In order to characterize the genetic architecture of epilepsy in a pediatric population from the Iberian Peninsula (including the Canary Islands), we conducted targeted exome sequencing of 246 patients with infantile-onset seizures with or without neurodevelopmental delay. We detected 107 variants in 48 different genes, which were implicated in neuronal excitability, neurodevelopment, synaptic transmission, and metabolic pathways. In 104 cases (42%) we detected variant(s) that we classified as pathogenic or likely pathogenic. Of the 48 mutated genes, 32 were dominant, 8 recessive and 8 X-linked. Of the patients for whom family studies could be performed and in whom pathogenic variants were identified in dominant or X-linked genes, 82% carried de novo mutations. The involvement of small copy number variations (CNVs) is 9%. The use of progressively updated custom panels with high mean vertical coverage enabled establishment of a definitive diagnosis in a large proportion of cases (42%) and detection of CNVs (even duplications) with high fidelity. In 10.5% of patients we detected associations that are pending confirmation via functional and/or familial studies. Our findings had important consequences for the clinical management of the probands, since a large proportion of the cohort had been clinically misdiagnosed, and their
families were subsequently able to avail of genetic counseling. In some cases, a more appropriate treatment was selected for the patient in question, or an inappropriate treatment discontinued. Our findings suggest the existence of modifier genes that may explain the incomplete penetrance of some epilepsy-related genes. We discuss possible reasons for non-diagnosis and future research directions. Further studies will be required to uncover the roles of structural variants, epimutations, and oligogenic inheritance in epilepsy, thereby providing a more complete molecular picture of this disease. In summary, given the broad phenotypic spectrum of most epilepsy-related genes, efficient genomic tools like the targeted exome sequencing panel described here are essential for early diagnosis and treatment, and should be implemented as first-tier diagnostic tools for children with epilepsy without a clear etiologic basis.

Keywords: epilepsy, genetic diagnosis, neurodevelopmental disorders, de novo mutations, incomplete penetrance, modifier genes

INTRODUCTION

Epilepsy is one of the most common neurological conditions, with a prevalence of ~1%. Etiological diagnosis of epilepsy in children using classical diagnostic tools is often a long and complex process. Moreover, most patients undergo multiple invasive and costly analyses but do not receive conclusive molecular diagnosis.

The emergence of next generation sequencing (NGS)-based tools has helped address the diagnostic challenge posed by epilepsy. These approaches allow genetic analyses to be performed faster, less expensively, and at much higher resolution. This technology has also helped identify many novel genes involved in epilepsy phenotypes.

In recent years several articles have described the use of targeted NGS or whole-exome sequencing (WES) to diagnose epileptic patients, with varying diagnostic yields. The results of these studies have helped characterize the molecular landscape of epilepsy, and indicate that infantile onset epilepsy is caused by a wide spectrum of genes, most of which are de novo variants in dominant genes (Epi4K Consortium et al., 2013; EuroEPINOMICS-RES Consortium, 2014; Allen et al., 2016), although a minority of recessive genes are also implicated. Mosaic mutations and copy number variants (CNVs) are other important sources of mutations in epileptic disorders (Gennaro et al., 2006; Vadlamudi et al., 2010; de Lange et al., 2017).

CNVs can cause epilepsy through the deletion or duplication of known epilepsy-related genes. Moreover, the presence of epilepsy-related genes in genomic regions affected by deletions can give rise to complex syndromic conditions (Wang et al., 2008; Dibbens et al., 2009; Helbig et al., 2009; de Kovel et al., 2010; Heinzlen et al., 2010; Dejanovic et al., 2014; Epilepsy Phenome/Genome Project Epi4K Consortium, 2015; Mefford, 2015; Borlot et al., 2017; Tsuchida et al., 2018).

Recently Oates et al. (2018) demonstrated that early gene panel screening of newborns with epilepsy could reduce the cost of subsequent tests from $9,362 to $2,838 and the median diagnostic delay from 3.43 years to 21 days. Gene panel testing for epilepsy has a high diagnostic yield among children with onset before 2 years of age, and an appreciable clinical, social, and financial impact.

In cases of children with epilepsy, establishing a molecular diagnosis in a clinical setting is essential to (1) establish the risk of recurrence in subsequent pregnancies; (2) end the all-too-common diagnostic odyssey endured by parents of undiagnosed children; (3) avoid unnecessary analyses and treatments; (4) provide an accurate prognosis; (5) optimize management; (6) provide a prenatal or preimplantation diagnosis for future pregnancies; (7) identify, at least in some cases, specific appropriate therapies and enable the application of precision medicine as targeted therapeutics emerge.

The importance of genomic analyses in epilepsy has been previously discussed in depth (Lemke et al., 2012; Mercimek-Mahmutoglu et al., 2015; Mei et al., 2017; Weber et al., 2017). In this study, we present a comprehensive description of the molecular signature of pediatric epilepsy patients of Iberian origin and discuss the results obtained.

MATERIALS AND METHODS

This study was approved by the Ethics Committee of the Hospital Santiago de Compostela (Spain). A total of 246 patients were recruited over a 5-year period from different neurologic units in Spain and Portugal, applying the following inclusion criteria: patients with any condition in which seizures are either the sole clinical expression or are part of more severe neurodevelopmental disorder.

Panel Design

Successive Epi-panels were constructed using OMIM-registered genes for which the associated clinical picture included seizures. In addition, we conducted a thorough review of the most recent scientific publications to identify genes associated with any form of epilepsy, with or without neurodevelopmental delay. We included genes for which only one published report was found, even if the involvement was suspected but not demonstrated. The selection of genes included the panel was regularly evaluated and updated throughout the study period to ensure the inclusion of
all novel genes reported in the literature. Each Epi-panel included all exons and at least 25 base pairs of the flanking intronic sequence of the selected genes. A list of the genes included in the consecutive versions of each panel can be provided upon request. The first panel consisted of 88 genes and final version 274 genes. This increase reflects the high rate of discovery of epilepsy-associated genes in recent years. Because the analyses conducted at the beginning of this project were done so using comparatively incomplete panels, all undiagnosed patients who provided consent were re-analyzed at the end of the study using the most up-to-date panel in order to rule out as many false negatives as possible.

**Targeted Next Generation Sequencing**

Enrichment was performed using in-solution hybridization technology (Sure Select XT; Agilent Technologies, Santa Clara, California) and subsequent sequencing using MiSeq or NextSeq platforms (Illumina, Santa Clara, California), as previously described (Fernández-Marmiesse et al., 2018). Image analysis and processing of the fluorescence intensities in sequences (“base calling”) was performed with Real Time Analysis (RTA) software v.1.8.70 (Illumina), and FastQC v0.11.8 program was used for data quality control (Babraham Bioinformatics-FastQC A Quality Control Tool for High Throughput Sequence Data, 2018). Reads were aligned to the reference genome GRCh37 with BWA v0.7.17 (Li and Durbin, 2009), and BEDTools 2.27.1 (Quinlan and Hall, 2010) and Picard v2.18.14 (Picard Tools-By Broad Institute, 2018) were used for intermediate steps. VarScan v.2.4.2 (Koboldt et al., 2009), SAMtools v1.9 (Li et al., 2009), GATK v4.0.10 (McKenna et al., 2010), Pindel v0.2.5 (Ye et al., 2009), and Platypus v0.8.1 (Rimmer et al., 2014) software were used for variant detection, SnpEff v4.3 (Cingolani et al., 2012) for variant annotation, and PattRec for CNV detection (Roca et al., 2019).

Varians that passed the quality control step were prioritized according to their minor allele frequencies (MAF < 0.01) in the following databases: 1000G, Exome Aggregation Consortium (ExAC), the Exome Variant Server (EV), the Genome Aggregation Database (gnomAD), and our in-house population database (onwards IberDB). Z-score was used to evaluate the conservation of genes which harbor rare/low-frequency variants (Roca et al., 2018). Possible pathogenicity of the missense variants detected was assessed using the in silico tools CONDEL (González-Pérez and López-Bigas, 2011), GERP++ (Davydov et al., 2010), and Human Splicing Finder (HSF 3.0) (Desmet et al., 2009). Variants were classified as “pathogenic” or “of uncertain significance” in accordance with the guidelines of the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015).

**Statistical Analyses**

The variants used for all calculations (unless otherwise stated) were those recorded in 2,504 individuals enrolled in Phase 3 of the 1000 Genomes Project (1000G) and in 125,748 individuals whose data are included in v2.1.1 of the Genome Aggregation Database (gnomAD). Variants classified as “pathogenic” or “likely pathogenic” in the ClinVar database were filtered out.

The mutation tolerance of each epilepsy-associated gene was quantified by determining the corresponding z-score (Roca et al., 2018). This score was calculated by first regression of the total number of missense variants against the total number of missense and synonymous variants observed for a given gene. The z-score is the corresponding studentized residual of the regression. Genes with a negative z-score have fewer missense variants than expected based on the expected mutation burden and are likely to be less tolerant of functional mutations.

To prioritize the rare missense variants and assign them a higher or lower weight in terms of their contribution to the patient’s phenotype we examined gene tolerance to missense variation (z-score) for all epilepsy-associated genes in Iberian patients. The results are shown in Figure 1.

**RESULTS**

The median (interquartile range) coverage of the samples analyzed with successive versions of the Epi-panel was 392X (309X–461X). The overall target coverage at 10X of the genes included in the successive versions of the Epi-panel ranged from 97.3 to 99.85%. However, mean target coverage increased over time, reaching 98.1–99.7% for the last 3 versions, indicating progressive optimization of the test. The regions missed were almost identical across the different samples, and shared a high GC content, repeat elements, or homology with other parts of the genome.

To prioritize the rare missense variants and assign them a higher or lower weight in terms of their contribution to the patient’s phenotype we examined gene tolerance to missense variation (z-score) for all epilepsy-associated genes in Iberian patients. The results are shown in Figure 1.

Figures 2A,B show the statistical comparison of GERP and CONDEL scores for variants detected in our cohort with those of other missense variants found in the same genes in controls (extracted from 1000G and gnomAD databases). It is important to highlight that, of the missense variants found in the genes in the databases, MAF values were <0.01 in 94.88% of cases, <0.005 in 92.57% of cases, and <0.001 in 85% of cases.

We detected variants in 52 different genes. Of these, 17 were involved in neuronal excitability, 30 in neurodevelopment and synaptic transmission, and 5 in metabolic pathways. Variants found in our cohort are shown in Tables 1A, B. For 32 genes (65%) variants were detected in 1 single patient, and for the following 7 genes variants were detected in more than 2 patients each: SCN1A (16 patients), KCNQ2 (13 patients), CDKL5 (7 patients), SCN2A (4 patients), MEC2P2 (4 patients), GNAO1 (4 patients), and FOXG1 (3 patients). In 42% of cases the mutation identified was considered pathogenic or likely pathogenic and the main cause of patient’s phenotype (Table 1A). In 10.5% of cases prioritized variants were classified as potentially involved (Table 1B), but further evidence will be required to confirm this association. It is important to note that in 9 cases variants could not be classified as pathogenic or non-pathogenic owing to a lack of variant analysis data from the progenitors required to confirm
or rule out \textit{de novo} events. Figure 3 shows the proportions of different types of inheritance and different types of variants.

The clinical features of the patients carrying the variants that appear in Tables 1A, B are listed in the Supplementary Table 1. The locations of the variants in some of the genes are shown in Figures 4, 5. The position of 2 CNVs in FOXG1 and GHFN relative to those previously reported is shown in Figure 6.

**DISCUSSION**

**Utility of Z-Score and GERP and CONDEL Scores for Variant Prioritization**

We found that z-score and GERP and CONDEL scores can be combined to provide an effective means of assessing the potential pathogenicity of variants detected in epilepsy patients.

A negative z-score implies a marked sensitivity to missense variation, and indicates that missense variations in these genes are much more likely to affect the function of the encoded proteins. The genes most sensitive to missense variation were \textit{KCNT1}, \textit{TSC2}, and \textit{MTOR} (Figure 1). However, the variants detected in these 3 genes in our cohort were classified as "of uncertain significance," as explained below. The next most sensitive genes were 3 ion-channel-encoding genes (\textit{SCN2A}, \textit{SCN8A}, \textit{KCNQ2}), mutations in which are found in a significant percentage of patients with neurodevelopmental problems. A higher z-score indicates a lower sensitivity to missense variation. For example, \textit{WWOX} (z-score > 3) is a recessive gene with low sensitivity to missense changes. In fact, most disease-associated mutations described for this gene are truncating mutations (Shaukat et al., 2018). \textit{SCN10A} and \textit{RFT1} are recessive genes and therefore had the highest z-scores.

As shown in Figures 2A, B, GERP and CONDEL scores constitute a useful means of discriminating between innocuous and deleterious variants. For some highly conserved genes (\textit{MTOR}, \textit{SCN2A}, \textit{SCN8A}, \textit{KCNA2}, \textit{SCN1A}, \textit{DNM1}, \textit{KCNCl}, \textit{GNAO1}, \textit{STXB1}, \textit{GRIN2A}, \textit{HCNI}, \textit{NPRl2}), we observed very high mean GERP values and low variability (Figure 2A). Any missense variation in these genes can be considered potentially deleterious. For other genes (\textit{KCNQ2}, \textit{SLC13A5}, \textit{KCNT1}, \textit{PROSC}) the range of GERP scores was much wider, implying that some nucleotides are much more sensitive to variation than others, and therefore that in addition to CONDEL score the location of the variant is important.

As shown in Figure 2B, CONDEL score allowed effective discrimination between deleterious and innocuous variants for the 2 most mutated genes in our cohort: \textit{KCNQ2} and \textit{SCN1A} ($p = 2.8 \times 10^{-5}$ and $1.12 \times 10^{-6}$, respectively; Wilcoxon test). In general, for all variants shown in Table 1A, CONDEL scores were above the upper limit of distribution of the control variants (yellow bars), except for 2 variants in \textit{RFT1} (recessive) and \textit{CDKL5} (X-linked). Interestingly, both these variants had very high GERP scores. This suggests that even though the amino acid change is functionally relatively innocuous, its location may have a significant impact on the protein (e.g., from regulatory point of view), since it is highly conserved throughout evolution.
Variants Classified as Pathogenic and Likely Pathogenic
Most Mutated Genes
We detected 23 variants in SCN1A (Tables 1A, B and Figure 3A), including a long duplication that encompassed the sodium channel cluster (2q24.3) and was confirmed by array comparative genomic hybridization (CGH). Of the 23 SCN1A variants detected 16 were considered highly likely to contribute to the patient’s phenotype (Table 1A). As previously noted in other studies, missense variants in this gene were associated with milder phenotypes, except for those located in the pore region of the channel. We also found that phenotype severity was inversely proportional to familial incidence (Meng et al., 2015).

Six patients carried SCN2A variants, but only 4 were classified as pathogenic or likely pathogenic (Figure 4B). As expected, the 2 SCN2A truncating variants were associated with a late-onset
| P  | S | GENE    | INH | ZIG | VFS | FH | ClinVar | dbSNP ID  | MAF   | Type | ISP | GERP | Ref Seq | Variant |
|----|---|---------|-----|-----|-----|----|---------|-----------|-------|------|-----|------|---------|---------|
| 62 | F | SCN1A   | AD  | HT  | DE NOVO | LP | rs121917945 | 1.2E+00   | 0.76  | 4.14 | 0.66 | 4.39 | NM_00165963.1 c.5694delA; p.R262Q |
| 63 | F | KCNQ2   | AD  | HT  | DE NOVO | P  | rs864321712  | 1.2E-08   | 0.66  | 3.98 | 0.67 | 3.97 | NM_00165963.1 c.2404delA; p.R801Q |
| 64 | F | SCN8A   | AD  | HT  | DE NOVO | LP | rs121918264 | 1.2E-10   | 0.66  | 5.11 | 0.67 | 4.41 | NM_00165963.1 c.4134delA; p.R1382Q |
| 65 | F | KCNT1   | AD  | HT  | TRANS   | -  | rs97976020  | 1.2E-10   | 0.64  | 5.07 | 0.67 | 5.07 | NM_00165963.1 c.2243delA; p.R748Q |
| 66 | F | HCN1    | AD  | HT  | DE NOVO | LP | rs1057519647 | 1.2E+00   | 0.69  | 5.42 | 0.69 | 5.42 | NM_00165963.1 c.6174delA; p.R2059Q |

(Continued)
| P  | S  | GENE     | INH | ZIG | VFS | FH  | ClinVar | dbsNP ID | MAF     | Type | ISP | GEREP | Ref Seq | Variant |
|----|----|----------|-----|-----|-----|-----|---------|----------|---------|------|-----|-------|---------|---------|
| 67 | M  | SLC13A5  | AR  | HT  | TRANS | –   | –       | rs1211773372 | –       | FS   | -    | 3.12  | NM_177550.4 | c.1611delT;p.(IL504fs*23) |
| 1030 | F  | FGFl2   | AD  | HT  | NA    | –   | P       | rs886039903  | –       | MISS | 0.54 | 5.15  | NM_177550.4 | c.1514C>T;p.P505L |
| 37  | F  | PRRT2   | AD  | HT  | NA    | +   | –       |          | CNV    | NC   | 0.00016.10  | g.?(29824300)_ (29826034_?)del |
| 207 | F  | PRRT2   | AD  | HT  | NA    | +   | P       | rs587778771  | 0.007  | FS   | -    | 3.9   | NM_001256442.1 | c.649_650insC;p.(R217Pfs*8) |
| 545 | F  | ARX     | XL  | HT  | DE NOVO | –   | –       |          | MISS   | 0.74 | 3.87  | NM_139058.2 | c.1039T>G;p.(F347V) |
| 385 | M  | MECP2   | XL  | HE  | DE NOVO | –   | –       |          | NON    | 5.67 | 5.15  | NM_001110792.1 | c.1386C>G;p.Y462X |
| 474 | F  | MECP2   | XL  | HT  | NA    | –   | P       | rs886041728  | –       | FS   | -    | 5.11  | NM_001110792.1 | c.1087_1088insC;p.(K363Thr*30) |
| 572 | M  | MECP2   | XL  | HE  | NA    | –   | P/LP    | rs61751443  | –       | MISS | 0.52 | 5.06  | NM_001110792.1 | c.917G>A;p.R306H |
| 965 | F  | MECP2   | XL  | HT  | NA    | –   | –       |          | NON    | 5.8   | 5.45  | NM_001110792.1 | c.1453G>A;p.E485* |
| 420 | F  | FOXG1   | AD  | HT  | NA    | –   | –       |          | NON    | 4.05 | 5.15  | NM_00052494.4 | c.746G>A;p.W256X |
| 427 | M  | FOXG1   | AD  | HT  | DE NOVO | –   | P/LP    | rs398124202 | –       | MISS | 0.68 | 3.99  | NM_00052494.4 | c.256C>T;p.Q86X |
| 872 | M  | FOXG1   | AD  | HT  | DE NOVO | –   | –       |          | MISS   | 0.50 | 5.07  | NM_001037343.1 | c.455G>A;p.C152Y |
| 47  | F  | CDKL5   | XL  | HT  | DE NOVO | –   | –       |          | FS     | -0.27 | 3.94  | NM_001037343.1 | c.2635_2636del;p.(L879Efs*30) |
| 310 | M  | HNRNPU  | AD  | HT  | DE NOVO | –   | –       |          | MISS   | 0.78 | 5.17  | NM_006545.4 | c.724C>T;p.R242C |
| 421 | M  | TBC1D24 | AR  | HO  | TRANS | –   | –       | rs936929276  | –       | MISS | 0.54 | 5.15  | NM_177550.4 | c.341G>A;p.R114H |
| 317 | F  | GNAO1   | AD  | HT  | DE NOVO | –   | P       | rs587770659  | –       | NON  | -    | 6.03  | NM_001037343.1 | c.2413C>T;p.(G805X) |
| 472 | M  | CDKL5   | XL  | HT  | DE NOVO | –   | P/LP    | rs267608433 | –       | FS   | -    | 4.59  | NM_001037343.1 | c.163_166del;p.(L879Efs*30) |
| 473 | F  | CDKL5   | XL  | HT  | DE NOVO | –   | P/LP    | rs267608433 | –       | NON  | -    | 6.03  | NM_001037343.1 | c.2413C>T;p.(G805X) |
| 1127 | F  | CDKL5   | XL  | HT  | NA    | –   | –       |          | MISS   | 0.75 | 5.98  | NM_001037343.1 | c.380A>T;p.H127L |
| 506 | M  | GNAO1   | AD  | HT  | NA    | –   | –       |          | MISS   | 0.52 | 5.07  | NM_001037343.1 | c.649_650insC;p.(R217Pfs*8) |
| 289 | M  | GNAO1   | AD  | HT  | NA    | –   | –       |          | MISS   | 0.50 | 4.27  | NM_031844.2 | c.401_402delAC;p.(D134Gfs*10) |
| 951 | M  | DEPDC5  | AD  | HT  | EN    | –   | –       |          | CNV    | NC   | 0.016991017.1 | c.724C>T;p.R242C |
| 164 | M  | SLC9A6  | XL  | HE  | DE NOVO | –   | –       |          | NON    | 4.88  | 4.88  | NM_013206.0 | c.124C>T;p.G42* |
| 1238 | F  | WWOX    | AR  | HT  | TRANS | –   | –       |          | NON    | -    | -    | NM_01425371.1 | c.1601C>G;p.S534* |
| 996 | M  | NPR2    | AD  | HT  | PAT   | +   | –       |          | MISS   | 0.78 | 5.17  | NM_006545.4 | c.949G>A;p.S317R |
| 1213 | M  | NPR2    | AD  | HT  | NA    | +   | –       |          | NON    | 5.15  | 5.15  | NM_001077350.2 | c.138delC;p.W46* |
| 951 | M  | DEPDC5  | AD  | HT  | EN    | –   | –       |          | CNV    | NC   | 0.034067.1 | c.1081_1082del;p.(G42_43del) |
| 782 | M  | SRGAP2  | AD  | HT  | NA    | –   | –       |          | CNV    | NC   | 0.034067.1 | c.1081_1082del;p.(G42_43del) |
| 519 | M  | RBFOX1  | AD  | HT  | PAT   | –   | –       |          | CNV    | NC   | 0.011881.1 | c.3_4del;p.(A42_43del) |
| 530 | M  | GPHN    | AD  | MAT | –    | –   | –       |          | CNV    | NC   | 0.008751.1 | c.143_144del;p.(G294_295del) |
| 819 | M  | GPHN    | AD  | MAT | –    | –   | –       |          | CNV    | NC   | 0.008751.1 | c.143_144del;p.(G294_295del) |
| 186 | F  | ARHGEF9 | XL  | HT  | DE NOVO | –   | –       |          | MISS   | 0.79 | 5.65  | NM_015185.2 | c.541G>C;p.G181R |

(Continued)
| P | S | GENE | INH | ZIG | VFS | FH | ClinVar | dbSNP ID | MAF | Type | ISP | GERP | Ref Seq | Variant |
|---|---|------|-----|-----|-----|----|---------|----------|-----|------|-----|------|-----------|---------|
| 1126 | M | SLC6A1 | AD | HT | NA | – | – | – | MISS | 0.69 | 4.21 | NM_0003042.3 | c.373G>A; p.V125M |
| 1156 | F | SLC6A1 | AD | HT | DE NOVO | – | – | – | MISS | 0.69 | 5.57 | NM_0003042.3 | c.373G>A; p.V125M |
| 1058 | M | IL1RAPL1 | XL | MOS | DE NOVO | – | – | – | CNV | - | - | NG_000062.1 | c.(–,–,1057+1_1058-1)del |
| 598 | M | STXB1 | AD | HT | DE NOVO | – | – | – | MISS | 0.72 | 5.71 | NM_003165.3 | c.469-4A>T; p.P157L |
| 1214 | M | STXB1 | AD | HT | NA | – | – | – | MISS | 0.78 | 4.75 | NM_003165.3 | c.469-4A>T; p.P157L |
| 1233 | F | STXB1 | AD | HT | NA | – | – | – | SPLA | - | - | NM_003165.3 | c.469-4A>T; p.P157L |
| 201 | F | DNM1 | AD | HT | DE NOVO | – | – | – | MISS | 0.50 | 1.83 | NM_0025014.1 | c.629_630del; p.W210fs |
| 481 | M | DNM1 | AD | HT | DE NOVO | – | – | – | MISS | 0.71 | 5.42 | NM_001288739.1 | c.208C>A; p.P69L |
| 740 | M | SYN1 | XL | HE | MAT | + | – | – | MISS | 0.59 | 4.01 | NM_0006960.3 | c.576_577+1del; p.G192fs |
| 218 | F | GRIK2A | AD | HT | DE NOVO | – | – | – | MISS | 0.45 | 5.65 | NM_003833.4 | c.636_657del; p.S212fs |
| 857 | F | ARHGEF15 | AD | HT | NA | – | – | – | NONFS | - | - | NM_0025014.1 | c.709_723del; p.G237fs |
| 685 | M | RHOB | AD | HT | DE NOVO | – | – | – | MISS | 0.62 | 5.08 | NM_001288739.1 | c.875G>A; p.R292Q |
| 904 | M | GABRG2 | AD | HT | PAT | + | – | – | NON | - | - | NM_001288739.1 | c.629_630del; p.W210fs |
| 76 | M | SYNGAP1 | AD | HT | DE NOVO | – | – | – | NON | - | - | NM_0006960.3 | c.208C>A; p.P69L |
| 859 | F | SLC35A2 | AD | HT | NA | – | – | – | FS | - | - | NM_0025014.1 | c.576_577+1del; p.G192fs |
| 561 | M | RFT1 | AR | HO | TRANS | – | US | rs148716754 | 0.003 | MISS | 0.54 | 6.03 | NM_0025014.1 | c.709_723del; p.G237fs |
| 1197 | M | DHO3S | AR | AD | HT | DE NOVO | – | – | – | MISS | 0.79 | 5.72 | NM_0025014.1 | c.709_723del; p.G237fs |
| 954 | M | PROSC | AR | HT | TRANS | – | – | – | MISS | 0.43 | 3.19 | NM_0007198.3 | c.208C>A; p.P69L |
| PROSC | HT | – | – | – | – | – | – | MISS | 0.40 | 4.94 | NM_0007198.3 | c.208C>A; p.P69L |
| 38 | M | PCCA | AR | HT | TRANS | – | – | – | SP | - | - | NM_000282.3 | c.629_630del; p.W210fs |
| PCCA | HT | – | – | – | – | – | – | MISS | 0.57 | 3.78 | NM_000282.3 | c.629_630del; p.W210fs |

P, patient; S, sex; M, male; F, female; INH, type of inheritance described for the gene up to now; AD, autosomal dominant; AR, autosomal recessive; XL, X-linked; ZIG, zygosity; HT, heterozygous; HO, homozygous; HE, hemizygous; VFS, variant family study; M, maternal; PAT, paternal. Abbreviations: NA, not available; FH, familial history; ClinVar, classification in the variant in ClinVar database; P, pathogenic; LP, likely pathogenic; US, uncertain significance; B, benign; LB, likely benign; dbSNP, ID in Single Nucleotide Polymorphism Database; MAF, minor allele frequency in public databases; TYPE, functional consequence; MISS, missense; NON, nonsense, SP, splicing; FS, frameshift; NONFS, nonframeshift; CNV, copy number variant; GERP, genomic evolutionary rate profiling score; ISP, in silico prediction with CONDEL software; RFSEQ, Reference Sequence.

ASD/ID phenotype that was unresponsive to Na+ channel blockers. Conversely, the missense variant in P759 behaved as a gain of function (GOF) variant and was associated with early onset and a severe seizure phenotype.

We detected 13 missense KCNQ2 variants in our cohort, 6 of which had been previously described by our group in an article providing extensive information on phenotype-genotype relationship (Hortigüela et al., 2017). As expected, the phenotypes of patients with variants located in the pore-forming domain were severe (Figure 4C).

Seven patients from our cohort carried variants in KCNT1. The lower z-score of this gene (Figure 1A) indicates high sensitivity to missense variation. However, owing to the variable penetrance of variants in this gene it is difficult to demonstrate a definitive association between these variants and the patients' phenotypes. Patient P41 was found to carry biallelic KCNT1 mutations; one caused altered splicing and the other was a missense variant (R262Q) previously described as a de novo variant in an individual with malignant migrating focal seizures of infancy. In line with the variable penetrance previously reported for this gene, the progenitor of the patient carrying this mutation (P41) was healthy. This is not the first case of recessive inheritance in this gene (Martin et al., 2014; Moller et al., 2015). The same missense pathogenic variant R262Q was detected in P957. However, in this case it was not possible to demonstrate a de novo event. While another missense heterozygous indel leading to the missense change S971N was detected in P1008, a familial study could not be performed for this particular patient, preventing classification of the variant. Finally, variant R474L was detected in P268. This variant was found to be paternally inherited. Nonetheless, we consider it highly likely that this variant contributes to the patient's phenotype, as 3 variants affecting the same amino acid (R474S, R474C, and R474H) are classified as disease-associated variants in the ClinVar database (Bércia et al., 2012). The incomplete penetrance previously described for this gene (Moller et al., 2015) remains to be explained.

Five variants were detected in FOXG1. Of the 3 missense variants, 2 were located in the conserved forkhead domain (FHD; Figure 5C) where most disease-associated missense variants have been found (Mitter et al., 2018). R219G was a paternally inherited variant located in this key region. We observed high GERP and CONDEL scores for this variant. However, the clinical phenotype did not match that typically associated with FOXG1.
| P | S | GENE | INH | ZYG | VFS | FH | ClinVar | dbSNP | MAF | TYPE | ISP | GERP | Ref seq | Variant |
|---|---|------|-----|-----|-----|----|---------|-------|-----|------|-----|------|----------|---------|
| **FUNCTIONAL VALIDATION REQUIRED** | | | | | | | | | | | | | | |
| 248 | F | SCN9A | AD | HT | TRANS | – | – | rs180949263 | 0.0002 | MISS | 0.66 | 3.34 | NM_002977.3 | c.5672G>A/p.R1891H |
| 198 | F | CHD2 | AD | HT | MAT | – | – | rs202084111 | 0.0024 | MISS | 0.53 | 4.47 | NM_002977.3 | c.4612T>C/p.W1588R |
| 918 | M | MTOR | AD | HT | PAT | – | – | rs190645065 | 8.9E-106 | MISS | 0.52 | 5.89 | NM_004958.3 | c.7249G>A/p.V2417M |
| 1084 | M | MTOR | AD | HT | PAT | – | – | rs201557303 | 0.0002 | MISS | 0.42 | 4.91 | NM_004958.3 | c.5197G>A/p.A1733T |
| 1144 | M | TSC2 | AD | HT | PAT | – | – | – | – | MISS | 0.59 | 5.51 | NM_006548.4 | c.1724T>C/p.E575P |
| **INCOMPLETE PENETRANCE** | | | | | | | | | | | | | | |
| 754 | M | GABRB1 | AD | HT | MAT | – | LP | rs1135401786 | – | MISS | 0.75 | 3.29 | NM_000812.3 | c.157C>T/p.R53W |
| 938 | F | GABRB1 | AD | HT | MAT | – | – | – | – | MISS | 0.53 | 4.47 | NM_000812.3 | c.775G>A/p.I259V |
| 1084 | M | MTOR | AD | HT | PAT | – | – | – | – | MISS | 0.59 | 5.51 | NM_006548.4 | c.1724T>C/p.E575P |
| 1260 | M | KCNT1 | AD | HT | PAT | – | – | – | – | MISS | 0.59 | 5.51 | NM_006548.4 | c.1724T>C/p.E575P |
| **PLAUSIBLE OLGOCENIC INHERITANCE** | | | | | | | | | | | | | | |
| 13 | M | SCN1A | AD | HT | MAT | – | B/LB | rs121917956 | 0.002 | MISS | 0.58 | 4.89 | NM_001165963.1 | c.5749C>G/p.R1917G |
| 141 | F | SCN1A | AD | HT | PAT | – | – | rs1226308717 | 0.0008 | MISS | 0.54 | 5.44 | NM_001165963.1 | c.5749C>G/p.R1917G |
| 369 | M | SCN1A | AD | HT | NA | – | – | rs121917956 | 0.0002 | MISS | 0.59 | 5.44 | NM_001165963.1 | c.5782C>T/p.R1928G |
| 268 | M | KCNT1 | AD | HT | PAT | – | – | – | – | MISS | 0.59 | 5.44 | NM_001165963.1 | c.126G>T/p.K42N |
| **DATA FROM PROGENITORS UNAVAILABLE: DE NOVO MUTATION STATUS UNCONFIRMED** | | | | | | | | | | | | | | |
| 1038 | M | SCN8A | AD | HT | NA | + | – | – | – | MISS | 0.50 | 4.96 | NM_014191.3 | c.4633A>G/p.T1545P |
| 981 | HCN1 | AD | HT | NA | – | – | – | – | MISS | 0.58 | 5.37 | NM_021072.3 | c.715G>T/p.V239L |
| 738 | F | FOXG1 | AD | HT | NA | – | – | – | – | MISS | 0.42 | 1.65 | NM_002545.4 | c.458G>A/p.G153V |
| 887 | M | SCN1A | AD | HT | NA | – | – | rs777120925 | 7.49E-05 | MISS | 0.54 | 4.75 | NM_001165963.1 | c.1457C>G/p.A486G |
| 523 | F | STXB1 | AD | HT | NA | – | – | rs746172968 | 8.2E-06 | MISS | 0.53 | 5.27 | NM_003165.3 | c.1756C>G/p.D586N |
| 676 | M | HECD1 | AD | HT | NA | – | – | – | – | MISS | 0.59 | 3.93 | NM_01320643.1 | c.1829G>T/p.R610L |
| 1029 | M | GRIA3 | XL | HT | HE | NA | – | – | – | – | MISS | 0.41 | 5.52 | NM_000828.4 | c.793G>A/p.A265T |
| 1099 | M | HECD2 | AD | HT | NA | – | – | – | – | MISS | 0.44 | 5.10 | NM_020760.2 | c.900C>T/p.S300R |
| 884 | F | ARHGEF9 | XL | HT | NA | – | – | – | – | MISS | 0.49 | 5.46 | NM_015185.2 | c.1454G>T/p.G485V |
mutations (no microcephaly, no stereotypical movements, and no regression) and therefore it was classified as of uncertain significance. Similarly, variant G153V was not located in the DNA-binding forkhead domain, in-silico predictions did not support its deleteriousness, and the patient’s clinical picture did not match that typically associated with FOXG1 mutations.

We detected 7 variants in CDKL5 (Figure 5D). The 4 missense CDKL5 mutations were located in the N-terminal kinase domain, which plays a particularly important role in brain function (Fehr et al., 2013). All were the result of de novo mutations, except for the variant detected in P1127, for which we observed very high GERP and CONDEL scores and an associated phenotype concordant with this disorder.

Four patients carried heterozygous missense variants in GNAO1. Two of these patients carried the same variant. Our analyses revealed that this gene was one of the best conserved, with a low z-score (Figure 1) and a mean GERP score > 5 (Figure 2A). The phenotype of these 4 patients perfectly matched that described for GNAO1-mutation carriers. Unfortunately, it was not possible to test the de novo status of variant V301L in P606. This variant is located in a highly conserved nucleotide. Because the phenotype fitted well with that described for GNAO1 mutations, we classified this variant as likely pathogenic, although familiar or functional studies will be required to support this classification.

Variants With Special Characteristics
P186 and P884 carried a heterozygous missense variant in the X-linked gene ARHGEF9. In P186, the variant was located in the RhoGEF domain, which encodes the guanine nucleotide exchange factor (GEF) activity of collybistin, while in P884 the variant was located in the C-terminal domain (Figure 5E). While the CONDEL score for this variant was low, Human Splicing Finder software revealed that this nucleotide change promotes the emergence of a cryptic donor splice site ∼65% stronger than the wild-type site, potentially affecting the splicing of the last exon. Due to a lack of patient RNA sample and samples from family members it was not possible to determine the true involvement of this variant in the patient’s phenotype.
P685 carried a de novo missense variant in the BTB2 domain of RHOBTB2 (Figure S1). This variant was located in an amino acid contiguous to that in which a de novo missense variant was previously described in a female patient (Lopes et al., 2016) who exhibited developmental stagnation at 6–9 months, coinciding with the onset of generalized epilepsy and additional clinical signs including Rett syndrome-like hand stereotypies, intense eye communication, and sleep problems.
CNVs Detected in Our Epilepsy Cohort
As mentioned in the introduction, genomic CNVs account for a substantial proportion of the genetic burden in about 3% of patients with idiopathic epilepsies, and increase the risk of idiopathic generalized epilepsy and a wide range of neurodevelopmental disorders (Helbig et al., 2009; de Kovel et al., 2010; Heinzen et al., 2010; Coe et al., 2012; Lal et al., 2013; Møller et al., 2013). We detected 9 CNVs in our cohort. One microdeletion in RBFOX1 encompassing exons 2 and 3 (Δ2+3) was detected in P519. Figure 6 shows several of the microdeletions reported for this gene. All affect exons at the 5' end of the gene. Some are proven de novo mutations, while others are inherited from an unaffected progenitor, indicating variable penetrance. We found 2 microdeletions in GPHN in patients P530 and P819, both of which were maternally inherited. Deletions affecting GPHN exons are extremely rare in the general population (Mefford et al., 2011; González et al., 2013; Lionel et al., 2013; Dejanovic et al., 2014). Previously described microdeletions in this gene are summarized in Figure 6. The common overlapping region across the 10 deletions encompasses exons 3 and/or 5, which encode the G domain of the gephyrin protein. G-domain trimerization is vital for the formation of the hexagonal gephyrin oligomer scaffolds required for stable GABA receptor clustering in postsynaptic inhibitory neurons.

Variants of Unknown Significance
Variants for Which Additional Evidence Is Required to Support Classification as Deleterious: Lack of Familial Studies and/or Functional Validation
P248 presented secondary partial epilepsy that was unresponsive to treatment, PMD, and Rett syndrome-like features, and was found to carry 2 biallelic mutations in SCN9A. In this case it would be necessary to demonstrate that the combined effects of 2 variants in opposite alleles give rise to the clinical phenotype. This is the first case in which biallelic mutations in SCN9A have been described.

In one (P523) of the 4 patients who carried variants in STXBP1 we detected a second missense variant in SIK1 (Table 1B). In silico analysis predicted that both variants were likely deleterious to the encoded protein. Unfortunately, it was not possible to determine the contribution of these 2 variants to the patient’s phenotype owing to the absence of DNA samples and clinical histories for the patient’s parents.

The uncertain significance of the variant in TSC2
In P1144, a male with neonatal seizures characteristic of tuberous sclerosis (TSC) who responded well to vigabatrin and is progressing toward normality, we detected a missense variant in TSC2. The very high GERP and CONDEL scores for this gene indicate a high degree of sensitivity to variation (this...
gene had the second lowest z-score of all genes studied). These findings support a functional impact of this variant. While the majority of TSC patients are diagnosed during the first 15 months of life, the disease often goes unnoticed owing to the wide phenotypic variability. Familial cases of TSC are caused by germline mutations, but 70% of cases are the result of somatic mutations (Lim et al., 2017). Familial transmission results in mild-to-moderate disease that may not meet all diagnostic criteria. In most such cases the first clinical sign is seizures. All types of seizures can be observed in TSC patients. Two-thirds of cases begin with focal refractory epilepsy. TSC patients have an increased risk of other neurocognitive deficits, including ASD, ID, and mood alterations. A response to vigabatrin is characteristic of these patients, but not distinctive. P1144 carried a paternally inherited missense variant (L575P) in TSC2. We received a blood sample from this patient a few days after her birth and the discovery of the variant surprised her pediatrician, since the patient’s seizures closely resembled those of TSC patients. However, to date no cutaneous stigmas or tubers have been detected. The patient’s father, who also carries the mutation, underwent magnetic resonance imaging (MRI) but no tubers were detected. P1144 has 2 siblings, both of whom were positive for the TSC2 variant. Neither of the siblings display clinical signs characteristic of TSC, but one is under study for short height with no apparent cause and the other was born with unilateral double ureter with grade IV reflux. MRI has been prescribed for both siblings. Recent studies (Caylor et al., 2018; Liu et al., 2018) have identified cases of TSC2 mutations in which the initial presentation consisted of seizures and describe the incidental diagnosis of asymptomatic family members. Further studies and follow-up are required to confirm the involvement of this variant in the phenotypes of the family members. Given the importance of a timely TSC diagnosis for appropriate clinical management, these cases highlight the potential benefits of an unbiased molecular diagnostic approach.

**FIGURE 6** | Microdeletions found in RBFOX1 and GPHN. Yellow, (Lionel et al., 2013); green, (Dejanovic et al., 2014); blue, Fernández-Marmiesse (present study).

---

**Variants in genes involved in TOR and GATOR complexes**

Six of our patients carried variants in genes involved in TOR and GATOR complexes. Brain somatic activating mutations in MTOR have been described in patients with epilepsy caused by focal cortical dysplasia (FCD) type II (Lim and Lee, 2016; Møller et al., 2016). The variants described are located along all exons of MTOR, although certain hotspots have been identified (Figure 4J). Recently, the phenotypic spectrum associated with germline variants in MTOR was extended to include milder phenotypes than previously reported (Møller et al., 2016). Møller et al. found 5 de novo germline mutations in MTOR in 6 individuals with variable epilepsy phenotypes (ranging from focal to generalized) and brain malformations (ranging from no malformation to macrocephaly). Moreover, they describe a variant found in a mother-daughter pair with nocturnal epilepsy of the frontal lobe. MTOR is therefore also a candidate gene for epilepsy without cortical malformation. Our analyses revealed that this gene was the second best conserved: its low z-score implies a high sensitivity to missense variants (Figure 1). Nonetheless, the CONDEL scores for the variants detected in our cohort do not indicate a high level of deleteriousness compared with controls (Figure 2). Two patients in our cohort carried MTOR variants (Table 1B). In the first patient (P918), we detected 2 biallelic missense variants in MTOR. The patient’s mother, maternal uncle, and brother presented FS. The probability of finding 2 missense rare variants in this gene is 3.98 e-4. To our knowledge, there are no reports of epilepsy-associated biallelic mutations in this gene. The second patient (P1084) carried the missense K42N variant, which affects a highly conserved nucleotide. A familial study revealed a paternal family history of seizures, and indicated that the variant was inherited from the patient’s father. However, we could not demonstrate perfect cosegregation of the variant with the seizure phenotype, indicating variable penetrance. Further genomic and functional studies will be required to corroborate a relationship between these variants and the patient’s epileptic phenotype.

**Incomplete Penetrance**

Except in cases of patients with a positive family history, variants detected in dominant genes inherited from unaffected progenitors or with a MAF > 0 were not classified as pathogenic or likely pathogenic, in accordance with the guidelines of the American College of Medical Genetics and Genomics (ACMG). These variants (listed in Table 1B) are thus pending classification. A role in the patient’s phenotype cannot be ruled out since incomplete penetrance is not unusual in autosomal dominant epileptic disorders, as described for example in patients with PRRT2 frameshift mutations (Cloarec et al., 2012; Lee et al., 2012) and mutations in SCN1A (Gennaro et al., 2003; Fukuma et al., 2004; Kimura et al., 2005; Mancardi et al., 2006; Depienne et al., 2009), SCN8A, KCNIT (Møller et al., 2015), SLC12A5, or DEPDC5 (Baulac, 2016). Potential explanations for incomplete penetrance include parental mosaicism, oligogenic inheritance, and the Knudson 2-hit mechanism (discussed below).

**The Knudson 2-hit mechanism**

Somatic mutational events in the brain are frequent (Hoang et al., 2016). An individual can inherit one germinal variant from one
progenitor and subsequent appearance during development of a
second variant in the same or another gene can have an additive
effect, giving rise to a clinical phenotype. In FCD, reports of brain
somatic mutations in genes involved in the mTORC1 pathway,
especially activating somatic MTOR variants, are increasingly
common (Poduri et al., 2013; Lim et al., 2015; Nakashima et al.,
2015; Mirzaa et al., 2016; Møller et al., 2016; Marsan and Baulac,
2018). The presence of loss-of-function mutations in DEPD5
is the most common cause of familial focal epilepsies. However,
only a subset of patients among families with DEPD5 mutations
develop FCD (Baulac, 2016); other family members appear to
present non-lesional epilepsy. Ribierre et al. (2018) demonstrated
that a biallelic 2-hit (brain somatic and germline) mutational
mechanism in DEPD5 causes focal epilepsy with FCD.

Digenic inheritance or modifier genes
Mutations in different ion-channel genes can exacerbate or
counteract epileptic phenotypes (Glasscock et al., 2007; Hawkins
et al., 2011; Klassen et al., 2011; Calhoun et al., 2017; Hasan
et al., 2017). These reports suggest that the co-occurrence of 2
mutations in distinct genes that are independently innocuous can
give rise to a clinical phenotype and help explain the incomplete
penetrance described for many epilepsy-related genes. Functional
studies and/or further reports of similar cases will obviously be
required to confirm this hypothesis.

Our cohort included several potential examples of digenic
inheritance. For example, 3 patients in our cohort carried variant
R1928G in SCN1A (frequency, 0.0026 in 1000G). In 2 cases (P13
with EIEE and P396 with Dravet syndrome and severe ID) this
variant co-existed with a rare missense variant in CLCN2 (R68H
and S123F; respectively). Neither of these CLCN2 variants are
recorded in PVDB or in 1344 Ib-chr, and the first was confirmed
to be in trans with the R1928G variant. The probability of finding
2 rare missense variants simultaneously in both genes is 3.99e−4.
The CLCN2 encodes the ClC-2 chloride channel. Underscoring
the importance of this channel in the brain, biallelic mutations
in this gene are associated with leukoencephalopathy with
ataxia (MIM_615651). Although there is some evidence linking
CLCN2 mutations to susceptibility to epilepsy, these data remain
controversial, and these cases could alternatively be explained by
undetected digenic inheritance (Sander et al., 2000; Haug et al.,
2003; Kleeu↵uss-Lie et al., 2009; Saint-Martin et al., 2009). In the
third case (patient P141), the SCN1A variant was present in cis
with a second SCN1A variant and in trans with a KCN71 variant
(Prob = 1.99e−3). This patient presented neonatal refractory
epilepsy (Othara syndrome), central coordination disturbance,
peculiar phenotype, hepatomegaly, and ventriculomegaly.

Another example was the patient P1113, a male with seizures
and GDD, carrying a missense KCN71 variant, with a frequency
in public variant databases that was incompatible with dominant
inheritance with full penetrance, accompanied by 2 other ion
channel variants in HCN1 and CACNA1H. In this case further
functional studies will be required to determine the variant’s role
in the patient’s phenotype (Prob ~ 0).

P676, in addition to carrying a variant in GABBR1,
carried another variant in GABBR2 (Table 1B), a gene also
implicated in EIEE (EuroEPINOMICS-RES Consortium, 2014).

Unfortunately, this patient lives in an institution and their
parents could not be contacted to perform a familial study.

Digenic inheritance was clearer in the case of 2 Portuguese
siblings, P501 and P968, who presented very different forms of
epilepsy. One had a clear molecular diagnosis for a recessive gene
(SLC12A5) and carried a maternally inherited SCNIA variant.
The patient’s sister, who had a milder phenotype consistent
with absence epilepsy that was well-controlled with valproate,
carried a combination of inherited heterozygous variants in
genes encoding ion channels: SCNIA (maternal), CACNA1A
(paternal), and SLC12A5. In individuals of French-Canadian
origin, the SLC12A5 missense variants R952H and R1049C
in heterozygosity have been shown to predispose carriers to
generalized epilepsy (Kahle et al., 2014; Puskarjov et al., 2014).
In these studies, in vitro functional expression studies revealed that
these variants impair SLC12A5 function. Moreover, in several
cases the variants were inherited from an unaffected parent,
consistent with incomplete penetrance. Further evidence from
functional studies in mouse models will be required to confirm
cases of proposed digenic inheritance.

Limitations of the Study
A limitation of the targeted resequencing approach is that,
unlike WES, it cannot detect novel epilepsy-associated genes.
Furthermore, unlike WGS it does not allow detection of
mutations in intronic sequences or non-coding RNAs with
transcriptional regulation roles which together correspond to
80% of human genome according to ENCODE Project
Consortium 2012 (The ENCODE Project Consortium, 2012).
Moreover, our approach can detect neither balanced structural
variants nor CNVs outside coding regions, both of which can
change the regulatory landscape of developmental genes and alter
the delicate balance between transcriptional enhancers, silencers,
and insulators (Conrad et al., 2010; Spielmann and Klopocki,
2013). Epimutations, which may account for disease in ~20%
of neurodevelopmental disorders and congenital anomalies
(Barbosa et al., 2018) are also undetectable with this technology.
A final limitation of our approach is that mutations in somatic
mosaicism may also go undetected in blood samples and it
has been shown that somatic mutational events in brain tissue
are frequent (Hoang et al., 2016) and a high frequency of
mosaic pathogenic variants in epilepsy-associated genes has been
demonstrated (Stosser et al., 2018).

CONCLUSIONS

- Well-optimized targeted exome sequencing of a large high
number of genes with high mean coverage can be highly
beneficial to clinical specialists and to pediatric epileptic
patients and their families. By using this approach in our
cohort, we were able to establish a genetic diagnosis in
~42% of cases. This in turn alleviates parental anxiety
and guilt, helps identify at-risk family members, and can
facilitates reproductive decision-making. Furthermore, it
provides parents with access to a community with shared
experiences, limits the need for counterproductive tests and
treatments, and enables selection of the most appropriate
antiepileptic therapy.
● The length of the mean diagnostic delay in our cohort (~6 years, measured from the time of onset of clinical criteria until molecular diagnosis) underscores the importance of implementing tests of this kind as first-tier diagnostic tools for epilepsy patients.

● To take into account the varying mutational sensitivities of epilepsy-related genes, their specific mutational architecture, the in-silico prediction of missense variants (GERP and CONDEL) and filtering variants using a large, well-characterized database of the patient’s population is essential to ensure optimal prioritization of the variants detected.

● CNVs account for a considerable percentage of the mutational load in epilepsies (as much as nonsense and frameshift variants). If the mean coverage achieved using our panel had been lower (as is the case with many diagnostic panels), this specific source of mutations would have gone unnoticed.

DATA AVAILABILITY STATEMENT

The data analyzed for this study can be found in the BioProject repository, accession number PRJNA551134: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA551134.

ETHICS STATEMENT

This study was approved by the Ethics Committee of the Hospital Santiago de Compostela (Spain). A total of 246 patients were recruited over a 5-year period from different neurologic units in Spain and Portugal, applying the following inclusion criteria: patients with any condition in which seizures are either the sole clinical expression or are part of more severe neurodevelopmental disorder.

AUTHOR CONTRIBUTIONS

AF-M designed and optimized successive versions of NGS based epilepsy panels, interpreted genetic data, generated reports, and wrote the manuscript. IR performed bioinformatic analyses, transforming raw data into an annotated table of prioritized variants, conducted statistical analyses of z-score and GERP and CONDEL scores, contributed to the writing of successive versions of the manuscript, compiled references, and managed variant annotation and optimized figures. The rest of the authors contributed with the clinical care of patients, blood sample collection from patients and their families, collected clinical and family history data, and collaborated in the edition of successive versions of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/finns.2019.01135/full#supplementary-material

## Supplementary Table 1 | Clinical features of patient listed in Table 1.

| Clinical Feature | Values |
|------------------|--------|
| Age at onset     | 6.5 y   |
| Sex              | M      |
| Seizure type     | Partial |
| Family history   | Positive |
| Diagnosis        | Epilepsy |
| Treatment        | Medication |

REFERENCES

Allen, N. M., Conroy, J., Shahwan, A., Lynch, B., Correa, R. G., Pena, S. D. J., et al. (2016). Unexplained early onset epileptic encephalopathy: Exome screening and phenotype expansion. Epilepsia 57, e12–e17. doi: 10.1111/epi.13250

Babraham Bioinformatics-FastQC A Quality Control Tool for High Throughput Sequence Data (2018). Available online at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed October 11, 2018).

Barbosa, M., Joshi, R. S., Garg, P., Martin-Trujillo, A., Patel, N., Jadhav, B., et al. (2018). Identification of rare de novo epigenetic variations in congenital disorders. Nat. Commun. 9:2064. doi: 10.1038/s41467-018-04540-x

Barcía, G., Fleming, M. R., Deligniere, A., Gazula, V.-R., Brown, M. R., Langouet, M., et al. (2012). De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. Nat. Genet. 44, 1255–1259. doi: 10.1038/ng.2441

Baulac, S. (2016). mTOR signaling pathway genes in focal epilepsies. Prog. Brain Res. 226, 61–79. doi: 10.1016/bbr.2016.04.013

Borlot, F., Regan, B. M., Bassett, A. S., Stavropoulos, D. J., and Andrade, D. M. (2017). Prevalence of pathogenic copy number variation in adults with pediatric-onset epilepsy and intellectual disability. JAMA Neurol. 74, 1301–1311. doi: 10.1001/jamaneurol.2017.1775

Calhoun, J. D., Hawkins, N. A., Zachwieja, N. J., and Kearney, J. A. (2017). Cacna1g is a genetic modifier of epilepsy in a mouse model of Dravet syndrome. Brain. 133, 23–32. doi: 10.1093/brain/awp262

Cloarec, R., Bruneau, N., Rudolf, G., Massacrier, A., Salmi, M., Bataillard, A., et al. (2012). PRRT2 links infantile convulsions and paroxysmal dyskinesia with migraine. Neurology 79, 2097–2103. doi: 10.1212/wnl.0b013e3182752c46

Coe, B. P., Girirajan, S., and Eichler, E. E. (2012). A genetic model for neurodevelopmental disease. Curr. Opin. Neurobiol. 22, 829–836. doi: 10.1016/j.conb.2012.04.007

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., et al. (2010). Origins and functional impact of copy number variation in the human genome. Nature 464, 701–712. doi: 10.1038/nature08856

Davydov, E. V., Goode, D. L., Sirotta, M., Cooper, G. M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PloS Comput. Biol. 6:e10001025. doi: 10.1371/journal.pcbi.1001025

de Kovel, C. G. F., Trucks, H., Helbig, I., Mefford, H. C., Baker, C., Leu, C., et al. (2010). Recurrent microdeletions at 15q11.2 and 16p13.1 predispose to idiopathic generalized epilepsies. Brain J. Neurol. 133, 23–32. doi: 10.1093/brain/awp262

de Lange, I. M., Rump, P., Neuteboom, R. F., Augustijn, P. B., Hodgès, K., Kistemaker, A. L., et al. (2017). Male patients affected by mosaic PCDH19 mutations: five new cases. Neurogenetics 18, 147–153. doi: 10.1007/s10048-017-0517-5

Dejanovic, B., Lal, D., Catarino, C. B., Arjune, S., Belaidi, A. A., Trucks, H., et al. (2014). Exonic microdeletions of the gephyrin gene impair GABAergic synaptic inhibition in patients with idiopathic generalized epilepsies. Neurobiol. Dis. 67, 88–96. doi: 10.1016/j.nbd.2014.02.001

Deprins, S., Pierre, R., Daviet, S., Vanhaecke, T., Calisse, S., and De Zegher, F. (2010). Spectrum of SCN1A gene mutations associated with Dravet syndrome: analysis of 333 patients. J. Med. Genet. 46, 183–191. doi: 10.1136/jmg.2008.062323

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., and Béroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37, e67–e67. doi: 10.1093/nar/gkp215
The image contains a page from a scientific document discussing genetic diagnosis in epilepsy. The text includes references to various studies and researchers in the field of epilepsy genetics. The page contains multiple references to studies on genetic causes of epilepsy, including copy number variations, mutations in synaptic transmission genes, and neuronal voltage-gated ion channels as genetic modifiers.

For example, the text mentions a study by Lim, J. S., et al. (2011) on the role of neuronal voltage-gated ion channels in epilepsy. It also references research by Grollman, A. P., et al. (2016) on genome-wide quantification of rare somatic mutations in normal human tissues using massively parallel sequencing.

The document highlights the importance of genetic heterogeneity in epilepsy, with studies showing that mutations in specific genes can contribute to different forms of epilepsy. It also discusses the use of next-generation sequencing technologies to identify novel genetic causes of epilepsy.

In summary, the page provides an overview of the current understanding of genetic diagnosis in epilepsy, emphasizing the role of genetic mutations in various epilepsy syndromes and the use of advanced genetic analysis techniques to identify these mutations.

This text is from a scholarly article discussing the Epi4K Consortium and the Epilepsy Phenome/Genome Project, highlighting key findings in genetic diagnosis in epilepsy. The references are cited throughout the text, covering a range of studies from 2004 to 2019, including publications in journals such as *Hum. Mol. Genet.*, *Am. J. Hum. Genet.*, and *Epilepsia*.

The text likely serves as a review or summary of recent research in the field, providing insights into the genetic basis of epilepsy and the methods used to identify these genetic causes.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. doi: 10.1101/gr.107524.110

Mefford, H. C. (2015). Copy number matters in epilepsy. Epilepsy Curr. 15, 180–182. doi: 10.3159/1535-7511-4.1.80

Mefford, H. C., Yendle, S. C., Hsu, C., Cook, J., Geraghty, E., McMahon, J. M., et al. (2011). Rare copy number variants are an important cause of epileptic encephalopathies. Ann. Neurol. 70, 974–985. doi: 10.1002/ana.22645

Mei, D., Parrini, E., Marini, C., and Guerrini, R. (2017). Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. Epilepsia 56, 707–716. doi: 10.1111/epi.12954

Mirzaa, G. M., Campbell, C. D., Solovieff, N., Goold, C., Jansen, L. A., Mirzaa, G. M., Campbell, C. D., Solovieff, N., Goold, C., Jansen, L. A., et al. (2016). Association of MTOR mutations with developmental brain disorders, including megalencephaly, focal cortical dysplasia, and pigmentary mosaicism. JAMA Neurol. 73, 836–845. doi: 10.1001/jamaneurol.2016.10363

Mitter, D., Pringsheim, M., Kaulisch, M., Plümacher, K. S., Schröder, S., Warthemann, R., et al. (2018). FOXG1 syndrome: genotype–phenotype association in 83 patients with FOXG1 variants. Genet. Med. 20, 98–108. doi: 10.1038/gim.2017.75

Möller, R. S., Heron, S. E., Larsen, L. H. G., Lim, C. X., Ricos, M. G., Bayly, M. A., et al. (2015). Mutations in KCNT1 cause a spectrum of focal epilepsies. Epilepsia 56, e114–e120. doi: 10.1111/epi.13071

Möller, R. S., Weber, Y. G., Klitten, L. L., Trucks, H., Muhle, H., Kunz, W. S., et al. (2013). Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. Epilepsia 54, 256–264. doi: 10.1111/epi.12078

Möller, R. S., Weckhuysen, S., Chipaux, M., Marsan, E., Taly, V., Bebin, E. M., et al. (2016). Germline and somatic mutations in the MTOR gene in focal cortical dysplasia and epilepsy. Neurot. Genet. 2, e118. doi: 10.1211/NXG.0000000000000118

Nakashima, M., Saito, H., Takei, N., Tohyama, J., Kato, M., Kitaura, H., et al. (2015). Somatic Mutations in the MTOR gene cause focal cortical dysplasia type Ilb. Ann. Neurol. 78, 375–386. doi: 10.1002/ana.24444

Oates, S., Tang, S., Rosch, R., Lear, R., Hughes, E. F., Williams, R. E., et al. (2018). Incorporating epilepsy genetics into clinical practice: a 360° evaluation. npj Genomic Med. 3:13. doi: 10.1038/s41525-018-0052-9

Picard Tools-By Broad Institute (2018). Available online at: http://broadinstitute.github.io/picard/ (accessed October 11, 2018).

Poduri, A., Evroney, G. D., Cai, X., and Walsh, C. A. (2013). Somatic mutation, genomic variation, and neurological disease. Science 341, 1237758. doi: 10.1126/science.1237758

Puskarjov, M., Seja, P., Heron, S. E., Williams, T. C., Ahmad, F., Iqbal, Z., Twigg, S. R. F., et al. (2014). Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. Nat. Genet. 46, 912–918. doi: 10.1038/ng.3036

Roca, I., Fernández-Marmiesse, A., Gouveia, S., Segovia, M., and Couce, M. (2018). Prioritization of variants detected by next generation sequencing according to the mutation tolerance and mutational architecture of the corresponding genes. Int. J. Mol. Sci. 19:1584. doi: 10.3390/ijms19061584

Roca, I., González-Castro, L., Maynou, J., Palacios, L., Fernández, H., Couce, M. L., et al. (2019). PattRec: An easy-to-use CNV detection tool optimized for targeted NGS assays with diagnostic purposes. Genomics. doi: 10.1016/j.ygeno.2019.07.011. [Epub ahead of print].

Shaukat, Q., Hertecant, J., El-Hattab, A. W., Ali, B. R., and Suleiman, J. (2018). West syndrome, developmental and epileptic encephalopathy, and severe CNS disorder associated with WWOX mutations. Epileptic Disord. Int. Epileyp. Videolape 20, 401–412. doi: 10.1684/epid.2018.1005

Stone-McCaffrey, N. J., et al. (2019). Genetic Diagnosis in Epilepsy. 180–182. doi: 10.1101/gr.107524.110

Tsuchida, N., Nakashima, M., Kato, M., Heyman, E., Inui, T., Hagino, K., et al. (2018). Detection of copy number variations in epilepsy using exome data. Clin. Genet. 93, 577–587. doi: 10.1111/cge.13144

Vanlambrouck, L., Dibbens, L. M., Lawrance, K. M., Iona, X., McMahon, J. M., Murrell, W., et al. (2010). Timing of de novo mutagenesis—a twin study of sodium-channel mutations. N. Engl. J. Med. 363, 1335–1340. doi: 10.1056/NEJMoat0910752

Wang, J., Kurahashi, H., Ishii, A., Kojima, T., Ohfu, M., Inoue, T., et al. (2008). Microchromosomal deletions involving SCN1A and adjacent genes in severe myoclonic epilepsy in infancy. Epilepsia 49, 1528–1534. doi: 10.1111/j.1528-1167.2008.01609.x

Weber, Y. G., Biskup, S., Helbig, K. L., von Spiczak, S., and Lerche, H. (2017). The role of genetic testing in epilepsy diagnosis and management. Expert Rev. Mol. Diagn. 17, 739–750. doi: 10.1080/14737159.2017.1335998

Ye, K., Schulz, M. H., Long, Q., Apweiler, R., and Ning, Z. (2009). Pindel: a pattern discovery tool for paired-end short reads. Bioinform. Oxf. Eng. 25, 2865–2871. doi: 10.1093/bioinformatics/btp394

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.