UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À

L'UNIVERSITÉ DU QUÉBEC À CHICOUTIMI

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN SCIENCES CLINIQUES ET BIOMÉDICALES

PAR

OLEKSANDR OLIFIR

IDENTIFICATION OF PATHOGENIC GENETIC VARIANTS USING WHOLE GENOME SEQUENCING IN PATIENTS WITH EPILEPSY

MARS 2025

Résumé

Trouver l'étiologie génétique de l'épilepsie est une tâche difficile en raison des mécanismes génétiques variables et du grand nombre de variants pathogènes qui n'ont pas été identifiés auparavant. Chaque nouveau variant potentiellement associé à l'épilepsie doit être soigneusement classifié selon les lignes directrices d'interprétation des variants du Collège américain de génétique et de génomique (ACMG). Dans notre étude, nous avons utilisé les séquences du génome entier de 422 patients atteints d'épilepsie et de 2166 témoins pour détecter de nouveaux variants rares et dommageables causant la maladie dans les gènes associés à l'épilepsie. Nous avons démontré que l'analyse computationnelle des variants, basée sur la fréquence dans population et les prédictions *in silico*, améliore l'identification de nouveaux variants délétères probablement pathogènes chez les patients atteints d'épilepsie, même lorsque l'accès à des trios et des études fonctionnelles est limité. De plus, nous avons comparé la charge des variants rares délétères de signification incertaine entre les cas et les témoins. L'analyse de la charge des variants rares de type missense délétères dans des gènes spécifiques associés au phénotype a montré un excès de ces variants dans la cohorte d'épilepsie par rapport aux témoins. Les variants rares de type missense délétères étaient significativement enrichis dans tous les cas d'épilepsie et dans l'épilepsie généralisée génétique (EGG), mais pas dans l'épilepsie focale non acquise (NAFE) par rapport aux témoins. Ces études soulignent l'importance des lignes directrices de l'ACMG pour l'interprétation des variants dans les cas d'épilepsie non diagnostiqués, en particulier pour la pratique clinique et les diagnostics en laboratoire.

Abstract

Finding the genetic etiology of epilepsy is a challenging task due to variable genetic mechanisms and the high number of pathogenic variants which were not identified previously. Every new variant potentially associated with epilepsy must be carefully classified according to variant interpretation guidelines from the American College of Medical Genetics and Genomics (ACMG). In our study, we used whole genome sequences of 422 patients with epilepsy and 2166 controls to detect novel disease-causative rare damaging variants in genes previously associated to epilepsy. We have demonstrated that computational analysis of variants, based on frequency in the population and *in silico* prediction programs, enhances the identification of novel loss-of-function likely pathogenic variants in patients with epilepsy, even when access to trio analysis and functional studies is limited. In addition, we compared the burden of rare deleterious missense variants of uncertain significance between cases and controls. Burden analysis of rare damaging missense variants in phenotype specific genes showed an excess of these variants in the epilepsy cohort compared to controls. Rare damaging missense variants were significantly enriched in all epilepsy cases and in Genetic Generalized Epilepsy (GGE), but not in Non-Acquired Focal Epilepsy (NAFE) when compared to controls. These studies highlight the importance of the ACMG guidelines for variant interpretation in underdiagnosed cases of epilepsy, particularly in clinical practice and laboratory diagnostics.

Table of Contents

List of Abbreviations, Symbols and Acronyms

ACMG……American College of Medical Genetics

CENet……..…… Canadian Epilepsy Network

CAE…… Childhood Absence Epilepsy

DEE…… Developmental Epileptic Encephalopathy

EEG……...…..……Electroencephalography

GGE……..…… Genetic Generalized Epilepsy

GWAS……..……Genome-Wide Association Study

ILAE…… International League Against Epilepsy

JAE…….… Juvenile Absence Epilepsy

JME……...…. Juvenile Myoclonic Epilepsy

LoF……...…...Loss of Function

MAF……..…… Minor Allele Frequency

MRI……Magnetic Resonance Imaging

NAFE…… Non-Acquired Focal Epilepsy

OR……..…… Odds Ratio

SNV…….……Single Nucleotide Variant

VUS……Variant of Unknown/ Uncertain Significance

WES……Whole Exome Sequencing

WGS…… Whole Genome Sequencing

Chapter 1:

Introduction

1. Epilepsy

1.1.Epidemiology

Epilepsy is a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures and is practically applied as having two unprovoked seizures >24h apart[1]. Seizures clustering within 24 hours carry nearly the same risk for subsequent seizures as a single isolated seizure[2]. The term "unprovoked" implies absence of a factor which may lower the seizure threshold at that point in time. Conversely, provoked seizures, may result from electrolyte disorders, toxins, head injury, infectious processes, vascular anomalies, tumors or other mass lesions, and many other causes[3]. In this study, we focus exclusively on epilepsies that do not have identifiable provocative factors.

Epilepsy is a prevalent neurological condition, impacting individuals across diverse demographics, including age, race, socioeconomic status, and geographical regions. International population-based studies established the prevalence of epilepsy is 6.38 per 1000 people and the incidence is 61.44 per 100 000 people every year[4]. Overall prevalence rate of epilepsy in male are higher than those in female. Epilepsy has a strong impact on quality of life, encompassing unpredictable seizures leading to injury, hospitalization, and mortality, as well as rising risk of mental health comorbidities like anxiety and depression[5,6]. The development of mental health comorbidities may be due to the stigma associated with seizures, which can induce distress and lower self-esteem[7]. In addition, epilepsy patients experience higher rates of unemployment and early retirement, along with imposing a significant economic burden on the healthcare system[8,9].

Epilepsy is a complex disease in terms of clinical presentation, and it is important to provide an overview of its classification to better understand the underlying genetic causes and mechanisms of the disease.

1.2. Classification of the epilepsies

The 2017 ILAE Seizure Classification introduces a multi-level approach to classification of the epilepsies,

At the first level, the identification of seizure types is performed. Seizures are classified into focal onset, generalized onset, and unknown onset[10]. Following this, the classification proceeds to diagnose the type of epilepsy, which encompasses focal, generalized, combined generalized and focal, as well as an undefined epilepsy category. Finally, the third level involves identifying specific epilepsy syndromes for a more detailed diagnosis[11].

On the second level, the type of epilepsy should be defined. Diagnosis is made on clinical presentation of seizure type, supported by electroencephalography (EEG) findings. The diagnosis of Generalized Epilepsy requires generalized spike-wave activity on EEG and one of the generalized onset seizures (absence, myoclonic, atonic, tonic or tonic–clonic seizures)[12]. The diagnosis of Focal Epilepsies requires focal epileptiform discharges on EEG and one of the focal onset seizures (focal aware seizures, focal impaired awareness seizures, focal motor seizures, focal non-motor seizures or focal to bilateral tonic–clonic seizures)[11]. If the patient has both generalized and focal seizures, the diagnosis of Combined Generalized and Focal Epilepsies can be established. When clinicians can not determine a specific epilepsy syndrome, identifying the epilepsy type may be the final diagnostic step.

The third level involves diagnosing epilepsy syndrome. Epilepsy syndrome is defined as a characteristic cluster of clinical and EEG features, often supported by specific etiological findings (structural, genetic, metabolic, immune, and infectious)[13]. Epilepsy syndromes usually have shared etiologic, prognostic, and treatment implications, as well as distinctive comorbidities, such as depression and generalized anxiety disorder[11].

In this research, our attention is drawn to two distinct categories of epilepsy: Genetic Generalized Epilepsy (GGE) and Non-Acquired Focal Epilepsy (NAFE). GGE is characterized by generalized epilepsy, where seizures spread to networks of neurons in both cerebral hemispheres, while NAFE constitutes a focal epilepsy, indicating that seizures arise from a network of neurons within a single hemisphere of the brain[14]. Both GGE and NAFE are not associated with structural brain lesions and demonstrate evidence of heritability.

1.3. Syndromes of genetic generalized epilepsy

Patients with GGE experience one or a combination of the following generalized seizure types: absence, myoclonic, tonic–clonic, and myoclonic–tonic–clonic seizures. Generalized tonic, atonic, myoclonic–atonic, and focal seizures and epileptic spasms exclude a diagnosis of GGE. The EEG shows the classical finding of generalized spike- wave discharges, typically 2.5– 5.5 Hz, which are often brought out during drowsiness, sleep, and on awakening. Discharges often appear fragmented during sleep and can have focal features. However, consistent focal epileptiform activity or focal slowing should not occur[15].

The GGE include the following syndromes: childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), epilepsy with generalized tonic–clonic seizures alone (GTCA), epilepsy with eyelid myoclonia (EEM),

epilepsy with myoclonic absences (EMA) and myoclonic epilepsy in infancy (MEI). CAE, JAE, JME, and GTCA form a distinct subgroup of the genetic generalized epilepsies, collectively termed Idiopathic Generalized Epilepsies (IGE)[16].

The classical syndromes of GGEs are classified based on seizure types and age of onset. In this section, we provide clinical details of the GGE syndromes observed in our cohort from the Canadian Epilepsy Network (CENet) project[17].

Childhood absence epilepsy (CAE). Seizures occur many times daily and consist of brief staring spells, sometimes with rhythmic eye blinking or motor automatisms, lasting seconds, with immediate return to the baseline level of awareness and activity[18]. Generalized spike-wave and absence seizures are provoked by hyperventilation in most untreated patients[19]. This epilepsy syndrome has good prognosis and remits in 60% of children within 2 years of onset or by early adolescence[15].

Juvenile absence epilepsy (JAE). Seizures characterized by absence seizures and generalized tonic-clonic seizures. In some patients, there may be additional myoclonic jerks as well[20]. It can be challenging to differentiate between JAE and CAE. One way to distinguish them is by noting the older age at onset and the lower frequency of absence seizures in JAE[15].

Juvenile myoclonic epilepsy (JME). This type of genetic generalized epilepsy is characterized by the presence of absence, myoclonic, and generalized tonic-clonic seizures[21]. Myoclonic seizures typically occur shortly after waking and when tired. Sleep deprivation is an important provoking factor. JME is usually considered a lifelong disorder, often requiring lifelong therapy[15].

Other genetic generalized epilepsy syndrome includes epilepsy with eyelid myoclonia (EEM). EEM (previously known as Jeavons syndrome) is characterized by frequent eyelid myoclonia, with or without absences, induced by eye closure and photic stimulation.

1.4. Syndromes of non-acquired focal epilepsies

The group of focal epilepsy syndromes, which can present at various ages, encompasses several different syndromes. According to International League Against Epilepsy classification, non-acquired focal epilepsies can be classified to the next major groups: focal self-limited epilepsy syndromes of infancy and childhood and focal epilepsy syndromes with onset at a variable age.

In this section, we provide clinical details of the NAFE syndromes observed in our cohort. Some syndromes were identified through genetic testing in cases where the NAFE phenotype was not initially specified.

The group of self-limited focal epilepsies share common features, typically beginning in early childhood and resolving before puberty. *Self-limited neonatal epilepsy* usually starts within the first week of life, presenting as focal tonic or clonic seizures in an otherwise healthy infant. This epilepsy syndrome typically resolves by 6 weeks of age. In contrast, *self-limited infantile epilepsy* begins after the neonatal period, typically before 18 months of age[22,23].

The group of focal epilepsy syndromes with onset at a variable age includes sleep-related hypermotor epilepsy (SHE), familial mesial temporal lobe epilepsy (FMTLE), familial focal epilepsy with variable foci (FFEVF), and epilepsy with auditory features.

Sleep-related hypermotor epilepsy (SHE). Syndrome that is characterized by clusters of nocturnal motor seizures that are often stereotyped and brief (<2 minutes). They vary from simple arousals from sleep to dramatic, often hyperkinetic events with tonic or dystonic features. Seizures occur only during wakefulness. Affected individuals may experience an aura. Age of onset ranges from infancy to adulthood[24].

Familial mesial temporal lobe epilepsy (FMTLE). Syndrome of adolescence or early adulthood that is characterized by focal aware seizures with prominent psychic or autonomic features which rarely progress to bilateral tonic-clonic seizures. These focal seizures most commonly manifest as either intense déjà vu, dream-like states, stereotyped flashbacks, fear, epigastric discomfort, olfactory sensations. Hippocampal sclerosis, characterized by neuronal loss and scarring, is the most common histopathological abnormality found in patients with drug-resistant temporal lobe epilepsy.[25,26].

Familial focal epilepsy with variablefoci (FFEVF). Familial partial epilepsy with variable foci is marked by focal seizures that arise from different regions of the brain. FFEVF has an association with focal cortical dysplasia.[27]

Some patients with focal epilepsy syndromes such as FMTLE-HS, SHE, and FFEVF may experience cognitive, neurologic, or psychiatric impairments, with a poor prognosis that can evolve into developmental and/or epileptic encephalopathy[28].

Most epilepsy syndromes have a genetic etiology, involving both monogenic and polygenic mechanisms. Identifying the genetic causes of epilepsy is crucial for the management and prognosis of the disease.

2. Genetics in epilepsy

Epilepsy without structural brain lesions or other underlying health conditions are known to have a genetic basis, typically attributed to monogenic and polygenic etiology, as well as complex interactions between genetic factors and environmental influences. Genetic testing is strongly recommended for all individuals with unexplained epilepsy with exome/genome sequencing or a multi-gene panel followed by chromosomal microarray[29]. Genetic testing in patients with epilepsy can guide clinical decision-making and improve patient outcomes. Identifying a genetic etiology for epilepsy can help avoid unnecessary tests and procedures, inform appropriate medication choices (indicated vs. contraindicated), and provide a clearer prognosis for the disease's progression.

Epilepsy genetic research has progressed considerably, from early linkage studies to the use of advanced sequencing technologies. *CHRNA4*, the first gene identified in connection with epilepsy, was discovered in 1995 and has been associated with sleep-related hypermotor epilepsy. Later, mutations in genes that regulate ion channels, such as *KCNQ2* in benign familial neonatal seizures and *SCN1A* in Dravet syndrome, have been discovered[30,31]. Now, over 900 genes have been identified as being linked to epilepsy[32]. Modern genomic technologies enable the detection of common genetic variants, and the analysis of copy number variants (CNVs) linked to epilepsy through genome-wide approaches. Furthermore, rare single nucleotide variants can be identified using methods like exome sequencing or whole-genome sequencing[33].

2.1.Monogenic epilepsies

Monogenic epilepsies are usually associated with ultra-rare variants in coding regions which are commonly investigated through exome sequencing. Ultra-rare *de novo* genetic

variants often cause Mendelian diseases and have a significantly greater impact on individual risk, but their contribution to overall population heritability is limited[34].

Most of the monogenic epilepsies identified to date are linked to developmental and epileptic encephalopathies (DEE)[35]. The monogenic etiology can be identified in up to 40% of patients with DEE[36]. Many monogenic DEEs include channelopathies, where pathogenic mutations result in a gain or loss of function of voltage-gated or ligand-gated ion channels as in *GRIN2A, SCN2A, SCN1A,* and *SCN8A* genes. Monogenic DEE also include metabolic conditions (*GLUT1* deficiency syndrome caused by *SLC2A1* mutations), synaptopathies (*VAMP2*-related neurodevelopmental disorders), alteration in neuronal proliferation, migration, and differentiation (*HECW2*-related DEE), and others[37].

A proportion of GGE phenotypes which are caused by variants in single genes is very limited with diagnostic yield up to 3 %[38]. Pathogenic variants in the *SLC2A1, GABRA1, GABRD, HCN2* and *GABRAG2* genes are responsible for rare sporadic or familial cases of generalized epilepsies[32,37,39]. Based on a literature review, childhood absence epilepsy has been reported in patients with pathogenic variants in the *GABRG2, SLC2A1, GABRA1*, and *GABRB3* genes. Juvenile absence epilepsy has been associated with variants in the *CLCN2* and *EFHC1* genes. The juvenile myoclonic epilepsy phenotype has been linked to pathogenic variants in the *CACNB4, CLCN2, EFHC1, GABRD*, and *GABRA1* genes[15,40].

Exome testing reveals a diagnostic yield of around 1 % in NAFE[41]. Examples of monogenic NAFEs include autosomal dominant sleep-related hypermotor epilepsy, caused by mutations in *DEPDC5, CHRNA4, CHRNA2, CHRNB2, NPRL2, NPRL3, KCNT1*; familial focal epilepsy occasionally reported with *DEPDC5, NPRL2, NPRL3*. Autosomal dominant

lateral temporal lobe epilepsy caused by *DEPDC5* mutations. *KCNQ2* and *KCNQ3* are involved in familial neonatal epilepsy, and *PRRT2, SCN2A,* and *SCN8A* pathogenic variants cause familial infantile epilepsy[25].

2.2.Large-scale sequencing studies on epilepsy

Advancements in genome sequencing technologies have opened up new opportunities for identifying genetic variants associated with neurological disorders[42]. Rare genetic variations have been studied in epilepsy through whole exome sequencing (WES) and whole genome sequencing (WGS). Studies on rare variants require cohorts that are as large as possible to allow new statistically significant discoveries in case-control studies.

A study by Feng *et al.*[43] involving 9,170 cases and 8,436 controls found excess of ultra-rare, deleterious variants in constrained genes and in genes previously associated with epilepsy. It showed a significant mutational burden in highly constrained epilepsy genes among individuals with GGE and NAFE.

More recent whole-exome sequencing study of epilepsy conducted by Epi25 Collaborative *et al.*[44] included 20,979 cases and 33,444 controls, to investigate rare variants that confer disease risk. This study focused on the enrichment of ultra-rare genetic variants in individual genes, gene sets, and copy number variants. It was shown that genes encoding ion channels show strong association with multiple epilepsy subtypes. Top candidate genes with higher rare variant enrichment were found in those that play roles in synaptic transmission and neuronal excitability. This study demonstrates that rare coding variants in genes highly expressed in the brain are risk factors for the epilepsy.

Current understanding suggests that the majority of epilepsy cases are likely to have a polygenic structure, involving not only rare but also common genetic variants[45]. The discovery of common variants associated with known complex diseases can be conducted with genome-wide association studies (GWAS)[46].

2.3.Genome-wide association studies and polygenic risk score

In recent years, genome-wide association studies (GWAS) in epilepsy have focused on identifying candidate common genetic variations linked to increased epilepsy risk.

The most recent meta-analysis conducted by the ILAE in 2022 included 29,944 cases and 52,538 controls[47]. Four loci were significantly associated with all epilepsies in *BCL11A, SCN1A, RORB* and *KCNIP2* genes. For GGE, 22 loci were significantly associated with the condition. Three genome-wide significant loci were associated specifically with JME in *GABRA2, CACNA1I* and *TMEM74* genes, and one locus with CAE in *BCL11A* gene. No locus reached significance for NAFE, JAE and GTCSA cases.

Another important GWAS was conducted by Song et al. in 2021[48]. They performed a meta-analysis using data from the 2018 ILAE GWAS[49], UK biobank[50], Japanese population[51] and FINNGEN[52], covering a total of 26,352 cases and 774,517 controls. They identified three significant risk loci, one at 2q24.3 within the SCN1A gene, which had been previously reported in other studies. Additionally, they found novel associations at 7q21.11 and 8p23.1. The genes associated with these loci include *GRM3* and *TNKS*.

Although GWAS have offered important insights into disease mechanisms, the contribution of individual GWAS loci to the disease is relatively small. It is well established that common genetic variants with small effects on specific diseases can be combined as

polygenic risk score (PRS). Recent studies have demonstrated that individuals with epilepsy possess a notably higher epilepsy PRS compared to those without the condition[53]. A study conducted by *Moreau et al.*[54] estimated that PRS explains more of the variance in idiopathic generalized epilepsy than in patients with nonacquired focal epilepsy. Despite the potential of PRS to quantify the cumulative effects of a number of genetic variants, it is not widely used for predicting epilepsy risk in clinical practice[55].

2.4.Other genetic mechanisms

New genetic mechanisms have been discovered to play a role in the development of epilepsy, such as effects of non-coding regulatory variants and repeat expansion[56,57]. In addition, somatic mosaicism has been shown to be responsible for certain focal epileptogenic lesions[58]. Moreover, epigenetic mechanisms, including DNA methylation, histone modifications, and non-coding RNAs, play pivotal roles in altering gene expression and synaptic plasticity, contributing to epileptogenesis[33].

Complex genetic etiology highlights the importance of continued research into the genetic mechanisms of epilepsy, particularly in uncovering pathogenic variants that can be used in clinical settings for the diagnosis and management of epilepsy. Since common genetic variation has been extensively studied in epilepsy, our studies focused exclusively on rare, damaging genetic variants in coding regions with the goal of identifying novel pathogenic variants, using the ACMG guidelines for variant interpretation.

3. ACMG guidelines overview

The issue with undiagnosed epilepsy cases lies in the high number of pathogenic variants that have not been previously identified. Every new variant potentially associated

with epilepsy must be carefully classified according to variant interpretation guidelines from the American College of Medical Genetics and Genomics (ACMG) and the American Association of Molecular Pathology (AMP)[59]. The approach for classifying variants is designed to be applicable to those associated with Mendelian diseases, regardless of whether they are identified using single-gene tests, multi-gene panels, exome sequencing, or whole-genome sequencing.

According to ACMG/AMP guidelines, detected variants can be classified, either as pathogenic (P), likely pathogenic (LP), variant of unknown/uncertain significance (VUS), likely benign (LB), or benign (B)[59]. In this framework, "likely" indicates a confidence level of over 90% that a variant is either pathogenic or benign. VUS is applied in two situations: variants with conflicting evidence, or variants with insufficient evidence regarding their functional characterization or clinical significance. P/LP variants can influence clinical decision-making for individual or familial screening and treatment recommendations[60].

The ACMG/AMP guidelines assign codes to each piece of evidence based on its category (population data, computational predictions, functional data, or segregation data, de novo data) and its strength: benign supporting (BP1–6), strong (BS1–4), or stand-alone (BA1), or pathogenic supporting (PP1–5), moderate (PM1–6), strong (PS1–4), or very strong (PVS1) (Figure 1). The evidence codes are then analyzed together using established rules for combining criteria, which classify sequence variants into a five-tier system ranging from benign to pathogenic[59](Figure 2).

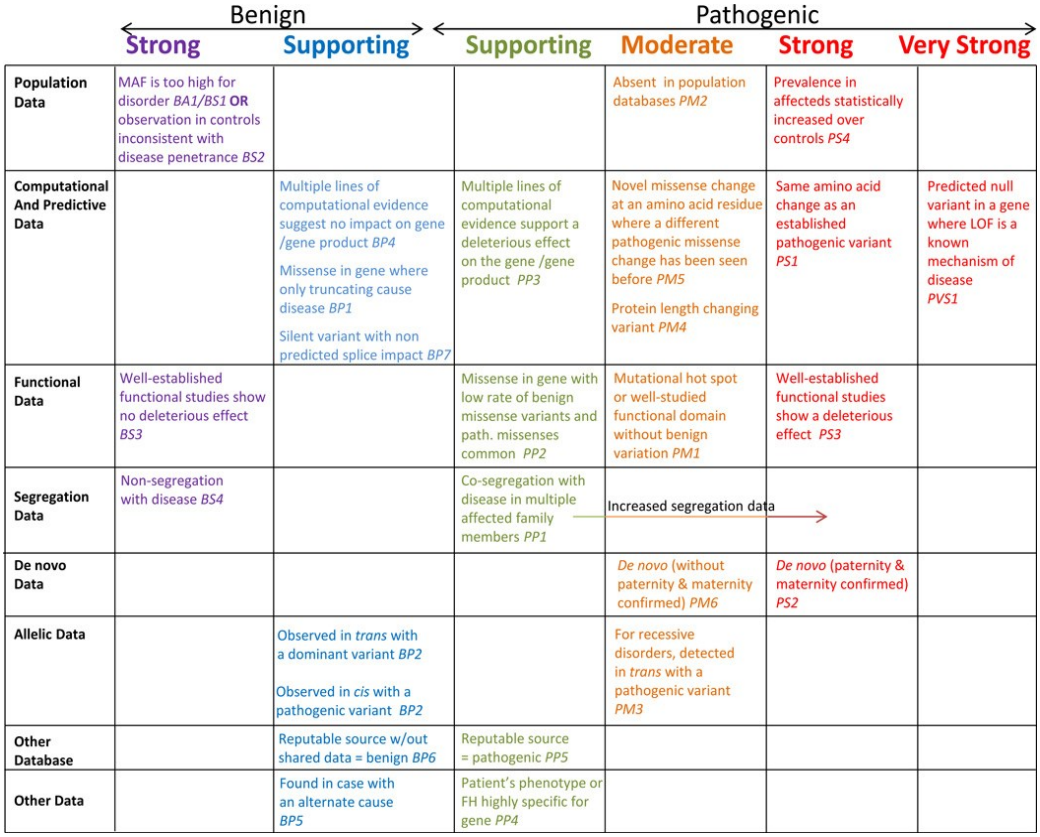

Figure 1. Evidence Framework with criteria from ACMG guidelines. From (with authorization): Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine. Volume 17, Issue 5, May 2015, Pages 405-424. Richards et al.

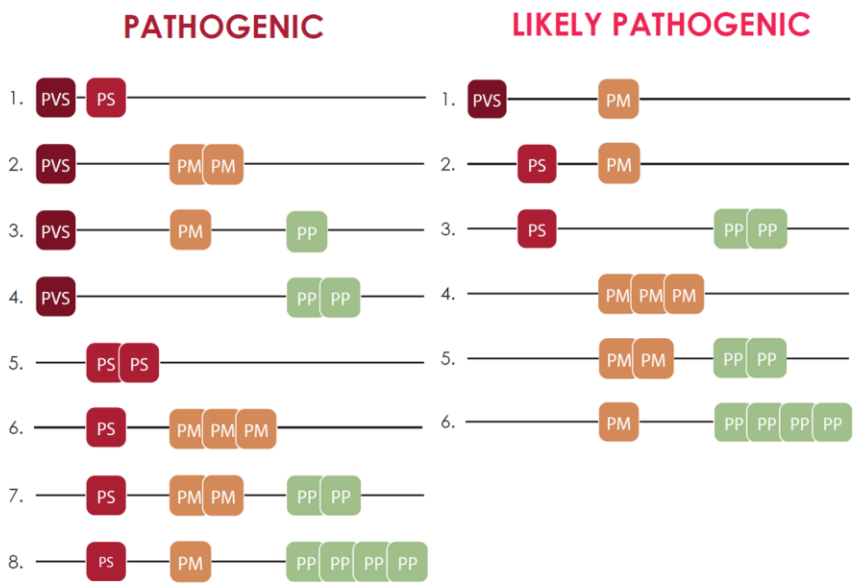

Figure 2. Combinations of criteria required to classify variants as Pathogenic and Likely Pathogenic. From (with authorization): Mastermind Variant Interpretation Cards, by Genomenon.

Since the ACMG guidelines were first published in 2015, they have undergone numerous modifications introduced by various laboratories and working groups. Specifically, ClinGen recommends using the original criteria code followed by an underscore and the updated strength level. For example, a criterion originally labeled as "PVS1" (Pathogenic Very Strong) could be updated to "PVS1_strong" or "PVS1_moderate," depending on new evidence or interpretation[61]. This approach establishes quantitative parameters for criteria by incorporating a point-based grading system. At the time of analysis, the ACMG had not published updated recommendations, so we are using the guidelines proposed in 2015.

The following sections outline the main categories of variant classification, focusing on each category and the associated pathogenic criteria.

3.1. Population frequency

Population frequency data help distinguish rare variants, which are more likely to cause Mendelian disorders, from the millions of common and mostly benign variants found in every human genome. When using population databases, it's important to check if healthy or disease cohorts were used, if multiple family members were included, and the subjects' age range.[62].

Two criteria may be applied in the population category:

PS4: This criterion is used when the prevalence of a variant is significantly higher in affected individuals compared to controls in case-control analysis. However, it is infrequently applied due to the small sample sizes in studies of rare variants, which often lack sufficient statistical power.[63].

PM2: This criterion applies when a variant is absent or rare in population databases. The gnomAD population database is often used to define allele frequency of variants[62].

3.2.Computational prediction

Computational (*in silico*) predictive programs determine the potential impact of the variant on the protein structure and function. To assess the pathogenicity of variants, various tools analyze the conservation of the region using a range of mathematical and biochemical approaches to determine the extent to which the variant disrupts a conserved region. While these tools are highly advanced, they should not be used alone to classify a variant[64].

Five criteria may be applied in the computational prediction category:

PVS1: Applied when a variant is predicted to cause a null protein product, typically in cases where loss-of-function (LoF) is the known disease mechanism, such as with nonsense mutations, frameshifts, or large deletions.[63].

PS1: Used when a variant causes the same amino acid change as a previously established pathogenic variant.

PM4: Applied for in-frame deletions or insertions in a non-repeat region or stop-loss variants. According to the new ClinGen recommendations, PM4 should not be used for variants where PVS1 is also applied to prevent double-counting of evidence.[61].

PM5: Used for novel missense variants where a different missense change has been reported at the same amino acid residue.

PP3: Applied when multiple lines of computational evidence indicate that a variant is likely to have a deleterious effect on the gene or its product.

3.3.Functional studies

Functional studies can provide powerful insight into the effect of a variant on protein function and have the capacity to reclassify variants[65]. Functional assays can vary significantly depending on the type of disease and variant being studied and may include splicing assays, animal and cellular models, as well as various in vitro systems[66].

Three criteria may be applied in the computational prediction category:

PS2: Applied when well-established functional studies show a deleterious effect.

PM1: Applies only to missense variants located in a mutational hot spot and/or a critical, well-established functional domain.

PP2: Used when a missense variant is found in a gene with a low rate of benign missense variation, indicating that the gene is more functionally constrained.

3.4.Trio analysis

Trio testing assesses the pathogenicity of detected variants through testing of both the parental and proband samples and has been widely used for the diagnosis of various hereditary disorders. Trio test is useful for variant interpretation only if the patient's phenotype aligns with the disease gene association[67].

Two criteria may be applied in the trio analysis category:

PS2: Applied when de novo evidence is assessed and both parental relationships are verified, either through trio exome/genome sequencing or a panel of informative genetic markers.

PM6: Applied if only one parental relationship has been tested.

3.5.Segregation analysis

Segregation data can be examined using information from a single family that has several affected individuals with a particular variant. The guidelines also indicate that as cosegregation with disease in multiple affected family members becomes, the strength of PP1 can be elevated from supporting evidence up to strong evidence[68].

3.6.Allelic data

Allelic data can be used when two different variants are found in a gene linked to a recessive disorder. If one of these variants is already confirmed as pathogenic, determining

that the second variant is in trans (on the opposite chromosome) can provide moderate evidence for its pathogenicity. This is referred to as PM3.

3.7.Other data

If the patient's phenotype or family history is highly specific for a disease with a single genetic etiology, the PP4 criterion can be applied. The key consideration is the strong correlation between the phenotype and the associated gene.

If a reputable source has recently reported a variant as pathogenic, but the evidence for independent evaluation is not available, the PP5 criterion may be applied. However, ClinGen recommends that PP5 should not be used[69].

4. Objectives

The main objective of this study was to determine the molecular genetic diagnostic rate of epilepsy using disease databases and identify novel pathogenic genetic variants using ACMG criteria in patients with genetic generalized epilepsy and non-acquired focal epilepsy. Specifically, the aim was to determine whether these variants exhibit genotype-phenotype relationships with known epilepsy syndromes. Another objective was to perform a burden analysis of rare, damaging missense variants in GGE and NAFE to investigate whether there is variant enrichment in one phenotype compared to the other. To accomplish these tasks, we used a method to prioritize variants of interest. The chosen methodology is based on ACMG criteria guidelines and implementation of *in silico* prediction programs.

Chapter 2:

Materials and Methods

1. Cohort

1.1.Overview

The epilepsy cohort represented in this study was taken from the Canadian Epilepsy Network (CENet) project.[17]. The epilepsy cohort consists of families with at least three affected individuals with Generalized Genetic Epilepsy (GGE) and with Non-Acquired Focal Epilepsy (NAFE). Among the genetic generalized epilepsy cases, we had patients with JME, CAE, and JAE (Table 1). These cases were previously collected from the CHUM Research Center in Montreal and the Hospital for Sick Children in Toronto. Genome sequencing was used to establish a genetic diagnosis that could not be determined using other diagnostic methods. For this project only WGS data from unrelated GGE and NAFE patients was used.

A control group composed WGS of 2166 individuals from the CARTaGENE project was used to perform burden analysis. The CARTaGENE is a population-based biobank and the largest ongoing prospective health study of men and women in Quebec[70].

Table 1. Number of individuals for each phenotype

Phenotype	Total
All Cases:	422
females	226
males	184
unknown	12
GGE, including:	208
JME	47
CAE	24
JAE	7
NAFE	214
Controls	2166

1.2.Phenotyping

The diagnosis of patients was carried out by a neurologist at the CHUM Research Center in Montreal and the Hospital for Sick Children in Toronto. The ILAE guidelines were followed for diagnosis and, when possible, for identifying the clinical syndrome. Patients with clinical and EEG characteristics meeting the GGE and NAFE syndromes were included in cohort. An MRI of the brain was used to detect any potentially epileptogenic lesions. Additional clinical information was available upon request from clinical charts, including details on comorbid conditions and any observed developmental or cognitive impairments.

Sequencing

DNA was extracted from the blood samples and was sent to Genome Quebec Innovation Center in Montreal for sequencing. WGS was made at 30X coverage. Genomic

DNA (gDNA) was cleaned using ZR-96 DNA Clean & ConcentratorTM-5 Kit (Zymo) prior to being quantified using the Quant-iTTM PicoGreen dsDNA Assay Kit (Life Technologies) and its integrity assessed on agarose gels. Libraries were generated using the TruSeq DNA PCR-Free Library Preparation Kit (Illumina) according to the manufacturer's recommendations. Libraries were quantified using the Quant-iTTM PicoGreen Double Stranded DNA (dsDNA) Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). The average size fragment was determined using a LabChip GX (PerkinElmer) instrument. The libraries were denatured in 0.05N NaOH and diluted to 8pM using HT1 buffer. The clustering was done on an Illumina cBot and the flowcell was run on a HiSeq 2500 for 2×125 cycles (paired-end mode) using v4 chemistry and following the manufacturer's instructions. A phiX library was used as a control and mixed with libraries at 0.01 level. The Illumina control software used was HCS 2.2.58 and the real-time analysis program used was RTA v. 1.18.64. bcl2fastq v1.8.4 was used to demultiplex samples and generate fastq reads[71]. Sequencing was carried out by the Genome Quebec team without the involvement of the authors and were transferred for further genomic analysis, including the one conducted in this project.

The DRAGEN Germline Pipeline v3.10.4 was used to align the FASTQ on the Human hg38 reference genome and to call variants. The GVCF files produced were then used in the DRAGEN Joint Genotyping Pipeline v3.10.4 to produce a VCF file for patients and controls separately. Variants were annotated using dbNSFP database and classified according to the ACMG standards.

2. Variant analysis

2.1.Variant database

Disease databases are used to identify genetic variants that have been previously reported as pathogenic and linked to specific disease phenotypes. Disease databases are useful for gathering information, but they should be used with caution due to the large number of variants with conflicting evidence of pathogenicity and cases of misclassification.

All variants were assessed for evidence of pathogenicity in the ClinVar database (version of January 10th, 2023). ClinVar is a freely accessible archive of reports of the relationships among genomic variants and phenotypes submitted by laboratories, researchers and clinicians. The total number of pathogenic variants extracted from ClinVar for epilepsy phenotype was 4,139 and included different types of variations (single nucleotide variants, deletions, duplications, insertions, and indels) associated with both isolated forms of epilepsy and syndromes in which seizures are combined with other symptoms. All these variants were assessed for their presence in our epilepsy cohort.

2.2.Filtering and validation of rare damaging variants

To detect novel pathogenic/likely pathogenic variants, we performed variant filtering of WGS from an epilepsy cohort using ACMG guidelines, focusing on population, computation, and functional criteria. (Figure 3). Variant filtering was performed using a Python script.

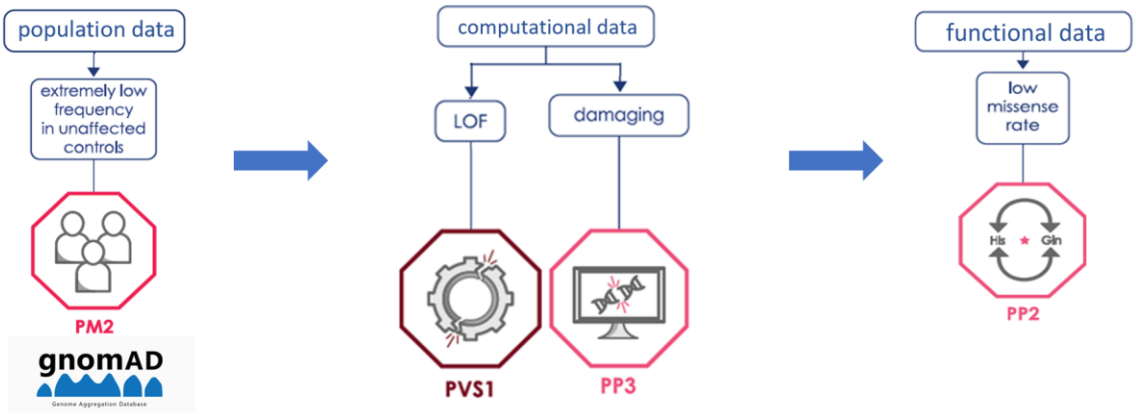

Figure 3. Criteria from ACMG guidelines that were used in the study.

We generated a list of 356 epilepsy-related genes based on commercially available epilepsy gene panel Invitae and review of literature[72]. Genes intolerant to loss-of-function (LoF) variants were identified with a pLI score > 0.9, and genes intolerant to missense variants were identified with a Z-score > 3.09 (functionally constrained genes). If LoF variants were found in genes intolerant to LoF, the PVS1 criterion was applied; if missense variants were found in functionally constrained genes, the PP2 criterion was applied. Next, only variants with a minor allele frequency (MAF) ≤ 0.01 in the Genome Aggregation Database (gnomAD) were retained, which aligns with the application of the PM2 criteria. Finally, missense variants were filtered by pathogenicity predictions of PolyPhen-2 damaging or probably damaging, SIFT damaging and CADD_PHRED scaled score ≥ 20[73]. PP3 was applied to variants predicted as deleterious by the in silico programs listed above.[59].

All novel identified variants were verified using variant description validation software, Variant Validator[74]. HGVS nomenclature were used for reporting sequence variants[75]. Transcript topographies were created using Protter[76]. The topography of pathogenic variants in the tolerance landscape was visualized using MetaDome[77].

2.3.Burden analysis

A burden analysis of rare damaging missense variants of uncertain significance (VUS) was conducted to look for any enrichment of these variants in the patients[46]. The criteria for inclusion in the burden analysis were as follows: the variant had to be predicted as damaging by in silico tools (PP3) and possess a low allele frequency in the general population (PM2). Variants that satisfy these criteria are referred to as "qualifying variants".

To investigate the presence of rare deleterious genetic variants, we conducted two separate burden analyses for two subsets of genes: one with a list of phenotype-specific genes and another with missense-intolerant epilepsy-related genes. In total, 18 phenotype-specific genes and 80 genes related to epilepsy with missense intolerance were selected (Table 2). A list of phenotype-specific genes was generated based on a literature review, including genes with previously reported pathogenic variants specifically linked to GGE and NAFE syndromes. The set of missense-intolerant epilepsy-related genes was generated based on a Z-score > 3.09 and includes all epilepsy-related genes, without a strict link to GGE and NAFE phenotypes, including those associated with epileptic encephalopathy.

The Human Protein Atlas was used to categorize proteins based on their functional roles, resulting in three functional groups. "Voltage-gated/Ligand-gated ion channels" and "Transporters, other than ion channels," were separated based on their mechanisms, as ion channels typically do not require energy and allow ions through the concentration, in contrast to transporters, which require energy and are unidirectional against the gradient[78]. The "Others" category included genes involved in transcriptional regulation, intracellular signaling, enzymatic activity, and other processes.

Table 2. Genes used in burden analyses

Phenotype-specific genes	
Phenotype	Genes
GGE genes	*CACNB4, CLCN2, GABRD, SLC2A1, HCN2, EFHC1, GABRA1, GABRG2, GABRB3, CASR, RORB*
NAFE genes	*CHRNA2, CHRNA4, KCNT1, NPRL2, NPRL3, DEPDC5, CHRNB2*
Missense intolerant epilepsy-related genes	
Functional category	Genes
Voltage-gated/ Ligand-gated ion channels	*CACNA1A, CACNA1B, CACNA1E, GRIN2B, GRIN2D, HCN1, KCNA1, KCNA2, KCNB1, KCNC1, KCND2, KCNH1, KCNH2, KCNQ2, KCNQ5, SCN1A, SCN3A, SCN8A, HCN2, GRIN2D, GABRA2, GABRB2, GABRA5, GRIN1, GRIN2B*
Transporters, other than ion channels	*AP2M1, ATP1A1, ATP1A2, ATP1A3, ATP6V0A1, CLCN4, CLTC, DDX3X, DNM1L, FLNA, GABBR2, GNB1, SLC2A1, SLC12A5, SLC6A1, SLC6A8, STXBP1, CASR, TUBA1A*
Other	*ARHGEF9, ATRX, ACTL6B, CAMK2B, CASK, CHD2, CTNNB1, CYFIP2, DNM1, DOCK7, DYRK1A, EEF1A2, FBXO11, GATAD2B, GNAO1, GPHN, HNRNPU, KIF2A, NACC1, NEDD4L, PACS1, PAFAH1B1, PLCB1, PPP2CA, PPP2R1A, PPP2R5D, PPP3CA, PTEN, PURA, RORB, SATB2, SMC1A, SPTAN1, STAG2, TBL1XR1, TUBB2A, NF1, TUBB2B, SYNGAP1*

The cohort allelic sums test (CAST) was used in our burden analysis to estimate the number of case and control subjects carrying at least one qualifying variant in each subset of genes. This test assumes that the presence of at least one qualifying variant increases disease risk. If an individual was carrying multiple qualifying variants, they were counted as one to prevent overestimation[79]. A two-sided Fisher's Exact Test was used to evaluate the association between the enrichment of rare variants across the cohort and control groups. Gene-focused burden was conducted with logistic regression model using glmnet library.

Chapter 3:

Results

1. Variant identification

Based on the results of variants identified in the epilepsy cohort, we grouped all variants into three categories: previously classified pathogenic variants from ClinVar, novel likely pathogenic variants, and variants of uncertain significance (VUS) (Figure 4). It is important to note that some VUS were found in individuals with pathogenic or novel pathogenic variants.

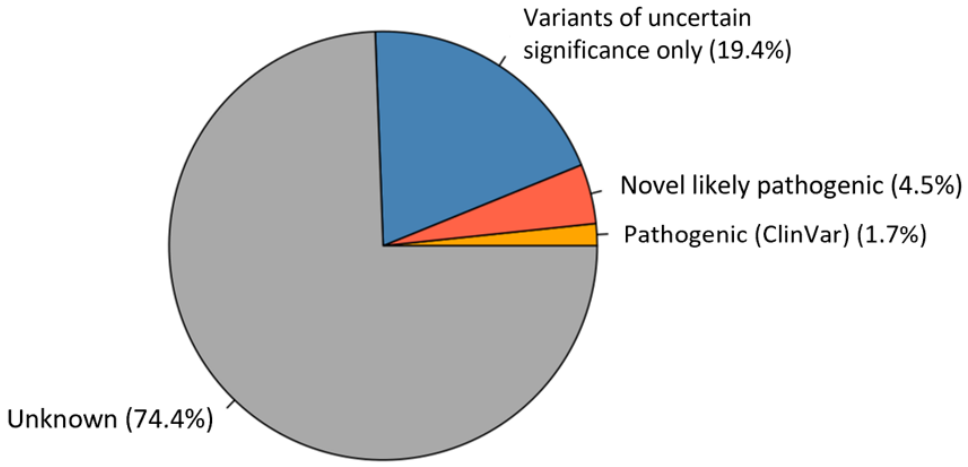

Figure 4. Diagnostic rate among cohort (% of patients carrying a specific group of variants).

1.1.Diagnostic rate according to ClinVar

Seven patients carried a pathogenic variant previously reported in ClinVar to be associated with seizures and that could explain the phenotype. Variants were found in heterozygous conditions and associated with epileptic syndromes with autosomal-dominant inheritance based on OMIM data. This accounts for 1.7 % of the 422 epilepsy cases (Table

3). Aside from this, five patients were carriers of pathogenic variants for the following autosomal recessive diseases: pyridoxal phosphate-responsive seizures NM_018129.4(PNPO): (c.674G>A (p. Arg225His)), progressive myoclonic epilepsy type 2b NM_198586.3(NHLRC1):(c.76T>A (p.Cys26Ser)), and congenital disorder of glycosylation type I in three patients NM_000303.2(PMM2): (c.422G>A (p. Arg141His)).

Table 3. Pathogenic variants in cohort identified with ClinVar Database

ClinVar Preferred name/ Zygosity	Phenotype in ClinVar (Inheritance) OMIM	Phenotype in cohort
NM_001242896.3(DEPDC5): c.2512C>T (p. Arg838Ter), het	Epilepsy, familial focal, with variable foci type 1 (AD) 604364	NAFE
NM_003042.4(SLC6A1): c.1024G>A (p. Val342Met), het	Myoclonic-atonic epilepsy (AD) 616421	GGE
NM_003042.4(SLC6A1): c.1024G>A (p. Val342Met), het	Myoclonic-atonic epilepsy (AD) 616421	GGE, CAE
NM_172107.4(KCNQ2): c.628C>T (p. Arg210Cys), het	Seizures, benign familial neonatal type 1 (AD) 121200	NAFE
NM_004287.3(GOSR2): c.430G>T (p. Gly144Trp), het	Progressive myoclonic epilepsy type 6 (AD) 614018	NAFE
NM_020822.3(KCNT1): c.2012C>T (p. Thr671Ile), het	Autosomal dominant nocturnal frontal lobe epilepsy type 5 (AD) 615005 Developmental and epileptic encephalopathy type 14 (AD) 614959	GGE, JME
NM_001127644.2(GABRA1): c.640C>T (p. Arg214Cys), het	Epilepsy, juvenile myoclonic type 5 (AD) 611136	NAFE
NAFE, non-acquired focal epilepsy; GGE, Genetic generalized epilepsy; CAE, Childhood absence epilepsy; JME Juvenile myoclonic epilepsy; het, heterozygous, AD, autosomal-dominant.		

1.2.Novel likely pathogenic variants

To detect novel likely pathogenic variants, we adopted an epilepsy-gene approach, focusing on 356 genes across all our proband cases. We detected 16 loss-of-function likely pathogenic variants in 19 probands, corresponding to 4.5% of patients in the epilepsy cohort.

These variants included multiple nonsense and frameshift variants identified in heterozygous condition (Table 4).

The identified LoF variants exhibited low allele frequency (PM2) in the general population based on GnomAD. When combined with the PVS1 criteria (null variant in epilepsy genes intolerant to loss-of-function), they were classified as likely pathogenic according to the ACMG rules for variant classification. Due to ACMG guidelines, the detection of variants in multiple unaffected individuals from the control group was considered strong evidence for a benign interpretation (BS2) and, therefore, these variants were not included in the list of novel likely pathogenic variants. None of the variants presented in Table 4 were found among unaffected control individuals.

Table 4. Novel likely pathogenic loss-of-function variants

Gene	Variant/Zygosity	GnomAD frequency	Age of seizure onset	Form of epilepsy
SYNGAP1	NM_006772.3:c.3963_3967del (p.Ala1322fs), het	0.0000065	9	NAFE
SYNGAP1	NM_006772.3:c.3962dup (p.Ala1322fs), het	0.000699	14	NAFE Focal activity: anterior regions of the brain.
SYNGAP1	NM_006772.3:c.1151_1152insTGGGGGGGGGGGGGGGGGCGGGGGGGGGGG (p.Ser385fs), het	NA	7	GGE, CAE
			NA	NAFE
			NA	NAFE
SYNGAP1	NM_006772.3:c.1065del(p.Arg356fs), het	NA	NA	NAFE
CACNA1A	NM_001127222.2:c.3829C>T (p.Arg1277*), het	NA	4.5	GGE
CACNA1A	NM_001127222.2:c.3408_3409insGG (p.Pro1137fs), het	NA	12	NAFE Focal activity: temporal lobes bilaterally

CACNA1B	NM_001243812.2:c.6558G>A (p.Trp2186*), het	0.000013	16	GGE, JAE
SLC12A5	NM_001134771.2:c.121+1090_121+1091insGGC (p.Arg86fs), het	NA	14	NAFE Focal activity: left temporal lobe.
SCN1A	NM_001165963.4:c.5797C>T (p.Arg1933*), het	NA	NA	GGE Jeavons Syndrome
NACC1	NM_052876.4:c.956_957del (p.Val319fs), het	NA	NA	GGE Jeavons Syndrome
KIF1A	NM_001244008.2:c.2840del (p.Leu947fs), het	0.0000065	9	GGE, CAE
			NA	NAFE
CHD2	NM_001271.4:c.1453C>T (p.Arg485*), het	NA	NA	GGE Jeavons Syndrome
DNM1	NM_001005336.3:c.2434del (p.Ala812fs), het	0	7	GGE
GPHN	NM_020806.5:c.68_71del (p.Ser23fs), het	NA	NA	GGE
PCLO	NM_033026.6:c.1499del (p.Pro500fs), het	NA	14	NAFE Focal activity: anterior regions of the brain
RAI1	NG_007101.2(NM_030665.4):c.5659+42C>T (p.Gln89*), het	0.0000065	12	GGE
NAFE, non-acquired focal epilepsy; GGE, Genetic generalized epilepsy; CAE, Childhood absence epilepsy; JAE, Juvenile absence epilepsy; het, heterozygous.				

Notably, two patients with novel likely pathogenic variants had anatomical changes detected on MRI. Specifically, the individual with p. Pro1137fs in *CACNA1A* showed minimal hippocampal asymmetry, while the patient with p. Trp2186* exhibited slight atrophy of the left temporal lobe. Anatomic changes are frequently associated with specific epilepsy syndromes and should be taken into consideration when investigating the underlying genetic causes of the epilepsy phenotype.

Another two patients had epilepsy in combination with other symptoms. The patient with the p. Ser23fs variant in *GPHN* presented with dyslexia, ADHD, and central hearing

loss. The patient with p. Gln89* in RAI had learning difficulties and Systemic Lupus Erythematosus. Epilepsy with marked psychiatric and systemic comorbidity may be other diseases that are accompanied by seizures.

LoF variants, especially when rare, are classified as likely pathogenic due to the strong PVS1 criterion, which has a high level of pathogenicity strength. On the other hand, rare missense variants in epilepsy genes are frequently classified as variants of uncertain significance (VUS), unless additional supporting criteria are met. We conducted a burden analysis of rare missense variants to evaluate whether there is significant enrichment of these variants in specific epilepsy gene sets.

2. Burden analysis of rare, damaging VUS

2.1.Burden analysis for phenotype-specific genes

When taken collectively, genes specific to the phenotypes of GGE and NAFE, the burden analysis revealed statistically significant enrichment in the cohort due to rare damaging missense variants (OR 1.99; 95% CI 1.29–3.0, p=0.0017), particularly in the GGE subgroup (OR 2.56; 95% CI 1.48–4.26, p=0.0008) (Figure 5A).

In a second analysis, we separated genes specific to GGE and conducted a burden analysis only on individuals with the GGE phenotype. The same approach was applied to NAFE genes using the NAFE phenotype subgroup. Patients with GGE showed a positive odds ratio in genes associated with generalized epilepsy syndromes. Conversely, there was no observed impact on patients with NAFE in genes associated with focal epilepsy syndromes. With the current sample size, it was not possible to achieve significance when separating epilepsy subtypes for corresponding gene subsets (Figure 5B).

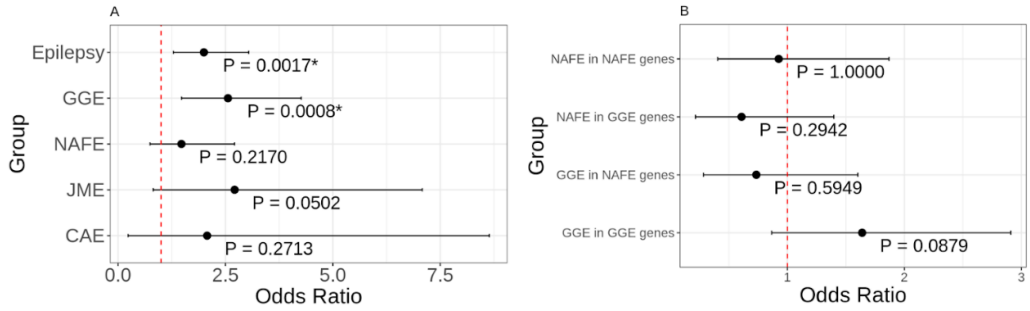

Figure 5: Odds ratio in patients vs controls in: A) Phenotype-specific genes (combining GGE and NAFE genes) and B) GGE and NAFE phenotypes within their respective phenotype-specific gene subgroups.

2.2.Burden analysis for missense intolerant epilepsy genes

In the second separate burden analysis, we investigated whether missense-intolerant epilepsy-related genes are enriched with rare damaging missense variants in the epilepsy cohort. It was shown that rare damaging variants are more prevalent in the epilepsy group, but this was not statistically significant (Figure 6A).

Next, we examined missense variants in gene sets that could more specifically underlie the observed enrichment in ion channels, transporters, and other genes. The study demonstrated statistically significant enrichment in genes of the transporter category (2.72; 95% CI, 1.27–5.58, p = 0.0062) (Figure 6B).

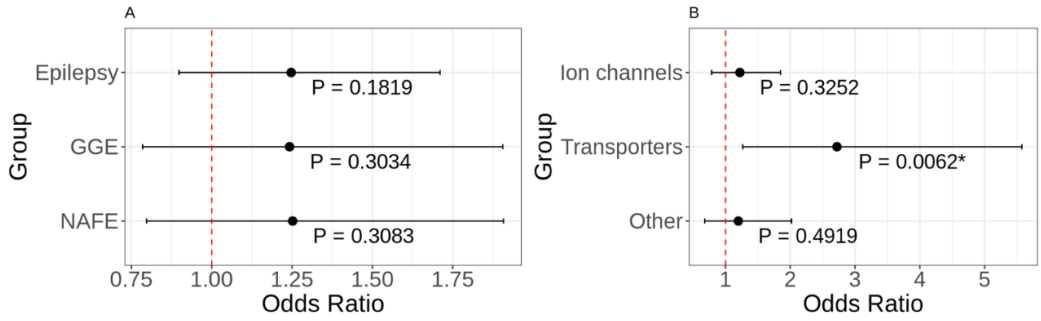

Figure 6. Burden analysis of epilepsy genes with Z-scores > 3.09, by phenotype (A) and by functional categories for epilepsy phenotype (B).

2.3.Gene-focused burden

A logistic regression approach was used to compare the enrichment of VUS in all genes (Table 2) between cases and controls. It was found that certain genes exhibit a higher frequency of rare damaging variants in cases compared to controls, as determined by the odds ratio. Specifically, voltage-gated ion channel genes (*CLCN2, CHRNA2*) and transporter genes (*SLC2A1, ATP6V0A1, ATP1A2*) were significantly enriched (Figure 7), (Table 7).

It was also shown that some genes had rare damaging VUS exclusively in epilepsy cases, specifically *GABRA1, GABRB2, GABRD, AP2M1, STXBP1, ACTL6B* and *KIF2A* (Table 8).

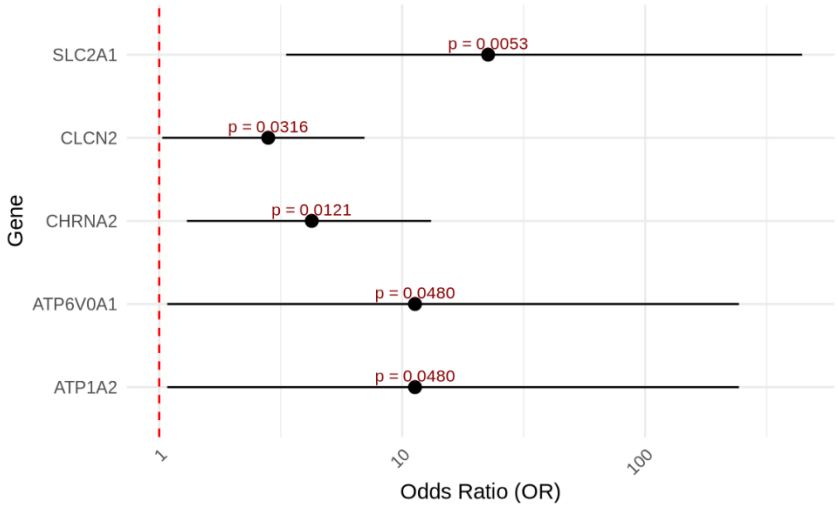

Figure 7. Genes with the most significant impact on the epilepsy group, determined based on odds ratio.

Table 7. Variant of Uncertain Significance in genes with significant odds ratios

Gene	Variant	Transcript	gnomAD AF	ACMG/AMP classification	Phenotype
CLCN2 (found in controls)	chr3:184352781: G:A p.Arg725Trp	ENST00000265593.9	0	PM2, PP3	GGE
CLCN2 (found in controls)	chr3:184357051: G:A p.Arg343Trp	ENST00000265593.9	0	PM2, PP3	GGE, JME
CLCN2 (found in controls)	chr3:184353722: C:T p.Asp599Asn	ENST00000265593.9	0.000875	PM2, PP3	GGE
CLCN2 (found in controls)	chr3:184357673: A:G p.Leu240Pro	ENST00000265593.9	0.0002915	PM2, PP3	GGE, CAE
CLCN2	chr3:184353404: C:G p.Gly625Ala	ENST00000265593.9	NA	PM2, PP3	NAFE
CLCN2	chr3:184357629: G:A p.Arg61Gln	ENST00000265593.9	NA	PM2, PP3	NAFE
CLCN2	chr3:184354293: G:A p.Ala493Val	ENST00000344937.11	0.000039	PM2, PP3	GGE
SLC2A1	chr1:42929606:G :A p.Ser285Phe	ENST00000426263.9	NA	PM2, PP2, PP3	GGE
SLC2A1	chr1:42930643:C: A p.Gly167Cys	ENST00000426263.9	NA	PM2, PP2, PP3	GGE, JME

SLC2A1	chr1:42930685:G:A p.Arg153Cys	ENST00000426263.9	0.0000065	PM2, PP2, PP3	GGE
SLC2A1	chr1:42930685:G:A p.Arg153Cys	ENST00000426263.9	0.0000065	PM2, PP2, PP3	GGE
CHRNA2	chr8:27471000:A:G p.Leu20Pro	ENST00000240132.7	NA	PM2, PP3	GGE
CHRNA2	chr8:27469841:G:A p.Arg72Cys	ENST00000240132.7	0.0001455	PM2, PP3	NAFE
CHRNA2	chr8:27469873:C:T p.Arg61Gln	ENST00000240132.7	0	PM2, PP3	GGE
ATP6V0A1	chr17:42460911:G:A p.Arg6Gln	ENST00000264649.10	0.0002365	PM2, PP2, PP3	GGE
ATP6V0A1	chr17:42495163:C:T p.Arg489Trp	ENST00000264649.10	0.0000195	PM2, PP2, PP3	NAFE
ATP1A2	chr1:160136941:C:T p.Thr917Met	ENST00000265593.9	NA	PM2, PP2, PP3	GGE, JME
ATP1A2	chr1:160135591:G:C p.Gly758Ala	ENST00000361216.8	0.0002035	PM2, PP2, PP3	NAFE

Table 8. Genes with Variant of Uncertain Significance exclusively in epilepsy cases

Gene	Variant	Transcript	gnomAD AF	ACMG/AMP classification	Phenotype
GABRA1	chr5:161882638:C:T p.Arg214Cys	ENST00000428797.7	NA	PM2, PP3	NAFE
GABRD	chr1:2029599:C:G p.Ser299Cys	ENST00000638771.1	0.000013	PM2, PP3	NAFE
GABRB2	chr5:161330899:C:T p.Arg354His	ENST00000393959.6	0.0000655	PM2, PP2, PP3	NAFE
AP2M1	chr3:184180229:C:T p.Thr134Met	ENST00000382456.7	0.000013	PM2, PP2, PP3	NAFE
STXBP1	chr9 :127663344:G:A p.Arg71Gln	ENST00000496504.3	NA	PM2, PP2, PP3	NAFE
ACTL6B	chr7:100647014:C:T p.Arg298Gln	ENST00000160382.10	0.0000065	PM2, PP2, PP3	NAFE
ACTL6B	chr7:100656336:C:G p.Gly7Arg	ENST00000160382.10	NA	PM2, PP2, PP3	NAFE
KIF2A	chr5:62363753:A:G p.Ile421Val	ENST00000381103.6	NA	PM2, PP2, PP3	GGE

Chapter 4:

Discussion

In our study, we used the ACMG guidelines for variant interpretation to identify novel pathogenic genetic variants in patients with genetic generalized epilepsy and non-acquired focal epilepsy. The application of ACMG criteria increased the diagnostic rate threefold compared to using only data on previously reported pathogenic variants in ClinVar (from 1.7% to 6.2%), enabling the identification of a broader range of variants potentially causative of epilepsy.

1. Variant interpretation

1.1.Pathogenic variants based on ClinVar

In this section we focus on phenotypes of patients with pathogenic variants from ClinVar. We compared the phenotypes of patients with epilepsy patients from our cohort with identified pathogenic variants to those reported in the ClinVar database.

The comparison of phenotypes of patients with epilepsy who carry the same pathogenic variants plays an important clinical role, as it helps to determine genotype–phenotype relationships and demonstrates the variability in clinical manifestations of the same genetic variant across different patients, a phenomenon known as phenotypic heterogeneity[80].

Among six pathogenic variants previously reported in ClinVar, two patients exhibited matching phenotypes. The variant (DEPDC5):c.2512C>T (p.Arg838Ter) was associated with familial focal epilepsy with variable foci, while (KCNQ2):c.628C>T (p.Arg210Cys) was linked with benign familial neonatal seizures.

Another four patients exhibited a different phenotype from that reported in the ClinVar database. Variant, (SLC6A1):c.1024G>A (p.Val342Met), was detected in two

patients diagnosed with GGE, with one exhibiting CAE. Despite ClinVar's association of this variant with myoclonic-atonic epilepsy, our clinical presentation in our patient was less severe. For the remaining three variants, the subclinical form of epilepsy differed from ClinVar's reports. For instance, the *GOSR2* gene variant c.430G>T (p.Gly144Trp), previously linked to progressive myoclonic epilepsy type 6, manifested as NAFE in our case. Similarly, the *KCNT1* gene variant c.2012C>T (p.Thr671Ile), associated with autosomal dominant nocturnal frontal lobe epilepsy and developmental epileptic encephalopathy, presented as JAE in our case. Furthermore, the *GABRA1* variant c.640C>T (p.Arg214Cys), typically associated with GGE or JME, was observed in an individual with NAFE in our study. This variability in phenotypic presentation may be attributed to variable expressivity and phenotypic heterogeneity for mentioned variants. The clinical phenotype associated with the pathogenic genetic variants, in particular its severity is influenced by the modifier variants as well as nongenetic factors. Modifier variants may contribute to the variable phenotypic expression of the disease among individuals.[81]

Phenotypic heterogeneity and variable expressivity can pose challenges in variant interpretation in clinical practice when the previously reported phenotype does not align with the clinical presentation of a patient carrying the same genetic variant. Expanding the variant metadata in public genetic databases by reporting phenotypic variability would enhance our ability to resolve conflicting interpretations of variants[82]. It is also important to report cases where pathogenic variants are found in unaffected individuals to address the issue of overestimating the pathogenicity of certain variants.

All the pathogenic variants mentioned previously are missense mutations, with the exception of c.2512C>T (p.Arg838Ter) in the *DEPDC5* gene, which is a nonsense mutation.

Most of the genes with these variants were missense-intolerant, except for *GOSR2* and *KCNT1*. This supports the idea that missense-tolerant genes, such as *GOSR2* and *KCNT1*, can also harbor disease-causing missense variants, particularly if they are in hotspot regions or functional domains critical to protein function.

Reviewing ClinVar for pathogenic variants alongside the regional missense constraint track can help identify hotspots or functional domains that lack benign variation. In this context, moderate evidence for variant pathogenicity (PM1) can be added, indicating that the variant is located in a critical or well-established functional domain or mutational hotspot with high regional constraint, making it more likely to be deleterious[62]. It was established that all pathogenic variants in epilepsy found based on ClinVar were in regions intolerant to missense mutations. (Figure 7).

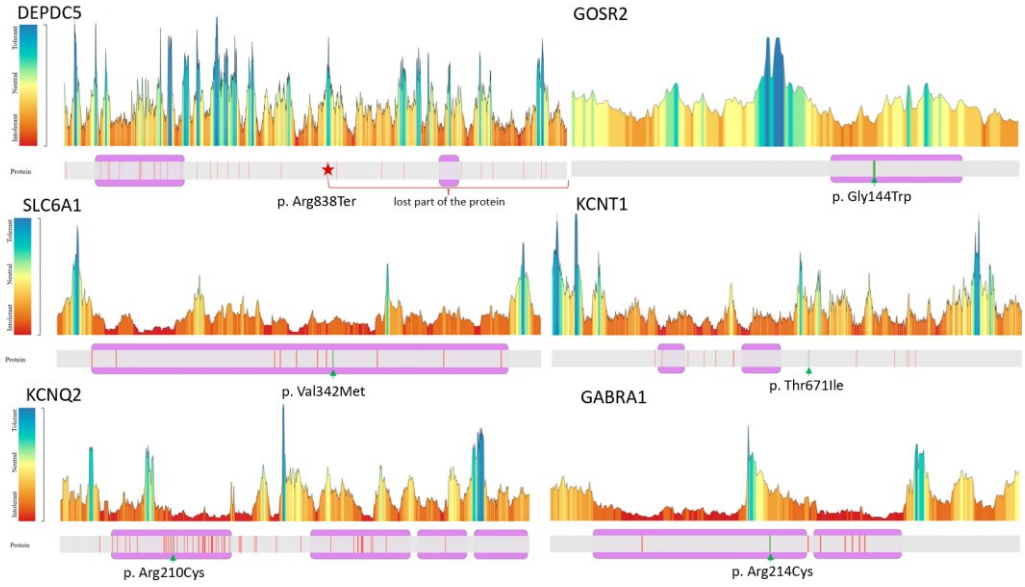

Figure 7. Topography of pathogenic variants from ClinVar on the corresponding protein, along with the tolerance landscape. Pathogenic variants found in epilepsy patient are shown with a red star for nonsense variants and with green arrows for missense variants. Red,

orange, and yellow regions represent hotspot areas intolerant to missense variants. Purple bars indicate functional domains of the protein. Red lines represent other ClinVar pathogenic variants in this protein.

1.2.Novel likely pathogenic variants

Identifying new likely pathogenic variants has the potential to deepen our understanding of the pathophysiological mechanisms underlying epileptogenesis and inform the development of more precise treatment strategies. Two distinct types of loss-of-function variants were identified among the novel likely pathogenic variants in our study: nonsense and frameshift mutations.

Nonsense mutations associated with premature termination codon which can have several different consequences, including exon skipping and decreased mRNA stability as well as protein truncation[83]. Any nonsense variant in loss-of-function-intolerant epilepsy genes has the potential to disrupt protein function, especially if it affects functional domains. In our study, the identified variant (CACNA1A):c.3829C>T (p.Arg1277*), impacts the function of ion transport protein domains PF00520 and PF16905 in the calcium voltage-gated channel subunit alpha1A (Figure 8A). These domains have previously shown multiple pathogenic missense variants associated with epilepsy phenotypes[84]. Variants in the *CACNA1A* gene have been linked to absence seizures in humans, and a patient with the c.3829C>T (p.Arg1277*) variant has been reported to experience absence seizures[85].

Another subtype of nonsense variants, potentially relevant in epilepsy etiology, are truncating variants located close to the N-terminus of the transcript. According to ACMG guidelines, interpreting these variants must be done cautiously, especially if the variant

affects the last 50 base pairs of the last exon[59]. For example, the truncating variant c.5797C>T (p.Arg1933*) in the sodium voltage-gated channel alpha subunit 1 (*SCN1A*) causes truncation of 76 terminal amino acids. Given its proximity to the N-terminus and the absence of apparent effects on known functional domains, it's plausible that this variant may impact transmembrane anchoring or protein folding processes (Figure 8B). Similarly, the variant c.6558G>A (p.Trp2186*) in the calcium voltage-gated channel Subunit Alpha1 B (*CACNA1B*) causes truncation of 152 terminal amino acids and may disrupt spatial folding of the transmembrane protein. For greater validity regarding the pathogenicity of these variants, functional assays should be performed, which was beyond the scope of our study.

Besides transmembrane proteins, we identified two nonsense variants in intracellular protein regulators: the truncating variant c.1453C>T (p.Arg485*) in the *CHD2* gene and the variant c.5659+42C>T (p.Gln89*) in the RAI1 gene. The chromodomain helicase DNA-binding protein 2 (*CHD2*), serves as an ATPase and plays a crucial role in neurodevelopment and pathogenic variations in the *CHD2* gene have been linked to adult-onset epilepsy[86]. Individual with the intronic variant c.5659+42C>T (p.Gln89*) in the *RAI1* gene exhibited a phenotype of genetic generalized epilepsy, learning difficulties and Systemic Lupus Erythematosus. Deletion of *RAI1* gene is typically associated with Smith-Magenis syndrome.[87]. It is possible that the complex clinical presentation in our patient could be explained by an underdiagnosed case of Smith-Magenis syndrome, which often includes generalized seizures as one of its symptoms. The adoption of WGS in genetic testing enables the detection of genetic variations in the intronic regions of genes, revealing new pathogenic variants that were missed by exome sequencing due to its inability to cover non-coding regions.

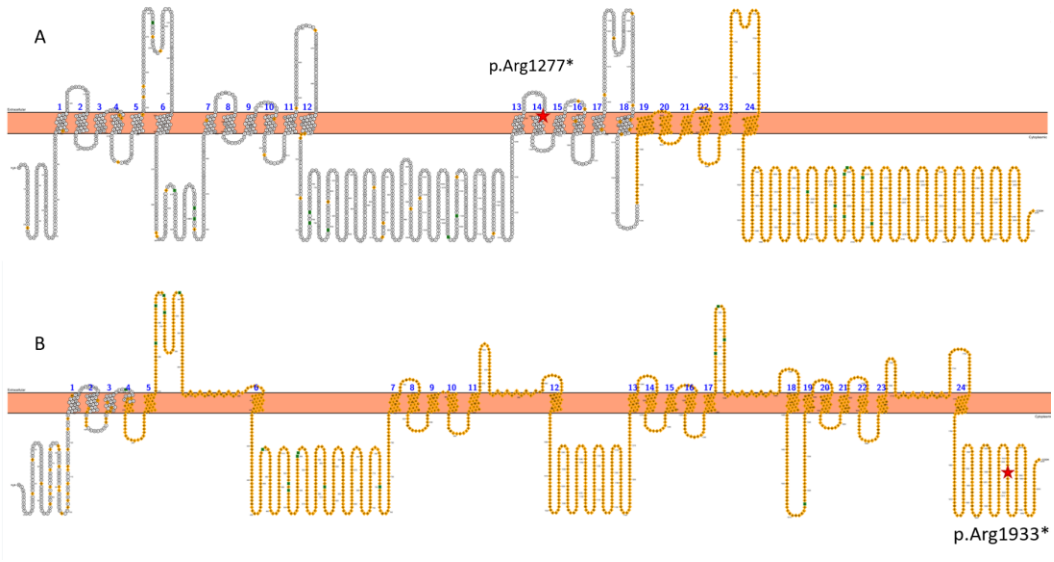

Figure 8. Nonsense variants in transmembrane proteins located at different proximities to the N-terminus of the transcript. A. Calcium voltage-gated channel subunit alpha1 A (CACNA1A). NM_001127222.2. B. Sodium voltage-gated channel alpha subunit 1 (SCN1A). NM_001165963.4. The red star represents the location of the stop codon on the protein.

Frameshift variants result from nucleotide insertions or deletions of a length not divisible by three, shifting the position of the reading frame. Frameshift indels result in a protein that significantly differs from the native protein, especially when the indel occurs early in the transcript sequence[88]. For example, the novel identified variant c.68_71del (p.Ser23fs), in the gephyrin alter protein sequence and length, potentially affecting function (Figure 9A). Exonic microdeletions in the gephyrin gene have been observed to disrupt GABAergic synaptic inhibition in individuals with idiopathic generalized epilepsy[89]. We posit that this genetic alteration represents a rare risk factor for GGE due to its potential to interfere with GABAergic inhibitory synaptic transmission.

We also found four frameshift variants in *SYNGAP1* gene, which code a SynGAP protein important in the regulation of excitatory synapse structure and function in glutamatergic neurons. The presence of a clear LoF variant in a patient usually leads to the diagnosis of a *SYNGAP1*-mediated developmental encephalopathy[90]. Interestingly, none of our patients with frameshift variants in the *SYNGAP1* gene exhibited symptoms of developmental encephalopathy. Instead, three individuals showed a phenotype consistent with NAFE, while one displayed features of CAE. Similarly, a patient with p.Ala812fs in *DNM1* gene presents a phenotype of GGE, while biallelic LoF variants usually are linked to *DNM1* encephalopathy[91]. This phenomenon could be explained by haploinsufficiency, where heterozygous loss-of-function variants are insufficient to fully maintain normal function, resulting in a less severe phenotype compared to homozygous variants.

Another frameshift variant c.121+1090_121+1091insGGC (p.Arg86fs) was found in *SLC12A5* gene in a patient with NAFE (Figure 9 B). The number '+1090' indicates that this insertion occurs 1090 bases after the exon 121 region of the gene, disrupting the normal splicing process. This could result in incorrectly spliced mRNA that includes extra nucleotides, leading to a frameshift in the downstream protein sequence. This is a case where even a three-nucleotide insertion in an intron can result in a frameshift mutation. *SLC12A5* encodes the K+-Cl– cotransporter 2, which is the primary Cl− extruder in neurons[92]. This damaging effect could disrupt chloride homeostasis in neurons and contribute to the development of seizures[93]. Biallelic pathogenic variants were previously associated with *SLC12A5*-related epilepsy of infancy with migrating focal seizures[94].

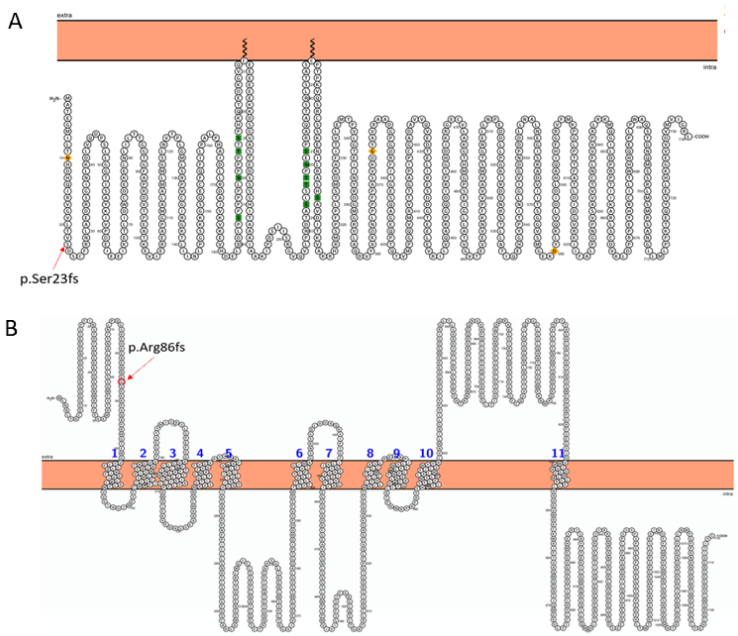

Figure 9. Frameshift variants in transmembrane proteins with proximity to the C-terminus of the transcript. A. Gephyrin (GPHN). NM_020806.5. B. Potassium-chloride transporter member 5 (SLC12A5). NM_001134771.2. The arrow shows the starting point of the reading frame shift on the protein.

The identification of novel likely pathogenic variants in our study serve as candidates for conducting functional studies on laboratory animal models.

2. Burden analysis on rare damaging VUS

In our analysis, we were unable to classify novel missense variants as likely pathogenic or pathogenic due to the lack of sufficient criteria. Every rare damaging missense variant was classified as VUS based on PP3 and PM2 criteria. Additionally, the PP2 criterion was used for genes known to be intolerant to missense variants.

In general, the frequency of detection of VUS increases in proportion to the amount of DNA sequenced[95]. It was shown that as multigene panel size increases, the number of

VUS per individual also rises progressively[96]. Burden analysis, conducted in the study, aimed to establish if the enrichment of VUS is higher in cases compared to controls in two subsets of genes. These subsets of genes demonstrate two different approaches for gene inclusion. In the first case, a phenotype-focused approach is used; in the second, genes related to epilepsy or seizures that are intolerant to missense mutations are included.

2.1.Phenotype-specific genes

The burden analysis of rare damaging VUS in NAFE and GGE genes combined revealed a significant difference between all epilepsy cases and controls. When comparing only GGE patients with controls in the same burden, the odds ratio was statistically significantly higher, indicating that the previously discussed trend was mainly driven by the GGE patients. This aligns with findings from previous studies on GGE, which have consistently demonstrated that GGE has higher heritability compared to NAFE[43,47,97]. This analysis also shows that careful gene selection, based on phenotype–genotype inference, can help narrow down the number of VUS to those most likely to be disease-causing.

It is important to mention that these significant results were observed when genes for NAFE and GGE were considered collectively on both phenotypes. In contrast, when separating GGE and NAFE and analyzing the corresponding phenotypes, we were unable to reach significant results due to the small sample size.

2.2.Missense intolerant epilepsy genes

We conducted a separate burden analysis for missense-intolerant genes to determine whether rare damaging VUS are more enriched in epilepsy-related genes with a Z-score >

3.09. Previous studies suggest that rare damaging variants in missense-intolerant genes and constrained regions are more highly enriched for epilepsy and cognitive performance[98,99].

The burden analysis of missense-intolerant epilepsy genes showed no significant difference for all cases, as well as for GGE and NAFE phenotypes. Although it was expected that variants in missense-intolerant epilepsy genes would show a greater burden of rare damaging variants compared to controls, it is not excluded that the small sample size in this study was insufficient to achieve significance.

Burden analysis of VUS in missense-intolerant epilepsy genes has revealed a notable contribution from genes categorized as transporters. In our study, this gene category originally encompassed those involved in ATP-dependent transport or cotransport, which are believed to strongly influence the epileptogenesis of GGE and NAFE[100]. It has been shown that *SLC2A1, ATP6V0A1*, and *ATP1A2*, all of which are missense-intolerant transporter genes, had significant enrichment by VUS. Increasing the number of participants in our study could enhance the detection of true positive associations for some other genes, where statistical significance was not initially reached.

3. Limitations

In our analysis, we exclusively employed genome sequencing of patients. The use of trio genome sequencing (sequencing of the patient and both biological parents) would have enhanced our diagnostic yield by detecting *de novo* variants, which could strengthen the level of pathogenicity by applying the PS2 criterion to such variants.

In our study, we did not conduct functional studies to validate the effect of deleterious variants on protein function. Model organisms or/and cellular and biochemical *in vivo* or *in*

vitro assays are usually used in laboratory evaluations of variants implicated interpretation[65]. Conducting such studies could lead to the reclassification of variants from likely pathogenic to pathogenic and decrease the number of variants of unknown significance by reclassifying them as pathogenic or benign.

Sanger sequencing was not conducted to eliminate technical artifacts associated with the identified loss-of-function variants. Previous studies have indicated that nonsense variants were most likely to be predicted as true LoF compared to other variant types with only 19.1% predicted as not LoF, whereas frameshift variants had a slightly elevated proportion of predicted not LoF (30.9%)[101]. The elimination of variants with low read depth and GC-rich regions was conducted to reduce the chance of including technical artifacts.

Chapter 5:

Conclusion

The objective of this study was to identify novel disease-causing genetic variants and to improve the diagnostic rate among patients with epilepsy. To do so, ACMG guidelines for variant interpretation and computational prediction programs were used. By using this approach, it was determined that the application of ACMG guidelines improved the diagnostic rate threefold by identifying numerous novel LoF likely pathogenic variants in epilepsy genes. Functional studies must be conducted on novel likely pathogenic variants to validate the effect of LoF variants on protein function. It was also observed that some of the pathogenic variants previously reported in ClinVar, with a link to specific epilepsy syndromes, exhibited distinct phenotypes in our cohort, likely due to phenotypic heterogeneity.

Another objective of this study was to conduct a burden analysis to determine the enrichment of rare damaging VUS in epilepsy genes by comparing the epilepsy cohort with the control group. The burden analysis of combined phenotype-specific set of genes for GGE and NAFE revealed that rare deleterious variants are more prevalent in epilepsy patients, particularly in those with GGE. In contrast, NAFE patients show no significant difference compared to controls, suggesting either lower heritability or a different underlying genetic mechanism.

Burden analysis of rare damaging variants among missense intolerant epilepsy genes, has demonstrated statistically significant enrichment in genes of the transporter category. Overall, it is evident that rare damaging missense variants play a significant role in the etiology of epilepsy.

References

1. Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia*. 2014;55(4):475-482. doi:10.1111/epi.12550

2. Neligan A, Hauser WA, Sander JW. The epidemiology of the epilepsies. In: *Handbook of Clinical Neurology*. Vol 107. Elsevier; 2012:113-133. doi:10.1016/B978-0-444-52898-8.00006-9

3. Huff JS, Murr NI. Seizure. In: *StatPearls*. StatPearls Publishing; 2024. Accessed January 4, 2025. http://www.ncbi.nlm.nih.gov/books/NBK430765/

4. Fiest KM, Sauro KM, Wiebe S, et al. Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. *Neurology*. 2017;88(3):296-303. doi:10.1212/WNL.0000000000003509

5. Gurgu R, Ciobanu A, Danasel R, Panea C. Psychiatric comorbidities in adult patients with epilepsy (A systematic review). *Exp Ther Med*. 2021;22(2):909. doi:10.3892/etm.2021.10341

6. Kerr MP. The impact of epilepsy on patients' lives. *Acta Neurol Scand*. 2012;126:1-9. doi:10.1111/ane.12014

7. Mayor R, Gunn S, Reuber M, Simpson J. Experiences of stigma in people with epilepsy: A meta-synthesis of qualitative evidence. *Seizure*. 2022;94:142-160. doi:10.1016/j.seizure.2021.11.021

8. Allers K, Essue BM, Hackett ML, et al. The economic impact of epilepsy: a systematic review. *BMC Neurol*. 2015;15(1):245. doi:10.1186/s12883-015-0494-y

9. Kyu HH, Abate D, Abate KH, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018;392(10159):1859-1922. doi:10.1016/S0140-6736(18)32335-3

10. Fisher RS, Cross JH, French JA, et al. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia*. 2017;58(4):522-530. doi:10.1111/epi.13670

11. Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*. 2017;58(4):512-521. doi:10.1111/epi.13709

12. Sarmast ST, Abdullahi AM, Jahan N. Current Classification of Seizures and Epilepsies: Scope, Limitations and Recommendations for Future Action. *Cureus*. Published online September 20, 2020. doi:10.7759/cureus.10549

13. Wirrell E, Tinuper P, Perucca E, Moshé SL. Introduction to the epilepsy syndrome papers. *Epilepsia*. 2022;63(6):1330-1332. doi:10.1111/epi.17262

14. McWilliam M, Asuncion RMD, Al Khalili Y. Idiopathic (Genetic) Generalized Epilepsy. In: *StatPearls*. StatPearls Publishing; 2024. Accessed January 4, 2025. http://www.ncbi.nlm.nih.gov/books/NBK546611/

15. Hirsch E, French J, Scheffer IE, et al. ILAE definition of the Idiopathic Generalized Epilepsy Syndromes: Position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia*. 2022;63(6):1475-1499. doi:10.1111/epi.17236

16. Wirrell EC, Nabbout R, Scheffer IE, et al. Methodology for classification and definition of epilepsy syndromes with list of syndromes: Report of the ILAE Task Force on Nosology and Definitions. *Epilepsia*. 2022;63(6):1333-1348. doi:10.1111/epi.17237

17. Moreau C, Rébillard RM, Wolking S, et al. Polygenic risk scores of several subtypes of epilepsies in a founder population. *Neurol Genet*. 2020;6(3):e416. doi:10.1212/NXG.0000000000000416

18. Kessler SK, McGinnis E. A Practical Guide to Treatment of Childhood Absence Epilepsy. *Pediatr Drugs*. 2019;21(1):15-24. doi:10.1007/s40272-019-00325-x

19. Ma X, Zhang Y, Yang Z, et al. Childhood absence epilepsy: Elctroclinical features and diagnostic criteria. *Brain and Development*. 2011;33(2):114-119. doi:10.1016/j.braindev.2010.02.004

20. Yadala S, Nalleballe K. Juvenile Absence Epilepsy. In: *StatPearls*. StatPearls Publishing; 2024. Accessed January 4, 2025. http://www.ncbi.nlm.nih.gov/books/NBK559055/

21. Amrutkar CV, Riel-Romero RM. Juvenile Myoclonic Epilepsy. In: *StatPearls*. StatPearls Publishing; 2024. Accessed January 4, 2025. http://www.ncbi.nlm.nih.gov/books/NBK537109/

22. Zuberi SM, Wirrell E, Yozawitz E, et al. ILAE classification and definition of epilepsy syndromes with onset in neonates and infants: Position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia*. 2022;63(6):1349-1397. doi:10.1111/epi.17239

23. Specchio N, Wirrell EC, Scheffer IE, et al. International League Against Epilepsy classification and definition of epilepsy syndromes with onset in childhood: Position paper by the ILAE Task Force on Nosology and Definitions. *Epilepsia*. 2022;63(6):1398-1442. doi:10.1111/epi.17241

24. Kurahashi H, Hirose S. Autosomal Dominant Sleep-Related Hypermotor (Hyperkinetic) Epilepsy. In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Amemiya A, eds. *GeneReviews®*. University of Washington, Seattle; 1993. Accessed January 6, 2025. http://www.ncbi.nlm.nih.gov/books/NBK1169/

25. Gooley S, Crompton DE, Berkovic SF, ILAE Genetics Literacy Taskforce. ILAE Genetic Literacy Series: familial focal epilepsy syndromes. *Epileptic Disorders*. 2022;24(2):221-228. doi:10.1684/epd.2021.1393

26. Nayak CS, Bandyopadhyay S. Mesial Temporal Lobe Epilepsy. In: *StatPearls*. StatPearls Publishing; 2024. Accessed January 6, 2025. http://www.ncbi.nlm.nih.gov/books/NBK554432/

27. Weckhuysen S, Marsan E, Lambrecq V, et al. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia*. 2016;57(6):994-1003. doi:10.1111/epi.13391

28. Riney K, Bogacz A, Somerville E, et al. International League Against Epilepsy classification and definition of epilepsy syndromes with onset at a variable age: position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia*. 2022;63(6):1443-1474. doi:10.1111/epi.17240

29. Smith L, Malinowski J, Ceulemans S, et al. Genetic testing and counseling for the unexplained epilepsies: An evidence-based practice guideline of the National Society of Genetic Counselors. *Journal of Genetic Counseling*. 2023;32(2):266-280. doi:10.1002/jgc4.1646

30. Singh NA, Charlier C, Stauffer D, et al. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. *Nat Genet*. 1998;18(1):25-29. doi:10.1038/ng0198-25

31. Claes L, Del-Favero J, Ceulemans B, Lagae L, Van Broeckhoven C, De Jonghe P. De Novo Mutations in the Sodium-Channel Gene SCN1A Cause Severe Myoclonic Epilepsy of Infancy. *The American Journal of Human Genetics*. 2001;68(6):1327-1332. doi:10.1086/320609

32. Oliver KL, Scheffer IE, Bennett MF, Grinton BE, Bahlo M, Berkovic SF. GENES4EPILEPSY : An epilepsy gene resource. *Epilepsia*. 2023;64(5):1368-1375. doi:10.1111/epi.17547

33. Dwivedi R, Kaushik M, Tripathi M, Dada R, Tiwari P. Unraveling the genetic basis of epilepsy: Recent advances and implications for diagnosis and treatment. *Brain Research*. 2024;1843:149120. doi:10.1016/j.brainres.2024.149120

34. Hernandez RD, Uricchio LH, Hartman K, Ye C, Dahl A, Zaitlen N. Ultrarare variants drive substantial cis heritability of human gene expression. *Nat Genet*. 2019;51(9):1349-1355. doi:10.1038/s41588-019-0487-7

35. Helbig I, Riggs ER, Barry C, et al. The ClinGen Epilepsy Gene Curation Expert Panel—Bridging the divide between clinical domain knowledge and formal gene curation criteria. *Human Mutation*. 2018;39(11):1476-1484. doi:10.1002/humu.23632

36. Stefanski A, Calle-López Y, Leu C, Pérez-Palma E, Pestana-Knight E, Lal D. Clinical sequencing yield in epilepsy, autism spectrum disorder, and intellectual disability: A systematic review and meta-analysis. *Epilepsia*. 2021;62(1):143-151. doi:10.1111/epi.16755

37. Guerrini R, Balestrini S, Wirrell EC, Walker MC. Monogenic Epilepsies: Disease Mechanisms, Clinical Phenotypes, and Targeted Therapies. *Neurology*. 2021;97(17):817-831. doi:10.1212/WNL.0000000000012744

38. Mullen SA, Berkovic SF, the ILAE Genetics Commission. Genetic generalized epilepsies. *Epilepsia*. 2018;59(6):1148-1153. doi:10.1111/epi.14042

39. Krey I, Platzer K, Lemke JR. Monogenetic epilepsies and how to approach them in 2022. *Medizinische Genetik*. 2022;34(3):201-205. doi:10.1515/medgen-2022-2143

40. Zhang MW, Liang XY, Wang J, et al. Epilepsy-associated genes: an update. *Seizure: European Journal of Epilepsy*. 2024;116:4-13. doi:10.1016/j.seizure.2023.09.021

41. Tsai MH, Chan CK, Chang YC, et al. Molecular Genetic Characterization of Patients With Focal Epilepsy Using a Customized Targeted Resequencing Gene Panel. *Front Neurol*. 2018;9:515. doi:10.3389/fneur.2018.00515

42. Sinha S, Rabea F, Ramaswamy S, et al. Long read sequencing enhances pathogenic and novel variation discovery in patients with rare diseases. *Nat Commun*. 2025;16(1):2500. doi:10.1038/s41467-025-57695-9

43. Feng YCA, Howrigan DP, Abbott LE, et al. Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. *The American Journal of Human Genetics*. 2019;105(2):267-282. doi:10.1016/j.ajhg.2019.05.020

44. Epi25 Collaborative, Chen S, Abou-Khalil BW, et al. Exome sequencing of 20,979 individuals with epilepsy reveals shared and distinct ultra-rare genetic risk across disorder subtypes. *Nat Neurosci*. 2024;27(10):1864-1879. doi:10.1038/s41593-024-01747-8

45. Perucca P, Bahlo M, Berkovic SF. The Genetics of Epilepsy. *Annu Rev Genom Hum Genet*. 2020;21(1):205-230. doi:10.1146/annurev-genom-120219-074937

46. Goswami C, Chattopadhyay A, Chuang EY. Rare variants: data types and analysis strategies. *Ann Transl Med*. 2021;9(12):961-961. doi:10.21037/atm-21-1635

47. International League Against Epilepsy Consortium on Complex Epilepsies, Stevelink R, Campbell C, et al. GWAS meta-analysis of over 29,000 people with epilepsy identifies 26 risk loci and subtype-specific genetic architecture. *Nat Genet*. 2023;55(9):1471-1482. doi:10.1038/s41588-023-01485-w

48. Song M, Liu J, Yang Y, Lv L, Li W, Luo XJ. Genome-Wide Meta-Analysis Identifies Two Novel Risk Loci for Epilepsy. *Front Neurosci*. 2021;15:722592. doi:10.3389/fnins.2021.722592

49. The International League Against Epilepsy Consortium on Complex Epilepsies, Abou-Khalil B, Auce P, et al. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun*. 2018;9(1):5269. doi:10.1038/s41467-018-07524-z

50. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779

51. Ishigaki K, Akiyama M, Kanai M, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet*. 2020;52(7):669-679. doi:10.1038/s41588-020-0640-3

52. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613(7944):508-518. doi:10.1038/s41586-022-05473-8

53. Leu C, Stevelink R, Smith AW, et al. Polygenic burden in focal and generalized epilepsies. *Brain*. 2019;142(11):3473-3481. doi:10.1093/brain/awz292

54. Moreau C, Rébillard RM, Wolking S, et al. Polygenic risk scores of several subtypes of epilepsies in a founder population. *Neurol Genet*. 2020;6(3):e416. doi:10.1212/NXG.0000000000000416

55. Heyne HO. Polygenic risk scores in epilepsy. *Medizinische Genetik*. 2022;34(3):225-230. doi:10.1515/medgen-2022-2146

56. Ellis CA, Petrovski S, Berkovic SF. Epilepsy genetics: clinical impacts and biological insights. *The Lancet Neurology*. 2020;19(1):93-100. doi:10.1016/S1474-4422(19)30269-8

57. Girard A, Moreau C, Michaud JL, Minassian B, Cossette P, Girard SL. Unraveling the role of non-coding rare variants in epilepsy. Taguchi Y h., ed. *PLoS ONE*. 2023;18(9):e0291935. doi:10.1371/journal.pone.0291935

58. Gooley S, Perucca P, Tubb C, Hildebrand MS, Berkovic SF. Somatic mosaicism in focal epilepsies. *Current Opinion in Neurology*. 2024;37(2):105-114. doi:10.1097/WCO.0000000000001244

59. Richards S, Aziz N, Bale S, et al. 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*. 2015;17(5):405-424. doi:10.1038/gim.2015.30

60. Pavlick DC, Frampton GM, Ross JR. Understanding variants of unknown significance and classification of genomic alterations. *The Oncologist*. 2024;29(8):658-666. doi:10.1093/oncolo/oyae149

61. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Human Mutation*. 2018;39(11):1517-1524. doi:10.1002/humu.23626

62. Gudmundsson S, Singer-Berk M, Watts NA, et al. Variant interpretation using population databases: Lessons from gnomAD. *Human Mutation*. 2022;43(8):1012-1030. doi:10.1002/humu.24309

63. Zhang J, Yao Y, He H, Shen J. Clinical Interpretation of Sequence Variants. *Curr Protoc Hum Genet*. 2020;106(1):e98. doi:10.1002/cphg.98

64. Garcia FADO, Andrade ESD, Palmero EI. Insights on variant analysis in silico tools for pathogenicity prediction. *Front Genet*. 2022;13:1010327. doi:10.3389/fgene.2022.1010327

65. On behalf of the Clinical Genome Resource Sequence Variant Interpretation Working Group, Brnich SE, Abou Tayoun AN, et al. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med*. 2020;12(1):3. doi:10.1186/s13073-019-0690-2

66. Kanavy DM, McNulty SM, Jairath MK, et al. Comparative analysis of functional assay evidence use by ClinGen Variant Curation Expert Panels. *Genome Med*. 2019;11(1):77. doi:10.1186/s13073-019-0683-1

67. Kim SH, Kwon SS, Lee JS, et al. Analysis of trio test in neurodevelopmental disorders. *Front Pediatr*. 2022;10:1073083. doi:10.3389/fped.2022.1073083

68. Biesecker LG, Byrne AB, Harrison SM, et al. ClinGen guidance for use of the PP1/BS4 co-segregation and PP4 phenotype specificity criteria for sequence variant pathogenicity classification. *The American Journal of Human Genetics*. 2024;111(1):24-38. doi:10.1016/j.ajhg.2023.11.009

69. Biesecker LG, Harrison SM. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genetics in Medicine*. 2018;20(12):1687-1688. doi:10.1038/gim.2018.42

70. Awadalla P, Boileau C, Payette Y, et al. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *International Journal of Epidemiology*. 2013;42(5):1285-1299. doi:10.1093/ije/dys160

71. Girard, Alexandre. *Unraveling the Role of Non-Coding Rare Variants in Epilepsy and Its Subtypes with Deep Learning*. Master's Thesis. Université du Québec à Chicoutimi; 2023.

72. Coppola A, Cellini E, Stamberger H, et al. Diagnostic implications of genetic copy number variation in epilepsy plus. *Epilepsia*. 2019;60(4):689-706. doi:10.1111/epi.14683

73. Pejaver V, Byrne AB, Feng BJ, et al. Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. *The American Journal of Human Genetics*. 2022;109(12):2163-2177. doi:10.1016/j.ajhg.2022.10.013

74. Freeman PJ, Hart RK, Gretton LJ, Brookes AJ, Dalgleish R. VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions. *Human Mutation*. 2018;39(1):61-68. doi:10.1002/humu.23348

75. Den Dunnen JT. Describing Sequence Variants Using HGVS Nomenclature. In: White SJ, Cantsilieris S, eds. *Genotyping*. Vol 1492. Methods in Molecular Biology. Springer New York; 2017:243-251. doi:10.1007/978-1-4939-6442-0_17

76. Omasits U, Ahrens CH, Müller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*. 2014;30(6):884-886. doi:10.1093/bioinformatics/btt607

77. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Human Mutation*. Published online May 22, 2019:humu.23798. doi:10.1002/humu.23798

78. Becchetti A, Aracri P. Ion Channels and Transporters. In: Bell E, ed. *Molecular Life Sciences*. Springer New York; 2014:1-22. doi:10.1007/978-1-4614-6436-5_190-2

79. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics*. 2014;95(1):5-23. doi:10.1016/j.ajhg.2014.06.009

80. Magrinelli F, Balint B, Bhatia KP. Challenges in Clinicogenetic Correlations: One Gene – Many Phenotypes. *Movement Disord Clin Pract*. 2021;8(3):299-310. doi:10.1002/mdc3.13165

81. Marian AJ. Clinical Interpretation and Management of Genetic Variants. *JACC: Basic to Translational Science*. 2020;5(10):1029-1042. doi:10.1016/j.jacbts.2020.05.013

82. Lazareva TE, Barbitoff YA, Nasykhova YA, Glotov AS. Major Causes of Conflicting Interpretations of Variant Pathogenicity in Rare Disease: A Systematic Analysis. *JPM*. 2024;14(8):864. doi:10.3390/jpm14080864

83. Mort M, Ivanov D, Cooper DN, Chuzhanova NA. A meta-analysis of nonsense mutations causing human genetic disease. *Hum Mutat*. 2008;29(8):1037-1047. doi:10.1002/humu.20763

84. Alehabib E, Esmaeilizadeh Z, Ranji-Burachaloo S, Tafakhori A, Darvish H, Movafagh A. Clinical and molecular spectrum of P/Q type calcium channel Cav2.1 in epileptic patients. *Orphanet J Rare Dis*. 2021;16(1):461. doi:10.1186/s13023-021-02101-y

85. Zamponi GW, Lory P, Perez-Reyes E. Role of voltage-gated calcium channels in epilepsy. *Pflugers Arch - Eur J Physiol*. 2010;460(2):395-403. doi:10.1007/s00424-009-0772-x

86. De Maria B, Balestrini S, Mei D, et al. Expanding the genetic and phenotypic spectrum of *CHD2* -related disease: From early neurodevelopmental disorders to adult-onset epilepsy. *American J of Med Genetics Pt A*. 2022;188(2):522-533. doi:10.1002/ajmg.a.62548

87. Boot E, Linders CC, Tromp SH, Van Den Boogaard M, Van Eeghen AM. Possible underreporting of pathogenic variants in *RAI1* causing SMITH–MAGENIS syndrome. *American J of Med Genetics Pt A*. 2021;185(10):3167-3169. doi:10.1002/ajmg.a.62380

88. Bermejo-Das-Neves C, Nguyen HN, Poch O, Thompson JD. A comprehensive study of small non-frameshift insertions/deletions in proteins and prediction of their phenotypic effects by a machine learning method (KD4i). *BMC Bioinformatics*. 2014;15(1):111. doi:10.1186/1471-2105-15-111

89. Dejanovic B, Lal D, Catarino CB, et al. Exonic microdeletions of the gephyrin gene impair GABAergic synaptic inhibition in patients with idiopathic generalized epilepsy. *Neurobiology of Disease*. 2014;67:88-96. doi:10.1016/j.nbd.2014.02.001

90. Kilinc M, Arora V, Creson TK, et al. Endogenous Syngap1 alpha splice forms promote cognitive function and seizure protection. *eLife*. 2022;11:e75707. doi:10.7554/eLife.75707

91. Yigit G, Sheffer R, Daana M, et al. Loss-of-function variants in *DNM1* cause a specific form of developmental and epileptic encephalopathy only in biallelic state. *J Med Genet*. 2022;59(6):549-553. doi:10.1136/jmedgenet-2021-107769

92. Fukuda A, Watanabe M. Pathogenic potential of human SLC12A5 variants causing KCC2 dysfunction. *Brain Research*. 2019;1710:1-7. doi:10.1016/j.brainres.2018.12.025

93. Blaesse P, Airaksinen MS, Rivera C, Kaila K. Cation-Chloride Cotransporters and Neuronal Function. *Neuron*. 2009;61(6):820-838. doi:10.1016/j.neuron.2009.03.003

94. Stödberg T, McTague A, Ruiz AJ, et al. Mutations in SLC12A5 in epilepsy of infancy with migrating focal seizures. *Nat Commun*. 2015;6(1):8038. doi:10.1038/ncomms9038

95. Shirts BH, Pritchard CC, Walsh T. Family-Specific Variants and the Limits of Human Genetics. *Trends in Molecular Medicine*. 2016;22(11):925-934. doi:10.1016/j.molmed.2016.09.007

96. Rehm HL, Alaimo JT, Aradhya S, et al. The landscape of reported VUS in multi-gene panel and genomic testing: Time for a change. *Genetics in Medicine*. 2023;25(12):100947. doi:10.1016/j.gim.2023.100947

97. Song M, Liu J, Yang Y, Lv L, Li W, Luo XJ. Genome-Wide Meta-Analysis Identifies Two Novel Risk Loci for Epilepsy. *Front Neurosci*. 2021;15:722592. doi:10.3389/fnins.2021.722592

98. Koko M, Krause R, Sander T, et al. Ultra-rare constrained missense variants in the epilepsies: Shared and specific enrichment patterns in neuronal gene-sets. Published online April 19, 2021. doi:10.1101/2021.04.18.440264

99. Chen CY, Tian R, Ge T, et al. The impact of rare protein coding genetic variation on adult cognitive function. *Nat Genet*. 2023;55(6):927-938. doi:10.1038/s41588-023-01398-8

100. Beamer E, Kuchukulla M, Boison D, Engel T. ATP and adenosine—Two players in the control of seizures and epilepsy development. *Progress in Neurobiology*. 2021;204:102105. doi:10.1016/j.pneurobio.2021.102105

101. Singer-Berk M, Gudmundsson S, Baxter S, et al. Advanced variant classification framework reduces the false positive rate of predicted loss-of-function variants in population sequencing data. *The American Journal of Human Genetics*. 2023;110(9):1496-1508. doi:10.1016/j.ajhg.2023.08.005

Supplementary 1. VUS in phenotype-specific genes, identified in the epilepsy cohort.

№	Gene	Variant	Transcript	SIFT	Polyphen 2	CADD	gnomAD AF	ACMG classification	Phenotype
S16868 c	CLCN2	chr3:184352781:G:A p.Arg725Trp	ENST00000265593.9	D	P	23.7	0	PM2, PP3	GGE
S16898 c	CLCN2	chr3:184357051:G:A p.Arg343Trp	ENST00000265593.9	D	D	27.9	0	PM2, PP3	GGE, JME
S17105 c	CLCN2	chr3:184353722:C:T p.Asp599Asn	ENST00000265593.9	D	D	26.1	0.000875	PM2, PP3	GGE
S24441 c	CLCN2	chr3:184357673:A:G p.Leu240Pro	ENST00000265593.9	D	D	28.3	0.0002915	PM2, PP3	GGE, CAE
S27708	CLCN2	chr3:184353404:C:G p.Gly625Ala	ENST00000265593.9	D	D	26.6	NA	PM2, PP3	NAFE
S27729	CLCN2	chr3:184357629:G:A p.Arg61Gln	ENST00000265593.9	D	D	24.6	NA	PM2, PP3	NAFE
S20209	CLCN2	chr3:184354293:G:A p.Ala493Val	ENST00000344937.11	D	P	24.8	0.000039	PM2, PP3	GGE
S22442 c	EFHC1	chr6:52438562:C:T p.Arg182Cys	ENST00000637353.1	D	P	25.3	0.00001	PM2, PP3	NAFE
S24289 c	EFHC1	chr6:52452743:A:T p.Asp210Val	ENST00000637353.1	D	D	27.4	0.0037175	PM2, PP3	NAFE
S17026	EFHC1	chr6:52454186:C:T p.Thr164Met	ENST00000635984.1	D	D	24.8	0.000013	PM2, PP3	GGE
S01573 c	CACNB4	chr2:152098386:T:G p.Ser31Arg	ENST00000539935.6	D	D	29.7	0.0001455	PM2, PP3	GGE, JME
S06309* c	CACNB4	chr2:152098386:T :G p.Ser31Arg	ENST00000539935.6	D	D	29.7	0.0001455	PM2, PP3	GGE
S15711 c	CACNB4	chr2:152098386:T :G p.Ser31Arg	ENST00000539935.6	D	D	29.7	0.0001455	PM2, PP3	GGE
S20210	CACNB4	chr2:151870539:C :T p.Gly231Ser	ENST00000539935.6	D	D	27.4	0	PM2, PP3	GGE
S24283	GABRD	chr1:2029599:C:G p.Ser299Cys	ENST00000638771.1	D	D	26.2	0.000013	PM2, PP3	NAFE
S24281*	GABRA1	chr5:161882638:C:T p.Arg214Cys	ENST00000428797.7	D	D	28.4	NA	PM2, PP3	NAFE
S16883	SLC2A1	chr1:42929606:G:A p.Ser285Phe	ENST00000426263.9	D	D	28.5	NA	PM2, PP2, PP3	GGE
S22402	SLC2A1	chr1:42930643:C:A p.Gly167Cys	ENST00000426263.9	D	D	26.8	NA	PM2, PP2, PP3	GGE, JME
S20204	SLC2A1	chr1:42930685:G:A p.Arg153Cys	ENST00000426263.9	D	D	31	0.0000065	PM2, PP2, PP3	GGE
S20205	SLC2A1	chr1:42930685:G:A p.Arg153Cys	ENST00000426263.9	D	D	31	0.0000065	PM2, PP2, PP3	GGE
S01565	HCN2	chr19:590258:G:T p.Gly105Cys	ENST00000251287.3	D	P	25.4	0.0000205	PM2, PP2, PP3	GGE, CAE
S06309*	HCN2	chr19:590258:G:T p.Gly105Cys	ENST00000251287.3	D	P	25.4	0.0000205	PM2, PP2, PP3	GGE
S24513	DEPDC5	chr22:31768830:A:G p.Tyr69Cys	ENST00000458532.2	D	D	27.6	0.0000525	PM2, PP3	GGE

S27704 c	DEPDC5	chr22:31797646:G:A p.Val272Ile	ENST00000645693.1	D	D	25.5	0.001558	PM2, PP3	NAFE
S14233	DEPDC5	chr22:31802725:A:G p.Asn323Ser	ENST00000645693.1	D	P	25.6	0.0002165	PM2, PP3	NAFE
S22455	DEPDC5	chr22:31802779:G:A p.Gly341Asp	ENST00000646701.1	D	D	27.6	NA	PM2, PP3	NAFE
S24282	KCNT1	chr9:135765705:C:T p.Arg409Trp	ENST00000488444.6	D	D	29.8	0.0000065	PM2, PP2, PP3	NAFE
S16893	KCNT1	chr9:135772718:C:T p.Thr652Ile	ENST00000488444.6	D	P	22.3	0.0000065	PM2, PP2, PP3	GGE JME
S20208	CHRNB2	chr1:154571869:T:C p.Met349Thr	ENST00000368476.4	D	P	25.9	0	PM2, PP3	GGE
S22413*	CHRNB2	chr1:154571191:C:T p.Ala123Val	ENST00000368476.4	D	D	26.4	NA	PM2, PP3	NAFE
S22406	CHRNA2	chr8:27471000:A:G p.Leu20Pro	ENST00000240132.7	D	P	23.7	NA	PM2, PP3	GGE
S24281*	CHRNA2	chr8:27469841:G:A p.Arg72Cys	ENST00000240132.7	D	D	28.8	0.0001455	PM2, PP3	NAFE
S24437	CHRNA2	chr8:27469873:C:T p.Arg61Gln	ENST00000240132.7	D	P	26.1	0	PM2, PP3	GGE
S12337 c	CHRNA4	chr20:63346797:C:T p.Val609Ile	ENST00000370263.9	D	D	25.3	0.0001455	PM2, PP3	GGE, JME
S22413*	CHRNA4	chr20:63350473:G:A p.Thr313Ile	ENST00000370263.9	D	D	24.9	NA	PM2, PP3	NAFE
S22456 c	CHRNA4	chr20:63356390:G:A p.Thr85Met	ENST00000370263.9	D	D	25.9	0	PM2, PP3	NAFE
S27728	CHRNA4	chr20:63349726:G:A p.Pro562Leu	ENST00000370263.9	D	D	25.3	0	PM2, PP3	NAFE

'*' is shown for individuals with more than one variant in the same set of genes.
'c' is shown for variants that were also found in the unaffected control group.

Supplementary 2. VUS in missense-intolerant epilepsy genes, identified in the epilepsy cohort.

№	Gene	Variant	Transcript	SIFT	Polyphen2	CADD	gnomADAF	ACMG classification	Phenotype
S12354*c	ATP1A2	chr1:160136941:C:T p.Thr917Met	ENST00000265593.9	D	D	26.8	NA	PM2, PP2, PP3	GGE, JME
S27645c	ATP1A2	chr1:160135591:G:C p.Gly758Ala	ENST00000361216.8	D	P	26	0.000203 5	PM2, PP2 PP3	NAFE
S13833*c	CACNA1E	chr1:181798748:C:T p.Arg2286Trp	ENST00000367573.7	D	D	24.9	0.00006 5	PM2, PP2 PP3	NAFE
S17095c	CACNA1E	chr1:181794991:G:A p.Arg2052Gln	ENST00000367573.7	D	P	25.1	0.00003 9	PM2, PP2, PP3	GGE, JME
S24499c	CACNA1E	chr1:181798362:G:A p.Arg2157Gln	ENST00000367573.7	D	D	26.5	0.000565 5	PM2, PP2, PP3	GGE, JME
S12373*	KCNT2	chr1:196398587:T:C p.Asn424Asp	ENST00000294725.13	D	D	27.1	NA	PM2, PP2, PP3	GGE, JME
S13544c	SCN3A	chr2:165130164:C:T p.Gly900Ser	ENST00000283254.12	D	D	27	0.000032 5	PM2, PP2, PP3	GGE, CAE
S22475	GABRB2	chr5:161330899:C:T p.Arg354His	ENST00000393959.6	D	P	26.8	0.000065 5	PM2, PP2, PP3	NAFE
S28118	KCNH2	chr7:150958241:G:T p.Pro245His	ENST00000262186.10	D	D	22.7	NA	PM2, PP2, PP3	NAFE
S22440c	KCNH2	chr7:150952757:C:T p.Val409Met	ENST00000262186.10	D	P	26.5	0.000019 5	PM2, PP2, PP3	NAFE
S22416c	KCNH2	chr7:150948476:C:T p.Arg887His	ENST00000262186.10	D	P	22.9	0.000006 5	PM2, PP2, PP3	NAFE
S16897c	KCNH2	chr7:150959602:G:A p.Arg148Trp	ENST00000262186.10	D	P	24.6	0.000644	PM2, PP2, PP3	GGE
S16896c	KCNH2	chr7:150959602:G:A p.Arg148Trp	ENST00000262186.10	D	P	24.6	0.000644	PM2, PP2, PP3	GGE, JME
S01565*	KCNH2	chr7:150957437:G:A p.Arg328Cys	ENST00000262186.10	D	D	31	0.000610 5	PM2, PP2, PP3	GGE, CAE
S01581c	KCND2	chr7:120274652:C:A p.Ala7Glu	ENST00000331113.9	D	D	25.6	0.000006 5	PM2, PP2, PP3	GGE, JME

S27657	ACTL6B	chr7:100647014:C:T p.Arg298Gln	ENST00000160382.10	D	P	26.7	0.0000065	PM2, PP2, PP3	NAFE
S27641	ACTL6B	chr7:100656336:C:G p.Gly7Arg	ENST00000160382.10	D	D	31	NA	PM2, PP2, PP3	NAFE
S27729*	SPTAN1	chr9:128584289:T:C p.Ile734Thr	ENST00000372731.8	D	D	24.1	NA	PM2, PP2, PP3	NAFE
S27729*	DNM1	chr9:128203525 :G:T p.Ala19Ser	ENST00000372923.7	D	P	26.2	NA	PM2, PP2, PP3	NAFE
S20201	GPHN	chr14:66824543:G:A p.Ala91Thr	ENST00000315266.9	D	P	27.2	0.0001905	PM2, PP2, PP3	NAFE
S22380*c	GPHN	chr:14:66508555:A:T p.Asn10Tyr	ENST00000478722.5	D	D	28.5	0.000164	PM2, PP2, PP3	GGE
S17120*c	GABRA5	chr15:26883398:G:C p.Gly113Ala	ENST00000335625.10	D	D	26	0.000039	PM2, PP2, PP3	GGE
S16887	GNAO1	chr16:56343893:T:G p.Phe336Leu	ENST00000262494.12	D	D	24.2	0.0000065	PM2, PP2, PP3	GGE
S16871	ATP6V0A1	chr17:42460911:G:A p.Arg6Gln	ENST00000264649.10	D	P	32	0.0002365	PM2, PP2, PP3	GGE
S22474	ATP6V0A1	chr17:42495163:C:T p.Arg489Trp	ENST00000264649.10	D	D	32	0.0000195	PM2, PP2, PP3	NAFE
S22485c	NF1	chr17:31340589:G:A p.Ala2336Thr	ENST00000358273.9	D	P	27.1	0.0000325	PM2, PP2, PP3	NAFE
S20217	NF1	chr17:31230358:C:T p.Ser1030Leu	ENST00000356175.7	D	D	23	NA	PM2, PP2, PP3	GGE
S01565*c	HCN2	chr19:590258:G:T p.Gly105Cys	ENST00000251287.3	D	P	25.4	0.0000205	PM2, PP2, PP3	GGE, CAE
S06309*	HCN2	chr19:590258:G:T p.Gly105Cys	ENST00000251287.3	D	P	25.4	0.0000205	PM2, PP2, PP3	GGE
S22421c	GRIN2D	chr19:48414553:C:T p.Pro461Ser	ENST00000263269.3	D	P	25.6	0.0000065	PM2, PP2, PP3	NAFE
S01542c	SLC12A5	chr20:46043701 :A:G p.Ile436Val	ENST00000243964.6	D	P	24.6	0.0000195	PM2, PP2, PP3	GGE, CAE
S06309*c	SLC12A5	chr20:46043701 :A:G p.Ile436Val	ENST00000243964.6	D	P	24.6	0.0000195	PM2, PP2, PP3	GGE
S24439c	SLC12A5	chr20:46043701 :A:G p.Ile436Val	ENST00000243964.6	D	P	24.6	0.0000195	PM2, PP2, PP3	GGE
S27714	AP2M1	chr3:184180229:C:T p.Thr134Met	ENST00000382456.7	D	P	26.2	0.000013	PM2, PP2, PP3	NAFE

S12326*c	ATP1A3	chr19:41981552:G:A p.Arg433Cys	ENST00000602133.5	D	P	26	0.000479 5	PM2, PP2, PP3	GGE
S16866	CLTC	chr17:59644286:T:A p.Leu18Gln	ENST00000621829.4	D	P	24.4	0.000039	PM2, PP2, PP3	GGE
S15718c	CACNA1B	chr9:137955780:C:G p.Leu385Val	ENST00000371372.6	D	D	25	0.000026	PM2, PP2, PP3	NAFE
S22444	CACNA1B	chr9:137984214:G:A p.Arg578Gln	ENST00000371372.6	D	D	29.9	NA	PM2, PP2, PP3	NAFE
S22376c	CACNA1B	chr9:138052136:T:C p.Val1252Ala	ENST00000371372.6	D	D	27.5	NA	PM2, PP2, PP3	GGE, JAE
S22369	CACNA1B	chr9:138096536:G:A p.Arg1716Gln	ENST00000371372.6	D	D	32	0.000013	PM2, PP2, PP3	GGE
S13833*	CACNA1B	chr9:138120844:C:T p.Ser2151Leu	ENST00000277551.6	D	D	25.2	0.000256	PM2, PP2, PP3	NAFE
S15720	CACNA1B	chr9:138120844:C:T p.Ser2151Leu	ENST00000277551.6	D	D	25.2	0.000256	PM2, PP2, PP3	NAFE
S24437c	CACNA1B	chr9:138121989:A:G p.His2337Arg	ENST00000371372.6	D	P	23	0.007619	PM2, PP2, PP3	GGE
S27648	PLCB1	chr20:8757109:G:T p.Gly762Trp	ENST00000625874.2	D	P	31	NA	PM2, PP2, PP3	NAFE
S16865	KIF2A	chr5:62363753:A:G p.Ile421Val	ENST00000381103.6	D	P	25.5	NA	PM2, PP2, PP3	GGE
S27655	SCN8A	chr12:51806635:C:T p.Arg1717Cys	ENST00000354534.11	D	P	26.2	0	PM2, PP2, PP3	NAFE
S24435	SLC6A1	chr3:11026305:G :A .Val342Met	ENST00000644803.1	D	D	28.5	0.000006 5	PM2, PP2, PP3	GGE
S24441	SLC6A1	chr3:11026305:G :A .Val342Met	ENST00000644803.1	D	D	28.5	0.000006 5	PM2, PP2, PP3	GGE, CAE
S27646	STXBP1	chr9 :127663344:G :A p.Arg71Gln	ENST00000496504.3	D	D	32	NA	PM2, PP2, PP3	NAFE
S17048	PPP3CA	chr4:101025902:G:A p.Thr278Met	ENST00000512215.5	D	D	25.3	0.00002	PM2, PP2, PP3	GGE
S17120*	PPP3CA	chr4:101025902:G:A p.Thr278Met	ENST00000512215.5	D	D	25.3	0.00002	PM2, PP2, PP3	GGE
S27660c	SCN1A	chr2:165991324:G:T p.Pro1890His	ENST00000641603.1	D	P	26	0.000683 5	PM2, PP2, PP3	NAFE
S27649*	SCN1A	chr2:166013761:G:A p.Leu1219Phe	ENST00000635776.1	D	D	30	NA	PM2, PP2, PP3	NAFE
S27659	SCN1A	chr2:166038128:C:T p.Arg854Gln	ENST00000635776.1	D	D	28.1	NA	PM2, PP2, PP3	NAFE

S27725	SCN1A	chr2:166046828:T:A p.Lys440Ile	ENST00000635776.1	D	D	29.7	NA	PM2, PP2, PP3	NAFE
S27640	CYFIP2	chr5:157389393:T:A p.Cys1138Ser	ENST00000618329.4	D	D	24.9	0.000575	PM2, PP2, PP3	NAFE
S27649*	DOCK7	chr1:62555950:A:G p.Ile824Thr	ENST00000635253.2	D	P	23.8	0	PM2, PP2, PP3	NAFE
S28124	KCNQ2	chr20:63444721:G:A p.Arg210Cys	ENST00000629676.2	D	D	32	NA	PM2, PP2, PP3	NAFE
S22434*	KCNQ2	chr20:63472358:G:C p.Arg36Gly	ENST00000629241.2	D	D	23.6	NA	PM2, PP2, PP3	NAFE
S01564c	KCNQ5	chr6:72622240:G:C p.Trp17Cys	ENST00000403813.6	D	P	26.4	0.000026	PM2, PP2, PP3	GGE
S12373*c	KCNQ5	chr6:73194725:G:A p.Ala695Thr	ENST00000403813.6	D	P	23	0.0001375	PM2, PP2, PP3	GGE, JME
S27662	NEDD4L	chr18:58342928:G:A p.Arg403His	ENST00000356462.10	D	P	25.3	0.0000195	PM2, PP2, PP3	NAFE
S12354*c	PPP2R5D	chr6:43007213:C:G p.Phe180Leu	ENST00000485511.6	D	D	25.6	0.0000065	PM2, PP2, PP3	GGE, JME
S06281	DYRK1A	chr21:37506134 :C:T p.Arg519Trp	ENST00000647188.2	D	D	22.6	NA	PM2, PP2, PP3	NAFE
S12326*c	DYRK1A	chr21:37511914:C:T p.Arg559Cys	ENST00000644942.1	D	D	27.4	0.000013	PM2, PP2, PP3	GGE
S22380*c	DYRK1A	chr21:37511914:C:T p.Arg559Cys	ENST00000644942.1	D	D	27.4	0.000013	PM2, PP2, PP3	GGE
S16881c	DYRK1A	chr21:37512274:G:C p.Ala679Pro	ENST00000644942.1	D	D	24.6	0.003976	PM2, PP2, PP3	GGE

'*'*' is shown for individuals with more than one variant in the same set of genes.
'c' is shown for variants that were also found in the unaffected control group.