Gene panels for epilepsy suggest that previously defined variants of unknown significance may play an important role in epilepsy and certain variants may be pathogenic when occurring together.

Authors: Yara Hussein1, Hila Weisblum-Neuman2, Bruria Ben-Zeev2,3,#, Shani Stern1,#

1Sagol Department of Neurobiology, Faculty of Natural Sciences, University of Haifa, Haifa, Israel
2 Pediatric Neurology Unit, The Edmond and Lily Safra Children’s Hospital, Sheba Medical Center, Israel
3 Faculty of Medicine, Tel Aviv University, Israel

#Corresponding Author
email: sstern@univ.haifa.ac.il

Abstract

Epilepsy is a chronic neurological disorder characterized by recurrent seizures and associated neurological, cognitive, psychological, and social effects. The prevalence of active epilepsy is estimated to be between 4-10 per 1000 individuals in the general population, with the highest incidence occurring during infancy and childhood. Genetic mutations play a significant role in epilepsy, and over 500 genes have been associated with the condition. Next-Generation Sequencing (NGS) panels are utilized for genetic testing, but a substantial proportion of results remain uncertain and are not considered directly causative of epilepsy. In this study, we reevaluated a subgroup of patients with inconclusive variant findings or multiple Variants of Uncertain Significance (VUSs) in their test results. We identified two unrelated variants c.2133G>C in SCN9A and c.316G>A in QARS1 to be potentially pathogenic variants. Additionally, we identified a frequent genetic combination, the RANBP2&RYR3 being prominent among other possible combinations. The RANBP2 gene consistently co-occurred with RYR3 variants in uncertain results, suggesting potential pathogenicity. We also analyzed unaffected parents’ data and observed certain combinations inherited from different parents, raising the possibility of specific gene combinations as risk factors for the disease.
Introduction

Epilepsy is a chronic neurological ailment distinguished by a persistent susceptibility to generate seizures and incur subsequent neurobiological, cognitive, psychological, and social ramifications arising from recurrent seizure occurrences\(^1,2\).

According to the world health organization, the prevalence of active epilepsy is estimated to be between 4-10 per 1000 individuals in the general population; On a global scale, approximately 5 million people are diagnosed with epilepsy yearly\(^3\) while the majority of epilepsy is diagnosed during pediatric years in which around 1 in 150 children is diagnosed with epilepsy, with the highest incidence rate occurring in infancy\(^4\).

Seizures are classified into three distinct categories: generalized, focal, and epileptic spasms. Generalized seizures exhibit involvement of both cerebral hemispheres, while focal seizures originate from specific regions within one hemisphere and may propagate further. Generalized seizures encompass absence seizures characterized by unresponsiveness, tonic-clonic seizures featuring convulsive movements accompanied by impaired consciousness, myoclonic seizures characterized by brief muscle contractions, and atonic seizures leading to a loss of muscular tone. On the other hand, focal seizures manifest diverse clinical presentations contingent upon the particular cerebral area affected, encompassing visual, motor, or sensory symptoms. Dyscognitive seizures, often associated with temporal lobe epilepsy, are characterized by compromised consciousness. Certain seizures may exhibit an aura, a preceding focal seizure marked by preserved awareness and distinctive symptoms. Epileptic spasms involve abrupt extension or flexion of the limbs, predominantly observed in infantile spasms (West syndrome)\(^5,6\).

Epilepsy can be attributed to various causes but the genetic component is strong: Over 50% of epilepsy cases have a genetic component\(^7,8\). The literature has described over 500 genes associated with epilepsy, suggesting their potential contribution to the development of the condition\(^9\). Various genes were found to be highly penetrant to epilepsy including sodium channel subunits, GABA receptor subunits, and others\(^10,11\) such as \(SCN1A\) and \(KCNQ2\)\(^12,13\). Electrophysiological studies on mutations associated with epilepsy show a neuronal hyperexcitability pattern\(^14-18\) and a reduction in GABA-positive neurons\(^14,15\).

One of the diagnostic tools used for epilepsy is Next-Generation Sequencing (NGS), which is a powerful DNA sequencing technology. It allows for the high-throughput analysis of DNA or RNA, generating sequencing data rapidly and cost-effectively compared to traditional methods. NGS involves several steps, including DNA fragmentation, library preparation, sequencing by synthesis, and data analysis, enabling the simultaneous sequencing of multiple DNA fragments\(^19\). The yield of positive molecular tests is relatively low, about 15%-25% are categorized as pathogenic or likely pathogenic\(^12,20\); On the other hand, most of the variants found in these epilepsy panels are categorized as variants of uncertain significance (VUS)\(^12\), posing a substantial challenge in contemporary genetic variation screening approaches and genetic counseling.

The aim of our work is to reevaluate a subgroup of tested individuals with inconclusive variant findings, or more than one VUS in their panel results, assuming that several rare variants in genetic hotspots can lead to epilepsy manifestation, yet not categorized as pathogenic.
Methods

Invitae Epilepsy testing
After obtaining approval from the institutional committee and written informed consent from all participants and their legal guardians, genetic testing was conducted on probands and selected unaffected biological parents using the Invitae epilepsy test, which utilizes a targeted gene panel based on NGS technology. Genomic DNA was extracted from blood or saliva samples and enriched for specific regions of interest using a hybridization-based protocol, followed by sequencing using Illumina technology. The sequencing depth for all targeted regions was set at a minimum of 50 times, and the resulting reads were aligned to the GRCh37 reference sequence. The analysis focused on the coding sequence of the indicated transcripts, including a 10-base pair flanking intronic sequence as well as specific genomic regions known to be causative of disease. The aim is to identify single-nucleotide variants (SNVs), short and long indels, exon-level deletions/duplications, and rare structural rearrangements that disrupt coding sequences. The panel targets genes associated with both syndromic and nonsyndromic causes of epilepsy, providing a comprehensive assessment of the genetic factors underlying the condition.

Variants Classification
Observed variants are classified as pathogenic (directly contribute to the development of a disease), likely pathogenic (likelihood of causing a known genetic condition), Benign (not known to cause genetic conditions, but can alter the protein products or change the gene’s expression), likely benign (is not expected to lead to a genetic condition, but the scientific evidence is not as strong as for the variants that are classified benign) or variant of uncertain significance (VUS) which is a variation in DNA that has an uncertain or unknown impact on health. These classifications are based on Sherloc-Invitae's variant classification algorithm, based on the initial American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) classification framework, and represent the industry standard among clinical genetic testing laboratories. In order to minimize uncertainty in genetic testing, follow-up testing for selected unaffected biological parents of patients previously tested was done.

Observed Variants analysis
A retrospective analysis was performed on a cohort of a single tertiary center for child neurology. The electronic tests of 156 probands and 74 unaffected biological parents performed in the years 2018-2022 in the Edmond and Lily Safra Children’s Hospital, Sheba medical center, were used and data were extracted using custom-written MATLAB scripts (R2023a, Mathworks) and anonymously analyzed (included and excluded data are shown in supplementary Fig S1a). Tests were grouped into 3 groups, Positive tests (in which there was a detection of at least one pathogenic/likely pathogenic variant that directly contributes to the emergence of epilepsy), Negative tests (no detected variant in the selected genes or only benign or likely benign detected variants) and Uncertain tests (at least one detected VUS). For each, the total number of detected variants was counted, and the most definitive leaders were found and compared to their population incidence based on gnomAD browser (https://gnomad.broadinstitute.org/). Protein network path analysis was performed to find significant pathways against the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) databases for observed pathogenic and uncertain variants separately. The Network Analyst web application was used to create a graphical representation of Protein network path analysis (https://www.networkanalyst.ca).

In the collective findings, the most prevalent combinations of VUS were identified across each subgroup. An identical analysis was conducted on probands who possessed data for both unaffected biological parents, aiming to identify notable combinations of genetic variants, genetic correlations were calculated using Correlation AnalyzeR (https://gccri.bishop-lab.uthscsa.edu/shiny/correlation-analyzet/).
Results

156 probands underwent testing using the Invitae epilepsy panel (Blood samples = 94%). Approximately half were females (53.2%) and the mean age at test requisition was 7.6 ± 5.3 years (range, 0-23 years). The epilepsy panel is occasionally supplemented with additional panels or add-ons; on average, a total of 227.4 ± 59.7 genes were tested (with the main panel consisting of 187 genes, as depicted in the latest version of Invitae until 2022, as shown in supplementary figure S1b).

A total of 18% of the probands who underwent testing received a negative result, indicating no detectable genetic variants related to epilepsy, 22% of the probands obtained a positive result, indicating the presence of at least one pathogenic variant associated with epilepsy. The majority, constituting 60% of the probands, received an uncertain result (see methods) as shown in Figure 1a and supplementary Figure S1c.

On average, there were 2.3 ± 1.7 observed genetic variants across the probands. Notably, 11% of the individuals showed no observed variants in the targeted genes at all, while almost 65% exhibited 2 or more observed variants, as shown in Figure 1b. Among all the observed variants, 20% were classified as benign or likely benign, 10% were classified as pathogenic or likely pathogenic, and the majority, constituting 69%, were classified as variants of uncertain significance as shown in Figure 1c.

Regulation of action potential and multiple neuronal developmental processes are dysregulated in child epilepsy

As previously mentioned, 22% of the tests were positive, out of these 15% exhibited the presence of two pathogenic variants, while 85% indicated the presence of a single pathogenic variant contributing to the onset of epilepsy (Figure 2a). Among the pathogenic genes identified in the probands, the SCN1A gene- coding for the α-subunit of a neuronal voltage-gated sodium channel- was detected in 11% of the positive results, followed by MECP2 and CDKL5 genes, each observed in 8% of the cases as shown in Figure 2b. Noteworthy the pathogenic variants c.1274_1277dup in the HEXA gene, the c.1385G>C in the PNKP gene, the c.649dup in the PRRT2 gene, and the c.671A>G in the GCH1 gene, were observed twice each in the positive tests (Figure 2c).

Based on the detected pathogenic genes in our tests, Cellular Components (CC) related to neuron projection, cell projection, axon, growth cone, and others showed significant GO: CC enrichment (FDR<0.05). In terms of biological processes (BP), the transmission of nerve impulses, neurological system processes, regulations for action and membrane potentials, and neuron development were enriched (FDR < 0.05), as shown in Figure 2d-e indicating several neuronal alterations contributing to the development of epilepsy.

Two genetic alterations in the SCN9A and QARS1 genes, currently classified as VUS’s have a potential implication for pathogenicity in child epilepsy

For the uncertain results with at least one detected VUS, an average of 2 ± 1.2 VUSs were present in the patients' panels, out of which 45% showed a presence of a single VUS, while 55% exhibited two or more VUSs, as presented in Figure 3a. SCN5A, FASN, and RYR3 were the most frequently observed VUS genes as shown in Figures 3b and 3c. Furthermore, the c.2133G>C variant in the SCN9A gene and the c.316G>A variant in the QARS1 gene were observed three times each in the uncertain tests, along with other variants shown in Figure 3d. These variants appear to have a high penetrance in epilepsy patients compared to their presence in the general population based on the gnomAD database (p=3.76e-07) although classified to be with an uncertain significance.

In terms of GO biological processes, the detected variants of uncertain significance showed enrichment in processes related to the transmission of nerve impulses, neurological system processes, regulations for action and membrane potentials, cell signaling, synaptic transmission, neuron development, and differentiation, among others. Furthermore, various neuronal components and
channels such as voltage-gated calcium and potassium channels were enriched in terms of GO: CC (FDR < 0.05), (Figure 3e and Figure 3g). Notably, KEGG pathways associated with different types of synapses, including GABAergic and dopaminergic synapses and the synaptic vesicle cycle, showed significant enrichment (FDR < 0.05) as shown in Figure 3f.

**RANBP2 and RYR3 variant combination as a risk factor for child epilepsy**

Upon analysing the frequent combinations of VUSs in the overall results, it became evident that the RANBP2&RYR3 variant combination occurred most frequently; It was observed three times in our data. Additional combinations of VUS are shown in Figure 4a. Similarly, the RANBP2 & RYR3 combination remains the most prevalent when focusing solely on uncertain results, as shown in Figure 4b; Interestingly, all detected variants of the RANBP2 gene were designated as VUSs and coexisted in diverse combinations with the RYR3 variants.

For a subgroup of selected probands (n=38) within the cohort, genetic testing was conducted on both unaffected biological parents (n=74) to assess the variants identified in their children. Remarkably, 91% of the observed proband variants were also found in one unaffected parent, while 4% of the variants were determined to be de novo, meaning they were not present in either parent, as illustrated in Figure 4c. To investigate potential combinations that may contribute to the risk of developing epilepsy, optional combinations in the parental data were analyzed. The combination of RANBP2&RYR3 remained notably prominent among the observed variants in the proband data, as depicted in Figure 3d. Additionally, other combinations in probands data exhibited potential risk factors compared to their unaffected biological parents (p= 6.7471e-04).
Discussion

In the past decade, significant progress has been made in identifying gene mutations associated with epilepsy and unraveling the molecular mechanisms that contribute to the clinical presentation of the disease. It is becoming increasingly evident that comprehending these mechanisms is crucial for selecting optimal treatment approaches for affected individuals. Consequently, genetic testing is done to determine whether these features are associated with a known genetic condition. Epilepsy, being a disorder with a notable genetic component, has been linked to several specific genes known to play a direct role in its development. To facilitate the identification of potential pathogenic variants in these epilepsy-associated genes, NGS panels have been developed. These panels are designed to probe genes highly relevant to epilepsy, allowing for a targeted and comprehensive analysis of genetic variations such as Invitae's epilepsy panel.

In this study, we analyzed genetic data from patients diagnosed with child epilepsy who had undergone Invitae's tests. These samples were obtained either from blood or saliva, some with a confirmed genetic alterations correlated with epilepsy, and some with no definite genetic condition. It is important to highlight that the tests' results exhibited partial ambiguity, with approximately 60% of the tests (and 69% of the detected variants) showing significant uncertainty. One relatively unique characteristic of VUSs, is that while the result itself may remain static, its meaning is often resolved over time, as more data are gathered. This complex task is crucial for the comprehension of the underlying pathogenic elements, enabling the implementation of appropriate treatments aimed at reducing seizures and effectively managing the disorder.

Consistent with previously reported data, we observed that the SCN1A gene was one of the leaders in the observed pathogenic variants resulting in epilepsy seizures; The SCN1A gene is primarily recognized for its association with Dravet syndrome (DS), a severe and rare myoclonic epilepsy characterized by impaired psychomotor and neurological development, typically emerging in the first year of life in seemingly healthy infants. Additionally, SCN1A mutations have been implicated in various other epilepsy syndromes, often accompanied by significant comorbidities. Numerous epilepsy syndromes in humans stem from mutations in genes encoding voltage-gated sodium channels, with most of these mutations occurring in the SCN1A gene, and following a dominant inheritance pattern in addition to other genes such as SCN5A and SCN9A.

We observed that the c.2133G>C (p.Leu711Phe) variant in the SCN9A gene was repeated in 3% of our tests. This sequence change replaces leucine, with phenylalanine, which are both neutral and non-polar amino acids at codon 711 of the SCN9A protein. Similarly, a repeated unrelated variant is the c.316G>A (p.Asp106Asn) variant in the QARS1 gene-inherited in an autosomal recessive pattern and associated with "Microcephaly, progressive, with seizures and cerebral and cerebellar atrophy" syndrome (OMIM #615760). We observed that this variant was repeated also in 3% of our tests. This sequence change replaces aspartic acid, which is acidic polar, with asparagine, which is a neutral polar amino acid, at codon 106 of the QARS protein. The consistent observations of these variants in multiple unrelated patients indicate their individual potential pathogenicity, leading to their consideration as a causative variant responsible for epilepsy development. It is noteworthy that these specific variants have not been reported as associated with epilepsy yet.

Furthermore, we observed dysregulation of several neurodevelopmental functions and GABAergic synapses. A decreased activity of the inhibitory circuitry is thus likely to be a major factor contributing to seizure generation in patients.

Nearly half of the patients with an uncertain result, showed the presence of two or more VUSs among their detected variants, highlighting the importance of precise interpretation regarding the clinical relevance of various variant combinations. Notably, the combination of RYR3&RANBP2 variants was observed multiple times in our data, standing out among all other possible combinations. Interestingly, the RANBP2 variants, when present as VUS, consistently co-occurred with a variant of...
RYR3, suggesting a potential pathogenic role for this combination in contributing to the manifestation of epilepsy.

The RANBP2 gene (on chromosome 2q11-13) encodes the nuclear pore component of RAN binding protein 2 and is associated with acute necrotizing encephalopathy, an autosomal dominant ailment characterized by brain damage that usually follows an acute febrile disease. The RYR3 gene, a ryanodine receptor predominantly expressed in the brain, functions as a key regulator of calcium release from intracellular reservoirs. Its involvement in synaptic plasticity is well-established, and studies with RYR3 knockout mice have demonstrated compromised spatial learning abilities.

Upon analyzing the data from unaffected biological parents, it was observed that two distinct combinations were inherited, with each variant originating from different parents, thus suggesting that the combination of genes may be the cause of the child epilepsy. However, there was a single instance where one parent possessed the exact same combination as their affected child, despite not having a diagnosis of epilepsy. The increased prevalence of this genetic combination, relative to other potential combinations, suggests its potential role as a trigger or risk factor for disease development. However, it is crucial to note that the genetic correlation between these genes is quite low (Correlation AnalyzeR, r= 0.029; p=0.00805). Despite this, the observed pattern raises the possibility that specific gene combinations may influence disease manifestation or susceptibility. Further investigations are warranted to fully comprehend the implications and significance of these findings.

Acknowledgments
This work was supported by Zuckerman STEM Leadership Program and Israel science foundation (ISF) grants 1994/21 and 352/21.

Figure Legends

Figure 1. A graphical description of the probands data
(a) Tests outcomes categorized according to the observed genetic variants. (b) Observed genetic variants in the targeted genes on the NGS Invitae panel ranged from zero to 8 per individual. (c) A classification of the observed variants.

Figure 2. Diagnostic Results and genes with pathogenic findings
(a) Among individuals with positive results, the observed pathogenic variants in the targeted genes on the NGS Invitae panel ranged from 1 pathogenic variant responsible for the condition to two pathogenic variants per individual. (b) The pathogenic genes repeated more than three times in the positive proband data. (c) The pathogenic variants repeated more than once in the positive proband data. (d) Significant enrichment of affected GO biological process in the set of pathogenic genes (FDR < 0.05). (e) Significant enrichment of affected GO cellular components in the set of pathogenic genes (FDR < 0.05).

Figure 3. Uncertain Results and genes with uncertain findings
(a) Among individuals with uncertain results, the observed uncertain variants in the targeted genes on the NGS Invitae panel range from one to six variants per individual. (b) The frequent genes in the uncertain data. (c) The frequent genes with uncertain yield. (d) The uncertain Variant yield frequencies. (e) Significant enrichment of affected GO cellular components in the set of VUS genes (FDR < 0.05). (f) Significant enrichment of affected KEGG pathways in the set of pathogenic genes (FDR < 0.05). (g) Significant enrichment of affected GO biological process in the set of VUS genes (FDR < 0.05).

Figure 4. Genetic Combinations as a hotspot for child epilepsy
(a) Frequent genetic combinations among affected children. (b) Frequent genetic combinations among uncertain results in the SCN5A, Ryr3, and RANBP2 genes. (c) Distribution of De novo, Detected and
not tested variants in parent’s data. (d) Frequent genetic combinations among individuals and unaffected biological parents

**Figure S1**
(a) Included and excluded data summary. (b) The distribution of targeted genes in the panels among individuals, the main epilepsy panel is consisting of 187 genes. (c) The observed uncertain variants in the targeted genes on the NGS Invitae panel range from zero to seven variants per individual.
References


Observed variants

- WWOX  c.1146C>T
- SCN9A  c.2133G>C
- QARS  c.316G>A
- PEX19  c.115C>T
- SLC13A5  c.1366G>A
- RYR3  c.1623C>A
- PTPN23  c.1903G>A
- DIAPH1  c.2902C>T
- PIGN  c.364G>C
- CNTNAP2  c.3925G>A
- SCN8A  c.4885C>T
- GABRG2  c.548G>T

<table>
<thead>
<tr>
<th>Variant</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45%</td>
</tr>
<tr>
<td>VUS classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29%</td>
</tr>
<tr>
<td>VUS finding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15%</td>
</tr>
<tr>
<td># of VUS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6%</td>
</tr>
</tbody>
</table>

VUS finding

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>GO: Cellular Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCN5A</td>
<td>Protein complex</td>
</tr>
<tr>
<td>FASN</td>
<td>Intrinsic to plasma membrane</td>
</tr>
<tr>
<td>RYR3</td>
<td>Integral to plasma membrane</td>
</tr>
<tr>
<td>QARS</td>
<td>Voltage gated calcium channel complex</td>
</tr>
<tr>
<td>PIGG</td>
<td>Plasma membrane part</td>
</tr>
<tr>
<td>SETZ2</td>
<td>Cell projection</td>
</tr>
<tr>
<td>PRDM8</td>
<td>Cell projection part</td>
</tr>
<tr>
<td>BRAT1</td>
<td>Axon</td>
</tr>
<tr>
<td>DIAPH1</td>
<td>Voltage gated potassium channel complex</td>
</tr>
<tr>
<td></td>
<td>Neuron projection</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>GO: Biological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCN5A</td>
<td>Neurosystem development</td>
</tr>
<tr>
<td>FASN</td>
<td>Sodium ion transport</td>
</tr>
<tr>
<td>RYR3</td>
<td>Regulation of action potential</td>
</tr>
<tr>
<td>QARS</td>
<td>Calcium transport</td>
</tr>
<tr>
<td>PIGG</td>
<td>Neuron projection development</td>
</tr>
<tr>
<td>SETZ2</td>
<td>Neuron development</td>
</tr>
<tr>
<td>PRDM8</td>
<td>Neuron differentiation</td>
</tr>
<tr>
<td>BRAT1</td>
<td>Learning or memory</td>
</tr>
<tr>
<td>DIAPH1</td>
<td>Cell-cell signaling</td>
</tr>
<tr>
<td></td>
<td>Regulation of membrane potential</td>
</tr>
<tr>
<td></td>
<td>System process</td>
</tr>
<tr>
<td></td>
<td>Synaptic transmission</td>
</tr>
<tr>
<td></td>
<td>Neurological system process</td>
</tr>
<tr>
<td></td>
<td>Transmission of nerve impulse</td>
</tr>
</tbody>
</table>

![Distribution across observed genes](image_url)
VUS Combinations

Counts

Combination

RANBP2&RYR3
PIGG&RYR3
FASN&RYR3
KCNMA1&ROGDI
DIAPH1&RANBP2
RANBP2&SCN9A
RYR3&SCN9A
CACNA2D2&MTOR
RYR3&SCN5A

Detected Combinations

RANBP2&RYR3
KCNMA1&ROGDI
KCNMA1&SZT2
ROGDI&SZT2
AMT&RYR3
DIAPH1&RANBP2

Variant Combination

0
5
10%

Probands
Parents

4%
91%
4%

De Novo
Detected in One parent
Not tested

Detected Combinations

Variant Combination

RANBP2&RYR3
KCNMA1&ROGDI
KCNMA1&SZT2
ROGDI&SZT2
AMT&RYR3
DIAPH1&RANBP2

%