


# Modifier genes in *SCN1A*-related epilepsy syndromes

Iris M. de Lange<sup>1</sup>  | Flip Mulder<sup>1</sup> | Ruben van 't Slot<sup>1</sup> | Anja C. M. Sonsma<sup>1</sup> | Marjan J. A. van Kempen<sup>1</sup> | Isaac J. Nijman<sup>1</sup> | Robert F. Ernst<sup>1</sup> | Nine V. A. M. Knoers<sup>1,2</sup> | Eva H. Brilstra<sup>1</sup> | Bobby P. C. Koeleman<sup>1</sup>

<sup>1</sup>Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>2</sup>Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands

## Correspondence

Iris M. de Lange, Lundlaan 6, 3584CG 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

## Abstract

**Background:** *SCN1A* is one of the most important epilepsy-related genes, with pathogenic variants leading to a range of phenotypes with varying disease severity. Different modifying factors have been hypothesized to influence *SCN1A*-related phenotypes. We investigate the presence of rare and more common variants in epilepsy-related genes as potential modifiers of *SCN1A*-related disease severity.

**Methods:** 87 patients with *SCN1A*-related epilepsy were investigated. Whole-exome sequencing was performed by the Beijing Genomics Institute (BGI). Functional variants in 422 genes associated with epilepsy and/or neuronal excitability were investigated. Differences in proportions of variants between the epilepsy genes and four control gene sets were calculated, and compared to the proportions of variants in the same genes in the ExAC database.

**Results:** Statistically significant excesses of variants in epilepsy genes were observed in the complete cohort and in the combined group of mildly and severely affected patients, particularly for variants with minor allele frequencies of <0.05. Patients with extreme phenotypes showed much greater excesses of epilepsy gene variants than patients with intermediate phenotypes.

**Conclusion:** Our results indicate that relatively common variants in epilepsy genes, which would not necessarily be classified as pathogenic, may play a large role in modulating *SCN1A* phenotypes. They may modify the phenotypes of both severely and mildly affected patients. Our results may be a first step toward meaningful testing of modifier gene variants in regular diagnostics for individual patients, to provide a better estimation of disease severity for newly diagnosed patients.

## KEYWORDS

Dravet, epilepsy, GEFS+, modifier genes, phenotypic variability, *SCN1A*

Eva H. Brilstra and Bobby P. C. Koeleman contributed equally to the manuscript.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

## 1 | INTRODUCTION

*SCN1A* (OMIM #182389) is one of the most important epilepsy-related genes, with pathogenic variants leading to a wide range of phenotypes with varying disease severity (Claes et al., 2003; Escayg & Goldin, 2010; Mulley et al., 2005; Sadleir et al., 2017). One of the most severe associated diseases is Dravet syndrome, which is characterized by intractable epileptic seizures, a diminishing psychomotor development that results in mild to severe intellectual disability (ID), and often walking difficulties and behavioral problems (Brunklau, Ellis, Reavey, Forbes, & Zuberi, 2012; Dravet, 1978; Rilstone, Coelho, Minassian, & Andrade, 2012). Milder phenotypes include GEFS+ syndrome and febrile seizures, in which seizures show a milder course and a normal intellect is expected (Catterall, Kalume, & Oakley, 2010; Escayg & Goldin, 2010).

*SCN1A* encodes for the  $\alpha$ -subunit of a neuronal sodium channel, Nav1.1. Different pathogenic variants in *SCN1A* can have different effects on channel function, which partly explains why the gene is associated with multiple phenotypes. Variants leading to a complete loss of function (LoF) of the channel are virtually always associated with severe phenotypes, whereas milder disturbances in channel function usually cause milder clinical pictures (Meng et al., 2015). However, a large part of the phenotypic variability of patients remains unexplained: there are several reports of families in which multiple members carry the exact same pathogenic *SCN1A* variant, but nevertheless show an intra-familial variability in phenotype severity (Depienne et al., 2010; Guerrini et al., 2010; Mahoney et al., 2009; Passamonti et al., 2015; Pineda-Trujillo et al., 2005; Suls et al., 2010). Furthermore, Dravet syndrome patients with similar loss of function variants may show important phenotypic differences, ranging from severely disabled individuals to patients that live much more independent lives (Akiyama, Kobayashi, Yoshinaga, & Ohtsuka, 2010; Harkin et al., 2007; Jansen et al., 2006). This variability makes it difficult to accurately predict clinical outcomes in newly diagnosed young patients, which is very important to parents.

Several factors have been suggested to modify the clinical outcome of *SCN1A*-related epilepsy and to explain these phenotypic differences. Mosaicism for a pathogenic *SCN1A* variant can have a major ameliorating impact on disease severity (Depienne et al., 2010; Gennaro et al., 2006; de Lange, Koudijs, et al., 2018; Marini, Mei, Helen Cross, & Guerrini, 2006). Furthermore, variants in regulatory regions of *SCN1A* may modulate disease severity (Long et al., 2008; Zeng et al., 2014). Additionally, clinical management and especially the use of contra-indicated medication can affect clinical outcomes (Ceulemans, 2011; Guerrini et al., 1998; de Lange, Gunning, et al., 2018).

Moreover, variants in modifier genes may influence *SCN1A*-related phenotypes. An important effect of modifier

genes has already been described for several other genetic disorders (Emond et al., 2013; Guo et al., 2015; Vélez et al., 2016), and there are strong indications that genetic background can modulate the clinical effects of pathogenic *SCN1A*-related phenotypes too, in human patients as well as in *Scn1a* knock-out mice (Catterall et al., 2010; Depienne et al., 2010; Guerrini et al., 2010; Hawkins & Kearney, 2016; Miller, Hawkins, McCollom, & Kearney, 2014; Pineda-Trujillo et al., 2005; Scheffer, Zhang, Jansen, & Dibbens, 2009; Singh, Scheffer, Crossland, & Berkovic, 1999; Suls et al., 2010; Yu et al., 2006). Several potential modifier genes have already been identified: variants in *SCN9A*, *SCN8A*, *SCN2A*, *HLF*, *POLG*, *KCNQ2*, *CACNB4*, *CACNA1G*, and *CACNA1A* might aggravate or partially rescue clinical outcomes (Calhoun, Hawkins, Zachwieja, & Kearney, 2017; Gaily et al., 2013; Hammer et al., 2017; Hawkins & Kearney, 2012, 2016; NA, Martin, Frankel, Kearney, & Escayg, 2011; Martin et al., 2007; Ohmori et al., 2013, 2008; Singh et al., 2009). Potential modifier loci, identified in *Scn1a* knock-out mice with different disease severities, also contain several candidate modifier genes, including GABA receptor subunit genes, ion channel genes and genes associated with seizures or neuronal hyperexcitability (Miller et al., 2014). Furthermore, an enrichment of rare variants in neuronal excitability genes in general has been identified in severely affected Dravet syndrome patients, compared to mildly affected Dravet syndrome patients (Hammer et al., 2017). However, these potential modifiers each account for only a small portion of the clinical variability of *SCN1A*-related phenotypes. Many only show an effect when studied in large groups of patients and different patients might be affected by different modifiers or by multiple modifiers simultaneously. Currently, no clinically relevant modifier genes have been identified for which diagnostic testing can be offered, and thus more research is needed to understand clinical variability and to improve the counselling of patients.

Here, we investigate the presence of rare and more common variants in epilepsy-related genes that could potentially modify disease severity, in a cohort of 87 patients with *SCN1A*-related epilepsy. We provide a descriptive overview of variants present in patients with phenotypes on the most extreme ends of the spectrum, and furthermore investigate variants in six families with multiple affected members that show varying disease severities.

## 2 | MATERIALS AND METHODS

### 2.1 | Editorial policies and ethical considerations

The study was approved by the Ethical Committee of the University Medical Center Utrecht. Informed consent was

obtained from participants or their legal caretakers, according to the Declaration of Helsinki.

## 2.2 | Cohort and clinical data

### 2.2.1 | Participants

A cohort of 87 participants with pathogenic *SCN1A* variants was evaluated, most of whom have been described previously (de Lange, Gunning, et al., 2018; de Lange, Koudijs, et al., 2018). Only participants with pathogenic variants (class V) or likely pathogenic variants (class IV) in *SCN1A* were included, according to the American College of Medical Genetics and Genomics criteria (Richards et al., 2015). All variants had been detected and classified in diagnostic laboratories. Patients that had previously been shown to be mosaic for their pathogenic *SCN1A* variant were excluded (de Lange, Koudijs, et al., 2018).

### 2.2.2 | Disease severity classification

For all participants detailed clinical data were collected from medical records and semi-structured telephone interviews. Patients were either part of families with multiple *SCN1A* variants carriers, or the only affected member in their family. In all patients absolute disease severity was defined as cognitive functioning at the age of 6 years. We assessed cognitive functioning retrospectively at the age of 6 years old as previously described (de Lange, Gunning, et al., 2018). This was done to limit the influence of an older age on cognitive outcomes, since average cognition declines with age in Dravet syndrome patients. IQ- and developmental assessment scores, established at different ages, were interpolated by linear regression, to obtain approximate scores at the age of 6 years. This age was chosen since cognitive decline is generally most severe in the first years following disease onset (Brunklau et al., 2012; Nabbout et al., 2013). Patients with an IQ or developmental quotient (DQ) of >70 (no or borderline ID) at age 6 were classified as “mild”, while patients with an IQ or DQ of <50 (moderate or severe to profound ID) were classified as “severe”. Patients with an IQ/DQ score of 50–70 were classified as “intermediate”. Participants under the age of 6 years old were not classified, unless they already showed an IQ/DQ of <50. Participants for whom a classification at age 6 could not be reliably obtained were also not classified. Furthermore, if clearly varying phenotypes were present in families with multiple variant carrying family members (different syndromes, or large differences in seizure frequencies or cognitive outcomes), disease severity was defined as “mild” or “severe” relative to other affected family members (e.g., a participant with Dravet syndrome and an unaffected

father, both carrying the same pathogenic *SCN1A* variant, would be classified as “relatively severe” and “relatively mild” respectively).

To compare subgroups, we then excluded “mild” patients that did not carry an *SCN1A* variant that was predicted to cause a loss of function (LoF), or a variant that has been described before in Dravet syndrome patients. This limits the influence of the different pathogenic *SCN1A* variants itself on the phenotypes and creates a group of patients for which we can be relatively certain that ameliorating modifiers play a role. The “severe” and “intermediate” groups included patients with all mutation types.

## 2.3 | Molecular analyses

### 2.3.1 | Exome sequencing

Whole-exome sequencing was performed on DNA from lymphocytes in all patients by the Beijing Genomics Institute (BGI), using the Agilent V5 50M exome kit enrichment, followed by paired-end sequencing on an Illumina HiSeq. The resulting data was processed using an in-house developed pipeline (Ernst et al., 2017), according to the best practices guidelines (Auwera et al., 2013). Briefly, sequencing reads were mapped using BWA-MEM v0.7.5a (Li & Durbin, 2009), duplicates were marked and lanes were merged. Next, using GATK IndelRealigner (v3.4–46) (McKenna et al., 2010) indels were realigned and the GATK HaplotypeCaller tool was used to create a GVCF per patient containing SNPs and indels. These GVCFs were jointly genotyped using GATK GenotypeGVCFs for the described cohort. Variants were flagged using GATK VariantFiltration if they did not meet the certain criteria. For SNPs the criteria were: QD <2.0, MQ <40.0, FS >60.0, HaplotypeScore >13.0, MQRankSum <−12.5, ReadPosRankSum <−8.0, snpclusters ≥3 in 35 bp. The criteria for indels were as follows: QD <2.0, FS >200.0, ReadPosRankSum <−20.0. Finally, variants were annotated using SnpSift (v4.3t) and dbNSFP (v3.5).

### 2.3.2 | Filtering of variants

We investigated variants in 422 genes that are all either associated with epilepsy, are implicated to modify epilepsy phenotypes, are associated with neuronal excitability, or function in the same pathway as *SCN1A*, based on epilepsy gene panels used in the University Medical Center Utrecht (EPI00v18.1), previous literature and the KEGG pathway database ([https://www.kegg.jp/kegg-bin/show\\_pathway?ko04728](https://www.kegg.jp/kegg-bin/show_pathway?ko04728), accessed June 2016) (further referred to as “epilepsy genes”; see Data S1 for the complete list,

and Data S2 for characteristics). A distinction was made between established monogenic epilepsy genes (when present in the diagnostic epilepsy gene panel of the University Medical Center Utrecht (EPI00v18.1) and candidate genes (all other genes). We filtered for PASS-variants that were predicted to alter protein function (frameshift, stop-gain, stop-loss, start-loss, in-frame deletion, in-frame insertion, splice donor, splice acceptor, and nonsynonymous missense variants). Five categories of variants were established (type A–E), based on different minor allele frequencies (MAF) of the variants (in both exomes and genomes in the gnomAD database, r2.0 (Exome Aggregation Consortium et al., 2015), all populations) and on their deleteriousness as predicted by CADD scores (Combined Annotation-Dependent Depletion, v1.3) (Kircher et al., 2014): Type A variants have a MAF of  $<0.01$  and a (PHRED-scaled) CADD-score of  $>20$  (representing the top 1% deleterious substitutions in the human genome); type B variants have a MAF of  $<0.01$  and a CADD-score of  $>10$  (representing the top 10% deleterious substitutions in the human genome), type C variants have a MAF of  $<0.01$  and any CADD score; type D variants have a MAF of  $<0.05$ ; and type E variants have a MAF of  $<0.1$ . The known pathogenic *SCN1A* variant of each patient was excluded.

The same categories of variants were established for variants in four sets of control genes (control 1: immunodeficiency-related genes,  $n = 360$ ; control 2: genes related to cardiovascular disease [excluding genes related to conduction abnormalities],  $n = 109$ ; control 3: genes related to kidney disease,  $n = 223$ ; control 4: genes related to either hemostasis, erythroid cell membrane defects, congenital diarrhea, neonatal erythroderma, or angioedema,  $n = 297$ ), and in genes associated with ID (excluding genes also present in the epilepsy gene-list,  $n = 659$ ), based on genes included in diagnostic gene panels of the University Medical Center Utrecht (version 9, <http://www.umcutrecht.nl/NGS>) (see Data S1 for the complete lists, and Data S2 for characteristics).

## 2.4 | Data analyses

### 2.4.1 | Proportions of variants in epilepsy genes and control genes compared to the ExAC database

We investigated whether groups of patients carry an excess of variants in our selection of epilepsy genes, as compared to the number of variants in the different sets of control genes (1–4). Since these control sets contain different numbers of genes than the set of epilepsy genes, with different lengths and mutation rates, we first investigated a healthy control population to establish the normal ratios of variants between

the epilepsy genes and the different control sets. For this, we extracted variants in the same genes from the ExAC database (Exome Aggregation Consortium et al., 2015), using the same filters as applied in our cohort (frameshift, stop-gain, stop-loss, start-loss, in-frame deletion, in-frame insertion, splice donor, splice acceptor and non-synonymous missense variants, with MAFs of either  $<0.01$ ,  $<0.05$  or  $<0.1$ ). Only variants present in non-Finnish Europeans were analyzed, as this population resembles the ethnicity of 97% of our own cohort. Directly comparing numbers of variants found in the ExAC database to other data can lead to incorrect results, as differences in sequencing methods, coverage and variant calling may lead to biases (Barrett et al., 2017). However, we expect the ratios of variants in epilepsy genes and control sets of genes to be roughly similar in both ExAC data and in our own sequencing data, as within each cohort the same protocols are used to analyse the various gene sets. We therefore compare these ratios, rather than absolute numbers of variants, in both cohorts. The ratio of variants in epilepsy genes and in the different sets of control genes in the ExAC database (=numbers of epilepsy gene variants divided by the number of control gene variants) was used to calculate the expected number of variants in epilepsy genes in our cohort (the established ExAC ratio times the number of variants in control genes in our cohort). We compared this expected number of variants in epilepsy genes to the actual number of variants found in our cohort, to obtain the percentage of over- or underrepresentation. Fishers' exact test was used to determine whether this over- or underrepresentation of epilepsy gene variants was statistically significant ( $p$ -value threshold for significance:  $<.05$  divided by the number of tests to the corrected for multiple testing). These analyses were performed for the ratio between epilepsy gene variants and variants in all four sets of control genes, for all three frequency thresholds ( $<0.01$ ,  $<0.05$  or  $<0.1$ ; type C, D and E variants) and for different groups of patients (the complete cohort, only patients on the extreme ends of the disease spectrum, and only intermediate patients). We hypothesized that patients with a phenotype on both the severe and mild ends of the disease spectrum would carry more variants in epilepsy genes than intermediate patients, as both groups are likely to have a modified phenotype.

### 2.4.2 | Differences between mild and severe patients

We then assessed the distribution of epilepsy gene variants present in our cohort in the different categories of patients (mild, severe and intermediate). The total number of alleles per group was calculated (=the number of genes in which at least one variant was found in at least one of the patient groups, multiplied by two alleles, multiplied by the number

of patients in the group, minus one for each X-linked gene for each male in the group). The number of found variants per group was then divided by the total number of alleles per group, to obtain a percentage of variants corrected for group size and for male/female ratio (since also variants in X-linked genes are present). Differences between groups were calculated using Fishers' exact test. Analyses were performed 5 times, once for each category of variants (type A–E). For each of these categories, we furthermore identified in which genes severe patients carried most variants compared to mild patients, and vice versa (Fishers' exact test, based on the numbers of variants and total alleles per gene in each group).

### 2.4.3 | Variants found in families with variable phenotypes and in patients with the most extreme phenotypes

In families with multiple affected family members with different disease severities, we report variants that were present in only severe patients and not in their milder family member(s) (=possible negative modifiers that could aggravate the phenotype) and variants that were present in mild patients but not in their severe family member(s) (=possible positive modifiers that could ameliorate the phenotype). Only the most predicted deleterious variants are described (type A), as it is difficult to prove the influence of more common and milder variants.

We furthermore report the most predicted deleterious variants in the patients with the most extreme phenotypes from the mild and severe groups (IQ at the age of six <30 and all the mild patients [IQ >70]). For each patient, the predicted most deleterious variant in an established epilepsy gene and the predicted most deleterious variant in a candidate gene is reported, based on the highest CADD score. When the variant with the highest CADD score was present in a recessive gene, the highest CADD score in a dominant gene was reported, if these were present.

## 3 | RESULTS

For 87 participants whole-exome data were obtained (see Table S1 for information on *SCN1A* pathogenic variants and clinical data). Coverage values of the analyzed gene sets differed between cohorts (ExAC and the described cohort), but was similar for sets within the same cohort (Data S3 and S4). In all patients, their known *SCN1A* pathogenic variants could be identified, except for large structural variants (e.g., deletions of the complete *SCN1A* gene), meaning no samples swaps had occurred.

Varying phenotypes were observed in six families (Data S5). For 69 participants an estimation of cognitive

functioning at the age of 6 years old could be made: 22 participants were severely affected, 29 were mildly affected, and 18 patients were categorized as intermediate. For 18 patients no estimation could be made, because they were either under the age of 6 and still mildly affected, or they were severely affected but no official IQ/DQ assessments were available close to the age of 6, meaning we cannot be sure when the exact decline happened. Ten of the mildly affected patients carried a LoF variant or a variant that was previously associated with a severe phenotype, and were included in the “mild” group. The mild group included both brothers from family 3; although one brother is significantly more severely affected than the other, both brothers were still categorized as mild at the age of 6 years old. The “mild” and “severe” patients combined are referred to as “extreme” patients.

### 3.1 | Proportion of variants in epilepsy genes and control genes as compared to the ExAC database

Table 1 depicts the numbers of variants found in epilepsy genes, ID genes and different sets of control genes, for each groups of patients, in this cohort and in the ExAC database. We observed a significant excess of variants in epilepsy genes in the complete cohort, most strongly for type D variants but also for type E variants, when compared to ratios of variants in the ExAC database (111%–126%,  $p < .0003$ ). A statistically significant overrepresentation of type D epilepsy gene variants was furthermore observed for extreme patients in relation to the control gene sets combined (118%,  $p < .0003$ ). Overall, extreme patients showed a two- to fourfold (type E) and five- to sevenfold (type D) greater excess of epilepsy gene variants than intermediate patients (Table 2; Figure 1a). This pattern was observed in relation to all sets of control genes except for set 2. No significant excess of variants in intellectual disability genes was observed (Table 2; Figure 1b). There was no significant excess of variants in control set 1 in relation to variants in control set 3 and 4, conform expectation and suggesting validity of these analyses (Table 2; Figure 1c).

### 3.2 | Differences between mild and severe patients

When assessing the distribution of variants present in epilepsy genes in our cohort between the different groups of patients (mild, severe and intermediate), no statistically significant differences were observed (Table 3; Figure 2). This is likely due to smaller sample sizes. We performed a power calculation for type D variants of intermediate patients versus mild and severe patients (the category in which the largest differences

were observed), to estimate the effect size that could be reliably detected in a cohort of this size. This analysis showed that, in order to reach statistical significance with a power of 0.8 and our sample size of 50 patients, the group of mild and severe patients combined would have to contain more than six times as much variant alleles as the group of intermediate patients (25% vs. 4.03%). Since only a 1.1 times more variant alleles were observed in the group of mild and severe patients (4.47% vs. 4.03%), our study is likely underpowered to assess the distribution of variants among groups of patients. For each category of variants, we report the genes in which severe patients carried most variants compared to mild patients, and vice versa in Table 4.

### 3.3 | Variants found in families with variable phenotypes

Detailed clinical data on the phenotypes of participants from the six families with varying phenotypes are described in Data S4. *Family 1* consists of a severely affected 10 year old proband with Dravet syndrome, and a father with mild epilepsy and normal cognitive functioning. *Family 2* is a large GEFS+ family in which exome sequencing was performed in two mildly affected family members (only febrile seizures/no seizures at all) and in one severely affected family member (severe epilepsy and a severe social-emotional delay, classified as Dravet syndrome). *Family 3* consists of two brothers with Dravet syndrome, one of whom is more severely affected than the other. *Family 4* consists of a proband with a phenotype on the border of Dravet syndrome and GEFS+, with regression over the years. His father has never had any seizures. *Family 5* consists of two brothers of whom the oldest has severe Dravet syndrome and the youngest has a much milder phenotype. *Family 6* consists of a proband with mild Dravet syndrome, a father with a milder epilepsy and better cognitive functioning but with many social- and behavioral problems, and a grandmother who only experienced three seizures in her life.

Type A variants that were only present in either the mild or the severe family members of each family are depicted in Table 5.

### 3.4 | Variants found in patients with the most extreme phenotypes

We further investigated type A variants in patients with the most extreme phenotypes in this cohort (IQ at the age of six <30 or >70): this comprised all 10 “mild” patients, and seven of the “severe” patients. For each patient, their most predicted deleterious variant in both an established epilepsy gene and in an epilepsy candidate gene is depicted in Table 6.

**TABLE 1** Number of variants found in the ExAC database and in the current cohort for different groups of patients and different categories of variants

Gene set	Type C variants (MAF <sup>a</sup> <0.01)			Type D variants (MAF <0.05)			Type E variants (MAF <0.1)					
	ExAC	Complete cohort	Extreme patients	Intermediate patients	ExAC	Complete cohort	Extreme patients	Intermediate patients	ExAC	Complete cohort	Extreme patients	Intermediate patients
Epilepsy genes	300,155	851	337	161	556,293	1,876	662	335	805,686	2,455	888	481
Control 1	299,750	858	305	176	641,440	1,778	626	371	950,857	2,608	893	546
Control 2	90,285	278	130	49	185,549	497	218	83	260,406	765	327	140
Control 3	260,659	707	270	147	513,655	1,378	501	300	812,452	2,165	798	467
Control 4	261,763	702	247	131	565,998	1,603	579	331	853,276	2,291	824	473
Control 1–4 (total)	912,457	2,545	952	503	1,906,642	5,256	1,924	1,085	2,876,991	7,829	2,842	1,626
ID genes	557,769	1,572	582	355	1,099,721	3,172	1,132	708	1,707,615	4,919	1,739	1,071

<sup>a</sup>Minor allele frequency; only variants with a frequency below this threshold in both the exomes and genomes in the gnomAD database are included.

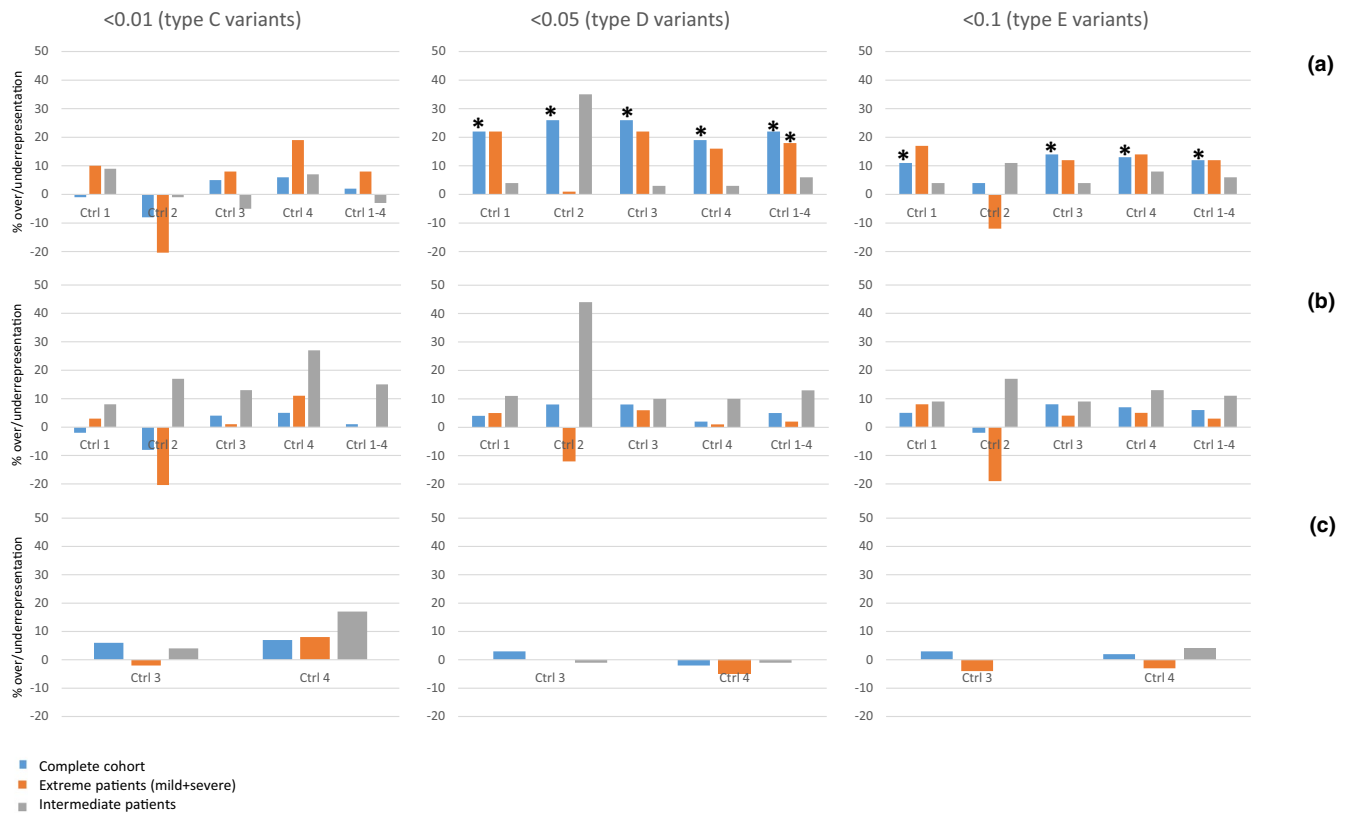
**TABLE 2** Overrepresentation of variants in epilepsy genes in the cohort, calculated based on different sets of control genes

Assessed group of genes	Over- or underrepresentation of epilepsy gene variants in cohort (% more or less than expected <sup>a</sup> [ <i>p</i> -values <sup>b</sup> ])											
	Type C variants (MAF <sup>c</sup> <0.01)				Type D variants (MAF <0.05)				Type E variants (MAF <0.1)			
	Based on control set	Complete cohort	Extreme patients	Intermediate patients	Complete cohort	Extreme patients	Intermediate patients	Complete cohort	Extreme patients	Intermediate patients	Complete cohort	Extreme patients
Epilepsy genes	1	-1 (.846)	10 (.221)	9 (.414)	<b>22 (&lt;.0003)</b>	22 (<.0005)	4 (.598)	<b>11 (&lt;.0003)</b>	17 (.001)	4 (.552)		
	2	-8 (.229)	-22 (.017)	-1 (.941)	<b>26 (&lt;.0003)</b>	1 (.904)	35 (.016)	4 (.388)	-12 (.044)	11 (.283)		
	3	5 (.387)	8 (.329)	-5 (.690)	<b>26 (&lt;.0003)</b>	22 (.001)	3 (.721)	<b>14 (&lt;.0003)</b>	12 (.019)	4 (.581)		
	4	6 (.285)	19 (.038)	7 (.597)	<b>19 (&lt;.0003)</b>	16 (.008)	3 (.727)	<b>13 (&lt;.0003)</b>	14 (.007)	8 (.257)		
1-4 (total)	2 (.677)	8 (.245)	-3 (.785)	<b>22 (&lt;.0003)</b>	<b>18 (&lt;.0003)</b>	6 (.374)	<b>12 (&lt;.0003)</b>	12 (.005)	6 (.292)			
ID genes	1	-2 (.718)	3 (.750)	8 (.042)	4 (.184)	5 (.288)	11 (.101)	5 (.045)	8 (.051)	9 (.097)		
	2	-8 (.178)	-28 (.001)	17 (.315)	8 (.127)	-12 (.073)	44 (.002)	-2 (.612)	-19 (.001)	17 (.092)		
	3	4 (.405)	1 (.940)	13 (.231)	8 (.025)	6 (.326)	10 (.166)	8 (.003)	4 (.407)	9 (.120)		
	4	5 (.280)	11 (.193)	27 (.020)	2 (.561)	1 (.918)	10 (.159)	7 (.006)	5 (.216)	13 (.025)		
1-4 (total)	1 (.784)	0 (1.000)	15 (.041)	5 (.045)	2 (.599)	13 (.011)	6 (.002)	3 (.320)	11 (.008)			
Control set genes	3	6 (.298)	-2 (.835)	4 (.738)	3 (.370)	0 (1.000)	-1 (.908)	3 (.323)	-4 (.367)	0 (1.000)		
	4	7 (.204)	8 (.393)	17 (.170)	-2 (.535)	-5 (.419)	-1 (.910)	2 (.465)	-3 (.578)	4 (.594)		

<sup>a</sup>Numbers represent the percentage of over- or underrepresentation of variants in epilepsy genes, based on a comparison of ratios of variants found in epilepsy genes and in different control groups, in the ExAC database and the current cohort (e.g. the first “7” means that 107% of the expected number of epilepsy genes was identified in our cohort, based on the ratio of variants present in epilepsy genes and control 1-genes in the ExAC database, and the number of variants in control-1 genes present in our cohort).

<sup>b</sup>*p*-values are based on Fishers' exact tests on the ratios of variants found in the ExAC database and the current cohort. Significant values (*p* < .0003968) are bolded.

<sup>c</sup>Minor allele frequency; only variants with a frequency below this threshold in both the exomes and genomes in the gnomAD database are included.



**FIGURE 1** Overrepresentation of variants in epilepsy genes in the cohort. Bars represent the percentage of over- or underrepresentation of variants in the different patient groups, based on the ratio of variants found in epilepsy genes and different control groups (ctrl 1, 2, 3, 4 and 1–4), compared to ratios in the ExAC database. (a) Variants in epilepsy genes compared to different control groups; (b) variants in intellectual disability genes compared to different control groups; (c) variants in control group 1 genes compared to control group 3 and 4 (negative control). Results are presented for categories of variants with different allele frequency cut-offs ( $<0.01$ ,  $<0.05$ ,  $<0.1$ ). Significant values are depicted by asterisks

### 3.5 | Comparison of current data and previous literature

A list of all previously implicated modifier genes for *SCN1A*-related epilepsy is shown in Data S6. The number of variants found in these genes in the current cohort is depicted for two different categories of variants: type A, representing the most deleterious variants with large effect size, and type D, as this is the category in which the largest overrepresentation of variants in epilepsy genes was found in patients with extreme phenotypes.

## 4 | DISCUSSION

Despite many efforts, we are still not able to fully explain variable phenotypes caused by similar pathogenic *SCN1A* mutations. More insight in modifying factors is essential for understanding genotype–phenotype relations and for accurate counselling of patients. Besides factors such as mosaicism, variants in regulatory regions, and clinical management, variants in modifier genes are suggested to modify phenotypes. We hypothesized that phenotypes of both severely and

mildly affected patients are influenced by modifier genes, as both are on the most extreme ends of the disease spectrum. However, different hypotheses are possible as to which kinds of variants can modify these phenotypes: rare variants with large effects, or multiple more common variants with smaller effects. Previously, rare and/or pathogenic variants in genes involved in neuronal excitability and other known epilepsy genes were suggested to be modifiers (Calhoun et al., 2017; Gaily et al., 2013; Hammer et al., 2017; Hawkins & Kearney, 2012, 2016; Hawkins et al., 2011; Martin et al., 2007; Miller et al., 2014; Ohmori et al., 2008, 2013; Singh et al., 2009). Although such variants may have large effects on phenotypes (a second hit in severely affected patients, or a compensating variant in mildly affected patients), they are unlikely to be present in all patients with extreme phenotypes: none of these single modifier genes has been shown to be clinically relevant in a large patient group (Hammer et al., 2017). Another possibility is the presence of (multiple) more common variants in modifier genes, that each have a smaller effect, but may simultaneously tip the balance over to a milder or more severe phenotype. Especially protective variants in mildly affected patients may not be rare, as they are not necessarily subject to negative selection.



**TABLE 3** Distribution of variants in the epilepsy genes between groups of patients, for different categories of genes

	Group of patients	Type A variants (CADD <sup>a</sup> >20/ MAF <sup>b</sup> <0.01)	Type B variants (CADD >10/ MAF <0.01)	Type C variants (all CADD/ MAF <0.01)	Type D variants (all CADD/ MAF <0.05)	Type E variants (all CADD/ MAF <0.10)
% of variant alleles (based on total number of alleles per group)	Mild ( <i>n</i> = 10)	1.79	2.33	2.76	4.38	5.70
	Severe ( <i>n</i> = 22)	2.04	2.33	2.76	4.52	5.81
	Mild + severe ( <i>n</i> = 32)	1.96	2.33	2.76	4.47	5.78
	Intermediate ( <i>n</i> = 18)	1.95	2.06	2.35	4.03	5.57
<i>p</i> -values Fishers' exact test	Mild versus severe	.456	1	.997	.73	.812
	Mild versus intermediate	.663	.379	.195	.334	.765
	Severe versus intermediate	.783	.294	.112	.109	.491
	Mild + severe versus intermediate	.998	.259	.089	.113	.727

<sup>a</sup>PHRED-scaled CADD (combined annotation dependent depletion). A score of >20 represents the top 1% deleterious substitutions in the human genome.

<sup>b</sup>Minor allele frequency; only variants with a frequency below this threshold in both the exomes and genomes in the gnomAD database are included.

To investigate both of the above described mechanisms, we explored in which categories of variants the largest excess was present in patients with extreme phenotypes, using variant data from the ExAC database. Although exact frequencies of variants may differ between the ExAC cohort and our own, we expected the ratios of variants in epilepsy genes versus unrelated genes to remain similar, which was confirmed when assessing the ratios observed in the different control genes (Figure 1c). We observed the largest overrepresentation of epilepsy gene variants in the MAF <0.05 category

(type D variants; Figure 1). This indicates that relatively common variants in epilepsy genes, which would not necessarily be classified as pathogenic, may have a large influence on phenotypes. These results are in line with recent findings (Niemi et al., 2018). Although we defined phenotype severity by cognitive capacities, no significant overrepresentation of variants in ID genes was observed, indicating that their role as modifier genes is limited. This implies that much of the cognitive phenotypic variability is driven by differences in seizure susceptibility, which argues for classifying Dravet

**TABLE 4** Top 5 genes with an overrepresentation of variants in mild or severe patients, per category of variants

	Type A variants (CADD <sup>a</sup> >20/ MAF <sup>b</sup> <0.01)	Type B variants (CADD >10/ MAF <0.01)	Type C variants (all CADD/ MAF <0.01)	Type D variants (all CADD/MAF <0.05)	Type E variants (all CADD/ MAF <0.10)
Genes with an excess of variants in mild patients (gene name [ <i>p</i> -value]) <sup>c</sup>	<i>SCN10A</i> (.027)	<i>SLC6A8</i> (.013)	<i>EFHC1</i> (.002)	<i>MOCS2</i> (.003)	<i>MOCS2</i> (.003)
	<i>ACTL6B</i> (.094)	<i>SCN10A</i> (.03)	<i>SCN10A</i> (.01)	<i>KCNH1</i> (.008)	<i>KCNH1</i> (.008)
	<i>COL3A1</i> (.094)	<i>RAI1</i> (.087)	<i>SLC6A8</i> (.013)	<i>EFHC1</i> (.01)	<i>DRD4</i> (.009)
	<i>DEPDC5</i> (.094)	<i>ACTL6B</i> (.094)	<i>DSC2</i> (.027)	<i>SLC6A8</i> (.013)	<i>SLC6A8</i> (.013)
	<i>KPNA7</i> (.094)	<i>COL3A1</i> (.094)	<i>RAI1</i> (.087)	<i>MYT1</i> (.027)	<i>CTSD</i> (.026)
Genes with an excess of variants in severe patients (gene name [ <i>p</i> -value])	<i>GPR98</i> (.049)	<i>GPR98</i> (.201)	<i>GPR98</i> (.193)	<i>RYR2</i> (.025)	<i>RYR2</i> (.013)
	<i>RYR2</i> (.419)	<i>CUX1</i> (.314)	<i>ANKRD11</i> (.3)	<i>CUX1</i> (.049)	<i>CUX1</i> (.049)
	<i>ITPR1</i> (.564)	<i>RYR2</i> (.417)	<i>ANK2</i> (.314)	<i>KCNB1</i> (.088)	<i>KCNB1</i> (.088)
		<i>SIK1</i> (.564)	<i>CUX1</i> (.314)	<i>ANK2</i> (.094)	<i>AKAP9</i> (.094)
		<i>TSCI</i> (.564)	<i>RYR2</i> (.417)	<i>AKAP9</i> (.094)	<i>ANK2</i> (.094)

<sup>a</sup>PHRED-scaled CADD (combined annotation dependent depletion). A score of >20 represents the top 1% deleterious substitutions in the human genome.

<sup>b</sup>Minor allele frequency; only variants with a frequency below this threshold in both the exomes and genomes in the gnomAD database are included.

<sup>c</sup>*p*-values are based on Fishers' exact test.

**TABLE 5** Rare and predicted deleterious variants present in only relatively mildly or severely affected members of families with varying *SCN1A*-related phenotypes

	Family <sup>a</sup>	Gene	Established epilepsy gene <sup>b</sup>	Variant	MAF <sup>c</sup>	CADD-phred score <sup>d</sup>
Relatively severely affected family members	1	<i>PRKACA</i>		c.452T > Clp.Ile151Thr (missense)	0	26.1
	1	<i>ITPR1</i>		c.1435G > Alp.Val479Ile (missense)	0.0047	22.2
	2	<i>GPR98</i>		c.3151G > Tlp.Asp1051Tyr (missense)	0.0021	26
	3	<i>DRD4</i>		c.1016G > Alp.Gly339Asp (missense)	0.0015	25
	3	<i>DEPDC5</i>	Yes	c.3551T > Alp.Leu1184Gln (missense)	0	32
	3	<i>DEPDC5</i>	Yes	c.3434C > Tlp.Ser1145Phe (missense)	0.00007785	21.6
	3	<i>CREB5</i>		c.685C > Alp.His229Asn (missense)	0.0003	25.5
	3	<i>DSG2</i>		c.166G > Alp.Val56Met (missense)	0.0019	27
	4	<i>ATF2</i>		c.977C > Tlp.Pro326Leu (missense & splice region)	0.0007	23.5
	5	<i>GNAS</i>		c.1648G > Alp.Ala550Thr (missense)	0.00002394	24.3
	5	<i>OGDHL</i>		c.2201T > Clp.Phe734Ser (missense)	0.0073	32
	6 (proband and father)	<i>CREB3</i>		c.359T > Clp.Leu120Pro (missense)	0.0003	23.6
	6 (proband and father)	<i>SNTA1</i>		c.566C > Tlp.Ser189Leu (missense)	0.0002	23.6
	6 (proband and father)	<i>GJA9</i>		c.22G > Alp.Gly8Arg (missense)	0.0017	29.3
	6 (proband and father)	<i>LGI2</i>		c.194C > Tlp.Ser65Phe (missense)	0.000005	28.4
	6 (only proband)	<i>TSC2</i>	Yes	c.275A > Tlp.Glu92Val (missense)	0.001	25.9
	6 (only proband)	<i>RAI1</i>	Yes	c.725C > Tlp.Pro242Leu (missense)	0.003	24.3
	6 (only proband)	<i>GABRA3</i>	Yes	c.766C > Tlp.Arg256Trp (missense)	0	29.3
Relatively mildly affected family members	1	<i>KCNB1</i>	Yes	c.2266A > Clp.Ile756Leu (missense)	0	23.3
	1	<i>PKP2</i>		c.76G > Alp.Asp26Asn (missense)	0.0084	33
	1	<i>GPR98</i>		c.9650C > Tlp.Ala3217Val (missense)	0.0091	23.6
	1	<i>SHANK3</i>	Yes	c.1379_1382delGAATlp.Arg460fs (frameshift)	0.0021	25.4
	1	<i>SYNGAP1</i>	Yes	c.3982_3983insCCCCCGlp.Arg1328fs (frameshift)	0	34
	1	<i>DNM3</i>		c.2171G > Alp.Arg724His (missense)	0.0059	23.3
	1	<i>GABRA6</i>		c.805G > Alp.Val269Ile (missens)	0.0025	28
	2 (father and brother of proband)	<i>RYR2</i>		c.4451A > Glp.Tyr1484Cys (missense)	0.000008183	25.5
	3	<i>PRRT2</i>	Yes	c.647C > Alp.Pro216His (missense)	0.0005	26.2
	3	<i>CHD5</i>		c.5074G > Tlp.Gly1692Trp (missense)	0.0002	34
	4	<i>SLC19A3</i>	Yes	c.388G > Alp.Val130Met (missense)	0.000004062	21
	4	<i>SZT2</i>	Yes	c.8384C > Glp.Thr2795Arg (missense)	0.000008129	27.5

(Continues)

TABLE 5 Continued

Family <sup>a</sup>	Gene	Established epilepsy gene <sup>b</sup>	Variant	MAF <sup>c</sup>	CADD-phred score <sup>d</sup>
6	<i>JUP</i>		c.1165C > Tlp.Arg389* (stop)	0.00003253	37
6	<i>GOSR2</i>	Yes	c.509A > Glp.Asn170Ser (missense)	0.0003	23.3
6	<i>DOCK3</i>		c.5446G > Alp.Val1816Met (missense)	0	22.5
6	<i>SLC6A1</i>	Yes	c.1243C > Alp.Leu415Ile (missense)	0.0025	20.1

\* indicates a stop mutation, as per the HGVS nomenclature guidelines that Molecular Genetics & Genomic Medicine requires to be used.

<sup>a</sup>Members from families 1–6, as described in Data S2. The upper part of the table represents the patients who are relatively severely affected, compared to their other family members; the lower part of the table represents the participants who are relatively mildly affected, compared to their other family members.

<sup>b</sup>Genes were considered established epilepsy genes when present in the diagnostic epilepsy gene panel of the University Medical Center Utrecht.

<sup>c</sup>Highest frequency of the variant observed in both exomes and genomes in the gnomAD database (Exome Aggregation Consortium et al., 2015).

<sup>d</sup>Combined annotation dependent depletion (Kircher et al., 2014): numbers represent PHRED-scaled CADD scores. CADD scores of >20 represent the top 1% deleterious substitutions in the human genome.

syndrome as an epileptic encephalopathy: a syndrome in which epileptiform activities contribute to a progressive cognitive dysfunction. However, genes that are associated with both epilepsy and ID were excluded from the ID gene set. We cannot exclude the possibility that variants in these genes are the most important modifiers of cognitive outcomes. These modifying effects may be caused by either changes in seizure susceptibility, or by a direct effect of these gene variants on cognitive functioning. Unfortunately, we did not have data on seizure severity at the age of six years old. Future prospective studies should include such data to further elucidate this relation. Similar outcomes were observed for comparisons to each control set, except for set 2; this may be due to the much smaller number of genes in this control set.

A surplus of variants in epilepsy genes was observed in patients with extreme phenotypes, but not for intermediate patients, indicating that indeed both the phenotypes of severe as well as mild patients may be under the influence of modifier genes. It is worth noting that our main outcome was cognition at the age of six years old, so a rapid or slow decline in cognition in the first years of disease. This measure may not necessarily completely correspond to long term cognitive outcomes.

The findings above suggest that it is difficult to draw conclusions from testing individual patients for variants in modifier genes: it is hard to prove whether a relatively common variant will have a substantial effect, and if so, what this effect will be, since variants are found in both mild and severely affected patients. For rare, pathogenic variants this may be easier. Although no significant excess of these variants was observed in mild or severe patients, they may still be present in several patients: only one extra variant with a large effect size may be necessary to drastically change outcomes, which is difficult to statistically detect. We therefore also report the type A variants (CADD >20, MAF <0.01) that were detected in the most extremely mild and severe patients, and those that were only present in either the mild or the severe members of

affected families, in a descriptive way. Studying families in which the same *SCN1A* variant leads to variable phenotypes has several advantages: not only is the primary influence of different *SCN1A* variants themselves removed from analyses, it also means that variants that are shared between both severe and mild family members can be excluded to have significant effects.

Statistically proving the modifying effects of single genes or even specific variants remains difficult; there are only small numbers of patients with extreme phenotypes and *SCN1A* mutations of which the effect can reliably be predicted, which consequently leads to low detection power. Our study may suffer from this. Furthermore, the variety of different possible modifier genes that may act simultaneously makes it difficult to attribute effects to specific variants. In addition, since some genes can carry both LoF and gain of function (GoF) variants, and also variants that cause no relevant effect at all, they may be incorrectly dismissed as modifier genes when variants are present in both severe, mild, and intermediate patients. Functional testing is required to conclusively prove or disprove any modifying effects of single variants, which is not feasible for all variants detected in this study. However, by presenting the most significant genes in each category of variants (Table 4) and variants that are likely to have the largest effects (Tables 5 and 6), we provide data for future reference. Combined with data from future studies similar to ours, trends in the cumulative data may be detected and groups of patients with similar genotype-phenotype correlations may be assembled for further research.

Despite a lack of statistical significance, some interesting results were observed in our study in relation to previous literature: a predicted deleterious *SCN8A* variant was detected in an extremely mild patient. *SCN8A* has previously been implicated to ameliorate *SCN1A* phenotypes by restoring normal seizure thresholds (Hawkins et al., 2011; Martin et al., 2007). Furthermore, several severe patients carried variants in genes that were previously described to worsen phenotypes

**TABLE 6** Predicted most deleterious, rare variants in extremely mild and extremely severe patients

	Established epilepsy genes <sup>a</sup>				Candidate modifier genes <sup>a</sup>				
	Patient (IQ at the age of 6) <sup>c</sup>	Gene	Variant	CADD <sup>d</sup>	MAF <sup>e</sup>	Gene	Variant	CADD <sup>d</sup>	MAF <sup>e</sup>
Mild patients <sup>b</sup>									
1 (89.1)	<i>SCN8A</i>	c.1925C > T p.Thr642Met	27.5	0	<i>EFHC1</i>	c.661C > T p.Arg221Cys	33	0.0009	
2 (72.8)	<i>MOC52</i>	c.367C > T p.His123Tyr	31	0.0036	<i>AIFM3</i>	c.496C > T p.Arg166Trp	34	0.0001	
3 (91.4)	<i>SCN9A</i>	c.3770A > G p.Asn1257Ser	23.9	0.0044	<i>ARNTL</i>	c.1376C > T p.Ser459Phe	29.3	0.00004	
Both 4 and 5 (brothers) (76.6 and 98.2)	<i>ACTL6B</i>	c.496G > A p.Val166Met	25.2	0	—	—	—	—	
Both 6 and 7 (twins) (83.7 and 93.3)	<i>KPNA7</i>	c.1353T > G p.Cys451Trp	26.7	0.00003964	<i>SCN10A</i>	c.4849G > T p.Val1617Phe	28.8	0.00007	
8 (100)	<i>RAI1</i>	c.5036C > T p.Ala1679Val	22.2	0.0006	—	—	—	—	
9 (75.7)	<i>NPRL3</i>	c.1123C > T p.Arg375Cys	24.7	0.00003229	<i>AIFM3</i>	c.1123C > T p.Arg375Cys	32	0.0044	
10 (81.9)	<i>KMT2A</i>	c.4972C > G p.Arg1658Gly	29.6	0.0001	<i>CACNA1H</i>	c.2470G > T p.Ala824Ser	22.8	0.00001	
Extremely severe patients <sup>f</sup>									
1 (18.9)	<i>KDM5C</i>	c.203G > A p.Arg68Gln	34	0.000005615	<i>PLCB2</i>	c.2716G > A p.Glu906Lys	28.8	0.0028	
2 (22.7)	<i>TSC2</i>	c.5383C > T p.Arg1795Cys	31	0.0015	<i>PHTF1</i>	c.1779G > T p.Trp593Cys	34	0	
3 (25.4)	<i>KCNQ3</i>	c.1706A > G p.Asp569Gly	29.4	0.00002031	<i>BCAN</i>	c.2117G > T p.Arg706Leu	32	0.0026	
4 (25.4)	<i>NBEA</i>	c.8350G > T p.Val2784Phe	32	0.0025	<i>CACNA1H</i>	c.6048 + 2_6048+5delTAGG	35	0.00003	
5 (26.7)	<i>BRAT1</i>	c.2353C > T p.Arg785Trp	24.4	0.0029	<i>GJA9</i>	c.22G > A p.Gly8Arg	29.3	0.0017	
6 (27.6)	—	—	—	—	<i>STXBP5L</i>	c.1135G > A p.Val379Met	28.6	0.0054	
7 (28.3)	<i>SLC6A5</i>	c.1735A > G p.Met579Val	25.1	0.0000488	<i>KCNH8</i>	c.1414A > G p.Ile472Val	26.8	0.0001	

<sup>a</sup>A distinction was made between established monogenic epilepsy genes (when present in the diagnostic epilepsy gene panel of the University Medical Center Utrecht) and candidate genes (all other genes in the epilepsy group).

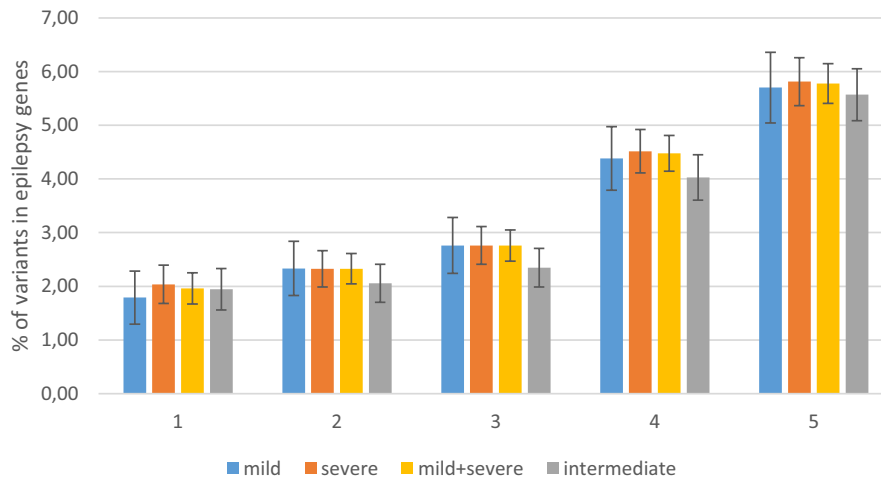
<sup>b</sup>Patients with an interpolated IQ at the age of six years old > 70 (all patients with an *SCN1A* variant that is predicted to cause complete LoF, or a variant that has been previously described in Dravet syndrome patients).

<sup>c</sup>IQ- and developmental assessment scores, conducted at different ages, were interpolated by linear regression, to obtain approximate scores at age 6 years of age as previously described (de Lange, Gunning, et al., 2018).

<sup>d</sup>Combined annotation dependent depletion (Kircher et al., 2014); numbers represent PHRED-scaled CADD scores (a score of > 20 represents the top 1% deleterious substitutions in the human genome).

<sup>e</sup>Highest frequency of the variant observed in both exomes and genomes in the gnomAD database (Exome Aggregation Consortium et al., 2015).

<sup>f</sup>Patients with an interpolated IQ at the age of six years old < 30.



**FIGURE 2** Distribution of variants in epilepsy genes among different groups of patients. Distribution of variants in epilepsy genes among different groups of patients, depicted for different types of variants (A–E). 1: (type A variants)  $CADD > 20/MAF < 0.01$ ; 2: (type B variants)  $CADD > 10/MAF < 0.01$ ; 3: (type C variants) all  $CADD/MAF < 0.01$ ; 4: (type D variants) all  $CADD/MAF < 0.05$ ; 5: (type E variants) all  $CADD/MAF < 0.10$ . Percentages of variants relative to the total number of alleles per group are shown

(*POLG*, *SCN2A*, *CACNA1A*, and *CACNA1G*), strengthening those associations (Calhoun et al., 2017; Gaily et al., 2013; Hawkins et al., 2011; Ohmori et al., 2013). *GPR98*, a gene implicated in myoclonic epilepsy (Myers et al., 2018), showed the highest overrepresentation of variants in severe patients in three categories of variants, and *SCN10A*, another sodium channel alpha-subunit gene, was most often implicated in mild patients. One relatively severe patient carried a *GABRA3* variant (family 6); several GABA receptor genes have already been suggested as potential *SCN1A* modifiers (Miller et al., 2014). Inhibition of *DOCK3*, in which a variant was found in a relatively mild patient (family 6), was previously shown to decrease epileptic activity (Li et al., 2016). Interestingly, frameshift variants in *SHANK3* and *SYNGAP1* were detected in a relatively mild patient (family 1). Both genes are associated with severe neurodevelopmental disorders (Carvill et al., 2013; Durand et al., 2007). The presence of these variants in a mildly affected patient may be explained by their location in the genes: the *SHANK3* variant resides in exon 11, which has previously been implicated to be absent from most or all *SHANK3* transcripts (Kolevzon et al., 2011). The *SYNGAP1* variant is at the 3' end of the gene, which may lead to less severe effects. Nevertheless, it remains a possibility that some of the presented variants are sequencing or calling errors, since it was not feasible to confirm all variants by Sanger sequencing. We however do not expect such variants to influence our main results, since similar error rates are to be expected between different groups of patients and categories of variants.

In conclusion, our results indicate that relatively common variants in epilepsy genes, which would not necessarily be classified as pathogenic by themselves, play a large role in modulating phenotypes, in both severely and mildly affected patients. Studies in larger cohorts, combined with functional

assessments, will be necessary to confirm or disprove the modifying effects of the genes implicated in this study. Our results may be a first step towards meaningful testing of modifier gene variants in regular diagnostics for individual patients, to provide a better estimation of disease severity for newly diagnosed patients.

#### ACKNOWLEDGEMENTS

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).

#### CONFLICTS OF INTEREST

None declared.

#### AUTHOR CONTRIBUTION

Conceived and designed the study: EHB, BPCCK; Data collection: IML, FM, RS, ACMS, MJAK, IJN, and FRE; Statistic analysis: IML; Wrote the paper: IML. All authors read and approved the final manuscript.

#### ETHICAL PUBLICATION STATEMENT

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  <https://orcid.org/0000-0003-2107-9224>

#### REFERENCES

Akiyama, M., Kobayashi, K., Yoshinaga, H., & Ohtsuka, Y. (2010). A long-term follow-up study of Dravet syndrome up to adulthood. *Epilepsia*, 51(6), 1043–1052. <https://doi.org/10.1111/j.1528-1167.2009.02466.x>

- Barrett, J., Buxbaum, J., Cutle, D., Daly, M., Devlin, B., Gratten, J., ... Wray, N. (2017). New mutations, old statistical challenges Based. *BioRxiv*.
- Brunklaus, A., Ellis, R., Reavey, E., Forbes, G., & Zuberi, S. (2012). Prognostic, clinical and demographic features in SCN1A mutation-positive Dravet syndrome. *Brain*, *135*(8), 2329–2336. <https://doi.org/10.1093/brain/awr151>
- Calhoun, J., Hawkins, N., Zachwieja, N., & Kearney, J. (2017). Cacna1g is a genetic modifier of epilepsy in a mouse model of Dravet syndrome. *Epilepsia*, *58*(8), 111–115. <https://doi.org/10.1111/epi.13811>
- Carvill, G. L., Heavin, S. B., Yendle, S. C., McMahon, J. M., Brian, O. J., Cook, J., ... Mefford, H. C. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in *CHD2* and *SYNGAP1*. *Nature Genetics*, *45*(7), 825–830. <https://doi.org/10.1038/ng.2646>. Targeted
- Catterall, W. A., Kalume, F., & Oakley, J. C. (2010). Na<sup>v</sup> 1.1 channels and epilepsy. *The Journal of Physiology*, *588*(11), 1849–1859. <https://doi.org/10.1113/jphysiol.2010.187484>
- Ceulemans, B. (2011). Overall management of patients with Dravet syndrome. *Developmental Medicine and Child Neurology*, *53*, 19–23. <https://doi.org/10.1111/j.1469-8749.2011.03968.x>
- Claes, L., Ceulemans, B., Audenaert, D., Smets, K., Löfgren, A., Del-Favero, J., ... De Jonghe, P. (2003). De novo *SCN1A* mutations are a major cause of severe myoclonic epilepsy of infancy. *Human Mutation*, *21*(6), 615–621. <https://doi.org/10.1002/humu.10217>
- de Lange, I. M., Gunning, B., Sonsma, A. C. M., van Gemert, L., van Kempen, M., Verbeek, N. E., ... Brilstra, E. H. (2018). Influence of contraindicated medication use on cognitive outcome in Dravet syndrome and age at first afebrile seizure as a clinical predictor in *SCN1A*-related seizure phenotypes. *Epilepsia*, *59*(6), 1154–1165. <https://doi.org/10.1111/epi.14191>
- de Lange, I. M., Koudijs, M. J., van 't Slot, R., Gunning, B., Sonsma, A. C. M., & van Gemert, L. J. J. M. ... Koeleman, B. P. C. (2018). Mosaicism of de novo pathogenic *SCN1A* variants in epilepsy is a frequent phenomenon that correlates with variable phenotypes. *Epilepsia*, *59*, 690–703. <https://doi.org/10.1111/epi.14021>
- Depienne, C., Trouillard, O., Gourfinkel-An, I., Saint-Martin, C., Bouteiller, D., Graber, D., ... LeGuern, E. (2010). Mechanisms for variable expressivity of inherited *SCN1A* mutations causing Dravet syndrome. *Journal of Medical Genetics*, *47*(6), 404–410. <https://doi.org/10.1136/jmg.2009.074328>
- Dravet, C. (1978). Les epilepsies graves de l'enfant. *Vie Med*, *8*(2), 543–548.
- Durand, C. M., Betancur, C., Boeckers, T. M., Bockmann, J., Chaste, P., Fauchereau, F., ... Bourgeron, T. (2007). Mutations in the gene encoding the synaptic scaffolding protein *SHANK3* are associated with autism spectrum disorders. *Nature Genetics*, *39*(1), 25–27. <https://doi.org/10.1038/ng.1933>
- Emond, M. J., Louie, T., Emerson, J., Zhao, W., Mathias, R. A., Michael, R., ... Bamshad, M. (2013). Exome sequencing of extreme phenotypes identifies *DCTN4* as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nature Genetics*, *44*(8), 886–889. <https://doi.org/10.1038/ng.2344>. Exome
- Ernst, R. F., van Roosmalen, M., de Ligt, J., Boymans, S., Janssen, R., & Nijman, I. J. (2017). UMCUGenetics/IAP: v2.6.1. <https://doi.org/10.5281/ZENODO.1040130>
- Escayg, A., & Goldin, A. (2010). Sodium channel *SCN1A* and epilepsy: Mutations and mechanisms. *Epilepsia*, *51*(9), 1650–1658. <https://doi.org/10.1111/j.1528-1167.2010.02640.x>
- Exome Aggregation Consortium, Lek, M., Karczewski, K., Minikel, E., Samocha, K., Banks, E., ... MacArthur, D. (2015). Analysis of protein-coding genetic variation in 60,706 humans, 1–26. <https://doi.org/10.1101/030338>
- Gaily, E., Anttonen, A.-K., Valanne, L., Liukkonen, E., Träskelin, A.-L., Polvi, A., ... Lehesjoki, A.-E. (2013). Dravet syndrome: New potential genetic modifiers, imaging abnormalities, and ictal findings. *Epilepsia*, *54*(9), 1577–1585. <https://doi.org/10.1111/epi.12256>
- Gennaro, E., Santorelli, F. M., Bertini, E., Buti, D., Gaggero, R., Gobbi, G., ... Zara, F. (2006). Somatic and germline mosaicism in Severe Myoclonic Epilepsy of Infancy. *Biochemical and Biophysical Research Communications*, *341*(2), 489–493. <https://doi.org/10.1016/j.bbrc.2005.12.209>
- Guerrini, R., Cellini, E., Mei, D., Metitieri, T., Petrelli, C., Pucatti, D., ... Zamponi, N. (2010). Variable epilepsy phenotypes associated with a familial intragenic deletion of the *SCN1A* gene. *Epilepsia*, *51*(12), 2474–2477. <https://doi.org/10.1111/j.1528-1167.2010.02790.x>
- Guerrini, R., Dravet, C., Genton, P., Belmonte, A., Kaminska, A., & Dulac, O. (1998). Lamotrigine and seizure aggravation in severe myoclonic epilepsy. *Epilepsia*, *39*(5), 508–512. <https://doi.org/10.1111/j.1528-1157.1998.tb01413.x>
- Guo, T., Chung, J. H., Wang, T., McDonald-McGinn, D. M., Kates, W. R., Hawula, W., ... Morrow, B. E. (2015). Histone modifier genes alter conotruncal heart phenotypes in 22q11.2 deletion syndrome. *American Journal of Human Genetics*, *97*(6), 869–877. <https://doi.org/10.1016/j.ajhg.2015.10.013>
- Hammer, M. F., Ishii, A., Johnstone, L., Tchourbanov, A., Lau, B., Sprissler, R., ... Hirose, S. (2017). Rare variants of small effect size in neuronal excitability genes influence clinical outcome in Japanese cases of *SCN1A* truncation-positive Dravet syndrome. *PLoS ONE*, *12*(7), 1–16. <https://doi.org/10.1371/journal.pone.0180485>
- Harkin, L. A., McMahon, J. M., Iona, X., Dibbens, L., Pelekanos, J. T., Zuberi, S. M., ... Scheffer, I. E. (2007). The spectrum of *SCN1A*-related infantile epileptic encephalopathies. *Brain*, *130*(3), 843–852. <https://doi.org/10.1093/brain/awm002>
- Hawkins, N., & Kearney, J. (2012). Confirmation of an epilepsy modifier locus on mouse chromosome 11 and candidate gene analysis by RNA-seq. *Genes Brain Behavior*, *11*(4), 452–460. <https://doi.org/10.1111/j.1601-183X.2012.00790.x>
- Hawkins, N. A., & Kearney, J. A. (2016). Hlf is a genetic modifier of epilepsy caused by voltage-gated sodium channel mutations. *Epilepsy Research*, *119*, 20–23. <https://doi.org/10.1016/j.eplepsyres.2015.11.016>
- Hawkins, N. A., Martin, M., Frankel, W., Kearney, J., & Escayg, A. (2011). Neuronal voltage-gated ion channels are genetic modifiers of generalized epilepsy with febrile seizures plus. *Neurobiology of Disease*, *41*(3), 655–660. <https://doi.org/10.1016/j.nbd.2010.11.016>
- Jansen, F. E., Sadleir, L. G., Harkin, L. A., Vadlamudi, L., McMahon, J. M., Mulley, J. C., ... Berkovic, S. F. (2006). Severe myoclonic epilepsy of infancy (Dravet syndrome): Recognition and diagnosis in adults. *Neurology*, *67*(12), 2224–2226. <https://doi.org/10.1212/01.wnl.0000249312.73155.7d>
- Kircher, M., Witten, D., Jain, P., O'Roak, B., Cooper, G., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315. <https://doi.org/10.1038/ng.2892>. A
- Kolevzon, A., Cai, G., Soorya, L., Takahashi, N., Grodberg, D., Kajiwar, Y., ... Buxbaum, J. D. (2011). Analysis of a purported *SHANK3* mutation in a boy with autism: Clinical impact of rare

- variant research in neurodevelopmental disabilities. *Brain Research*, 1380, 98–105. <https://doi.org/10.1016/j.brainres.2010.11.005>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, J., Mi, X., Chen, L., Jiang, G., Wang, N. A., Zhang, Y., ... Wang, X. (2016). Dock3 participate in epileptogenesis through rac1 pathway in animal models. *Molecular Neurobiology*, 53(4), 2715–2725. <https://doi.org/10.1007/s12035-015-9406-9>
- Long, Y. S., Zhao, Q. H., Su, T., Cai, Y. L., Zeng, Y., Shi, Y. W., ... Liao, W. P. (2008). 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. *Journal of Neuroscience Research*, 86(15), 3375–3381. <https://doi.org/10.1002/jnr.21790>
- Mahoney, K., Moore, S. J., Buckley, D., Alam, M., Parfrey, P., Penney, S., ... Young, T.-L. (2009). Variable neurologic phenotype in a GEFS+ family with a novel mutation in *SCN1A*. *Seizure*, 18(7), 492–497. <https://doi.org/10.1016/j.seizure.2009.04.009>
- Marini, C., Mei, D., Helen Cross, J., & Guerrini, R. (2006). Mosaic *SCN1A* mutation in familial severe myoclonic epilepsy of infancy. *Epilepsia*, 47(10), 1737–1740. <https://doi.org/10.1111/j.1528-1167.2006.00675.x>
- Martin, M. S., Tang, B., Papale, L. A., Yu, F. H., Catterall, W. A., & Escayg, A. (2007). The voltage-gated sodium channel *SCN8A* is a genetic modifier of severe myoclonic epilepsy of infancy. *Human Molecular Genetics*, 16(23), 2892–2899. <https://doi.org/10.1093/hmg/ddm248>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... DePristo, M. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110.20>
- Meng, H., Xu, H. Q., Yu, L., Lin, G. W., He, N., Su, T., ... Liao, W. P. (2015). The *SCN1A* mutation database: Updating information and analysis of the relationships among genotype, functional alteration, and phenotype. *Human Mutation*, 36(6), 573–580. <https://doi.org/10.1002/humu.22782>
- Miller, A., Hawkins, N., McCollom, C., & Kearney, J. (2014). Mapping genetic modifiers of survival in a mouse model of Dravet syndrome. *Genes Brain Behavior*, 13(2), 163–172. <https://doi.org/10.1111/gbb.12099>
- Mulley, J. C., Scheffer, I. E., Petrou, S., Dibbens, L. M., Berkovic, S. F., & Harkin, L. A. (2005). *SCN1A* mutations and epilepsy. *Human Mutation*, 25(6), 535–542. <https://doi.org/10.1002/humu.20178>
- Myers, K. A., Nasioulas, S., Boys, A., McMahon, J. M., Slater, H., Lockhart, P., & Ingrid, S. (2018). ADGRV1 is implicated in myoclonic epilepsy. *Epilepsia*, 59, 381–388. <https://doi.org/10.1111/epi.13980>
- Nabbout, R., Chemaly, N., Chipaux, M., Barcia, G., Bouis, C., Dubouch, C., ... Chiron, C. (2013). Encephalopathy in children with Dravet syndrome is not a pure consequence of epilepsy. *Orphanet Journal of Rare Diseases*, 8(1), 176. <https://doi.org/10.1186/1750-1172-8-176>
- Niemi, M. E. K., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M., ... Barrett, J. C. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*, 562(7726), 268–271. <https://doi.org/10.1038/s41586-018-0566-4>
- Ohmori, I., Ouchida, M., Kobayashi, K., Jitsumori, Y., Mori, A., Michiue, H., ... Matsui, H. (2013). *CACNA1A* variants may modify the epileptic phenotype of Dravet syndrome. *Neurobiology of Disease*, 50(1), 209–217. <https://doi.org/10.1016/j.nbd.2012.10.016>
- Ohmori, I., Ouchida, M., Miki, T., Mimaki, N., Kiyonaka, S., Nishiki, T., ... Matsui, H. (2008). A *CACNB4* mutation shows that altered Cav2.1 function may be a genetic modifier of severe myoclonic epilepsy in infancy. *Neurobiology of Disease*, 32(3), 349–354. <https://doi.org/10.1016/j.nbd.2008.07.017>
- Passamonti, C., Petrelli, C., Mei, D., Foschi, N., Guerrini, R., Provinciali, L., & Zamponi, N. (2015). A novel inherited *SCN1A* mutation associated with different neuropsychological phenotypes: Is there a common core deficit? *Epilepsy and Behavior*, 43, 89–92. <https://doi.org/10.1016/j.yebeh.2014.11.009>
- Pineda-Trujillo, N., Carrizosa, J., Cornejo, W., Arias, W., Franco, C., Cabrera, D., ... Ruíz-Linares, A. (2005). A novel *SCN1A* mutation associated with severe GEFS+ in a large South American pedigree. *Seizure*, 14(2), 123–128. <https://doi.org/10.1016/j.seizure.2004.12.007>
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5), 405–423. <https://doi.org/10.1038/gim.2015.30>
- Rilstone, J. J., Coelho, F. M., Minassian, B. A., & Andrade, D. M. (2012). Dravet syndrome: Seizure control and gait in adults with different *SCN1A* mutations. *Epilepsia*, 53(8), 1421–1428. <https://doi.org/10.1111/j.1528-1167.2012.03583.x>
- Sadleir, L. G., Mountier, E. I., Gill, D., Davis, S., Joshi, C., DeVile, C., ... Scheffer, I. E. (2017). Not all *SCN1A* epileptic encephalopathies are Dravet syndrome. *Neurology*, 89(10), 1035–1042. <https://doi.org/10.1212/WNL.0000000000004331>
- Scheffer, I. E., Zhang, Y. H., Jansen, F. E., & Dibbens, L. (2009). Dravet syndrome or genetic (generalized) epilepsy with febrile seizures plus? *Brain and Development*, 31(5), 394–400. <https://doi.org/10.1016/j.braindev.2009.01.001>
- Singh, N. A., Pappas, C., Dahle, E. J., Claes, L. R. F., Pruess, T. H., De Jonghe, P., ... Leppert, M. F. (2009). A role of *SCN9A* in human epilepsies, as a cause of febrile seizures and as a potential modifier of Dravet syndrome. *PLoS Genetics*, 5(9), 1–12. <https://doi.org/10.1371/journal.pgen.1000649>
- Singh, R., Scheffer, I. E., Crossland, K., & Berkovic, S. (1999). Generalized epilepsy with febrile seizures plus: A common childhood-onset genetic epilepsy syndrome. *Annals of Neurology*, 45(1), 75–81. [https://doi.org/10.1002/1531-8249\(199901\)45:1<75:AID-ART13>3.0.CO;2-W](https://doi.org/10.1002/1531-8249(199901)45:1<75:AID-ART13>3.0.CO;2-W)
- Suls, A., Velizarova, R., Yordanova, I., Deprez, L., Van Dyck, T., Wauters, J., ... De Jonghe, P. (2010). Four generations of epilepsy caused by an inherited microdeletion of the *SCN1A* gene. *Neurology*, 75(1), 72–76. <https://doi.org/10.1212/WNL.0b013e3181e62088>
- Van Der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Levy-Moonshine, A., Jordan, T., ... DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*, 43, 11.10.1–11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>
- Vélez, J. I., Lopera, F., Sepulveda-Falla, D., Patel, H. R., Johar, A. S., Chuah, A., ... Arcos-Burgos, M. (2016). APOE\*E2 allele delays age of onset in PSEN1 E280A Alzheimer's disease. *Molecular Psychiatry*, 21(7), 916–924. <https://doi.org/10.1038/mp.2015.177>

- Yu, F. H., Mantegazza, M., Westenbroek, R. E., Robbins, C. A., Kalume, F., Burton, K. A., ... Catterall, W. A. (2006). Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. *Nature Neuroscience*, 9(9), 1142–1149. <https://doi.org/10.1038/nn1754>
- Zeng, T., Dong, Z.-F., Liu, S.-J., Wan, R.-P., Tang, L.-J., Liu, T., ... Long, Y.-S. (2014). 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. *Human Genetics*, 133(6), 801–811. <https://doi.org/10.1007/s00439-014-1422-8>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** De Lange IM, Mulder F, van 't Slot R, et al. Modifier genes in *SCN1A*-related epilepsy syndromes. *Mol Genet Genomic Med*. 2020;8:e1103. <https://doi.org/10.1002/mgg3.1103>