Accepted: 16 January 2018

DOI: 10.1111/epi.14021

FULL-LENGTH ORIGINAL RESEARCH

Epilepsia®

Mosaicism of de novo pathogenic *SCN1A* variants in epilepsy is a frequent phenomenon that correlates with variable phenotypes

Iris M. de Lange[1] | Marco J. Koudijs[1] | Ruben van 't Slot[1] | Boudewijn Gunning[2] | Anja C. M. Sonsma[1] | Lisette J. J. M. van Gemert[3] | Flip Mulder[1] | Ellen C. Carbo[1] | Marjan J. A. van Kempen[1] | Nienke E. Verbeek[1] | Isaac J. Nijman[1] | Robert F. Ernst[1] | Sanne M. C. Savelberg[1] | Nine V. A. M. Knoers[1] | Eva H. Brilstra[1] | Bobby P. C. Koeleman[1]

[1]Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, the Netherlands

[2]The Epilepsy Institutes of the Netherlands Foundation (SEIN)

[3]Epilepsy Center Kempenhaeghe, Heeze, the Netherlands

Correspondence
Iris M. de Lange, Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, the Netherlands.
Email: I.M.deLange-2@umcutrecht.nl

Funding information
Stichting Vrienden WKZ, Grant/Award Number: 1614054; Dutch Epilepsy Foundation, Grant/Award Number: 2017-01

Summary

Objective: Phenotypes caused by de novo *SCN1A* pathogenic variants are very variable, ranging from severely affected patients with Dravet syndrome to much milder genetic epilepsy febrile seizures plus cases. The most important determinant of disease severity is the type of variant, with variants that cause a complete loss of function of the SCN1A protein (α-subunit of the neuronal sodium channel Nav1.1) being detected almost exclusively in Dravet syndrome patients. However, even within Dravet syndrome disease severity ranges greatly, and consequently other disease modifiers must exist. A better prediction of disease severity is very much needed in daily practice to improve counseling, stressing the importance of identifying modifying factors in this patient group. We evaluated 128 participants with de novo, pathogenic *SCN1A* variants to investigate whether mosaicism, caused by postzygotic mutation, is a major modifier in *SCN1A*-related epilepsy.

Methods: Mosaicism was investigated by reanalysis of the pathogenic *SCN1A* variants using single molecule molecular inversion probes and next generation sequencing with high coverage. Allelic ratios of pathogenic variants were used to determine whether mosaicism was likely. Selected mosaic variants were confirmed by droplet digital polymerase chain reaction and sequencing of different tissues. Developmental outcome was classified based on available data on intelligence quotient and school functioning/education.

Results: Mosaicism was present for 7.5% of de novo pathogenic *SCN1A* variants in symptomatic patients. Mosaic participants were less severely affected than non-mosaic participants if only participants with truncating variants are considered (distribution of developmental outcome scores, Mann-Whitney U, $P = .023$).

Significance: Postzygotic mutation is a common phenomenon in *SCN1A*-related epilepsies. Participants with mosaicism have on average milder phenotypes, suggesting that mosaicism can be a major modifier of *SCN1A*-related diseases.

E.H.B. and B.P.C.K. contributed equally to this article.

Detection of mosaicism has important implications for genetic counseling and can be achieved by deep sequencing of unique reads.

KEYWORDS

Dravet syndrome, epilepsy, mosaicism, postzygotic mutation, SCN1A

1 | INTRODUCTION

De novo pathogenic *SCN1A* variants are found in the majority of Dravet syndrome patients.[1–4] Dravet syndrome is characterized by onset in the first year of life, with generalized or unilateral clonic seizures triggered by fever, illness, or vaccination as first symptoms. Other seizure types often develop at a later stage, and prolonged status epilepticus can occur. Psychomotor development slows, usually in the second year of life, resulting in mild to severe intellectual disability (ID) in most patients. In addition, many patients experience walking difficulties and behavioral problems.[5–7] Pathogenic variants in *SCN1A* are also found in patients with the much milder genetic epilepsy febrile seizures plus (GEFS+) syndrome or febrile seizures only.[4]

SCN1A encodes for the α-subunit of a neuronal sodium channel, Nav1.1 (Figure 1). The main disease mechanism of *SCN1A*-related phenotypes is haploinsufficiency, caused by complete or partial loss of function. Truncating variants are expected to lead to a complete absence of expression of the mutant allele and thus to complete haploinsufficiency of Nav1.1. These variants are virtually always associated with severe phenotypes[8–10]; 97.7% of genomic rearrangements and splice site, nonsense, and frameshift *SCN1A* variants are mainly associated with Dravet syndrome.[10] The effect of pathogenic missense variants is more difficult to predict. Functional studies have shown varying degrees of loss of function through a lack of sodium current when pathogenic variants are located in critical regions of the gene (voltage sensor and/or ion-pore regions).[8–12] Most pathogenic missense variants associated with Dravet syndrome lead to a complete loss of sodium current, whereas missense pathogenic variants associated with milder phenotypes, such as GEFS+, lead to milder disturbances of channel function.[10] The location of missense variants is a strong indicator for the severity of channel disruption. However, it still cannot predict the effect of the variant on channel function and disease severity fully.[8–10,12] The same is true for physicochemical property changes due to missense variants.[8,11]

Although type and location of pathogenic *SCN1A* variants are a major determinant of disease severity, a significant disease variability remains unexplained; a wide phenotypic variability between patients with Dravet syndrome exists,[13–15] and variable phenotypes have even been associated with the same pathogenic variant.[2,12,16] These different *SCN1A*-related phenotypes may be indistinguishable at their first presentation, often leaving parents in great uncertainty about the future of their children when a pathogenic *SCN1A* variant is found early in life. A better prediction of disease severity is very much needed in daily practice to improve counseling, stressing the importance of identifying modifying factors in this patient group. Several modifying factors have already been suggested, such as variants in the *SCN1A* promoter region and in 5′- and 3′-untranslated regions, and variants in other genes.[13,17–20] Moreover, parental mosaicism for the pathogenic *SCN1A* variant has been well recognized in cases were mosaic parents of Dravet children show a mild epilepsy phenotype. This observation suggests that postzygotic pathogenic variants may be present in a significant percentage of carriers of de novo pathogenic *SCN1A* variants and that the percentage of mosaicism can affect the severity of the disease.[21–24] However, these are results from single case reports or studies aimed at the detection of low-grade parental mosaicism. A study to investigate the occurrence and effect of high-grade mosaicism in patients with de novo *SCN1A* pathogenic variants themselves has never been reported to our knowledge. Because mosaicism has been known to influence expression in other diseases,[25] we here investigate whether mosaicism is a common phenomenon in de novo pathogenic *SCN1A* variants, and whether it has a substantial effect on the severity of the disease. If both are true, then mosaicism might be a major modifier for *SCN1A*-related phenotypes, which can help predict the disease course. We here describe testing for mosaicism in 128 participants with de novo pathogenic *SCN1A* variants, using single molecule

Key Points

- Mosaicism is present in 7.5% of symptomatic patients with de novo pathogenic *SCN1A* variants
- Patients with mosaicism of truncating variants have on average milder phenotypes than patients with heterozygous truncating variants, which makes mosaicism an important modifier in *SCN1A*-related phenotypes
- Detection of mosaicism has important implications for genetic counseling and can be achieved by deep sequencing of unique reads

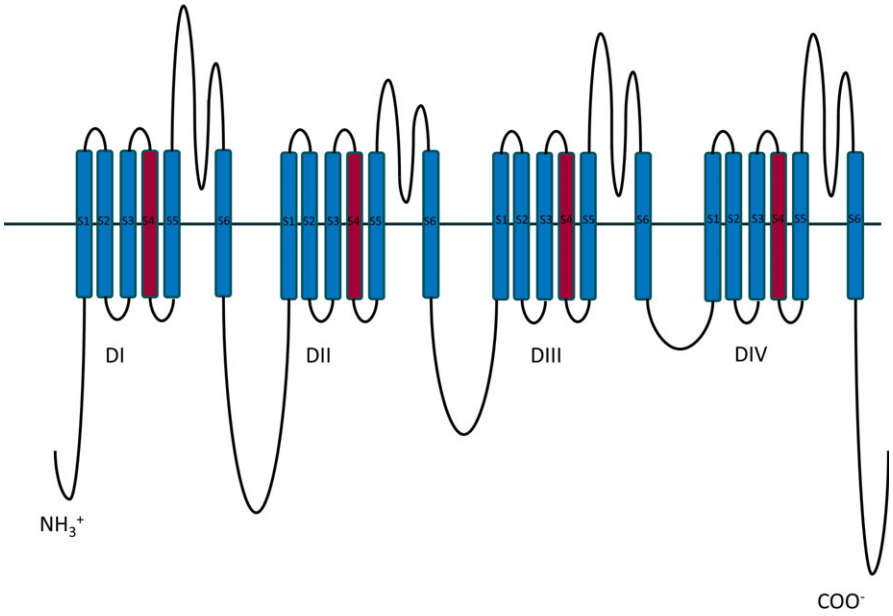

FIGURE 1 Schematic overview of the SCN1A protein (alpha unit of the neuronal voltage-gated sodium channel Nav1.1). SCN1A consists of 4 domains (DI-DIV), connected by intracellular loops. Each domain consists of 6 transmembrane segments (S1-S6). The S5 and S6 segments of all domains make up the pore of the channel, and the connecting loop between S5 and S6 is the pore loop. The S4 segment is the positively charged voltage sensor of the protein

molecular inversion probes (smMIPs) and next generation sequencing (NGS), to determine the incidence and clinical effects of mosaicism.

2 | MATERIALS AND METHODS

2.1 | Participants and clinical data

2.1.1 | Participants

A cohort of 128 participants with de novo *SCN1A* pathogenic variants was evaluated. Only participants with (likely) pathogenic *SCN1A* variants (class 4 and 5) were included (see Data S1 for more details). Multiplex ligation-dependent probe amplification (MLPA) was performed in all patients to exclude deletions or duplications. Because mosaicism is not expected in inherited pathogenic variants, we included only participants with de novo pathogenic variants (confirmed de novo by Sanger sequencing or NGS, n = 107; presumed de novo [negative family history] in the absence of parental DNA, n = 21). Because we aim to improve counseling in patients with *SCN1A* pathogenic variants detected in clinical care, all individuals with pathogenic variants detectable by standard diagnostic procedures were considered for inclusion. One hundred twenty-four participants exhibited epilepsy or febrile seizures (Dravet syndrome, n = 106, according to criteria published previously[26]). The other 4 participants were fathers of children with Dravet syndrome (n = 3) and the oldest member of a GEFS+ family (n = 1), who carried the same causal *SCN1A* pathogenic variant as their children but did not have any seizures themselves. In 2 of these 4 asymptomatic participants and in 2 others, mosaicism was already suspected based on their diagnostic results (Sanger sequencing or MLPA), 1 of whom has been described earlier.[27] Diagnostic testing in the 4 asymptomatic participants was only performed after their children were diagnosed, so one could argue that they should not have been included. Our main analyses are therefore performed without these participants, to achieve results applicable to symptomatic patients, which is clinically most meaningful. We however repeated the analyses including them, because asymptomatic carriers of pathogenic *SCN1A* variants have a higher a priori probability of being mosaic. By including them in our analyses, we achieve a more complete understanding of the effects of mosaicism as a modifier, as it gives the opportunity to investigate at what grade of mosaicism symptoms arise.

All eligible individuals known to the University Medical Center Utrecht were approached, to avoid any selection bias for milder phenotypes or known mosaic cases. Informed consent was obtained from participants or their legal caretakers according to the Declaration of Helsinki. The study was approved by the Ethical Committee of the University Medical Center Utrecht.

2.1.2 | Clinical data and statistical analyses

Detailed clinical data were collected from medical records for all participants, and a semistructured telephone interview was conducted when possible (n = 118). Furthermore, for a subset of participants, based on age, the PedsQL Measurement Model questionnaire was completed, to measure health-related quality of life for participants aged 0-25 years.[28] A classification of the developmental outcome was made, rated in a consensus meeting by a

child neurologist, neuropsychologist, and clinical geneticist who were blinded to the outcome of the mosaicism assessment. Developmental outcome was rated on a 5-point scale based on available data on intelligence quotient (IQ) and developmental level, adjusted for age at assessment (1 = no ID [IQ or developmental quotient (DQ) > 85], 2 = borderline ID [IQ or DQ = 70-85], 3 = mild ID [IQ or DQ = 50-70], 4 = moderate ID [IQ or DQ = 30-50], 5 = severe or profound ID [IQ or DQ < 30]). When no (recent) IQ or DQ was available, the assessment was made based on school functioning, communication, and adaptive behavior. Differences in outcomes between mosaic participants and nonmosaic participants were analyzed (Mann-Whitney *U* test for cognitive development and age at seizure onset, Fisher's exact test for seizure severity and age at first notice of developmental delay, independent samples *t* test for PedsQL results), for symptomatic participants as well as for the complete group of participants. We furthermore performed a second analysis for both groups in which only participants with truncating variants were assessed, to control for variation in outcomes due to the variant types themselves. This group comprises patients with frameshift and nonsense variants, large rearrangements, and splice site variants leading to frameshifts, all of which can be expected to lead to a similarly severe channel dysfunction of SCN1A. This in contrast to missense variants, for which the precise effects on channel function cannot be accurately predicted.

2.2 | Molecular analyses

2.2.1 | Mosaicism screening by smMIPs and NGS

All *SCN1A* exons were captured by smMIPs, as described earlier,[29] and sequenced (see Data S2 for more details).

The resulting data were analyzed using commercial software (SeqNext module of Sequence Pilot; JSI Medical Systems, Ettenheim, Germany; see Data S3 for more details). Reads with the same single molecule tag were assembled into one consensus read, to correct for polymerase chain reaction (PCR) and sequencing artifacts. In addition, the molecular tag discriminates unique reads from PCR duplicates, allowing the determination of quantitative sequence coverage of reads originating from unique DNA molecules. *SCN1A* pseudogene reads were removed from alignment and analysis. The earlier identified pathogenic variants were located and the percentages of mutated reads were used to determine whether mosaicism for these pathogenic variants was likely based on a binomial distribution. Only pathogenic variants with a coverage of at least 20X were initially analyzed. Two types of technical artifacts were identified, and pathogenic variants with a deviating alternative allele frequency (AAF) influenced by these were discarded as possible mosaics (see Data S4 for more details).

2.2.2 | Statistical analysis of smMIP data

We expect true heterozygous variants to follow a binomial distribution, in which variants with a higher coverage (number of observations) will have an AAF closer to 0.5. For each pathogenic variant, a *P*-values was calculated (chi-square test), to test whether the AAF deviated significantly from 0.5 (see Data S5 for more details).

2.2.3 | Confirmation by droplet digital PCR

Some of the samples with screening results compatible with mosaicism were reevaluated by droplet digital PCR (ddPCR; see Data S6).

2.2.4 | Confirmation in other tissues

DNA from buccal cells, saliva, urine, and/or brain tissue could be obtained for 5 participants who were suspected of mosaicism based on smMIP screening, and was analyzed with smMIPs and NGS or ddPCR as described above. No tissue samples were available for the other mosaic participants.

3 | RESULTS

3.1 | Statistical analysis of mosaicism screening by smMIPs

NGS results were obtained for 122 of 128 participants (Figure 2). The pathogenic variants of 6 participants could not be detected by smMIPs due to the nature of the variants (deletions or duplications spanning >1 smMIP, with no heterozygous single nucleotide polymorphisms [SNPs] present in that region to deduct possible mosaicism). Nine other participants did not meet the threshold of 20X coverage at the site of their pathogenic variant. The average unique read depth on the location of the known pathogenic variant of the 113 remaining participants was 1599X (ranging from 20X to 7320X, median = 1281X).

Fifteen pathogenic variants had *P*-values below the mosaicism cutoff. Six were discarded as possible mosaic variants due to the technical artifacts described earlier. Among the remaining 9 were 2 pathogenic variants in a repetitive region with an AAF < 0.5 that were selected for further testing by ddPCR. This confirmed true mosaicism in one and disproved it in the other, leaving 8 true mosaic pathogenic variants. All mosaic participants carried at least 1 neutral SNP with an AAF close to 0.5, excluding

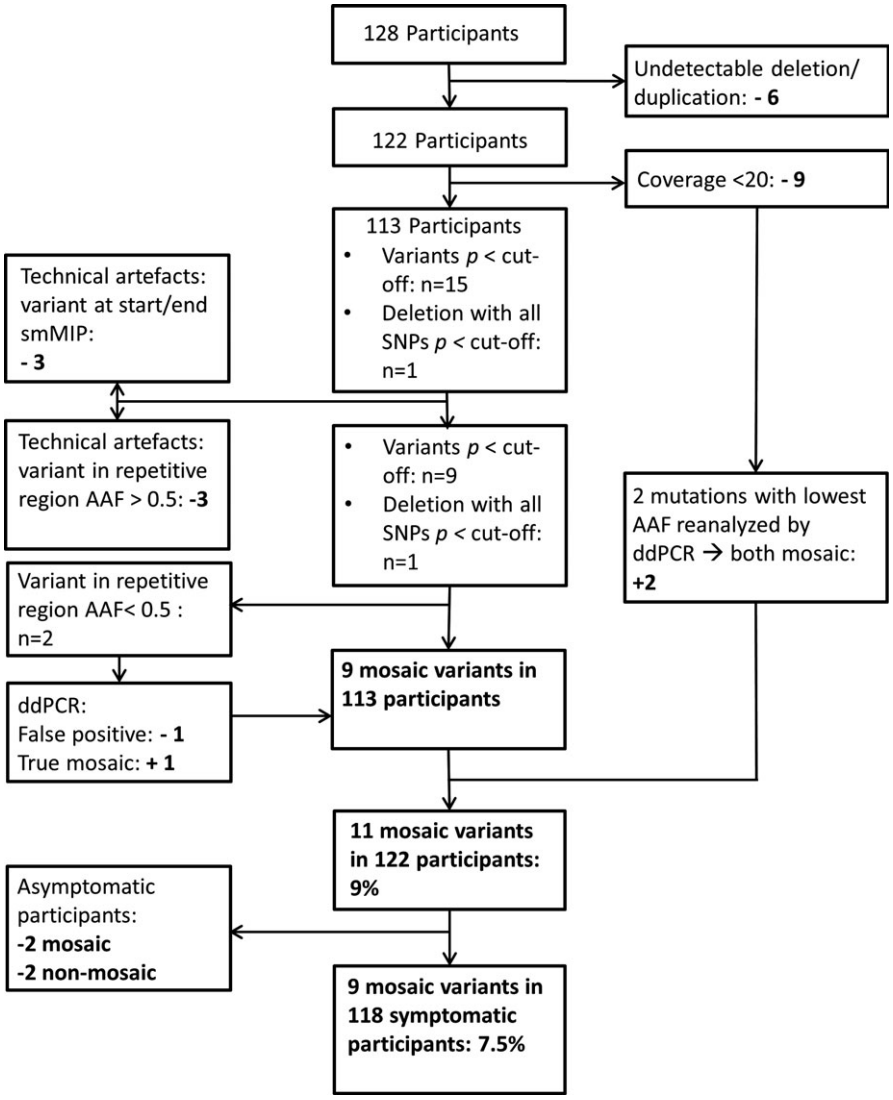

FIGURE 2 Flowchart of detected mosaic pathogenic variants. AAF, alternative allele frequency; ddPCR, droplet digital polymerase chain reaction; smMIP, single molecule molecular inversion probe; SNP, single nucleotide polymorphism

additional deletions or duplications of *SCN1A* as the cause for their deviating AAF.

The pathogenic variant in 1 of the participants for whom the SNP data were regarded as outliers was an *SCN1A* deletion of exon 2-23. Results of diagnostic genetic testing by MLPA had already suggested mosaicism for this deletion in 50% of cells. The significantly deviating AAFs of the SNPs present in this region confirmed mosaicism (AAF = 0.72-0.75, $P = 1.75 \times 10^{-34}$ to 3.63×10^{-55}, Figure 3), thereby raising the number of mosaic participants detectable by smMIPs to 9 of 113 (7.9%). AAFs of the mutated alleles ranged between 0.10 and 0.39 (Table 2).

As mentioned above, for 9 participants insufficient coverage was reached (<20X), and 14 others had a coverage <100X. With these read depths, high-grade mosaicism is virtually impossible to detect at our level of significance; at 100X an AAF of <0.25 is necessary to drop below the *P*-value threshold for significance, and at 20X even zero alternative allele reads only give a *P*-value of 7.7×10^{-6}, still above the threshold. We therefore reanalyzed 2 low-coverage pathogenic variants with the lowest AAFs (<0.25) with ddPCR.

3.2 | ddPCR

ddPCR was used to reanalyze 7 likely mosaic pathogenic variants detected by NGS, including the 2 poorly covered pathogenic variants described above. It was not feasible to design functional ddPCR essays for each likely mosaic pathogenic variant, so a selection was made based on the expected specificity of the probes. As mentioned earlier, 1 pathogenic variant in a repetitive region was disproved (Figure 2). The other 6, including the 2 low-coverage variants, were confirmed as mosaics (Table 1). This raises the final number of mosaic participants in our complete cohort to 11 of 122 (9%). For symptomatic patients only, the incidence is 9 of 118 (7.5%).

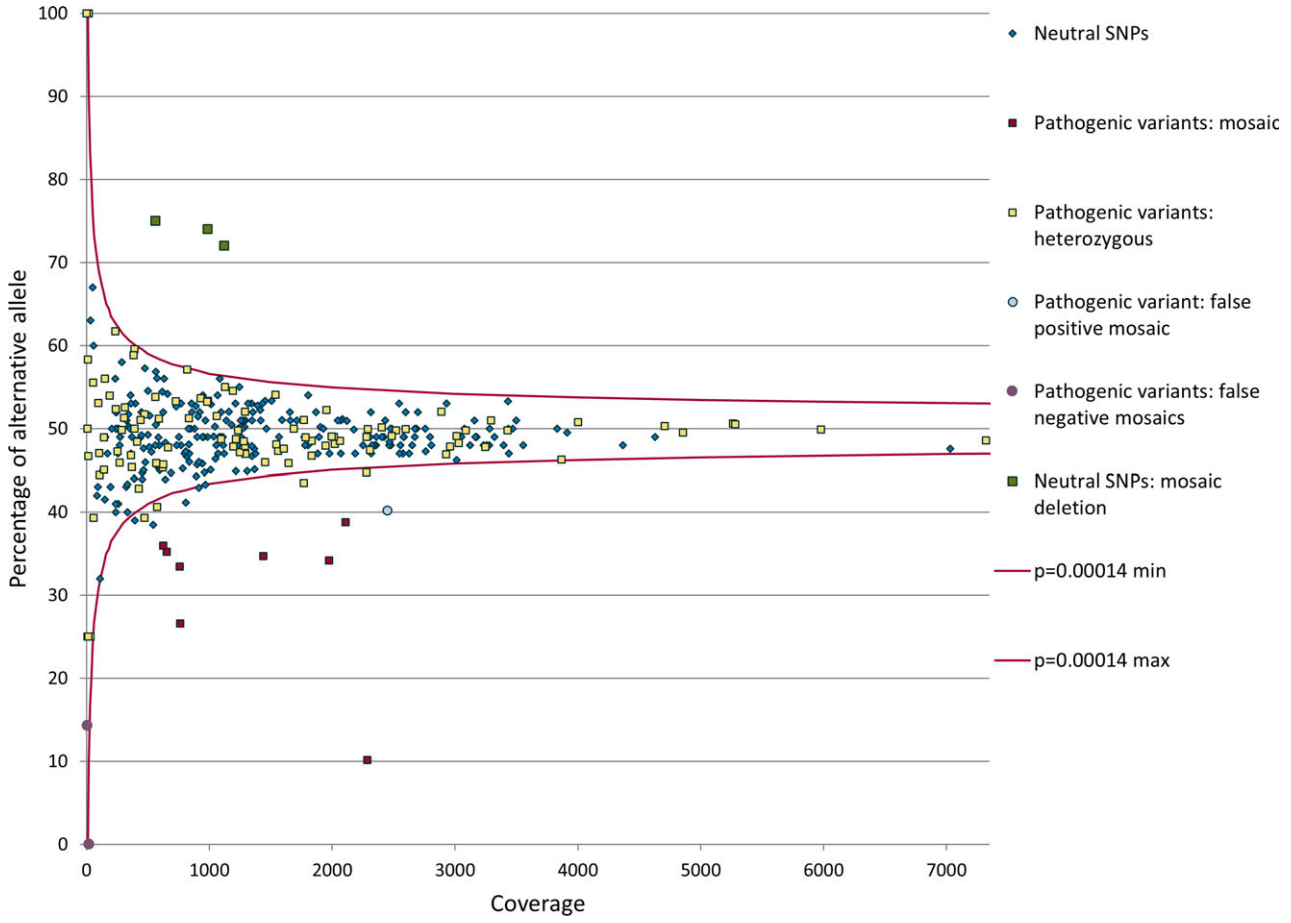

FIGURE 3 Overview of neutral single nucleotide polymorphisms (SNPs) and pathogenic variants. For each analyzed neutral SNP or pathogenic variant the percentage of alternative allele and the achieved unique coverage at its location is depicted.

3.3 | Sequencing of other tissues

DNA from buccal cells, saliva, urine, and/or brain tissue was analyzed with smMIPs in 5 mosaic participants. In all 5, mosaicism could be confirmed, in most with similar AAFs as in blood (Table 1).

3.4 | Read-frame restoring mosaic pathogenic variant

In 1 participant (Participant 6), not only her known frameshift pathogenic variant (insG) was seen at the variant site, but also a missense variant 3 base pairs upstream, that had not been identified in diagnostic testing. The known frameshift pathogenic variant was present in ~32% of the reads, and the second variant in ~13% of the reads (Figure S1A), together accounting for ~50%. No reads were present with both variants, proving independent haplotypes. We hypothesize that at first a heterozygous insG pathogenic variant was present, and that in a subpopulation of cells this allele acquired another variant, a delT 3 base pairs upstream. This leads to restoring of the original reading frame along with a T>G missense variant, resulting in mosaicism for both the frameshift and the new missense variant in different cell populations (Figure S1B). The missense variant is estimated to be less severe than the frameshift pathogenic variant. ddPCR confirmed that both variants were present in <50% of alleles, with more insG than delT alleles, although no exact percentages can be given because the Poisson statistics implemented in the software are not appropriate to calculate this in a 3-allelic situation.

3.5 | Phenotypic features and statistical analysis of clinical outcomes

Clinical features and molecular details of the 11 mosaic participants are summarized in Table 1. Two participants were clinically unaffected. Overall, 54 participants carried truncating variants, of whom 10 belonged to the mosaic participants. No statistically significant differences were seen in disease severity in comparing all symptomatic participants (Table 2). However, the mosaic group contained a much higher percentage of truncating variants (89% vs only 40% in the nonmosaic group). Therefore, we also compared

TABLE 1 Clinical features of mosaic participants

	Participant										
	Symptomatic patients									Asymptomatic patients	
Feature	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Age, y	10	60	10	6.5	25	15	44	39	7	41	43
Gender	Male	Male	Female	Female	Female	Female	Male	Male	Male	Male	Male
Pathogenic variant	c.622_657del insT, p.Asp208fs	c.C3430-3G, splicing	c.5193delA, p.Ile1733fs	c.C5348T, p.Ala1783Val	Deletion exon 2-23	c.982insG, p.Glu328fs and c.T980G, p.Leu327Arg	c.1537delG, p.Glu513fs	c.992delT, p.Leu331*	c.4262_4275del14, p.Gly1421fs	c.G3880-1A, splicing	c.G603-1T, splicing
Truncating variant	Yes	Yes	Yes	No	Yes	Yes and no	Yes	Yes	Yes	Yes	Yes
smMIPs, % of alternative allele in blood (total read depth at variant location)	35% (653X)	35% (1440X)	33% (760X)	36% (626X)	n.a.: ~50% of cells (MLPA), 72%-75% (SNPs) (559X-1119X)	32% and 13% (1968X)	27% (764X)	39% (2210X)	14% (7X)	0% (20X)	10% (2285X)
ddPCR, % of mosaicism in blood			38%			Both mosaic	27%	38%	28%	14%	
% of alternative allele in other tissues			28% in buccal cells (smMIPs: 354X)	38% in buccal cells (smMIPs: 32X)		34% and 12% in buccal cells (smMIPs: 176X)	8% in saliva (26X) and 0% in urine (10X; smMIPs)		16% in brain tissue (ddPCR)		
Syndrome diagnosis	Dravet	GEFS+	Dravet	Dravet	Dravet	Dravet	Dravet	Dravet	Dravet	Unaffected	Unaffected
ID	+	–	+	+	+	+	+	+	+	–	–
Developmental outcome[a]	3	1	3	5	5	3	2	5	2	1	1
Age at first notice of developmental delay, mo	30	n.a.	38	18	12	36	72	36	60	n.a.	n.a.
Seizure severity[b]	3	0	1	4	2	Major sz: 1; minor sz: 4	0	Major sz: 3; minor sz: 4	Major sz: 2; minor sz: 4	0	0
Status epilepticus, number of experienced episodes	1	0	1	2	12	1	0	2	2	0	0
Seizure types[c]	GTCS, hemiconvulsions, absences, tonic, myoclonias	GTCS	GTCS, atonic, CP, absences	GTCS, CP, focal, hemiconvulsions, absences, tonic, myoclonias	GTCS, focal, CP, tonic	GTCS, hemiconvulsions, focal, CP, absences, myoclonias, tonic	GTCS, absences (possibly other types)	GTCS, CP, absences, myoclonias	GTCS, CP, hemiconvulsions, atonic, absences, myoclonias, tonic	n.a.	n.a.
Current treatment	VPA, TPM	PB	VPA, LEV	TPM, LEV, STP	LEV, OXC, TPM	TPM, VPA	–			n.a.	n.a.

(Continues)

TABLE 1 (Continued)

	Participant										
	Symptomatic patients									Asymptomatic patients	
Feature	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Age at seizure onset, mo	5	24	12	3	5	6	8	3	11	n.a.	n.a.
Behavior difficulties	+	–	+ (ASD)	+	+ (ASD)	+ (aggressive, impulsive)	+	+ (ADHD, autism, aggressive, often angry)	+ (ADHD)	–	–
Walking difficulties	Ataxia, endorotation feet, pes plano valgus (surgery planned)	–	Pes plano valgus, orthopedic shoes, hypotonia, cannot run well	Ataxia, wears walking frame outside	Ataxia, mostly wheelchair dependent	Genu valgum, arch support necessary but can play sports	–	Ataxia, orthopedic shoes	–	–	–
Comorbidities					Acute metabolic encephalopathy (Reye syndrome) at age 7 years caused regression			SUDEP at age 39 years	SCAD, SUDEP at age 7 years		

[a]1 = no learning disability, 2 = borderline learning disability, 3 = mild learning disability, 4 = moderate learning disability, 5 = severe or profound learning disability.

[b]Currently. 0 = seizure-free, 1 = yearly seizures, 2 = monthly seizures, 3 = weekly seizures, 4 = daily seizures. Minor sz = short absences, short focal seizures, or myoclonias. Major sz = all other seizure types with loss of consciousness, or prolonged seizures (specified only if big and small seizures do not have the same score).

[c]Currently and/or in the past.

ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; CP, complex partial; ddPCR, droplet digital polymerase chain reaction; fs, frameshift; GEFS+, genetic epilepsy febrile seizures plus; GTCS, generalized tonic–clonic seizures; ID, intellectual disability; LEV, levetiracetam; MLPA, multiplex ligation-dependent probe amplification; n.a., not applicable; OXC, oxcarbazepine; PB, phenobarbital; SCAD, short-chain acyl-CoA dehydrogenase deficiency; smMIP, single molecule molecular inversion probe; SNP, single nucleotide polymorphism; STP, stiripentol; SUDEP, sudden unexpected death in epilepsy; TPM, topiramate; VPA, valproic.

TABLE 2 Comparison of outcomes between mosaic and nonmosaic participants in clinically affected patients

	All symptomatic patients			Symptomatic patients with truncating variants		
Group	**Mosaic patients**	**Nonmosaic patients**	**Statistics[a]**	**Mosaic patients**	**Nonmosaic patients**	**Statistics[a]**
Number of participants	9	109	—	8	44	—
Mean age, y	25	19	—	27	15	—
Developmental outcome[b]						
1, n	1	17	P = .393 (Mann-Whitney U, U = 409, z = −0.854)	1	1	**P = .023** (Mann-Whitney U, U = 87, z = −2.408)
2, n	2	7		2	2	
3, n	3	16		3	5	
4, n	0	27		0	14	
5, n	3	42		2	22	
Median score	3	4		3	4.5	
Age at first notice of developmental delay, mo[c]						
Young, <24 months	2 (25%)	52 (58%)	P = .170 (Fisher's exact text)	1 (14%)	26 (60%)	**P = .022** (Fisher's exact test)
Older, ≥24 months	6 (75%) (no delay, n = 1)	38 (42%) (no delay, n = 16)		6 (85%) (no delay, n = 1)	17 (40%) (no delay, n = 1)	
Average age at seizures onset, mo						
Mean (range)	8.67 (3-24)	6.98 (1-48)	P = .294 (Mann-Whitney U = 571, z = 1.050)	9.25 (3-24)	5.20 (1-11)	P = .054 (Mann-Whitney U = 251, z = 1.941)
Median	6	5		7	5	
Seizure severity[d]						
Major seizures often (2-4)	5 (56%)	77 (71%)	P = .452 (Fisher's exact test)	4 (50%)	37 (84%)	P = .051 (Fisher's exact test)
Major seizures rarely (0-1)	4 (44%)	32 (29%)		4 (50%)	7 (16%)	
Minor seizures often (2-4)	5 (56%)	63 (58%)	P = 1.000 (Fisher's exact test)	4 (50%)	32 (73%)	P = .231 (Fisher's exact test)
Minor seizures rarely (0-1)	4 (44%)	46 (42%)		4 (50%)	12 (27%)	
Quality of life[e]						
Completed questionnaires	n = 4	n = 66	P = .260 (independent samples t test, t_{68} = 1.136)	n = 4	n = 46	P = .087 (independent samples t test, t_{32} = 1.765)
Average total score (range)	67.68 (57-86)	56.02 (13-99)		67.68 (57-86)	52.04 (26-98)	

[a]All reported tests were performed 2-tailed with an alpha-level for significance of $P < .05$. Significant P-values are in boldface.

[b]Based on available data on intelligence quotient (IQ) and developmental level, adjusted for age at assessment. 1 = no intellectual disability (ID; IQ or developmental quotient [DQ] > 85), 2 = borderline ID (IQ or DQ = 70-85), 3 = mild ID (IQ or DQ = 50-70), 4 = moderate ID (IQ or DQ = 30-50), 5 = severe or profound ID (IQ or DQ < 30). When no (recent) IQ or DQ was available, the assessment was made based on school functioning, communication, and adaptive behavior.

[c]Numbers of participants and statistical test results are given for dichotomized scores (young vs older).

[d]Currently. 0 = seizure-free, 1 = yearly seizures, 2 = monthly seizures, 3 = weekly seizures, 4 = daily seizures. Numbers of participants are given for dichotomized scores (score 0-1 = rarely, score 2-4 = often).

[e]Quality of life total score, based on results of the PedsQL Measurement Model questionnaire.

the outcomes of participants with mosaic and nonmosaic truncating variants only (Table 2), so our analysis is not biased by differences in severity based on the variant type itself. In this group, statistically significant differences regarding developmental outcome were seen in favor of mosaic patients (median score = 3 [mild ID] in mosaic participants vs median score = 4.5 [moderate to severe ID] in nonmosaic participants, P = .023 [Mann-Whitney U test]).

TABLE 3 Comparison of outcomes between mosaic and nonmosaic participants (complete cohort including unaffected participants)

	All patients			Patients with truncating variants		
Group	**Mosaic participants**	**Nonmosaic participants**	**Statistics[a]**	**Mosaic patients**	**Nonmosaic patients**	**Statistics[a]**
Number of participants	11	111	—	10	44	—
Mean age, y	28	20	—	30	15	—
Developmental outcome[b]						
1, n	3	19	$P = .126$ (Mann-Whitney U, $U = 445$. $z = -1.530$)	3	1	**$P = .002$** (Mann-Whitney U, $U = 88$, $z = -3.114$)
2, n	2	7		2	2	
3, n	3	16		3	5	
4, n	0	27		0	14	
5, n	3	42		2	22	
Median score	3	4		2.5	4.5	
Age at first notice of developmental delay, mo[c]						
Young, <24 months	2 (25%)	52 (58%)	$P = .135$ (Fisher's exact test)	1 (14%)	26 (60%)	**$P = .039$** (Fisher's exact test)
Older, ≥24 months	6 (75%) (no delay, n = 3)	38 (42%) (no delay, n = 18)		6 (86%) (no delay, n = 3)	17 (40%) (no delay, n = 1)	
Average age at seizures onset, mo						
Mean (range)	8.67 (3-24)	6.98 (1-48)	$P = .294$ (Mann-Whitney $U = 571$, $z = 1.050$)	9.25	5.20	$P = .054$ (Mann-Whitney $U = 251$, $z = 1.941$)
Median	6; (no seizures, n = 2)	5; (no seizures, n = 2)		7; (no seizures, n = 2)	5; (no seizures, n = 0)	
Seizure severity[d]						
Major seizures often (2-4)	5 (45%)	77 (69%)	$P = .174$ (Fisher's exact test)	4 (40%)	37 (84%)	**$P = .008$** (Fisher's exact test)
Major seizures rarely (0-1)	6 (55%)	34 (31%)		6 (60%)	7 (16%)	
Minor seizures often (2-4)	5 (45%)	63 (57%)	$P = .535$ (Fisher's exact test)	4 (40%)	32 (73%)	$P = .067$ (Fisher's exact test)
Minor seizures rarely (0-1)	6 (55%)	48 (43%)		6 (60%)	12 (27%)	
Quality of life[e]						
Completed questionnaires	n = 4	n = 66	$P = .260$ (independent samples t test, $t_{68} = 1.136$)	n = 4	n = 30	$P = .087$ (independent samples t test, $t_{32}1.765$)
Average total score (range)	67.68 (57-86)	56.02 (13-99)		67.68 (57-86)	52.04 (26-98)	

[a]All reported tests were performed 2-tailed with an alpha-level for significance of $P < .05$. Significant P-values are in boldface.

[b]Based on available data on intelligence quotient (IQ) and developmental level, adjusted for age at assessment. 1 = no intellectual disability (ID; IQ or developmental quotient [DQ] > 85), 2 = borderline ID (IQ or DQ = 70-85), 3 = mild ID (IQ or DQ = 50-70), 4 = moderate ID (IQ or DQ = 30-50), 5 = severe or profound ID (IQ or DQ < 30). When no (recent) IQ or DQ was available, the assessment was made based on school functioning, communication, and adaptive behavior.

[c]Numbers of participants and statistical test results are given for dichotomized scores (young vs older).

[d]Currently. 0 = seizure-free, 1 = yearly seizures, 2 = monthly seizures, 3 = weekly seizures, 4 = daily seizures. Numbers of participants are given for dichotomized scores (score 0-1 = rarely, score 2-4 = often).

[e]Quality of life total score, based on results of the PedsQL Measurement Model questionnaire.

Mosaic participants furthermore had a later onset of developmental delay (86% after 24 months vs 40% of nonmosaic participants, $P = .022$ [Fisher's exact test]). A trend for later seizure onset (median 7 months vs 5 months, $P = .054$ [Mann-Whitney U test]) and for lower seizure frequency of major seizures (50% "rarely" vs 16% "rarely," $P = .051$ [Fisher's exact test]) was observed in the mosaic participants compared to nonmosaic participants. Repeated analyses, including the 4 asymptomatic participants, showed similar results (Table 3).

4 | DISCUSSION

4.1 | Prevalence of mosaicism in *SCN1A* pathogenic variants

Although mosaicism has been detected in many other genetic diseases caused by de novo heterozygous pathogenic variants, only a few studies have tried to estimate its prevalence, which ranges from 3.3% to 30%.[30–33] However, these studies all had different designs and inclusion criteria, which makes it difficult to compare results. Two studies are most similar to ours. The first[33] found mosaicism in 6.5% of de novo pathogenic variants in severe ID patients. smMIPs were used in this study as well, but only to confirm possible mosaic pathogenic variants with a deviating AAF, which were detected by other methods. The other study[34] found mosaicism in 0.6%-12.5% of pathogenic variants in genes related to epilepsy-related neurodevelopmental disorders, and in 1.3% of *SCN1A* variants specifically. Conclusions were based on the results of multigene epilepsy panels or whole exome sequencing. Differences in sequencing methods could explain the lower prevalences compared to the current study (mosaicism in 7.5% of symptomatic patients); by not screening with smMIPs, it is likely that some (high-grade) mosaic pathogenic variants were missed. The added value of the single molecule tag to deduplicate PCR copies is highlighted by our results; without deduplication, 4 mosaic participants show AAFs that differ significantly from frequencies after deduplication (5, 8, 10, and 15 percentage point differences). At least 2 mosaic participants would not have been identified if not for the single molecule tag.

By using MIPs with single molecule tags and high coverage, we were able to make very accurate estimates of the percentages of mosaicism, which were confirmed by ddPCR. This makes it possible to distinguish high-grade mosaicism from heterozygous pathogenic variants. Three of 4 participants in whom mosaicism was already suspected or shown, based on earlier clinical results, were detected as mosaics. The fourth participant did not meet our criteria for mosaicism because of low coverage at his variant site, although sequencing results were still very suggestive of mosaicism, and further testing by ddPCR confirmed this. It is possible that more mosaic pathogenic variants were missed for this reason, because high coverage could not be reached for all regions of *SCN1A*. We recommend reanalyzing variants that have a strongly deviating AAF but a coverage that is too low to reach statistical significance, as we did for our 2 variants with an AAF < 0.25. It is also possible to miss mosaics with relatively high coverage, as they require higher coverage to be reliably detected. For example, a pathogenic variant with an AAF of 0.4 can only fall below the confidence interval (Figure 2) at ~1000X or higher. By assuming that these variants are heterozygous although there is still a small possibility that they are actually mosaic, our result of mosaicism in 7.5% of symptomatic participants might be an underestimation. Furthermore, it is estimated that 7%-10% of parents of Dravet syndrome patients are low-grade mosaics for the *SCN1A* pathogenic variants of their offspring.[35] Therefore, it is likely that not all tested pathogenic variants were truly de novo. By excluding cases with parental mosaicism, the percentage of mosaic participants may increase even more.

Conversely, smMIPs may generate false-positive mosaics as well. Skewed AAFs at the ends of each smMIP can be due to the use of a NextSeq, which has lower quality base calls in the first few cycles. Future studies should take this into consideration during their experimental design, because smMIPs always start and end at the same coordinate. Skewed AAFs in repetitive regions should also be interpreted with caution, because this can be due to technical artifacts. By reanalyzing 2 pathogenic variants in repetitive regions by ddPCR, we could confirm mosaicism in 1 (Participant 8), while disproving it in the other.

4.2 | Mosaicism as a modifier in *SCN1A*-related epilepsy

Overall, participants with mosaic truncating variants were less severely affected than participants with heterozygous truncating variants, suggesting that mosaicism is an important modifier for *SCN1A*-related phenotypes. This is in line with previously published results.[10,36] Remarkably, in 2 of the 11 mosaic participants (18%), sudden unexpected death in epilepsy (SUDEP) occurred, whereas this occurred in none of the other 111 patients. This may be pure coincidence, but further studies on a possible association between SUDEP and mosaic *SCN1A* pathogenic variants are warranted. Two of the mosaic participants were clinically unaffected, leading to the estimation that symptoms arise at between 14% and 27% of mutated allele in blood. This is in line with earlier studies.[10,37] The effects of mosaicism would be underestimated if we did not take into account participants with such low AAFs that they do not show

any symptoms. However, by including them in our clinical analysis, our results might not be applicable to the group of symptomatic patients for whom we want to improve clinical counseling. Therefore, these participants were excluded from our main analyses (Table 2), but were included in repeated analyses (Table 3). Cognitive abilities usually decline as patients with Dravet syndrome age, which could create a bias in the analyses. However, the mosaic participants are older on average in all analyzed (sub)groups, so the effect of mosaicism on cognitive outcome might even be underestimated.

Interestingly, we also found mosaicism in 3 severely affected participants (Participants 4, 5, and 8). To our knowledge, this is the first report of mosaic *SCN1A* pathogenic variants that do not lead to a relatively mild phenotype. This suggests that other modifiers, such as variants in other genes, have a large influence as well. Furthermore, it is likely that the degree of mosaicism is of more importance than just its presence or absence. In addition, levels of mosaicism in lymphocytes do not necessarily correspond to the levels in brain. Clinical factors such as medication use and comorbidities should also be taken into account. For example, Participant 5 was affected by the unrelated Reye syndrome, which caused major developmental regression. Participants 4, 5, and 8 were all treated with sodium channel blockers, which are contraindicated drugs in Dravet syndrome. This emphasizes the importance of accurate clinical management for cognitive development; even patients with a favorable genotype can ultimately have a poor outcome.

4.3 | Implications for genetic counseling

Besides partially explaining and predicting differences in phenotype severity, our results also reveal other important consequences for genetic counseling. Patients who are only mildly affected because they are mosaic for a pathogenic variant should be made aware that their children are at risk to be more severely affected, when they inherit the variant in a heterozygous state. Furthermore, mosaicism in a proband virtually rules out germline mosaicism in the parents, which lowers their recurrence risk to zero. Nevertheless, Participant 7 shows that simply assessing mosaicism is not enough for accurate counseling. At first sight, she seemed mosaic for her known insG pathogenic variant. However, because we also discovered her mosaic T>G variant 3 base pairs upstream, we could reason that this was most likely a mosaic read frame-restoring variant of an originally heterozygous insG pathogenic variant. Similar read-frame restoring mosaic variants and human reverse mutations in general have previously been described.[38–41] This is the first report of a reverse mutation in Dravet syndrome. Several hypotheses about the occurrence of reverse mutations have been proposed.[41,42] In Dravet syndrome, however, these do not seem applicable, so this might be pure coincidence. Nonetheless, by deducing that the insG pathogenic variant only seemed mosaic but was probably originally heterozygous, we now cannot exclude that one of the patient's parents might be germline mosaic for this variant, with an inherent increased recurrence risk.

Overall, mosaicism is present in 7.5% of de novo pathogenic *SCN1A* variants in clinically affected patients, which implicates that postzygotic mutation is a common phenomenon in *SCN1A*-related epilepsies. Patients with mosaicism of truncating variants have on average milder phenotypes, which makes mosaicism an important modifier in *SCN1A*-related phenotypes. However, mosaicism is also seen in severely affected patients, implicating an important role of other modifiers, including accuracy of clinical management. Detection of mosaicism has important implications for genetic counseling regarding recurrence risk and phenotype prediction, and can be achieved by deep sequencing of unique reads. Our results stress the importance of implementing high-coverage NGS with attention for possible mosaic pathogenic variants in standard diagnostics.

ACKNOWLEDGMENTS

This study was supported by the Stichting Vrienden WKZ (project 1614054) on behalf of Stichting Panta Rhei, and the Dutch Epilepsy Foundation (project 2017-01).

DISCLOSURE

None of the authors has any conflict of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Iris M. de Lange http://orcid.org/0000-0003-2107-9224

REFERENCES

1. Claes L, Ceulemans B, Audenaert D, et al. De novo SCN1A mutations are a major cause of severe myoclonic epilepsy of infancy. Hum Mutat. 2003;21:615–21.
2. Mulley JC, Scheffer IE, Petrou S, et al. SCN1A mutations and epilepsy. Hum Mutat. 2005;25:535–42.
3. Depienne C, Bouteiller D, Keren B, et al. Sporadic infantile epileptic encephalopathy caused by mutations in PCDH19 resembles Dravet syndrome but mainly affects females. PLoS Genet. 2009;5:e1000381.
4. Escayg A, MacDonald BT, Meisler MH, et al. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet. 2000;24:343–5.

5. Dravet C. Les epilepsies graves de l'enfant. Vie Med. 1978;8:543–8.
6. Brunklaus A, Ellis R, Reavey E, et al. Prognostic, clinical and demographic features in SCN1A mutation-positive Dravet syndrome. Brain. 2012;135:2329–36.
7. Rodda JM, Scheffer IE, McMahon JM, et al. Progressive gait deterioration in adolescents with Dravet syndrome. Arch Neurol. 2012;69:873–8.
8. Zuberi SM, Brunklaus A, Birch R, et al. Genotype-phenotype associations in SCN1A-related epilepsies. Neurology. 2011;76:594–600.
9. Ceulemans BPGM, Claes LRF, Lagae LG. Clinical correlations of mutations in the SCN1A gene: from febrile seizures to severe myoclonic epilepsy in infancy. Pediatr Neurol. 2004;30:236–43.
10. Meng H, Xu HQ, Yu L, et al. The SCN1A mutation database: updating information and analysis of the relationships among genotype, functional alteration, and phenotype. Hum Mutat. 2015;36:573–80.
11. Kanai K, Yoshida S, Hirose S, et al. Physicochemical property changes of amino acid residues that accompany missense mutations in SCN1A affect epilepsy phenotype severity. J Med Genet. 2009;46:671–9.
12. Ishii A, Watkins JC, Chen D, et al. Clinical implications of SCN1A missense and truncation variants in a large Japanese cohort with Dravet syndrome. Epilepsia. 2017;58:282–90.
13. Akiyama M, Kobayashi K, Yoshinaga H, et al. A long-term follow-up study of Dravet syndrome up to adulthood. Epilepsia. 2010;51:1043–52.
14. Harkin LA, McMahon JM, Iona X, et al. The spectrum of SCN1A-related infantile epileptic encephalopathies. Brain. 2007;130:843–52.
15. Jansen F, Sadleir L, Harkin L, et al. Severe myoclonic epilepsy of infancy (Dravet syndrome): recognition and diagnosis in adults. Neurology. 2006;67:2224–6.
16. Parihar R, Ganesh S. The SCN1A gene variants and epileptic encephalopathies. J Hum Genet. 2013;58:573–80.
17. Ohmori I, Ouchida M, Kobayashi K, et al. CACNA1A variants may modify the epileptic phenotype of Dravet syndrome. Neurobiol Dis. 2013;50:209–17.
18. Martin MS, Tang B, Papale LA, et al. The voltage-gated sodium channel Scn8a is a genetic modifier of severe myoclonic epilepsy of infancy. Hum Mol Genet. 2007;16:2892–9.
19. Long YS, Zhao QH, Su T, et al. Identification of the promoter region and the 5'-untranslated exons of the human voltage-gated sodium channel Nav1.1 gene (SCN1A) and enhancement of gene expression by the 5'-untranslated exons. J Neurosci Res. 2008;86:3375–81.
20. Zeng T, Dong ZF, Liu SJ, et al. A novel variant in the 3′ UTR of human SCN1A gene from a patient with Dravet syndrome decreases mRNA stability mediated by GAPDH's binding. Hum Genet. 2014;133:801–11.
21. Depienne C, Arzimanoglou A, Trouillard O, et al. Parental mosaicism can cause recurrent transmission ofSCN1A mutations associated with severe myoclonic epilepsy of infancy. Hum Mutat. 2006;27:389.
22. Marini C, Mei D, Helen Cross J, et al. Mosaic SCN1A mutation in familial severe myoclonic epilepsy of infancy. Epilepsia. 2006;47:1737–40.
23. Selmer KK, Eriksson AS, Brandal K, et al. Parental SCN1A mutation mosaicism in familial Dravet syndrome. Clin Genet. 2009;76:398–403.
24. Gennaro E, Santorelli FM, Bertini E, et al. Somatic and germline mosaicisms in severe myoclonic epilepsy of infancy. Biochem Biophys Res Commun. 2006;341:489–93.
25. Gottlieb B, Beitel LK, Trifiro MA. Somatic mosaicism and variable expressivity. Trends Genet. 2001;17:79–82.
26. Verbeek NE, Van Der Maas NAT, Jansen FE, et al. Prevalence of SCN1A-related Dravet syndrome among children reported with seizures following vaccination : a population-based ten-year cohort study. PLoS One 2013;8:e65758.
27. Verbeek NE, Van Kempen M, Gunning WB, et al. Adults with a history of possible Dravet syndrome: an illustration of the importance of analysis of the SCN1A gene. Epilepsia. 2011;52:23–5.
28. Varni JW, Limbers CA. The pediatric quality of life inventory: measuring pediatric health-related quality of life from the perspective of children and their parents. Pediatr Clin North Am. 2009;56:843–63.
29. Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, et al. Reliable next-generation sequencing of formalin-fixed, paraffin-embedded tissue using single molecule tags. J Mol Diagnostics. 2016;18:851–63.
30. Qin L, Wang J, Tian X, et al. Detection and quantification of mosaic mutations in disease genes by next-generation sequencing. J Mol Diagnostics. 2016;18:446–53.
31. Ballif B, Rorem E, Sundin K, et al. Detection of low-level mosaicism by array CGH in routine diagnostic specimens. Am J Med Genet A .2006;140:2757–67.
32. Jamuar SS, Lam A-TN, Kircher M, et al. Somatic mutations in cerebral cortical malformations. N Engl J Med. 2014;371:733–43.
33. Acuna-Hidalgo R, Bo T, Kwint M, et al. Post-zygotic point mutations are an underrecognized source of novel genomic variation. Am J Hum Genet. 2015;97:67–74.
34. Stosser MB, Lindy AS, Butler E, et al. High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. Genet Med. 2017;1–8.
35. Xu X, Yang X, Wu Q, et al. Amplicon resequencing identified parental mosaicism for approximately 10% of "de novo" SCN1A mutations in children with Dravet syndrome. Hum Mutat. 2015;36:861–72.
36. Depienne C, Trouillard O, Gourfinkel-An I, et al. Mechanisms for variable expressivity of inherited SCN1A mutations causing Dravet syndrome. J Med Genet. 2010;47:404–10.
37. Shi YW, Yu MJ, Long YS, et al. Mosaic SCN1A mutations in familial partial epilepsy with antecedent febrile seizures. Genes Brain Behav. 2012;11:170–6.
38. Ellis NA, Ciocci S, German J. Back mutation can produce phenotype reversion in Bloom syndrome somatic cells. Hum Genet. 2001;108:167–73.
39. Waisfisz Q, Morgan NV, Savino M, et al. Spontaneous functional correction of homozygous Fanconi anaemia alleles reveals novel mechanistic basis for reverse mosaicism. Nat Genet. 1999;22:379–83.
40. Darling TN, Yee C, Bauer JW, et al. Revertant mosaicism: partial correction of a germ-line mutation in COL17A1 by a frame-restoring mutation. J Clin Invest. 1999;103:1371–7.

41. Hirschhorn R. In vivo reversion to normal of inherited mutations in humans. J Med Genet. 2003;40:721–8.
42. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. Nat Rev Genet. 2002;3:748–58.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: de Lange IM, Koudijs MJ, van't Slot R, et al. Mosaicism of de novo pathogenic *SCN1A* variants in epilepsy is a frequent phenomenon that correlates with variable phenotypes. *Epilepsia*. 2018;59:690–703. https://doi.org/10.1111/epi.14021