) was run to confirm all family relationships as correct. Exomiser (Robinson et al., 2014) version 12.1.0 was run on the joint-genotyped vcf file using the HPO term HP:0000729. Two different assemblers [HiCanu, Canu version 2.0 (Nurk et al., 2020) and Hifiasm, version 0.13-r307 (Cheng et al., 2021)] were utilized to generate de novo assemblies for each individual.

2.6 | Analysis of PacBio HiFi genome data from the Pangenome project

There were 49 individuals from the HPRC with publicly available PacBio HiFi data. Individual CCS bam files were downloaded from https://s3-us-west-2.amazonaws.com/human-pangenomics/index.html?prefix=submissions/ and converted to fastq files using PacBio bam2fastq version 1.3.0. Each CCS fastq file was then aligned to build 38 of the human genome (GRCh38_full_analysis_set_plus_decoy_hla.fa) using pbmm2 version 1.3.0 align. Structural variants were called using pbsv version 2.3.0. SNVs/indel GVCFs were called using DeepVariant version 1.0.0 and the GVCFs were joint genotyped using GLNexus version 1.2.7.

2.7 | Sanger confirmation of the KCNC2 DNV

Primers (Forward primer: GATCTGTTATGTTCCAGAAGTCGAT, Reverse primer: TAGTGAGCACACACAGTTCAAAAAC) were designed to target the exon (hg38: chr12:75050392-75050761) containing the Val473Ala variant using Primer3 (https://bioinfo.ut.ee/primer3-0.4.0/). The region was amplified using the Phusion High-Fidelity PCR Master Mix (ThermoFisher) and Sanger sequencing was performed at GeneWiz.

2.8 | Evolutionary assessment of the KCNC2 exon containing the DNV

We assessed the conservation of the *KCNC2* exon containing the DNV by examining the phyloP (Pollard et al., 2010) hg38 100-way scores from http://hgdownload.cse.ucsc.edu/goldenpath/hg38/phyloP100way/hg38.phyloP100way.bw. The scores were converted from a bigwig to a bedgraph (https://hgdownload.cse.ucsc.edu/admin/exe/linux.x86_64/bigWigToBedGraph) and the total number of positions in the file was calculated as well as the number of positions with a score less than the score of the DNV position (phyloP score = 9.29). Of the total 2661892195 positions in the file, 2660671278 had a score lower than the DNV site indicating this site is in the top 0.046% most conserved sites in the genome.

2.9 | KCNC2 DNVs identified in other individuals with neurodevelopmental disorders

The literature (Kaplanis et al., 2020; Rademacher et al., 2020; Satterstrom et al., 2020; Vetri et al., 2020) was searched for DNVs in individuals with neurodevelopmental disorders and identified seven additional individuals with DNVs in *KCNC2* (Table 1). To test whether these DNVs combined with the DNVs in the present study were enriched in the *KCNC2* gene, the denovolyzeR (Samocha et al., 2014; Ware et al., 2015) and chimpanzee-human (Coe et al., 2019) tests were run on the data. The DNVs were plotted on the protein (UniProtKB Q96PR1 [KCNC2_HUMAN]) using PROTTER (Omasits et al., 2013).

2.10 | Expression assessment of KCNC2

The expression of the KCNC2 gene was examined in different adult tissues in the GTEX database (https://gtexportal.org/home/gene/KCNC2). The expression of this gene across different timepoints and brain regions was visualized in the Brainspan database (https://www.brainspan.org/). Cell-type specific expression was visualized in the Allen Brain Map data (https://portal.brain-map.org) in the whole cortex and hippocampus from an 8-week-old mouse.

2.11 | Conservation to Drosophila melanogaster ortholog and Homo sapiens paralog

Initially, we observed a similar amino acid sequence (PVPVIV) in the human paralog Kv1.1 that is encoded by the *KCNA1* gene in Peters et al. (2011, fig. 1). In that study, the authors compared the *Drosophila melanogaster* ortholog called Shaker to the human protein called Kv1.1 and found there was a highly conserved region that includes the PVPVIV amino acids. Shaker is an ortholog of Kv3.2 and Kv1.1 is a paralog of Kv3.2. The study by Peters et al. (2011) further functionally assessed a variant in these two proteins and it is the same specific amino acid change we observe in Kv3.2. It is Kv3.2 Val473Ala, Shaker Val478Ala, and Kv1.1 Val408Ala. To quantitate the conservation formally we performed BLAST (Altschul et al., 1990) two sequences (https://blast.ncbi.nlm.nih.gov/Blast.cgi?BLAST_SPEC=blast2seq&LINK_LOC=align2seq&PAGE_TYPE=BlastSearch) comparing Shaker

TABLE 1 Individuals with neurodevelopmental disorders with DNVs in the *KCNC2* gene

Sample	Publication	Sex	Autism	Epilepsy	Developmental delay or intellectual disability	Protein HGVS	SIFT prediction for variant	PolyPhen prediction for variant	CADD score (PHRED)
PB.100.p1	Current study	F	Y	Y	Y	NP_631875.1: p.Val473Ala	deleterious (0)	Probably damaging (0.994)	26.7
PB.100.p2	Current study	F	Y	Y	Y	NP_631875.1: p.Val473Ala	deleterious (0)	Probably damaging (0.994)	26.7
26726	Kaplanis et al. (2020)	NA	NA	NA	Y	NP_631875.1: p.Thr437Ala	deleterious (0)	Probably damaging (1)	26.2
73575	Kaplanis et al. (2020)	NA	NA	NA	Y	NP_631875.1: p.Thr464Ile	deleterious (0)	Possibly damaging (0.897)	26.7
98034	Kaplanis et al. (2020)	NA	NA	NA	Y	NP_631875.1: p. Thr579Met	deleterious low confidence (0.05)	Possibly damaging (0.772)	26.9
DDD13k.00707	Kaplanis et al. (2020)	NA	NA	NA	Y	NP_631875.1: p. Val462Met	deleterious (0)	Probably damaging (0.999)	25.8
G01-GEA-114-HI	Satterstrom et al. (2020)	M	Y	NA	NA	NP_631875.1: p.Gly16Asp	deleterious (0)	Probably damaging (0.999)	26.4
Patient 3	Rademacher et al. (2020)	F	NA	Y	Y	NP_631875.1: p. Asp167Tyr	deleterious (0.01)	Possibly damaging (0.805)	27.4
Vetri_individual	Vetri et al. (2020)	M	NA	Y	Y	NP_631875.1: p.Val471Leu	deleterious (0)	Probably damaging (0.987)	25.4

Abbreviations: NA, not available; Y, yes; N, no; F, female; M, male.

and Kv3.2 as well as comparing Kv3.2 and Kv1.1. Each of these formal tests showed high conservation of the S6 transmembrane domain in which the amino acid change resides.

3 | RESULTS

3.1 | Genomic assessment of PB.100

Three advanced genomic technologies were applied to family PB.100 (Figure 1a) including PacBio HiFi long-read sequencing, 10× Genomics, and Bionano optical mapping to comprehensively characterize genomic variants in all five family members. By focusing on rare (private, inherited) and de novo variants only one potentially relevant variant was identified regarding the phenotype of autism in this family; the missense variant in the *KCNC2* gene was identified in both the 10× Genomics and PacBio HiFi technologies. Bionano optical mapping is not able to detect single-nucleotide variants; therefore, it was not possible to identify the variant in the Bionano data.

3.2 | Missense variant in KCNC2 gene

Exomiser (Robinson et al., 2014) was run in the Genomiser implementation to assess all variants in the genome for potential relevance to the autism phenotype of the affected females. One variant was prioritized as de novo in both females with autism but was not present in any other family member (Figure 1b). This variant chr12: g.75050587A>G (human genome build 38) encoded for a missense change of a Valine to Alanine at amino acid 473 (KCNC2: NM_001260497.1:c.1418T>C:p.(Val473Ala)). The Exomiser score was 0.8, the phenotype score was 0.5, and the variant score was 1.0. Several variant assessment programs rated this variant as severe including a Polyphen2 (Adzhubei et al., 2013) damaging score of 0.999, SIFT (Kumar et al., 2009) damaging score of 0.010, MutationTaster (Schwarz et al., 2014) pathogenic score of 1.000, and a CADD (Kircher et al., 2014) score of 26.700. The variant was not present in the gnomAD (Karczewski et al., 2020) database and was also not present in our joint-called dataset of 49 publicly available PacBio HiFi genomes. The reference nucleotide and amino acid are completely conserved across all 100 vertebrates underlying the 100-way

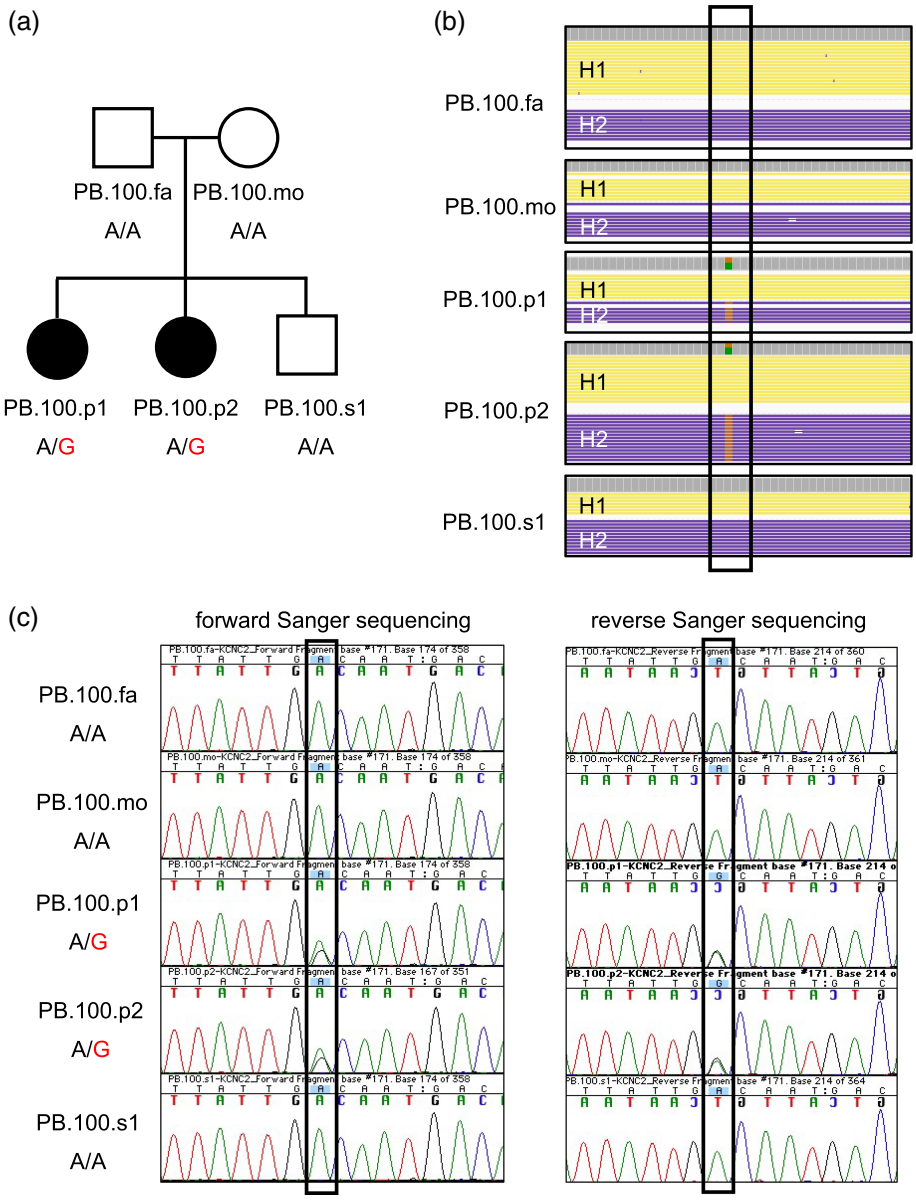

FIGURE 1 Family assessed in this study and identification of de novo missense variant in the *KCNC2* gene. (a) Pedigree of family PB.100 with unaffected parents, two female children with autism and epilepsy, and one unaffected male child. Shown below each individual is their genotype for the *KCNC2* variant (chr12:g.75050587A>G). As can be seen, the variant is de novo and only seen in the two affected individuals (PB.100.p1, PB.100.p2). (b) PacBio read data at and around the de novo *KCNC2* variant position (shown in box). Physically phased read data is shown and is labeled with H1 (haplotype 1) and H2 (haplotype 2) for each individual. The de novo *KCNC2* variant is only identified in the two affected individuals (PB.100.p1, PB.100.p2). (c) Sanger confirmation of the *KCNC2* variant detected only in the two affected individuals (PB.100.p1, PB.100.p2). The confirmation is seen in both the forward and reverse Sanger sequencing data

vertebrate alignment available in the UCSC browser (Kent et al., 2002). Regarding phyloP (Pollard et al., 2010) scores, this nucleotide position was in the top 0.05% of all bases in the genome indicating it is an extremely conserved base. We confirmed that the variant was present as de novo in both the PacBio and 10× data. Sanger sequencing was also performed on all family members to confirm the variant as de novo in the two affected females and was not present in any other family member (Figure 1c).

3.3 | Read-backed phasing of the data

Read-backed phasing of the PacBio data was performed using the Whatshap (Martin et al., 2016) tool on our DeepVariant (Poplin et al., 2018) calls. Through this analysis, a phase-informative single-nucleotide variant was identified 4880 bp away from the de novo variant (Figure 2). The variant was present on the same physical chromosome as the de novo variant in both females with autism. The informative variant has been inherited on the paternal chromosome suggesting the de novo variant arose as a germline mosaic variant in the paternal germline.

3.4 | Functional consequence of the Val473Ala missense variant

The *KCNC2* gene encodes the Kv3.2 potassium channel. The Kv3.2 potassium channel has six transmembrane domains and the Val473Ala missense variant resides at the last residue of the sixth transmembrane domain (S6) near to the intracellular space (Figure 3). This amino acid is highly conserved across orthologous (Figure 4a) and paralogous potassium channels (Figure 4b). In particular, the *exact same amino acid change* was identified in the Kv1.1 protein in a family with episodic ataxia / myokymia syndrome (Browne et al., 1994). Family 1 from that study was identified to have the amino acid change; it was present in all individuals

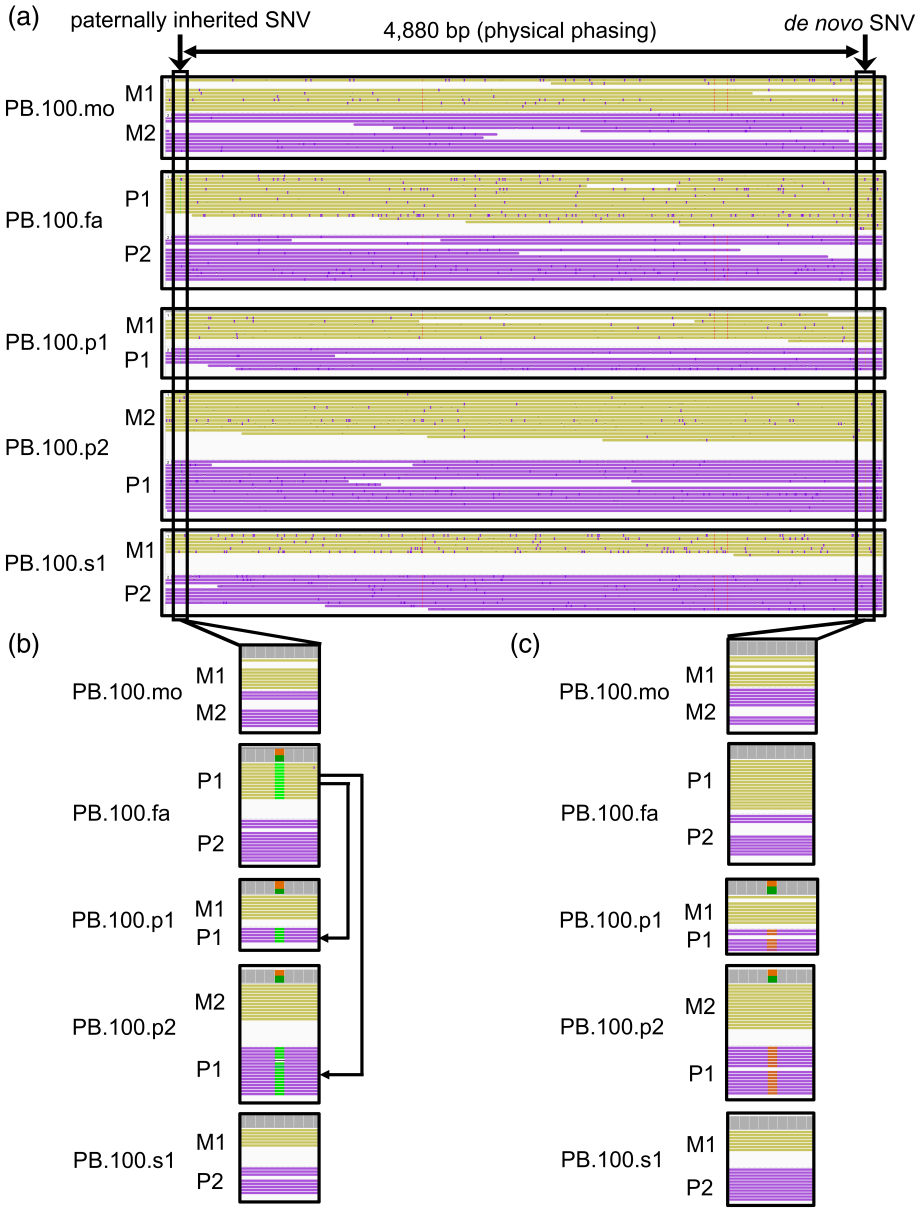

FIGURE 2 Physical phasing reveals the de novo variant arose on the paternal chromosome. (a) Shown is an extended window of the phased chromosomes. There is a phase-informative variant (inheritance from paternal chromosome 1 [P1]) 4880 bp away from the de novo KCNC2 variant and that resides on the same physical chromosome as the de novo variant in both PB.100.p1 and PB.100.p2. (b) Phase by transmission is shown where PB.100.p1 and PB.100.p2 inherit the phase-informative variant from PB.100.fa P1 chromosome. (c) The de novo variant is shown. Chromosome color is based on read-backed phasing result. M1, maternal chromosome 1; M2, maternal chromosome 2; P1, paternal chromosome 1; P2, paternal chromosome 2

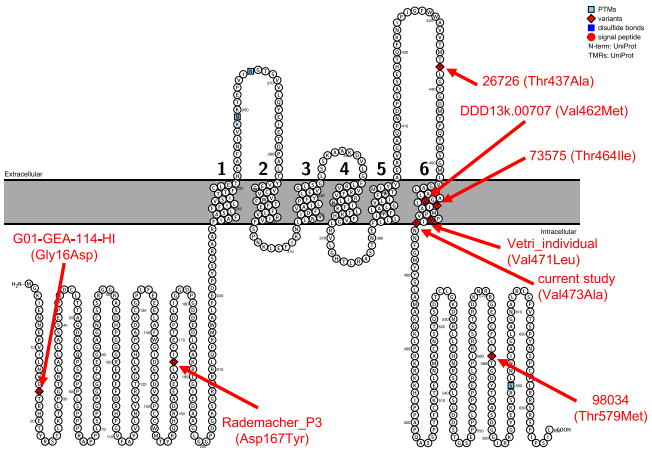

FIGURE 3 Protein plot of identified Kv3.2 variants in individuals with neurodevelopmental disorders. Shown in red is each variant with sample name from the original publications and the amino acid change. The two individuals (PB.100.p1 and PB.100.p2) are represented by the "current study (Val473Ala)" label. As can be seen, four of the six variants reside in the S6 transmembrane domain of the protein

with the syndrome and not in individuals without the syndrome. There does not appear to be a phenotypic overlap between the individuals in our study and those in Browne et al. (1994). This is likely due to differences in expression of the two proteins (Kv1.1 and Kv3.2). By functional modeling of both the human Kv1.1 protein and the orthologous Shaker protein in Drosophila it has been shown that *this specific amino acid* change affects the channel function by accelerating current decay, shifting the voltage dependence of activation, and preventing the channel from entering a long-lasting open state (Peters et al., 2011).

3.5 | Other KCNC2 variants in neurodevelopmental disorders

We searched through the literature and identified seven additional individuals with neurodevelopmental disorders that had a missense variant in this gene (de novo missense p-value $= 1.03 \times 10^{-5}$). Four of the variants reside within the S6 transmembrane domain (Figure 3). Some

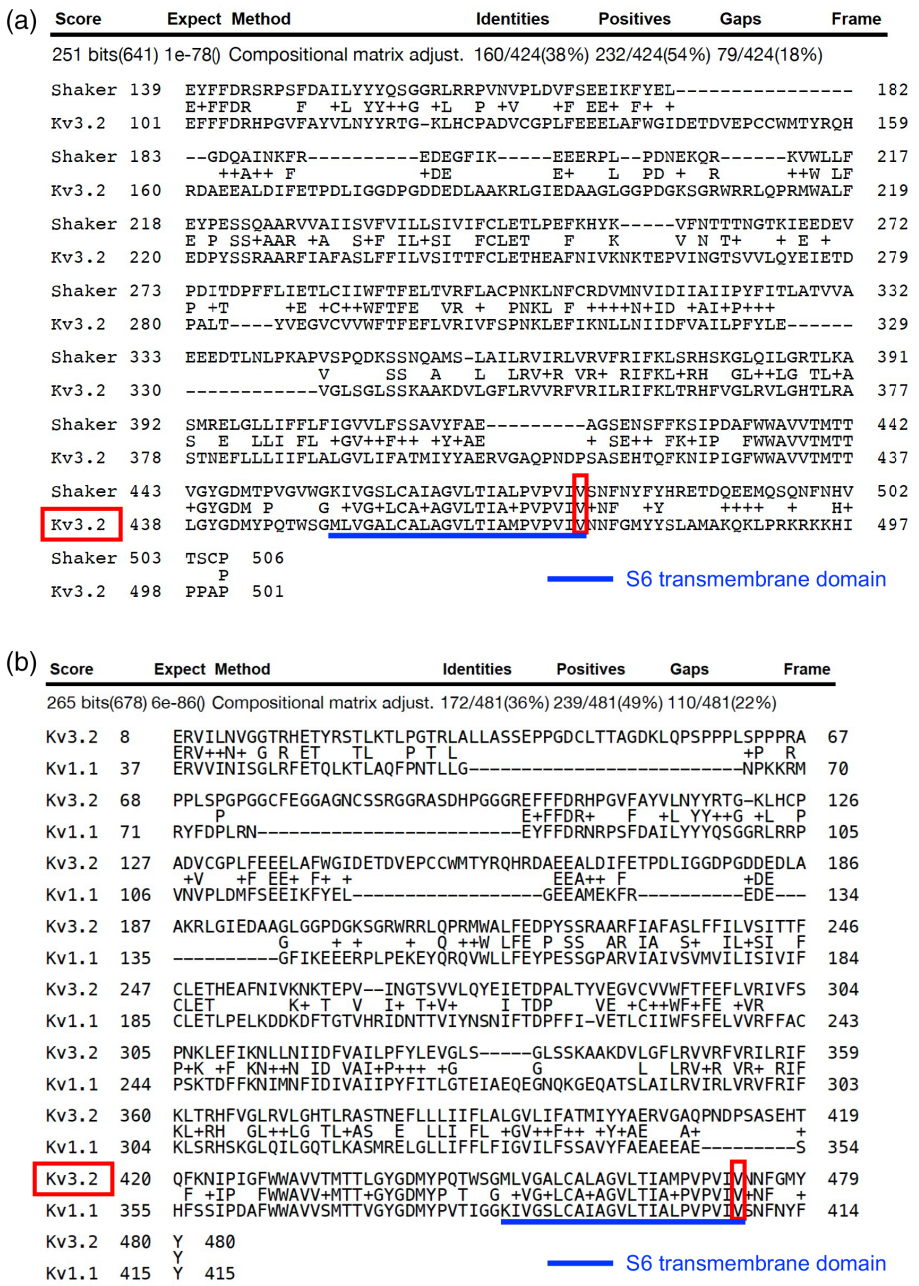

FIGURE 4 Kv3.2 conservation of the S6 transmembrane domain in the Shaker ortholog and Kv1.1 paralog. (a) BLAST of the Drosophila melanogaster Shaker ortholog to the Kv3.2 protein reveals high conservation of the S6 transmembrane domain and complete conservation at the 473 Valine amino acid position. (b) BLAST of the human Kv1.1 paralog to the Kv3.2 protein reveals high conservation of the S6 transmembrane domain and complete conservation at the 473 Valine amino acid position. These two proteins were compared since the Valine to Alanine change at the same protein location as in Kv3.2 has already revealed functional consequences of the change in both the Kv1.1 and Shaker proteins

individuals had phenotypic data available. There were some shared phenotypic features in those individuals (e.g., epilepsy) (Table 1).

3.6 | KCNC2 expression

In human tissues, the *KCNC2* gene is expressed highly in the brain and expressed in the pituitary (https://gtexportal.org/home/gene/KCNC2) (Figure 5a). It is lowly expressed in other human tissues. Over the lifetime in the brain, the *KCNC2* gene is expressed most highly after birth and is very highly expressed in the thalamus (https://www.brainspan.org/) (Figure 5b). Further delineation of its expression pattern by single-cell analysis reveals specificity for this channel in inhibitory neurons (https://portal.brain-map.org) (Figure 5c).

4 | DISCUSSION

In this study, utilization of highly accurate PacBio HiFi long-read sequencing provided genetic answers for a family whose prior clinical genetic test results were negative. By applying a family design of including an unaffected male sibling we identified a de novo missense variant, phased to the paternal chromosome, in the *KCNC2* gene in both females with autism that was not present in the unaffected brother. This variant has several effects on the function of the potassium channel. The channel is expressed in GABAergic neurons in the brain with the highest expression in the thalamus. This finding could lead to potential therapeutics regarding the seizure and neurocognitive phenotypes of the affected females.

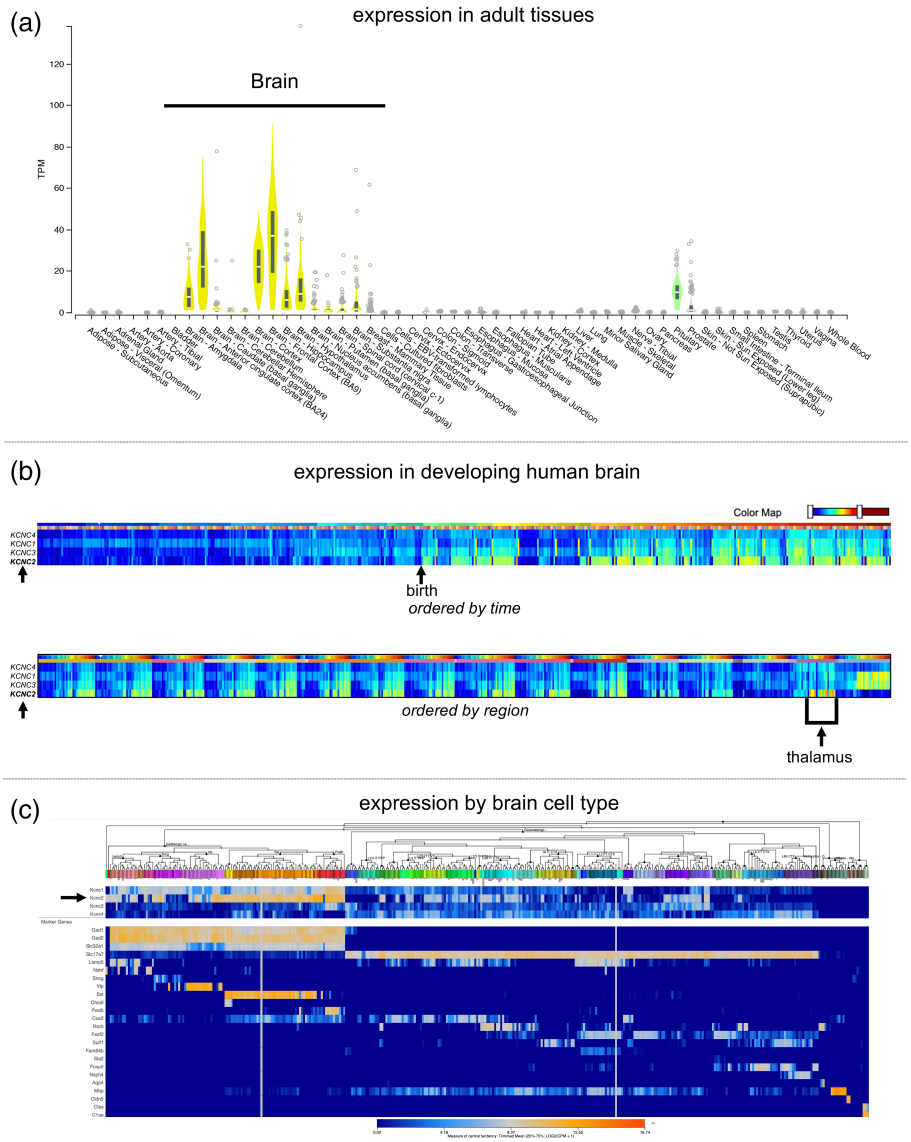

FIGURE 5 Expression of the *KCNC2* gene. (a) The *KCNC2* gene is expressed primarily in the brain from adult human tissues (data from https://gtexportal.org/home/gene/KCNC2). (b) In the developing human brain, the *KCNC2* gene is most highly expressed after birth (top) and has the highest expression in the thalamus (bottom) (data from https://www.brainspan.org/). Single-cell expression data reveals the highest expression in GABAergic neurons in the whole cortex and hippocampus from an 8-week-old mouse (data from https://portal.brain-map.org)

The main reason why this family was previously unexplained in the clinic is because the *KCNC2* gene was not currently in the list of OMIM disease genes in 2013. We reached out to the diagnostic laboratory of record in July of 2021, and the variant was still not considered in its clinical interpretation because OMIM does not consider it an established disease gene likely because only single case reports have been published. Since identified previous studies had functionally tested *precisely the same missense variant* in both a human paralog and Drosophila ortholog of this gene (Peters et al., 2011), we think that the effect is relevant for the phenotype we see in this family including both autism and epilepsy in both affected individuals. We also note that a recent preprint came out with evidence supporting the role of *KCNC2* in neurodevelopmental disorders (Schwarz et al., 2021) and so, it is possible this gene may soon be added to the OMIM disease gene list. Finally, long-read sequencing afforded the opportunity to phase the variant and show that in both females with autism the variant arose on the paternal chromosome providing evidence of germline mosaicism.

This study provided genetic answers for this family and serves as a prototypic framework by which long-read sequencing can be utilized in future clinical genomic studies. We recommend that for unexplained families that all family members, including unaffected siblings, be sequenced using long-read sequencing and that both variant detection as well as phasing be performed for each family member. Long-read sequencing enabled "precision genomics" in this case and warrants consideration as a clinical standard, especially in enigmatic cases in which genetic causes are strongly suspected, on our pathway to precision medicine. Finally, we note that careful reassessment of published literature should expand beyond the study of the gene alone. Rather, when there are gene family members, such as the case with potassium channels, those gene family members can also be examined for whether they have similar variants. This was also a lesson learned in a previous study of variants in glutamate receptors with regions of high conservation between paralogs (Geisheker et al., 2017) and should be more systematically tested in other gene families.

ACKNOWLEDGMENTS

Thank you to the family for participating in this study. Thank you to Pacific Biosciences for a SMRT grant to sequence this family. This work was also supported by grants from the National Institutes of Health (R00MH117165, P50HD103525). For Figure 5a, the Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from: https://gtexportal.org/home/gene/KCNC2 the GTEx Portal on 08/30/21. Data produced in this study are available through dbGaP study phs002698.v1.p1.

CONFLICT OF INTEREST

Jenny Ekholm, Aaron Wenger, William Rowell, Ari Grudo, and Jonas Korlach are employees of Pacific Biosciences.

AUTHOR CONTRIBUTIONS

Tychele N. Turner and John N. Constantino designed the study. Tychele N. Turner and Elvisa Mehinovic wrote the article. Tychele N. Turner, Teddi Gray, Meghan Campbell, and John N. Constantino met with the family and collected samples. John N. Constantino and Christina Gurnett assessed clinical phenotypes. Jane Grimwood performed long-read sequencing. Tychele N. Turner, Elvisa Mehinovic, Jenny Ekholm, Aaron Wenger, William Rowell, Ari Grudo, and Jonas Korlach analyzed the genomic data. All authors read and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available through dbGaP study phs002698.v1.p1. The data are not publicly available due to privacy or ethical restrictions.

ORCID

William Rowell https://orcid.org/0000-0002-7422-1194
John N. Constantino https://orcid.org/0000-0001-8888-4186
Tychele N. Turner https://orcid.org/0000-0001-8246-6477

REFERENCES

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*, *2013*, 7.20.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

An, J. Y., Lin, K., Zhu, L., Werling, D. M., Dong, S., Brand, H., Wang, H. Z., Zhao, X., Schwartz, G. B., Collins, R. L., Currall, B. B., Dastmalchi, C., Dea, J., Duhn, C., Gilson, M. C., Klei, L., Liang, L., Markenscoff-Papadimitriou, E., Pochareddy, S., … Sanders, S. J. (2018). Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. *Science (New York, N.Y.)*, *362*, 6420.

Browne, D. L., Gancher, S. T., Nutt, J. G., Brunt, E. R., Smith, E. A., Kramer, P., & Litt, M. (1994). Episodic ataxia/myokymia syndrome is associated with point mutations in the human potassium channel gene, KCNA1. *Nature Genetics*, *8*(2), 136–140.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*, 170–175.

Coe, B. P., Stessman, H. A. F., Sulovari, A., Geisheker, M. R., Bakken, T. E., Lake, A. M., Dougherty, J. D., Lein, E. S., Hormozdiari, F., Bernier, R. A., & Eichler, E. E. (2019). Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nature Genetics*, *51*(1), 106–116.

Doan, R. N., Lim, E. T., De Rubeis, S., Betancur, C., Cutler, D. J., Chiocchetti, A. G., Overman, L. M., Soucy, A., Goetze, S., Freitag, C. M., Daly, M. J., Walsh, C. A., Buxbaum, J. D., & Yu, T. W. (2019). Recessive gene disruptions in autism spectrum disorder. *Nature Genetics*, *51*(7), 1092–1098.

Garrison, E. M. G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907 [q-bio.GN].

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J., Ripke, S., & Sandin, S. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–885.

Geisheker, M. R., Heymann, G., Wang, T., Coe, B. P., Turner, T. N., Stessman, H. A. F., Hoekzema, K., Kvarnung, M., Shaw, M., Friend, K., Liebelt, J., Barnett, C., Thompson, E. M., Haan, E., Guo, H., Anderlid, B. M., Nordgren, A., Lindstrand, A., Vandeweyer, G., … Eichler, E. E. (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nature Neuroscience*, *20*(8), 1043–1051.

Iossifov, I., Levy, D., Allen, J., Ye, K., Ronemus, M., Lee, Y.-h., Yamrom, B., & Wigler, M. (2015). Low load for disruptive mutations in autism genes and their biased transmission. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(41), E5600–E5607.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., Stessman, H. A., Witherspoon, K. T., Vives, L., Patterson, K. E., Smith, J. D., Paeper, B., Nickerson, D. A., Dea, J., Dong, S., Gonzalez, L. E., Mandell, J. D., Mane, S. M., Murtha, M. T., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, *515*, 216–221.

Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., Beckmann, J. S., Rosenfeld, J. A., & Eichler, E. E. (2014). A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *American Journal of Human Genetics*, *94*(3), 415–425.

Kaplanis, J., Samocha, K. E., Wiel, L., Zhang, Z., Arvai, K. J., Eberhardt, R. Y., Gallone, G., Lelieveld, S. H., Martin, H. C., McRae, J. F., Short, P. J., Torene, R. I., de Boer, E., Danecek, P., Gardner, E. J., Huang, N., Lord, J., Martincorena, I., Pfundt, R., … Retterer, K. (2020). Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*, *586*(7831), 757–762.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., … MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006.

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315.

Krumm, N., Turner, T. N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B. P., Stessman, H. A., He, Z. X., Leal, S. M., Bernier, R., & Eichler, E. E. (2015). Excess of rare, inherited truncating mutations in autism. *Nature Genetics*, *47*(6), 582–588.

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(7), 1073–1081.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y. H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K., Buja, A., Krieger, A., Yoon, S., Troge, J., Rodgers, L., Iossifov, I., & Wigler, M. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*, *70*(5), 886–897.

Martin, M., Patterson, M., Garg, S., OFischer, S., Pisanti, N., Klau, G. W., Schöenhuth, A., & Marschall, T. (2016). WhatsHap: Fast and accurate read-based phasing. bioRxiv:085050.

Neale, B. M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K. E., Sabo, A., Lin, C. F., Stevens, C., Wang, L. S., Makarov, V., Polak, P., Yoon, S., Maguire, J., Crawford, E. L., Campbell, N. G., Geller, E. T., Valladares, O., Schafer, C., Liu, H., … Daly, M. J. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*, *485*(7397), 242–245.

Nurk, S., Walenz, B. P., Rhie, A., Vollger, M. R., Logsdon, G. A., Grothe, R., Miga, K. H., Eichler, E. E., Phillippy, A. M., & Koren, S. (2020). HiCanu: Accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Research*, *30*(9), 1291–1305.

Omasits, U., Ahrens, C. H., Müller, S., & Wollscheid, B. (2013). Protter: Interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics (Oxford, England)*, *30*(6), 884–886.

Padhi, E. M., Hayeck, T. J., Cheng, Z., Chatterjee, S., Mannion, B. J., Byrska-Bishop, M., Willems, M., Pinson, L., Redon, S., Benech, C., Uguen, K., Audebert-Bellanger, S., Le Marechal, C., Férec, C., Efthymiou, S., Rahman, F., Maqbool, S., Maroofian, R., Houlden, H., … Turner, T. N. (2021). Coding and noncoding variants in EBF3 are involved in HADDS and simplex autism. *Human Genomics*, *15*(1), 44.

Peters, C. J., Werry, D., Gill, H. S., Accili, E. A., & Fedida, D. (2011). Mechanism of accelerated current decay caused by an episodic ataxia type-1-associated mutant in a potassium channel pore. *Journal of Neuroscience*, *31*(48), 17449–17459.

Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R., & Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Research*, *20*(1), 110–121.

Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., Newburger, D., Dijamco, J., Nguyen, N., Afshar, P. T., Gross, S. S., Dorfman, L., McLean, C. Y., & DePristo, M. A. (2018). A universal SNP and small-indel variant caller using deep neural networks. *Nature Biotechnology*, *36*(10), 983–987.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Rademacher, A., Schwarz, N., Seiffert, S., Pendziwiat, M., Rohr, A., van Baalen, A., Helbig, I., Weber, Y., & Muhle, H. (2020). Whole-exome sequencing in NF1-related west syndrome leads to the identification of KCNC2 as a novel candidate gene for epilepsy. *Neuropediatrics*, *51*(5), 368–372.

Robinson, P. N., Köhler, S., Oellrich, A., Wang, K., Mungall, C. J., Lewis, S. E., Washington, N., Bauer, S., Seelow, D., Krawitz, P., Gilissen, C., Haendel, M., & Smedley, D. (2014). Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Research*, *24*(2), 340–348.

Samocha, K. E., Robinson, E. B., Sanders, S. J., Stevens, C., Sabo, A., McGrath, L. M., Kosmicki, J. A., Rehnstrom, K., Mallick, S., Kirby, A., Wall, D. P., MacArthur, D. G., Gabriel, S. B., DePristo, M., Purcell, S. M., Palotie, A., Boerwinkle, E., Buxbaum, J. D., Cook, E. H., Jr., … Daly, M. J. (2014). A framework for the interpretation of de novo mutation in human disease. *Nature Genetics*, *46*(9), 944–950.

Sams, E. I., Ng, J. K., Tate, V., Claire Hou, Y. C., Cao, Y., Antonacci-Fulton, L., Belhassan, K., Neidich, J., Mitra, R. D., Cole, F. S., Dickson, P., Milbrandt, J., & Turner, T. N. (2022). From karyotypes to precision genomics in 9p deletion and duplication syndromes. *HGG Advances*, *3*(1), 100081.

Sanders, S. J., He, X., Willsey, A. J., Ercan-Sencicek, A. G., Samocha, K. E., Cicek, A. E., Murtha, M. T., Bal, V. H., Bishop, S. L., Dong, S., Goldberg, A. P., Jinlu, C., Keaney, J. F., 3rd, Klei, L., Mandell, J. D., Moreno-De-Luca, D., Poultney, C. S., Robinson, E. B., Smith, L., … State, M. W. (2015). Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron*, *87*(6), 1215–1233.

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., … Buxbaum, J. D. (2020). Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*, *180*(3), 568–584.

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, *11*(4), 361–362.

Schwarz, N., Seiffert, S., Pendziwiat, M., Rademacher, A., Brünger, T., Hedrich, U. B. S., Augustijn, P. B., Baier, H., Bayat, A., Bisulli, F., Buono, R. J., Bruria, B. Z., Doyle, M. G., Guerrini, R., Heimer, G., Iacomino, M., Kearney, H., Klein, K. M., Kousiappa, I., … Weber, Y. (2021). Heterozygous variants in KCNC2 cause a broad spectrum of epilepsy phenotypes associated with characteristic functional alterations. medRxiv:2021.2005.2021.21257099.

Turner, T. N., Coe, B. P., Dickel, D. E., Hoekzema, K., Nelson, B. J., Zody, M. C., Kronenberg, Z. N., Hormozdiari, F., Raja, A., Pennacchio, L. A., Darnell, R. B., & Eichler, E. E. (2017). Genomic patterns of De novo mutation in simplex autism. *Cell*, *171*(3), 710–722.

Turner, T. N., Hormozdiari, F., Duyzend, M. H., McClymont, S. A., Hook, P. W., Iossifov, I., Raja, A., Baker, C., Hoekzema, K., Stessman, H. A., Zody, M. C., Nelson, B. J., Huddleston, J., Sandstrom, R., Smith, J. D., Hanna, D., Swanson, J. M., Faustman, E. M., Bamshad, M. J., … Eichler, E. E. (2016). Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory DNA. *American Journal of Human Genetics*, *98*(1), 58–74.

Turner, T. N., Sharma, K., Oh, E. C., Liu, Y. P., Collins, R. L., Sosa, M. X., Auer, D. R., Brand, H., Sanders, S. J., Moreno-De-Luca, D., Pihur, V., Plona, T., Pike, K., Soppet, D. R., Smith, M. W., Cheung, S. W., Martin, C. L., State, M. W., Talkowski, M. E., … Chakravarti, A. (2015). Loss of δ-catenin function in severe autism. *Nature*, *520*(7545), 51–56.

Turner, T. N., Wilfert, A. B., Bakken, T. E., Bernier, R. A., Pepper, M. R., Zhang, Z., Torene, R. I., Retterer, K., & Eichler, E. E. (2019). Sex-based analysis of de novo variants in neurodevelopmental disorders. *American Journal of Human Genetics*, *105*(6), 1274–1285.

Vetri, L., Calì, F., Vinci, M., Amato, C., Roccella, M., Granata, T., Freri, E., Solazzi, R., Romano, V., & Elia, M. (2020). A de novo heterozygous mutation in KCNC2 gene implicated in severe developmental and epileptic encephalopathy. *European Journal of Medical Genetics*, *63*(4), 103848.

Ware, J. S., Samocha, K. E., Homsy, J., & Daly, M. J. (2015). Interpreting de novo variation in human disease using denovolyzeR. *Current Protocols in Human Genetics*, *87*, 7.25.1–7.25.15.

Weiner, D. J., Wigdor, E. M., Ripke, S., Walters, R. K., Kosmicki, J. A., Grove, J., Samocha, K. E., Goldstein, J. I., Okbay, A., Bybjerg-Grauholm, J., Werge, T., Hougaard, D. M., Taylor, J., Skuse, D., Devlin, B., Anney, R., Sanders, S. J., Bishop, S., Mortensen, P. B., … Robinson, E. B. (2017). Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nature Genetics*, *49*(7), 978–985.

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P. C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A.,

Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C. S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., … Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, *37*(10), 1155–1162.

Wilfert, A. B., Turner, T. N., Murali, S. C., Hsieh, P., Sulovari, A., Wang, T., Coe, B. P., Guo, H., Hoekzema, K., Bakken, T. E., Winterkorn, L. H., Evani, U. S., Byrska-Bishop, M., Earl, R. K., Bernier, R. A., Zody, M. C., & Eichler, E. E. (2021). Recent ultra-rare inherited variants implicate new autism candidate risk genes. *Nature Genetics*, *53*, 1125–1134.

Yun, T., Li, H., Chang, P.-C., Lin, M. F., Carroll, A., & McLean, C. Y. (2020) Accurate, scalable cohort variant calls using DeepVariant and GLnexus. bioRxiv:2020.2002.2010.942086.

Zhou, J., Park, C. Y., Theesfeld, C. L., Wong, A. K., Yuan, Y., Scheckel, C., Fak, J. J., Funk, J., Yao, K., Tajima, Y., Packer, A., Darnell, R. B., & Troyanskaya, O. G. (2019). Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. *Nature Genetics*, *51*(6), 973–980.

How to cite this article: Mehinovic, E., Gray, T., Campbell, M., Ekholm, J., Wenger, A., Rowell, W., Grudo, A., Grimwood, J., Korlach, J., Gurnett, C., Constantino, J. N., & Turner, T. N. (2022). Germline mosaicism of a missense variant in *KCNC2* in a multiplex family with autism and epilepsy characterized by long-read sequencing. *American Journal of Medical Genetics Part A*, *188*A:2071–2081. https://doi.org/10.1002/ajmg.a.62743