

GENOTYPE-PHENOTYPE CORRELATIONS IN COMPLEX CHILDHOOD EPILEPSIES AND DEVELOPMENTAL EPILEPTIC ENCEPHALOPATHIES IN AN INDIAN COHORT

ALFIYA F

PhD THESIS

2024



**SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND
TECHNOLOGY, TRIVANDRUM**

An Institution of National Importance established by an Act of the Indian Parliament
(Act No.52 of 1980)

Dept. of Science and Technology, Govt. of India
www.sctimst.ac.in

**GENOTYPE-PHENOTYPE CORRELATIONS IN
COMPLEX CHILDHOOD EPILEPSIES AND
DEVELOPMENTAL EPILEPTIC
ENCEPHALOPATHIES IN AN INDIAN COHORT**

A THESIS SUBMITTED BY

ALFIYA F

TO

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND
TECHNOLOGY, TRIVANDRUM.

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE AWARD OF

DOCTOR OF PHILOSOPHY

2024

DECLARATION BY THE STUDENT

CERTIFICATE

I, **Alfiya F** hereby certify that I had personally carried out the work depicted in the thesis titled, “*Genotype-phenotype correlations in complex childhood epilepsies and developmental epileptic encephalopathies in an Indian cohort*”, except where due acknowledgment has been made in the text.

No part of this thesis has been submitted for the award of any other degree or diploma prior to this date.



Signature

Alfiya F

Date 29.01.2025

श्री चित्रा तिरुनाल आयुर्विज्ञान और प्रौद्योगिकी संस्थान

तिरुवनन्तपुरम - 695 011, केरल, भारत

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY

THIRUVANANTHAPURAM - 695 011, INDIA

(An Institute of National importance under Govt. of India)



CERTIFICATE BY THE RESEARCH GUIDE

Name of the Guide: Dr. Ramshekhar N Menon

Division/Department: Neurology

This is to certify that **ALFIYA F**, department/division of Neurology of this institute has fulfilled the requirements prescribed for the Ph.D. degree of the Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum.

The thesis entitled, "*Genotype-phenotype correlations in complex childhood epilepsies and developmental epileptic encephalopathies in an Indian cohort*" was carried out under my direct supervision. No part of the thesis was submitted for the award of any degree or diploma prior to this date.

Clearance was obtained from the Institutional Ethics Committee for carrying out the study.

Signature

Dr. Ramshekhar N Menon

Date 29.01.2025

श्री चित्रा तिरुनाल आयुर्विज्ञान और प्रौद्योगिकी संस्थान
तिरुवनन्तपुरम - 695 011, केरल, भारत
SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY
THIRUVANANTHAPURAM - 695 011, INDIA
(An Institute of National importance under Govt. of India)



CERTIFICATE BY THE RESEARCH CO-GUIDE

Name of the Co-guide: Dr. Ashalatha R

Division/Department: Neurology

This is to certify that **ALFIYA F**, department/division of Neurology of this institute has fulfilled the requirements prescribed for the Ph.D. degree of the Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum.

The genotype-phenotype correlation analysis work under the thesis entitled, "*Genotype-phenotype correlations in complex childhood epilepsies and developmental epileptic encephalopathies in an Indian cohort*" was carried out under my direct supervision. No part of the thesis was submitted for the award of any degree or diploma prior to this date.

Clearance was obtained from the Institutional Ethics Committee for carrying out the study.

Signature

Dr Ashalatha R

Date 29.01.2025

तार : चित्रमेट
Grams : Chitramet

फोन
Phone : 2443152

फाक्स
Fax : (91)471-2446433
2550728

ई-मेल
E-mail : sct. @sctimst.ac.in
वेबसाइट
Website : www.sctimst.ac.in

Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram 695014, Kerala State, India.
An Autonomous National Institute for Discovery, Innovation & Translation
in Biotechnology and Disease Biology,
Government of India, Ministry of Science & Technology, Department of Biotechnology.

राजीव गाँधी जैव प्रौद्योगिकी केन्द्र, तिरुवनन्तपुरम 695014, केरल, भारत.
जैवप्रौद्योगिकी और रोग जीवविज्ञान में आविष्कार, नवीनता एवं अनुवाद
की स्वायत्त राष्ट्रीय संस्थान,
भारत सरकार विज्ञान एवं प्रौद्योगिकी मंत्रालय, जैवप्रौद्योगिकी विभाग.

**PhD program affiliated to the Sree Chitra Tirunal Institute for Medical Sciences
and Technology, Trivandrum**

CERTIFICATE BY THE RESEARCH CO-GUIDE

Name of the Co-guide: Dr. Moinak Banerjee

Division/Department: Human Molecular genetics laboratory,
Division of Neurobiology and Genetics,
Rajiv Gandhi Centre for Biotechnology,

This is to certify that **ALFIYA F**, department/division of Neurology of this institute has fulfilled the requirements prescribed for the Ph.D. degree of the Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum.

The genotype-phenotype correlation analysis work under the thesis entitled, "*Genotype-phenotype correlations in complex childhood epilepsies and developmental epileptic encephalopathies in an Indian cohort*" was carried out under my direct Co-supervision. No part of the thesis was submitted for the award of any degree or diploma prior to this date.

Signature



Dr Moinak Banerjee

Date: 28.1.2025

Dr. MOINAK BANERJEE
Scientist-G
Human Molecular Genetics Laboratory
Rajiv Gandhi Centre for Biotechnology
Thiruvananthapuram - 695 014, Kerala State

श्री चित्रा तिरुनाल आयुर्विज्ञान और प्रौद्योगिकी संस्थान

तिरुवनन्तपुरम - 695 011, केरल, भारत

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY

THIRUVANANTHAPURAM - 695 011, INDIA

(An Institute of National importance under Govt. of India)



APPROVAL OF THE THESIS

The thesis entitled

Genotype-phenotype correlations in complex childhood epilepsies and
developmental epileptic encephalopathies in an Indian cohort

Submitted by

Alfiya F

for the degree of

Doctor of Philosophy

of

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND
TECHNOLOGY, TRIVANDRUM

is evaluated and approved by

Dr. Ramshekhar N Menon

(Name & Signature of the Guide)

.....

(Name & Signature of thesis examiner)

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Ramshekhar N Menon, for his invaluable guidance, unwavering enthusiasm, and the freedom and encouragement he provided throughout the duration of this study. I consider myself extremely fortunate to have had the privilege of working under his mentorship. It was his understanding, personal support, and patience that enabled me to complete my research work with a sense of accomplishment. This thesis represents not only my efforts but also the invaluable guidance he has provided, for which I am profoundly grateful.

I would like to express my deepest gratitude to my advisor, Dr. Moinak Banerjee, whose unwavering support, patience and guidance over the past six years have been instrumental in the completion of this study. From the very beginning of this journey, he has provided me with the perfect balance of freedom to explore my ideas and constructive feedback to shape them into a meaningful contribution to this field. His support during challenging times and his belief in my capabilities have served as a source of motivation and inspiration, driving me to expand the limits of my own potential.

The Director, SCTIMST is sincerely acknowledged for the resources made available on institute to facilitate the research activities. I extend my heartfelt thanks to the Registrar, the Deputy Registrar and the Dean for their invaluable academic support and assistance throughout this endeavor. I would also like to express my sincere gratitude to ICMR for their financial support in making this study possible.

I express my gratitude to Dr. Ashalatha R, Head of the Department of Neurology, for her encouragement and assistance, which significantly contributed to my work. I express my gratitude to the members of my Doctoral Advisory Committee, Dr. Dinesh Gupta, Dr. Sheela Nampoothiri and Dr. Madhusoodanan U K., for their invaluable time, support, insights, and corrections, which significantly enhanced the quality of my thesis and presentation.

I would also like to express my heartfelt gratitude to the Director, RGCB, for providing all the necessary facilities and resources to carry out my work.

I wish to express my heartfelt gratitude to Dr. Manna Jose for the invaluable support provided in collecting the clinical data, as well as for the assistance offered throughout my research. I sincerely thank Dr. Soumya Sundaram for her valuable time, support, and suggestions in my work which helped to improve my thesis. The support and help given by PhD students, post doctoral fellows, residents, and staff of SCTIMST are greatly acknowledged. I offer my sincere thanks to all the scientists and PhD students of BMT wing for their support and suggestions.

I want to express my heartfelt thanks to my previous and current HMG lab members, Dr. Shafeeqe C M, Dr. Shabeesh Balan, Dr. Aswathy P.M, Dr. Sindura K P, Dr. Anil Prakash, Binithamol K Polakkattil, Dr. Krishna Priya E K, Neethu Mohan, Rashmi Sukumaran, Samyukta Bhass, and Ardra M, for the incredible support they have provided throughout this journey. Their encouragement during tough moments, and the everyday kindness have truly been a pillar of strength for me. They have transformed the lab into a second home, and I feel so fortunate to work with such an amazing group of individuals. I would also wish to thank all my trainees for their invaluable assistance in the completion of my study.

My deepest and heartfelt thanks go to my husband and my son for their unwavering love, patience, and encouragement. They have been my greatest source of strength throughout this journey, standing by me through every challenge and making sacrifices so I could pursue my passion.

I thank my parents, my brother, my in laws, and all my family members who supported me throughout my PhD programme to make it a reality and dream come true. I thank the Almighty for blessing me with the opportunity to do research, for giving me the health, courage, and inner strength which kept me moving in all the ups and downs during the entire study period.

TABLE OF CONTENTS

DECLARATION BY THE STUDENT	i
CERTIFICATE BY THE RESEARCH GUIDE.....	ii
CERTIFICATE BY THE RESEARCH CO-GUIDE.....	iii
CERTIFICATE BY THE RESEARCH CO-GUIDE.....	ii
ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS.....	viii
LIST OF FIGURES	xi
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xv
SYNOPSIS.....	xvi
1 INTRODUCTION	1
1.1 Background	1
1.1.1 Genetics of DEE.....	3
1.1.2 Genetic Testing in DEE	6
1.1.3 Trio whole exome sequencing in the diagnosis of DEE	8
1.1.4 Functional analysis of disease-causing variants.....	9
1.1.5 Variants of uncertain significance	11
1.2 Hypothesis.....	12
1.3 Objectives.....	13
2 LITERATURE REVIEW.....	15
2.1 Childhood epilepsy and Developmental epileptic encephalopathies.....	15
2.2 Genetics of epilepsy and DEE.....	24
2.3 Disease mechanisms in childhood epilepsy and DEE.....	31
2.3.1 Voltage-gated ion channels	32
2.3.2 Ligand-gated ion channels	39
2.3.3 mTOR pathway	40
2.3.4 Metabolic pathways	41
2.3.5 Synaptic transmission	42
2.3.6 Chromatin remodeling and transcription regulation	44
2.4 Genetic testing in DEE.....	45
2.4.1 Next generation sequencing	45
2.4.2 Chromosomal microarray.....	50
2.4.3 Karyotyping	50
2.4.4 Multiplex-ligase dependent probe amplification	51
2.4.5 Sanger sequencing.....	51
2.5 Precision medicine	51
2.6 Functional analysis.....	53
3 MATERIALS AND METHODS.....	57

3.1	Subjects/study design	57
3.1.1	Inclusion criteria.....	59
3.1.2	Exclusion criteria	61
3.1.3	Phenotyping	61
3.2	DNA isolation	63
3.3	WES and Bioinformatics analysis.....	64
3.3.1	Functional prediction tools.....	65
3.3.2	Variant prioritization and classification	67
3.4	ACMG classification criteria for SNVs and INDELs.....	68
3.5	Copy number variant calling.....	71
3.6	Sanger sequencing validation for pathogenic/likely pathogenic variants ..	73
3.7	<i>In silico</i> functional analysis.....	76
3.7.1	Protein network analysis	76
3.7.2	Protein stability prediction	77
3.7.3	Protein structural prediction.....	78
3.7.4	Pathway analysis	78
3.7.5	Functional prediction for ion channel variants.....	79
3.7.6	AlphaMissense	81
3.8	Grouping of VUS into promising VUS/incidental VUS.....	81
3.9	Molecular dynamics simulation	82
3.9.1	Root mean square deviation	83
3.9.2	Root Mean Square Fluctuation.....	83
3.9.3	Radius of gyration	84
3.9.4	Protein Secondary Structure of Protein (DSSP) analysis.....	84
3.10	Multidisciplinary team meeting	85
3.11	Re-analysis of negative cases.....	85
4	RESULTS AND DISCUSSION	88
4.1	Genetic diagnostic yield	88
4.1.1	Samples and diagnosis	88
4.1.2	Diagnostic yield	91
4.1.3	Epilepsy syndromes and associated genes	101
4.1.4	Copy number variants	102
4.1.5	Precision medicine	102
4.1.6	Sanger sequencing validation.....	103
4.1.7	IGV visualization	104
4.2	Variants of uncertain significance.....	105
4.3	Re-analysis of negative cases.....	114
4.4	<i>In silico</i> functional analysis of disease causing variants.....	115
4.4.1	Protein interaction network.....	115
4.4.2	Human phenotype ontology enrichment analysis	117
4.4.3	Functional annotation and gene ontology analysis	120

4.4.4	Functional prediction of voltage-gated sodium ion channel variants.....	124
4.4.5	<i>In silico</i> functional prediction of a novel variant in STXBP1 gene .	128
4.4.6	<i>In silico</i> functional prediction of a homozygous recessive variant in UFSP2 gene.....	136
4.5	Functional and structural analysis of VUS with potential phenotype match.....	142
4.5.1	<i>In silico</i> functional analysis of SHQ1 wilf type and mutant protein	144
5	SUMMARY AND CONCLUSION.....	152
6	REFERENCES.....	159
ANNEXURES		

LIST OF FIGURES

Figure No	Figure Caption	Page No
Fig 1.1	The genetic framework of epilepsies	4
Fig 2.1	ILAE 2017 Classification of the Epilepsies	16
Fig 2.2	Classification of neonatal and infant epilepsy syndromes	18
Fig 2.3	Pathways implicated in Childhood epilepsy and DEE	31
Fig 2.4	Representation of alpha subunit of voltage-gated sodium ion channel: Two dimensional	36
Fig 2.5	Diagnostic tests in refractory childhood epilepsies and DEE	46
Fig 3.1	Workflow	58
Fig 3.2	Diagram that categorizes each criterion based on the type of evidence and the robustness of the criteria, distinguishing between benign assertions on the left side and pathogenic assertions on the right side	70
Fig 3.3	Illustration of variant categorization as either pathogenic or likely pathogenic, employing various tiers of evidence (Fasaludeen et al., 2024).	71
Fig 3.4	Grouping of VUS into Promising VUS/Incidental VUS	82
Fig 4.1	Study profile	89
Fig 4.2	Type of disease causing variants	92
Fig 4.3	Types of variants in diagnostic cases	92
Fig 4.4	Diagnostic yield comparison between trios, duos and singletons	95
Fig 4.5	Inheritance pattern of P/LP variants	96
Fig 4.6	Genetic diagnosis of epilepsy syndromes	97
Fig 4.7	Comparison of diagnostic yield (in %) and age of onset of epilepsy	99
Fig 4.8	Correlation analysis between age of onset and diagnostic yield	99

Fig 4.9	Functional mechanisms of pathogenic variants identified in our cohort	100
Fig 4.10	Epilepsy syndromes and associated disease causing genes	101
Fig 4.11	Representative Sanger sequencing validation image of a <i>de novo</i> variant <i>CHRNA4</i> (c.1105C>T)	104
Fig 4.12	Representative IGV visualization image of a <i>de novo</i> variant in <i>WDR45</i> (c.827+1G>A)	104
Fig 4.13	Gene-gene interaction of causative genes predicted by STRING	116
Fig 4.14	Phenotype enrichment analysis	118
Fig 4.15	Human phenotype enrichment analysis of causative genes using David functional database	119
Fig 4.16	Functional annotation of causative genes	120
Fig 4.17	Gene ontology analysis (Biological process)	122
Fig 4.18	Gene ontology analysis (Molecular function)	123
Fig 4.19	Gene ontology analysis (Cellular component)	123
Fig 4.20	Causative variants in SCN1A labeled in primary structure	126
Fig 4.21	Gene expression of STXBP1	129
Fig 4.22	Protein interaction network of STXBP1 protein	130
Fig 4.23	RMSD analysis of STXBP1 wild type (black) and mutant protein (red)	132
Fig 4.24	Changes in the radius of gyration of the wild type structure (black) and mutant structure (red)	133
Fig 4.25	RMSF analysis of the wild type (black) and mutant protein (red) during the 100ns MD simulation	134
Fig 4.26	Definition Secondary Structure of Protein analysis of wild-type (left) and mutant structure (right) at different time period	135
Fig 4.27	Gene expression data of UFSP2 protein	137
Fig 4.28	Protein interaction network of UFSP2 protein	138
Fig 4.29	RMSD analysis of wild type (black) and mutant type (red) protein structures of UFSP2 during 100 nanoseconds MD simulation	139
Fig 4.30	Rg analysis of wild type (black) and mutant type (red) UFSP2 protein	140

Fig 4.31	RMSF analysis wild type (black) and mutant type (red) UFSP2 protein	140
Fig 4.32	Definition Secondary Structure of Protein analysis of wild-type (left) and mutant structure (right) at different time period	141
Fig 4.33	Gene expression analysis of SHQ1	145
Fig 4.34	Protein interaction network analysis of SHQ1 protein	145
Fig 4.35	RMSD analysis of SHQ1 wild type (black) and mutant type (red) protein	148
Fig 4.36	RMSF analysis SHQ1 wild type (black) and mutant type (red) protein	149
Fig 4.37	Rg analysis of SHQ1 wild type (black) and mutant type (red) protein	149
Fig 4.38	Time evolution of secondary structural elements of wild type (left) and mutant type protein (right)	150

LIST OF TABLES

Table No	Table Caption	Page No
Table 2.1	Genes associated with complex childhood epilepsy and DEE	25
Table 2.2	Comparison of Genetic tests in DEE	48
Table 2.3	Examples of precision medicine applied in the treatment of epilepsy	52
Table 3.1	Functional prediction tools used for variant filtration and prioritization	66
Table 4.1	Demographic and clinical characteristics of probands	90
Table 4.2	List of pathogenic/likely pathogenic variants identified	93
Table 4.3	Clinical characteristics and diagnostic yield	98
Table 4.4	List of P/LP copy number variants identified	102
Table 4.5	Precision medicine applied	103
Table 4.6	List of VUS	105
Table 4.7	<i>In silico</i> functional prediction of causative variants in <i>SCN1A</i> gene	125
Table 4.8	Functional prediction of voltage-gated sodium channel variants	126
Table 4.9	<i>In silico</i> functional prediction of promising VUS	143

LIST OF ABBREVIATIONS

S No	Abbreviation	Full Form
1	DEE	Developmental and Epileptic Encephalopathies
2	ILAE	International League Against Epilepsy
3	EIEE	Early Infantile Epileptic Encephalopathy
4	NGS	Next-generation sequencing
5	WGS	Whole genome sequencing
6	WES	Whole exome sequencing
7	SNV	Single nucleotide variants
8	INDELs	Insertions/Deletions
9	CNV	Copy-number variants
10	ACMG	American College of Medical Genetics and Genomics
11	P/LP	Pathogenic or Likely pathogenic
12	VUS	Variants of Uncertain significance
13	VUS-D	VUS with potential phenotype match or Promising VUS
14	LKS/CSWS	Landau-Kleffner syndrome/Continuous spike-and-wave during sleep
15	LGS	Lennox-Gastaut Syndrome
16	EIMFS	Epilepsy of infancy with migrating focal seizures
17	EMAS	Epilepsy with myoclonic-atonic seizures
18	DSSP	Definition Protein Secondary Structure of Protein
19	RMSD	Root-mean-square-deviation
20	RMSF	Root-mean-square-fluctuation
21	Rg	Radius of gyration

SYNOPSIS

The thesis entitled ‘Genotype-Phenotype Correlations in Complex Childhood Epilepsies and Developmental Epileptic Encephalopathies in an Indian Cohort’ deals with the genetic diagnostic yield and *in silico* functional/structural analysis of deleterious variants associated with Developmental and Epileptic Encephalopathies (DEE).

DEE describes rare neurodevelopmental disorders with early infantile onset, characterized by the co-occurrence of epilepsy, intellectual disability, autism spectrum disorder, and developmental delay. Different aetiologies, such as structural, viral, metabolic, immunological, and genetic variables, can cause DEEs. It is challenging to understand the neurobiological mechanisms underlying the pathogenesis of DEE due to the etiological heterogeneity. Genetics plays a much greater role in determining the etiology of epilepsy. 70% of epilepsy cases are caused by genetic factors which can be single variants or multiple variants acting along with various environmental factors. Different twin studies have revealed significantly higher concordances of epilepsy for monozygotic twins than dizygotic twins. The recurrence chances of generalized epilepsy are five to ten times higher for first-degree relatives.

Genetic DEEs have been associated with mutations in several genes that are involved in various functions such as cell migration, proliferation, organization, synaptic transmission and plasticity, and brain excitability. The identification of genes has uncovered numerous disease mechanisms underlying severe epilepsies including channelopathies, synaptic dysfunction, transporter defects, transcriptional

dysregulation, impaired DNA repair and chromatin remodeling, as well as metabolic defects.

The advent of Next-generation sequencing (NGS) techniques allowed to identify a number of single gene mutations as well as copy number variants that can cause seizures and disrupt neurodevelopment. NGS has the potential to uncover causal mutations associated with epileptic syndromes and significantly enhance molecular diagnosis. NGS has started to replace Sanger sequencing because of its ability to sequence a large number of genes using targeted gene panels, whole exome sequencing (WES), and whole genome sequencing (WGS). Among the NGS techniques, whole-exome sequencing (WES) is the preferred method for the molecular diagnosis of these diseases due to their clinical and genetic heterogeneity. WES concentrates on the protein-coding regions of the genome, which cover 1- 2% of the total genome and account for about 85% of mutations that cause disease. Due to its low cost, high diagnostic yields, and outstanding benefits for the study of novel genes and subsequent research, WES is being used extensively in clinical practice.

With this background study, we hypothesize that WES gives a high yield of pathogenic variants associated with DEE as the variant profile is likely to be similar to other countries given the rarity of the disorder. The pathogenic variants identified potentially involve more than one functional pathway and the functional or structural analysis of these variants can be used to predict the variant pathogenicity.

The objectives of the works are:

1. To identify the risk variants in childhood epilepsy and developmental epileptic encephalopathy using WES
2. To investigate the genetic networks and gene ontology pathways by correlating risk variants and their phenotypes
3. To predict the functional and structural analysis of Variants of uncertain significance with potential phenotype match

The thesis is presented in six chapters. Chapter 1 begins with an introduction to the prevalence, types, and genetics of DEE. It also briefly introduces the importance of genetic diagnosis of DEE and *in silico* functional analysis of deleterious variants identified from NGS techniques. Chapter 2 includes a detailed literature review, concentrating on the genetic background of DEE, disease mechanism in childhood epileptic encephalopathies, diagnostic genetic testing in epilepsy, the importance of WES in the genetic diagnosis, and functional enrichment analysis of deleterious variants.

Chapter 3 covers the materials and methods utilized to accomplish the aforementioned objectives. Patients with complex childhood epilepsy (<12 years age of onset) associated with/without developmental delay, refractory epilepsy of unknown cause, or Epileptic Encephalopathy who attended at the pediatric epilepsy clinic conducted at SCTIMST, Trivandrum were included in this study. Results from magnetic resonance imaging (MRI), video electroencephalography (EEG), and developmental age were assessed. The first section includes the bioinformatics pipeline and various databases used for the WES analysis. Genomic DNA was isolated from the peripheral blood

samples of proband and his parents using the salting out method of DNA extraction. Sequencing libraries were generated using Agilent SureSelect Human All ExonV6 kit (Agilent Technologies) and paired-end sequenced on Illumina HiSeq X Ten (Illumina, San Diego, USA) for 150 cycles according to manufacturer's instructions. Raw fastq files were aligned to the human reference genome (hg19/GRCh37) using BWA (Burrows-Wheeler Aligner) and processed for duplicates removal and base recalibration. The SNPs and INDELs were called using Genome Analysis Toolkit's (GATK) best practice guidelines. The variants identified were annotated using ANNOVAR against allele frequency data, *in silico* prediction scores, and clinical relevance data. According to the guidelines of American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP), the variants were classified in a five-class scheme (Pathogenic/Likely Pathogenic/Benign/Likely Benign/Variants of Uncertain significance). The pipeline used for Copy Number Variant (CNV) calling, annotation and annotation was also explained.

The second section covers the tools and pipelines used for the functional analysis of pathogenic or likely pathogenic variants. To assess the pathogenicity of variants on protein function and structure stability, various *in silico* analysis tools were used based on protein stability, protein interaction pattern, and protein 3-dimensional structural analysis.

To identify the interactions of protein with other relevant proteins, STRING (Search Tool for the Retrieval of Interacting Genes) was used. Online support vector machine based tools, IMutant and MUpro were used to predict protein stability changes based on single site missense mutations. For performing pathway analysis, functional

annotation and enrichment analysis of genes identified with pathogenic variants, Wiki pathways and DAVID (Database for Annotation, Visualization and Integrated Discovery) functional annotation tools were used. FunNCion, a machine learning method which predicts the functional impact of voltage-gated sodium and calcium channel missense variants and SCN viewer, a web tool to facilitate the functional interpretation of sodium channel variants were used to detect the functional effect of pathogenic voltage-gated sodium channel missense variants identified. The pathogenicity of variants of uncertain significance with potential phenotype match was predicted using AlphaMissense, a deep learning model that predicts the pathogenicity of missense variants and classifies it as either likely benign, likely pathogenic or uncertain.

The third section includes the workflow of Molecular dynamics simulation (MDS). MDS is a computer simulation technique for analyzing the physical movements of atoms and molecules. AlphaFold predicted protein structure was extracted for the structural prediction analysis and the structures were validated using Ramachandran plot. The mutant residue was introduced using Mutagenesis wizard tool of PyMOL molecular graphics system.

A 100 nanosecond molecular dynamics simulation was carried out with periodic boundary conditions for wild-type and mutant protein models using GROMACS 2022.4. The root-mean-square-deviation (RMSD), root-mean-square-fluctuation (RMSF), and radius of gyration (Rg) were calculated. The secondary structure transitions of the wild-type and mutant protein were monitored using DSSP (Define Secondary structure of protein) algorithm.

Chapter 4 involves the results and discussion of the studies and is divided into four major sections. The first section includes the diagnostic yield of DEE and a comparative analysis of yield based on age of onset of epilepsy, various syndromes, functional consequences, biological mechanisms, and inheritance patterns. A significant diagnostic yield was achieved for trios and quads. There was a notable change ($P\text{-value} < 0.006$; Fisher's exact test) in the diagnostic yield was noted for patients with age of onset less than 3 years and patients with age of onset between 3 to 12 years.

Most of the disease causing variants identified were *de novo* and the remaining variants followed homozygous recessive, autosomal dominant, and compound heterozygous pattern of inheritance. The majority of disease-causing variants were linked to neurometabolism and ion channel genes. The second section covers the *in silico* functional and structural analysis of pathogenic variants associated with DEE. Protein interaction network analysis of the disease causing variants using STRING database revealed that most of the deleterious variants were enriched in ion channel genes; particularly in voltage-gated sodium channel genes. The gene ontology pathway analysis using DAVID functional annotation tool showed strong interactions of epilepsy genes associated with ion channel genes and neurometabolic transmitters. The human phenotype ontology enrichment analysis using wikipathways database showed that most of the disease causing variants were enriched in focal seizures and epileptic encephalopathy.

The third section includes the MDS analysis of STXBP1 (c.A434C:p.Tyr145Ser) and in UFSP2 (c.T344A:p.Val115Glu) variants. To evaluate the extent to which mutation

affects protein structure, RMSD, RMSF and Rg values were calculated for the wild and mutant protein structures. The results revealed the effect of mutation on protein stability, flexibility, compaction and structure of the protein.

The fourth section covers the functional and structural prediction of variants of uncertain significance (VUS) with potential phenotype matches. The analysis was done using stability prediction tool- MuPro, structural prediction tool- Missense3D, pathogenicity prediction tool- AlphaMissense and Molecular dynamics simulation.

After performing deep phenotyping and primary *in silico* functional and structural predictions of VUSs with potential phenotype match, the variant SETD1B:c.931C>T was re-classified as a pathogenic variant and SHQ1:c.850T>C as a promising VUS. For further analysis of SHQ1:c.850T>C variant, MDS was performed for 100ns. RMSD, RMSF and Rg values showed significant changes in the mutant protein when compared to the wild-type protein.

Chapter 5 provides the summary and conclusion of the study. This is the first study from India which demonstrates the effect of genetic diagnosis and clinical utility of patients with complex epilepsy and DEE through phenotyping, whole exome sequencing and interpretation. Establishing a correct diagnosis is crucial for guiding recurrence risk counseling, regardless of whether the variations are inherited or *de novo*. This helps families make informed decisions about reproduction. In order to further advance our knowledge of complex childhood epilepsy and enhance care for people who suffer from the condition, further genomic techniques including transcriptomes, metabolomes, and the expanded application of whole genome sequencing will likely be included in the future of epilepsy genetics. The study also

describes the functional and structural predictions of deleterious variants and VUSs identified from WES. This chapter also describes the further functional investigations needed to elucidate the molecular underpinnings of deleterious variants in epileptic encephalopathy. The thesis ends with bibliography, list of publications and other research outcomes.

1 INTRODUCTION

1.1 Background

Epilepsy is one of the most prevalent heterogeneous brain disorders that affects 70 million people worldwide (Singhi and Gupta, 2021). The prevalence rate is 5% in India with 12 million people affected, contributing to one-sixth of the global burden. In children in India, the prevalence rate ranges from 4.07 to 7.5 per 1000 people (Amudhan et al., 2015). According to the 2017 review of the Global Burden of Disease (GBD) study, the majority of epilepsy cases were concentrated in high-population, low-socioeconomic-status Asian and African nations, with India having the highest number of disability-adjusted life years (DALYs) followed by China, Nigeria, Pakistan, and the United States (Hu et al., 2021).

ILAE (International League Against Epilepsy) classification of epilepsies in 2017 presents three levels including seizure type, epilepsy type (focal epilepsy, generalized epilepsy, combined focal and generalized epilepsy, and unknown epilepsy), and epilepsy syndrome (Scheffer et al., 2017). In 2022, ILAE proposed a framework for classifying epilepsy syndromes with onset in the neonatal period and infancy (Zuberi et al., 2022). The epileptic syndrome has been defined as "a characteristic cluster of clinical and EEG features, often supported by specific etiological findings (structural, genetic, metabolic, immune, and infectious)". The age principle is used to classify epileptic syndromes into three categories: onset in childhood, onset in neonates and infants, and onset at a variable age. The term "epileptic encephalopathy" refers to a

presumed causal connection between developmental delay and epilepsy. An epileptic encephalopathy is characterized by recurrent episodes of epileptiform activity that adversely impact development, frequently resulting in the slowing down or regression of developmental skills and being linked to numerous seizures. Nowadays, it is recognized that a significant number of epileptic encephalopathies have a distinct molecular genetic origin. The word "developmental" was coined because affected individuals may experience developmental impairment due to many factors. Developmental and Epileptic Encephalopathies (DEE) describe rare neurodevelopmental disorders with early infantile onset, characterized by the co-occurrence of epilepsy, intellectual disability, autism spectrum disorder, and developmental delay.

Complex childhood epilepsy and DEEs are classified in our study as Early myoclonic encephalopathy, Ohtahara syndrome, Early Infantile Epileptic Encephalopathy (EIEE) (< 6 months of age of onset with focal or multifocal epilepsy), West syndrome, Dravet syndrome, Myoclonic atonic epilepsy, Lennox-Gastaut syndrome, Epilepsy of infancy with migrating focal seizures, Landau Kleffner syndrome/EE with spike and wave activation in sleep, Other focal/multifocal, Refractory generalized epilepsy and Progressive myoclonic epilepsy (Scheffer et al., 2017; Scheffer and Liao, 2020; Zuberi et al., 2022).

1.1.1 Genetics of DEE

DEEs result from various etiologies, including structural, infectious, metabolic, immune, and genetic factors. It is challenging to understand the neurobiological mechanisms underlying the pathogenesis of DEE due to the etiological heterogeneity. Understanding the root cause of a person's epilepsy or correctly identifying their specific epilepsy syndrome is crucial for guiding clinical management and anticipating outcomes.

Epidemiological research indicates that between 17% and 49% of children diagnosed with epilepsy also exhibit intellectual disabilities (Poke et al., 2023). In infancy, it is estimated that as much as 40% of epilepsy cases are linked to epileptic encephalopathies, with West syndrome being the predominant type, representing roughly one-third of these instances (Syrbe, 2022).

Genetics plays a much greater role in determining the etiology of epilepsy. 70% of epilepsy cases are caused by genetic factors which can be single variants or multiple variants acting along with various environmental factors (Thakran et al., 2020). Different twin studies have revealed significantly higher concordances of epilepsy for monozygotic twins than dizygotic twins (Kjeldsen et al., 2003; Vadlamudi et al., 2014). The recurrence chances of generalized epilepsy are five to ten times higher for first-degree relatives (Peljto et al., 2014).

The development of high-throughput sequencing has significantly impacted the genetic landscape of DEEs, identifying *de novo* pathogenic mutations and discovering additional DEE-associated genes. Genetic DEEs have been associated with mutations in several genes involved in various functions such as cell migration, proliferation, organization, synaptic transmission and plasticity, and brain excitability. The identification of genes has uncovered numerous disease mechanisms underlying severe epilepsies including channelopathies, synaptic dysfunction, transporter defects, transcriptional dysregulation, impaired DNA repair and chromatin remodeling, as well as metabolic defects (McTague et al., 2016). The genetic framework of epilepsies is illustrated in **Figure 1.1**, highlighting that a considerable portion of DEE phenotypes is linked to monogenic causes.

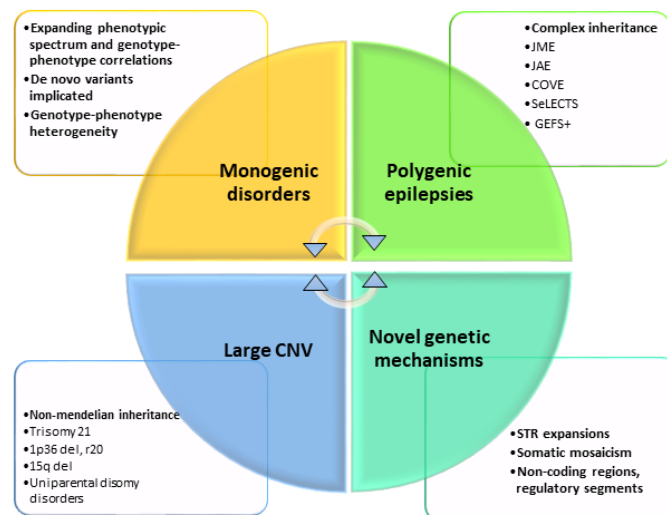


Figure 1.1: The genetic framework of epilepsies (Adapted from Fasaludeen et al., 2024)

The advent of Next-generation sequencing (NGS) techniques allowed us to identify several single-gene mutations as well as copy number variants that can cause seizures and disrupt neurodevelopment. NGS has the potential to uncover causal mutations associated with epileptic syndromes and significantly enhance molecular diagnosis. NGS has started to replace Sanger sequencing because of its ability to sequence a large number of genes using targeted gene panels, whole exome sequencing (WES), and WGS.

The application of genetic testing in pediatric epilepsy remains challenging in developing nations due to both technical limitations and financial barriers. Additionally, the catalog of identified epilepsy-related genes continues to grow daily. Nevertheless, given its significant influence on precision medicine and genetic counseling, genetic testing has become a crucial component of epilepsy research.

The American College of Medical Genetics (ACMG) guidelines categorize genetic variants into five distinct groups: benign, likely benign, variant of unknown significance (VUS), likely pathogenic, and pathogenic (Richards et al., 2015). This classification system is informed by several factors, including the type of variant (its prior documentation in affected individuals), familial segregation patterns, instances of *de novo* mutations, frequency in control populations, outcomes from functional analyses, anticipated impacts on protein function, positioning within functional domains, and evolutionary conservation across different species.

Up to 50% of patients with DEEs now have an aetiological diagnosis based on the identification of a pathogenic variant through molecular investigations such as gene panels, exome, and genome sequencing (Hamdan et al., 2017a; Mitta et al., 2020;

Palmer et al., 2018; Scheffer et al., 2023). To date, over 400 genes have been recognized as contributors to epileptic encephalopathies (EE) and DEE (Møller et al., 2019), with numerous additional genes proposed as potential candidates.

Many DEEs are linked to a specific genetic cause, with the pathogenic variant frequently arising *de novo* in the individual affected. A majority of patients present with *de novo* dominant mutations; however, there are also instances of X-linked and autosomal recessive etiologies (Balestrini et al., 2021). Identifying the pathogenic variant is essential, as several may be amenable to treatment. Furthermore, understanding the underlying disorder can guide pharmacological interventions or indicate the necessity for surgical assessment. The characteristics of the pathogenic variant, along with the genetic context, play a significant role in determining the phenotypic expression.

1.1.2 Genetic Testing in DEE

An accurate genetic diagnosis of epilepsy holds significant importance for both individuals and their families, offering substantial clinical and personal benefits. This is especially relevant in the context of developmental and epileptic encephalopathies, where early genetic testing has demonstrated cost-effectiveness and can eliminate the need for invasive investigations to identify a cause. Given the significant genetic diversity observed in the majority of epilepsies, NGS is widely regarded as the preferred approach for diagnostic evaluation and should be implemented as the primary investigative tool. Furthermore, NGS offers the advantage of facilitating thorough identification of both single nucleotide variants (SNVs) and copy number

variations (CNVs), and it is typically more economical than alternative methodologies. NGS techniques encompass exome sequencing (ES), genome sequencing (GS), and specialized gene panels that target epilepsy-related genes.

WES concentrates on the protein-coding regions of the genome, which cover 1- 2% of the total genome and account for about 85% of mutations that cause disease. Due to its low cost, high diagnostic yields, and outstanding benefits for the study of novel genes and subsequent research, WES is being used extensively in clinical practice.

The diagnostic yield of WES in DEE ranges from 24% to 46.6% (Fasaludeen et al., 2024; Sun et al., 2021; Vetri et al., 2024). A comprehensive study aimed at identifying both shared and unique ultra-rare genetic risk factors across various epilepsy types utilized WES on a cohort comprising 9,170 individuals affected by epilepsy and 8,436 control subjects of European descent. The findings revealed that, in comparison to the control group, individuals with any form of epilepsy exhibited a higher prevalence of ultra-rare, deleterious variants located in constrained genes, as well as in genes previously linked to epilepsy, with the most significant enrichment observed in those diagnosed with DEEs (Feng et al., 2019). In a nationwide initiative 'Care4Rare Canada' aimed at identifying genes responsible for rare pediatric epilepsies, a total of 82 WESs were conducted, resulting in the identification of 35 pathogenic variants, which corresponds to a detection rate of 43% (Dyment et al., 2015). A recent exome sequencing study conducted in Australia on patients diagnosed with developmental epileptic encephalopathy identified pathogenic or likely pathogenic variants in 29% of the cases (Scheffer et al., 2023).

Targeted or Epilepsy gene panel sequencing involves a targeted analysis of exome sequencing (ES) data related to genes linked to epilepsy. This approach is particularly beneficial in cases where multiple genes are identified as contributors to a similar phenotype across different individuals, especially when the clinical manifestations of the patient are notably unique. The diagnostic yield of targeted gene panel sequencing in individuals diagnosed with epileptic encephalopathy varies between 19% and 48.5% (Kothur et al., 2018; Mercimek-Mahmutoglu et al., 2015; Sánchez Fernández et al., 2019).

WGS, which covers the complete genome, can identify mutations in coding, non-coding, mitochondrial and major structural rearrangements, showing a diagnostic yield between 34% and 48% in epilepsy cases (Fasaludeen et al., 2024). The identification of pathogenic or potentially pathogenic *de novo* point mutations, insertions/deletions, and structural rearrangements linked to EIEE illustrates the effectiveness of WGS in minimizing both the time and expenses associated with the clinical diagnosis of EIEE (Ostrander et al., 2018).

Diagnostic Ultra Rapid Genome Sequencing and Rapid Genome Sequencing are clinical-grade methods that identify all major types of DNA variation in at least 90% of the nucleotides of an affected individual, with an accuracy rate of no less than 95% (Kingsmore et al., 2024). A multicenter, international cohort study known as Gene-STEPS (D’Gama et al., 2023) identified genetic diagnoses in infants presenting with new-onset epilepsy or complex febrile seizures, achieving a diagnostic rate of 43% through rapid genome sequencing. The median duration from the onset of seizures to the receipt of rapid genome sequencing results was recorded at 37 days.

1.1.3 Trio whole exome sequencing in the diagnosis of DEE

Among the NGS techniques, WES is the preferred method for the molecular diagnosis of DEE due to its clinical and genetic heterogeneity. Studies on parent-child or trio exome sequencing in cases of DEE have shown that *de novo* mutations (DNMs) are a major factor in the causes of DEE (Epi4K Consortium and Epilepsy Phenome/Genome Project, 2013; Hamdan et al., 2017a). WES of 10 trios with unaffected parents and their child revealed that seven children had *de novo* mutations in genes related to neuronal excitability suggesting trio-WES could aid in diagnosing sporadic pediatric epilepsies, especially in early-onset, difficult-to-manage cases (Veeramah et al., 2013). A retrospective cost-effectiveness analysis of trio exome sequencing method in comparison to a standard diagnostic strategy, for 32 patients diagnosed with epileptic encephalopathy, who had not received a diagnosis following initial "first-tier" testing, revealed that the trio exome sequencing approach was cost-effective (Palmer et al., 2018). Recent research employing trio-WES on a group of epilepsy patients in China has revealed that over 50% of individuals with unexplained epilepsy and associated seizures, along with developmental delays or intellectual disabilities, possess pathogenic or likely pathogenic variants (Chengyan et al., 2024; Jiang et al., 2021).

1.1.4 Functional analysis of disease-causing variants

Understanding the fundamental causes of epilepsy is essential for the effective clinical management of patients and for progressing neurobiological research that will guide future treatment possibilities. The identification of a genetic etiology in epilepsy plays

a crucial role in informing treatment strategies. A definitive diagnosis is essential for the patient to gain clarity regarding the underlying cause of the disorder, for the clinician to deliver the most effective treatment, and for predicting the progression of the disease. Precision medicine advocates for customized therapies based on specific genetic mutations. For example, the ketogenic diet is effective for GLUT1-deficiency syndrome, while sodium channel blockers are avoided in Dravet syndrome due to SCN1A mutations but used in SCN8A gain-of-function cases. mTOR pathway inhibitors are also used for epilepsy linked to tuberous sclerosis complex (Sisodiya, 2021; Striano and Minassian, 2020). Additionally, it is important for the clinical geneticist to offer appropriate genetic counseling to both the patient and their family. Consequently, conclusive evidence supporting the pathogenicity of genetic variants is therefore crucial.

The rise of newly identified missense variants complicates pathogenicity assessment, as not all mutations affect protein function, and those that do can have varying effects based on the disease context. Consequently, before conducting thorough analyses, it is challenging to categorize novel variants as either deleterious or neutral due to the uncertainty surrounding their functional and phenotypic effects, necessitating further research to confirm genetic diagnoses for new missense variants. *In vitro* characterization of new variants is ideal but often hindered by practical constraints. An alternative is using animal models with homologous mutations to confirm pathogenicity, but this method is also time-consuming, labor-intensive, and costly, making it impractical to assess all newly identified variants. In order to address the aforementioned limitations and to establish rapid and effective approaches (Al Mehdi

et al., 2019; Seifi and Walter, 2018) for forecasting the functional impact of nonsynonymous variants on protein stability, structure, function, and interacting pathways, a variety of computational tools have been developed.

The stability and structural integrity of proteins are crucial for their functionality and regulatory mechanisms. Many proteins need conformational changes to function properly, and disease-causing variants can disrupt protein folding and stability. Missense mutations may negatively affect protein structure by interfering with folding, interactions, solubility, or overall stability. The effects of these mutations can be studied using *in silico* methods, including 3D structural data, sequence alignments, and molecular dynamics simulations. Thus, initial *in silico* analysis of sequence data, followed by experimental validation of predicted deleterious mutations, is vital in modern genetic research.

The identification of epileptic encephalopathies, including West syndrome and Lennox–Gastaut syndrome, as being notably genetically diverse indicates that the underlying genetic anomalies may align with shared biological pathways. This revelation not only opens avenues for exploring treatments aimed at counteracting the effects of specific mutations but also motivates a deeper comprehension of the disrupted networks at genetic, cellular, and neurophysiological dimensions. Consequently, it prompts the development of both pharmacological and non-pharmacological intervention strategies.

1.1.5 Variants of uncertain significance

A large number of VUS have been found in epileptic encephalopathies, however classification of VUS remains a significant challenge (Burke et al., 2022). Although the use of multiple *in silico* variant predictors can help prioritise potentially deleterious variants, a thorough understanding of the disease's molecular mechanisms, as well as the structure, function, and biological network of the proteins involved in its pathogenesis, is still required. Protein structural analysis is a powerful method for identifying and prioritising genetic variations, and it should be utilized more extensively, especially when evaluating VUSs.

Genetic mutations can lead to the dysfunction of proteins by affecting their stability and modifying their interactions with other biological molecules. Deleterious genetic variations can alter the electrostatic characteristics of protein surfaces, leading to adverse effects including alterations in stability or folding, modifications in the selectivity and affinity for binding partners, and changes in the functional capabilities of the protein. Missense mutations can disrupt function through disrupting interactions with other biological components such as proteins, nucleic acids or polysaccharides, small molecule ligands, and metal ions, as well as by modifying protein stability. Investigating the ways in which mutations influence protein structure and stability represents a critical inquiry in the exploration of the relationship between mutation occurrence and disease manifestation.

Computational techniques can be used to anticipate how mutations affect protein function by affecting protein stability and interactions with other molecules such as

proteins, nucleic acids, and tiny molecules (Divi et al., 2023; Hollingsworth and Dror, 2018). Molecular dynamic simulations can disclose the positions of all atoms at femtosecond temporal precision, allowing them to capture a wide range of critical biomolecular phenomena such as conformational change, ligand binding, and protein folding. Importantly, such simulations can anticipate how biomolecules will respond to perturbations like mutation, phosphorylation, protonation, or the addition or removal of a ligand at the atomic level.

1.2 Hypothesis

We hypothesize that WES in trios gives a high yield of pathogenic variants associated with DEE, given the rarity of the disorder. The pathogenic variants and VUS with phenotype match identified potentially involve more than one functional pathway and the functional or structural analysis of these variants can be used to predict the variant pathogenicity.

1.3 Objectives

1. To identify the risk variants in childhood epilepsy and developmental epileptic encephalopathy using WES
2. To investigate the genetic networks and gene ontology pathways by correlating risk variants and their phenotypes
3. To predict the functional and structural analysis of Variants of uncertain significance with potential phenotype match

The thesis is presented in six chapters. **Chapter 1** begins with an introduction to the prevalence, types, and genetics of DEE. It also briefly introduces the importance of genetic diagnosis of DEE and *in silico* functional analysis of deleterious variants identified from NGS techniques. **Chapter 2** includes a detailed literature review, concentrating on the genetic background of DEE, disease mechanism in childhood epileptic encephalopathies, diagnostic genetic testing in epilepsy, the importance of WES in the genetic diagnosis, and functional enrichment analysis of deleterious variants.

Chapter 3 covers the materials and methods utilized to accomplish the objectives mentioned. The first section of chapter 3 includes the bioinformatics pipeline and various databases used for the WES analysis. The second section covers the tools and pipelines used for the functional analysis of pathogenic or likely pathogenic variants. The third section includes the workflow of Molecular dynamics simulation (MDS).

Chapter 4 involves the results and discussion of the studies and is divided into four major sections. The first section includes the diagnostic yield of DEE and a comparative analysis of yield based on age of onset of epilepsy, various syndromes, functional consequences, biological mechanisms, and inheritance patterns. The second section covers the *in silico* functional and structural analysis of pathogenic variants associated with DEE. The third section includes the MDS analysis and the fourth section covers the functional and structural prediction of VUS with potential phenotype match.

Chapter 5, which includes the Summary and Conclusion, summarizes the entire research conducted and highlights potential avenues for future investigation. Citation

of references, list of publications, and papers presented at conferences are also included.

2 LITERATURE REVIEW

2.1 Childhood epilepsy and Developmental epileptic encephalopathies

The International League Against Epilepsy defines epilepsy as a long-term predisposition to experience epileptic seizures, along with the neurobiological, cognitive, psychological, and social consequences that accompany this condition (Fisher et al., 2005). **Figure 2.1** illustrates the classification of the Epilepsies, a hierarchical framework intended to facilitate the categorization of epilepsy across various clinical settings (Scheffer et al., 2017). The classification framework for epilepsy begins with the categorization of seizure types. Seizures are divided into three main categories: focal onset, generalized onset, and unknown onset. The second level classifies Epilepsy Types, introducing a new category called "Combined Generalized and Focal Epilepsy," alongside the existing Generalized and Focal Epilepsies, as well as a category for Unknown. In diagnosing Generalized Epilepsy, it is common for the patient to exhibit generalized spike-wave patterns on an electroencephalogram (EEG). Focal Epilepsies encompass both unifocal and multifocal conditions, in addition to seizures that are localized to a single hemisphere. Combined Generalized and Focal

Epilepsies identifies patients with both seizure types, diagnosed mainly through clinical evaluation and supported by EEG results. The third level focuses on diagnosing an Epilepsy Syndrome, which is defined by a set of features that typically occur together. These features include particular types of seizures, specific patterns observed in electroencephalograms (EEGs), and imaging characteristics.

The etiology of seizures and epilepsy can be broadly classified into six categories- structural, genetic, infectious, metabolic, and immune, as well as an unknown group. A structural etiology involves identifiable abnormalities seen in neuroimaging, which, when combined with electroclinical evaluations, suggest that these abnormalities likely cause the patient's seizures. Structural etiologies can be classified as acquired (e.g., stroke, trauma, infections) or genetic (e.g., malformations of cortical development).

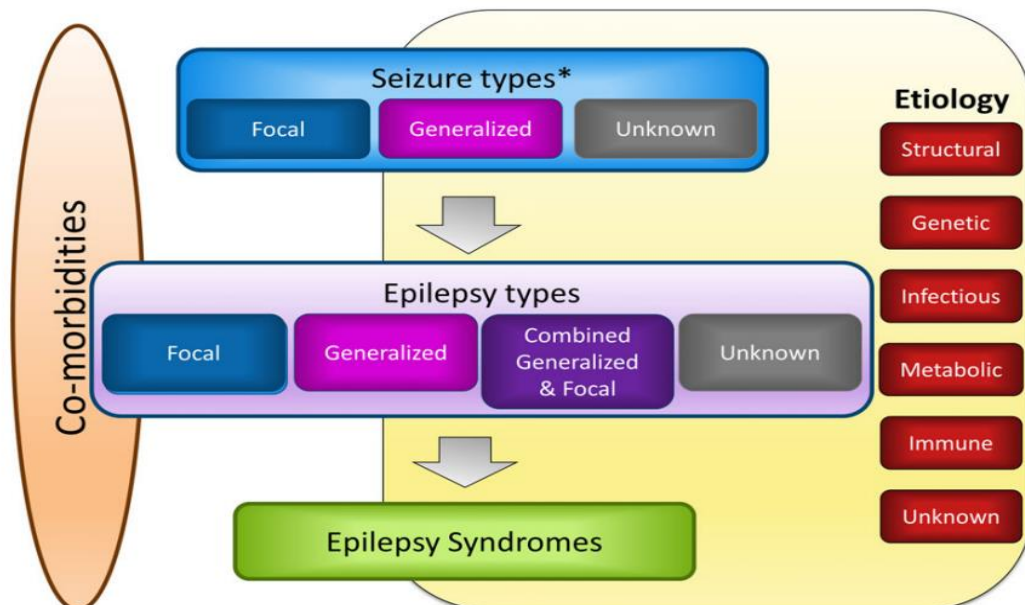


Figure 2.1: ILAE 2017 Classification of the Epilepsies (Adapted from Scheffer et al., 2017)

The concept of genetic epilepsy pertains to disorders that stem directly from a known or suspected genetic mutation, with seizures being a key feature of the condition. There is a wide variety of epilepsies linked to genetic factors, and in many cases, the specific genes responsible have yet to be identified. The concept of an infectious etiology suggests that seizures are a primary symptom of a condition arising from a known infection. This term specifically applies to individuals diagnosed with epilepsy, setting them apart from those who have seizures due to acute infections such as meningitis or encephalitis. Metabolic epilepsy refers to seizures that occur as a primary symptom of a recognized or suspected metabolic disorder, characterized by specific metabolic defects causing widespread biochemical changes. Examples include porphyria, uremia, aminoacidopathies, and pyridoxine-dependent seizures. Most metabolic epilepsies likely have a genetic origin, though some, like cerebral folate deficiency, may be acquired. Immune epilepsy is believed to stem from immune dysfunction, with seizures as a key symptom. This condition is characterized by evidence of autoimmune-induced inflammation in the central nervous system. The term "unknown" indicates that the etiology of the epilepsy has not been identified. A significant number of individuals with epilepsy continue to have an undetermined cause. Within this group, it is not feasible to establish a precise diagnosis beyond the fundamental electroclinical semiology.

In 2022, the International League Against Epilepsy (ILAE) introduced a classification system for epilepsy syndromes that manifest in neonates and infants (Zuberi et al., 2022). These syndromes are primarily categorized into two groups: Self-Limited Epilepsies, which are characterized by a high likelihood of spontaneous remission, and Developmental and Epileptic Encephalopathies (**Figure 2.2**). Syndromes that are

specific to certain etiologies arise from distinct genetic, structural, metabolic, immune, or infectious causes, and they exhibit consistent electroclinical characteristics, treatment approaches, and prognostic outcomes. The majority of etiology-specific syndromes that emerge during the neonatal or infantile stages are classified as Developmental and Epileptic Encephalopathies.

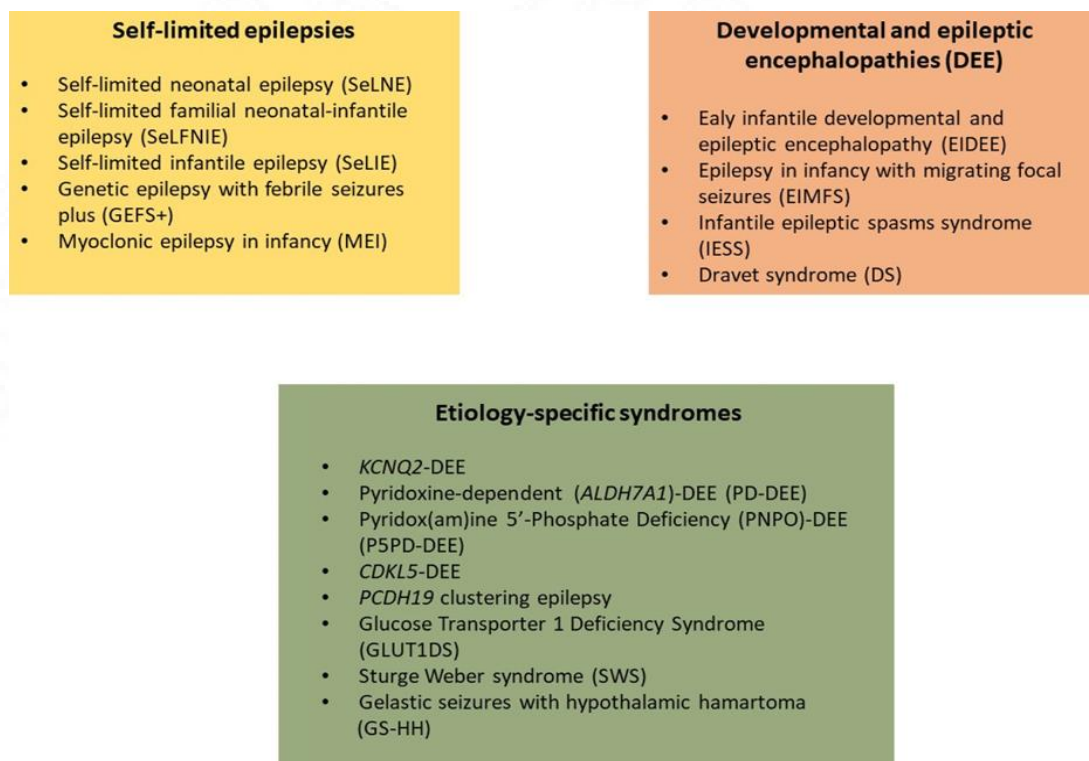


Figure 2.2: Classification of neonatal and infant epilepsy syndromes (Adapted from Zuberi et al., 2022)

Developmental epileptic encephalopathy (DEE)

Epileptic encephalopathy is characterized by the occurrence of frequent epileptiform discharges that adversely affect developmental progress, often leading to a slowdown or regression in developmental skills, and is usually associated with recurrent seizures. Many of these conditions have been associated with specific molecular genetic factors.

The use of the term 'developmental' highlights the multiple factors that contribute to the developmental deficits experienced by those affected. The genetic abnormalities can cause developmental delays on their own, while the epileptic encephalopathy further intensifies these developmental difficulties (Scheffer and Liao, 2020).

In the 2010 proposal for classification by the International League Against Epilepsy (ILAE), the term 'epileptic encephalopathy' was revised to 'developmental and epileptic encephalopathy' in the official 2017 update of the classification system (Scheffer et al., 2017). Developmental encephalopathies, which encompass conditions such as intellectual disability and autism spectrum disorder, are often linked to infantile epilepsy. The term "epileptic encephalopathy" is employed to denote a presumed causal connection between epilepsy and developmental delays. However, the pathogenesis of developmental encephalopathies appears to be largely independent of epilepsy, as evidenced by the discovery of various gene variants that are associated with both developmental encephalopathies and epilepsy. Additionally, there is a possibility of gene-related developmental encephalopathies occurring in the absence of epilepsy, and developmental encephalopathies may continue to progress even when seizures are effectively managed. Therefore, the designation "developmental and epileptic encephalopathy" may be a more fitting term than "epileptic encephalopathy."

The group of severe epilepsies that manifest during infancy and childhood encompasses various debilitating disorders characterized by recurrent epileptic seizures, often associated with developmental delays or regressions. This classification includes a wide range of conditions known as epileptic encephalopathies, where affected infants or children usually experience multiple types of seizures and exhibit

significant epileptiform activity on electroencephalograms (EEGs). These disorders are often accompanied by a slowdown or regression in developmental milestones, which may arise after the onset or exacerbation of seizures.

DEEs encompass a wide range of conditions characterized by the early onset of often severe seizures and irregularities in electroencephalogram (EEG) readings, all occurring alongside developmental deficits that usually worsen as a result of epilepsy. The most epileptic encephalopathies of childhood include:

Early myoclonic encephalopathy

This syndrome manifests during the neonatal stage with irregular myoclonus that transitions between different regions of the body. The myoclonus may intensify over time, and the majority of infants also exhibit mild partial seizures characterized by eye deviation and autonomic symptoms (Wirrell et al., 2005). The pathogenesis of early myoclonic encephalopathy exhibits considerable variability, with contributions from structural, metabolic, and genetic factors. The overarching clinical manifestation in early myoclonic encephalopathy appears to reflect a widespread process, particularly affecting the brainstem and white matter, which may result in deafferentation and increased cortical excitability (Beal et al., 2012).

Early Infantile Epileptic Encephalopathy (EIEE) or Ohtahara Syndrome

Ohtahara syndrome is a severe form of epilepsy that can manifest from the intrauterine stage up to three months of age. Infants typically experience sudden onset of tonic spasms, which may be generalized or lateralized, occurring either as isolated events or in clusters, and these spasms are not influenced by the sleep-wake cycle (Beal et al.,

2012; Jain et al., 2013). Many infants have a malformation in cortical development, leading to high mortality rates. Survivors often develop West syndrome or refractory partial seizures, which are usually hard to treat. However, surgical intervention may benefit infants with cortical dysplasia (Wirrell et al., 2005).

West Syndrome

West syndrome typically manifests during infancy, with a typical onset between 3 and 12 months of age. It is characterized by a triad of symptoms, which includes infantile spasms, mental retardation, and hypsarrhythmia (Jain et al., 2013; Sands and Gelinas, 2024). Infants with epileptic spasms often show developmental regression and an EEG with high-amplitude disorganization and multifocal epileptic activity. The causes of West syndrome are diverse, but for some patients, the etiology cannot be determined.

Dravet syndrome

Dravet Syndrome or Severe Myoclonic Epilepsy in Infancy, typically manifests within the first year of life. It is characterized by recurrent focal seizures or those that progress to secondary generalization, frequently precipitated by mild hyperthermia. Affected infants often experience prolonged febrile unilateral clonic convulsions that exhibit an alternating pattern, emerging in a child who was previously developing normally. Nonfebrile seizures may also occur. This phase is followed by various seizure types, such as myoclonic seizures, atypical absences, and complex focal seizures, which can progress to status epilepticus and cause significant psychomotor decline (Jain et al., 2013). Individuals often exhibit developmental delays accompanied by significant

deficits in neurocognitive functioning, encompassing impairments in motor skills, language abilities, and intellectual capacity (Wheless et al., 2020).

Lennox-Gastaut Syndrome (LGS)

Lennox-Gastaut syndrome (LGS) is a severe epileptic encephalopathy that usually appears between ages 1 and 8, mainly between 2 and 5. It is characterized by persistent polymorphic seizures, including tonic, atypical absence, atonic, and myoclonic seizures. About 50% of children experience "drop attacks," which can lead to injuries, and around two-thirds may suffer from nonconvulsive status epilepticus (Camfield, 2011; Hancock and Cross, 2013). Language abilities are often impaired, causing delays in thought and speech, along with motor function challenges. Severe behavioral issues like hyperactivity, aggression, and autism traits are common, and psychotic symptoms may develop over time. The long-term outlook is poor; while epilepsy may improve, complete seizure freedom is rare, and mental health conditions typically worsen.

Epilepsy of infancy with migrating focal seizures (EIMFS)

EIMFS is an age-related epileptic encephalopathy that exhibits a severe progression, typically beginning within the first six months of life. It is marked by an initial phase of normal early development, which is then succeeded by nearly constant migrating polymorphous focal seizures resistant to treatment. This is often accompanied by psychomotor retardation and, in the majority of affected children, a progressive decrease in head circumference percentile. Seizures exhibit significant resistance to drugs, frequently resulting in severe consequences (Coppola, 2009). The interictal EEG demonstrates the presence of multi-focal spikes and a slowing pattern, while the

ictal EEG indicates the emergence of shifting (migrating) foci at the onset of seizures (McTague et al., 2013).

EE with Continuous spike and wave during sleep (CSWS) including Landau Kleffner syndrome (LKS)

Landau-Kleffner Syndrome (LKS) usually begins in children aged 5 to 7, presenting as verbal auditory agnosia in otherwise healthy kids. Language skills decline progressively, with possible gradual or fluctuating symptoms. Various aphasia forms and mutism may occur, along with mild behavioral issues. About 75% of children with LKS experience infrequent seizures, mainly at night (Jain et al., 2013).

EE with CSWS typically emerges between 2 months and 12 years, peaking at ages 4 to 7. About 50% of cases have a normal neurodevelopmental history before onset. Seizures are the initial symptom in 80% of affected children, while others show a neuropsychological decline. Seizures are usually infrequent, occurring at night, and can be simple or complex focal, generalized tonic-clonic, or myoclonic. Interictal EEG during wakefulness shows focal or multifocal epileptiform discharges, which intensify during NREM sleep (Loddenkemper et al., 2011).

Epilepsy with myoclonic-atonic seizures (EMAS)

EMAS presents with a spectrum of seizure types, including myoclonic-atonic, absence, tonic-clonic, and, in some cases, tonic seizures, appearing in previously normal children between the ages of 18 and 60 months. The incidence of EMAS is approximately 1 in 10,000 children, representing about 1% to 2% of epilepsies that begin in childhood. It is observed more frequently in males, with 94% of cases

initiating within the first five years of life, predominantly between the ages of 3 and 4. Notably, around 25% of affected individuals experience their initial seizure during the first year of life (Caraballo et al., 2013).

2.2 Genetics of epilepsy and DEE

Genetic factors are anticipated to play a major role in different types of epilepsy, including classic idiopathic (genetic) generalized epilepsies, epileptic encephalopathies, and focal epilepsies. Advancements in NGS have enhanced our understanding of the complex neurobiological mechanisms of epilepsy. These insights contributed to the 2017 classification framework by the International League Against Epilepsy, highlighting the significance of genetic aetiologies in different epilepsy forms.

In 2018, the ILAE Consortium on Complex Epilepsies conducted the most extensive genome-wide association study (GWAS) in the field of epilepsy, encompassing 15,212 cases and 29,677 controls. This study identified 16 loci that reached genome-wide significance for epilepsy, with 11 of these being previously unreported (The International League Against Epilepsy Consortium on Complex Epilepsies et al., 2018). Currently, the Online Mendelian Inheritance in Man database contains over 200 genes or loci that are associated with epilepsy. The genetic determinants associated with DEE are outlined in **Table 2.1**.

Epilepsies with suspected complex inheritance pose challenges in understanding their genetic mechanisms, as multiple gene variants are thought to collectively increase disease risk. Various models of genetic risk exist, which may overlap, and factors like

epigenetics and environmental influences may also contribute. The International League Against Epilepsy (ILAE) categorizes genetic epilepsies into two types: those

Table 2.1: Genes associated with complex childhood epilepsy and DEE (Jose et al., 2022; McTague et al., 2016; Viswanathan et al., 2024)

Type of DEE	Genes
Early myoclonic encephalopathy	<i>PIGA, SETBP1, ERBB4, SIK1, SLC25A22</i>
EIEE or Ohtahara syndrome	<i>KCNQ2, KCNT1, PIGO, STXBP1, SCN2A, ARX, AARS, BRAT1, CACNA2D2, GNAO1, NECAP1, PIGA, SCN8A, SIK1, SLC25A22</i>
West syndrome	<i>CDKL5, STXBP1, ARX, ALG13, DOCK7, DNMI, GABRB1, GABRB3, GNAO1, GRIN1, GRIN2A, GRIN2B, MAGI2, MEF2C, NEDDL4, NDP, NRXN1, PIGA, PLCB1, PTEN, SCA2, SCN1A, SETBP1, SIK1, SLC25A22, SLC35A2, SPTAN1, ST3Gal3, TBC1D24, TCF4, WWOX</i>
Dravet syndrome	<i>SCN1A, PCDH19, GABRA1, GABRG2, HCN1, STXBP1</i>
Lennox-Gastout syndrome	<i>ALG13, CACNA1A, CDKL5, CHD2, DNM1, FLNA, GABRB3, GRIN2B, HNRPN, HNRNPH1, IQSEC1, IQSEC2, KCNQ3, MTOR, SCN1A, SCN2A, SCN8A, STXBP1</i>
EE with CSWS including Landau-Kleffner syndrome	<i>GRIN2A, CNKSR2, ATP1A2, CACNA1A, FOXP1, GRIN1, KCNMA1, KCNQ3, PPFIA3, PUF60, SETD1B, ZBTB18</i>
Epilepsy with myoclonic atonic seizures	<i>SLC2A1, SLC6A1, CHD2, GABRA1, GABRG2, SCN1A, SCN1B, KCNA2</i>
EIMFS	<i>KCNT1, SCN2A, SCN1A, PLCB1, QARS, SCN8A, SLC25A22, TBC1D24, SLC12A5</i>
Progressive myoclonic epilepsy	<i>CLN3, CLN5, TPP1, CLN2, CLN8, MFSD8, EPM2A, GBA, POLR3B, KCTD7,</i>
Myoclonic/generalized epilepsy	<i>EEF1A2, MEF2C, SCN1A, SLC2A1, SPTAN1, SYNGAP1, TBC1D24, CHD2, MEF2C, SYNGAP1, UBE3A</i>
Other focal or multifocal epilepsy	<i>ARHGEF9, DEPDC5, SCN1A, TBC1D24, PNKP, SLC2A1, FOXG1, MBD5, PIGO, SLC13A5, PCDH19, POLG, TNK2, ZEB2</i>

characterized by monogenic inheritance, which may involve either inherited or *de novo* pathogenic variants and those with complex inheritance, which can be polygenic and may or may not involve environmental influences (Scheffer et al., 2017).

In contrast to the phenotypically diverse forms of epilepsy, some types display unique and characteristic clinical features, as well as specific findings on diagnostic imaging. Monogenic epilepsies represent a small portion of genetic epilepsies but are often associated with earlier onset and more severe symptoms. Some studies suggest that a monogenic cause may be identified in up to 40% of individuals with severe epilepsy (Rastin et al., 2023). Oligogenic models represent genetic frameworks in which more than one allele is necessary to affect genetic susceptibility. These models encompass modifier genes and the concept of epistasis, wherein neither of the two variants alone impacts genetic risk; instead, the combination of both variants is essential for altering the level of risk. In contrast, the polygenic model suggests that many variants, possibly numbering in the hundreds or thousands, each play a relatively minor role in increasing the overall risk of disease. The omnigenic model extends the polygenic model, proposing that complex interactions within gene regulatory networks can explain why genes without a direct link to disease may still show significant associations, as they influence essential disease-related genes (Ellis et al., 2020; Perucca et al., 2020).

The first identification of a genetic link to epileptic encephalopathy occurred in 2001, when researchers found that all seven children in a study on Dravet syndrome had a *de novo* mutation in the *SCN1A* gene (Claes et al., 2001). The advent of advanced molecular techniques, such as chromosomal microarray analysis and next-generation

sequencing of multiple genes, has greatly enhanced the discovery of genes related to epileptic encephalopathies.

Mode of inheritance

Genetic mutations associated with epileptic encephalopathies frequently occur *de novo* and exhibit a pattern of dominant inheritance; nonetheless, certain patients may present with autosomal recessive, X-linked, or mitochondrial inheritance patterns.

Autosomal recessive:

Autosomal recessive (AR) inheritance in DEEs was initially seen as rare in outbred populations, with early gene identifications like SLC25A22, PLCB1, and PNPO primarily found in consanguineous populations (Happ and Carvill, 2020). A study examining diagnostic exome sequencing in 89 unselected patients diagnosed with "epileptic encephalopathy" found that approximately 43% of the cases resulted in a definitive diagnosis, with around 10% attributed to recessive inheritance (Helbig et al., 2016). In a comprehensive analysis that integrated whole exome sequencing and copy number variation assessment, it was determined that autosomal or X-linked recessive disorders were present in 38% of the cases diagnosed with early onset epileptic encephalopathy (Papuc et al., 2019). A recent study revealed that homozygous recessive variants in several genes, including *EPM2A*, *NHLRC1*, *CLN1*, *CLN2*, *TPP1*, *CLN3*, *CLN5*, *CLN7*, *MFSD8*, *ACOX1*, *GBA*, *POLR3B*, *KCTD7*, *L2HGDH*, *SLC35A2*, *ITPA* and *ARG1*, are associated with metabolic epilepsies (Jose et al., 2024).

Compound heterozygous:

Compound heterozygous (CH) variants occur when an individual inherits two distinct recessive alleles from different loci within the same gene. One variant is inherited from each heterozygous parent, contributing to various recessively inherited disorders. Several autosomal recessive genes showing a compound heterozygous inheritance pattern have been identified in cases of epileptic encephalopathies. Novel compound heterozygous variants reported in genes *WWOX*, *UBA5*, *GOT2*, *SPATA5*, *SLC13A5*, *etc.*, associated with developmental epileptic encephalopathies. Recently, Papuc et al. demonstrated the significant impact of primarily compound heterozygous or hemizygous recessive inheritance patterns, which contribute to a heightened recurrence risk in developmental epileptic encephalopathies (Papuc et al., 2019). A recent study on neonatal developmental and epileptic encephalopathy linked to autosomal recessive variants in the *SLC13A5* gene found that over 50% of participants had compound heterozygous variants (Matricardi et al., 2020).

Autosomal dominant:

Autosomal dominant mutations are inherited from affected parents and make up about 10% of DEE cases (Esmaeeli Nieh and Sherr, 2014) and are mainly missense mutations. Furthermore, the syndrome can also be inherited through somatic or germline mosaic mutations found in unaffected parents, resulting in familial clustering among their offspring. A study on Dravet syndrome found an inherited *SCN1A* mutation in the parental blood DNA of 12 cases, with mosaicism levels between 0.04% and 85.00%. The results showed a direct correlation between mosaicism degree and

the severity of the parent's condition, indicating a significant link between mutation burden and affected status (Depienne et al., 2010).

De novo:

Recent research has highlighted *de novo* variants as a significant contributing factor to epilepsy, indicating a generally low risk of recurrence. 30% to 50% of DEEs are associated with a pathogenic variant, with the majority of these variants arising *de novo* (Epi4K Consortium and Epilepsy Phenome/Genome Project, 2013) in either autosomal dominant or X-chromosome genes. A study of WGS in 197 individuals with DEE and their unaffected parents identified 53 *de novo* SNVs and 4 *de novo* CNVs, with *de novo* missense variants being particularly prevalent among DEE cases (Hamdan et al., 2017a).

X-linked:

The inheritance of X-linked mutations can follow either a dominant or recessive pattern. X-linked dominant inheritance is marked by vertical transmission, meaning that all daughters of affected males will inherit the condition. On the other hand, X-linked recessive inheritance also involves vertical transmission, but here, carrier females pass the trait on to their affected sons. Pathogenic variants in X-linked genes have been discovered in both males and females diagnosed with DEE, potentially attributed to an absence of skewed X-chromosome inactivation. Pathogenic variants in the X-linked genes *CDKL5*, *ARX*, *MECP2*, and *PCDH19* are the well-described causes of DEE (Happ and Carvill, 2020). Certain genes like *GABRA3*, *CASK*, *IQSEC2*, and *NEXMIF* show that hemizygous males often have more pronounced phenotypes than females. This is likely due to females having an active X chromosome with a

normal allele, while X-chromosome inactivation (XCI) silences one allele, balancing gene dosage. In many X-linked disorders, XCI skews towards the mutant allele, allowing females to express the non-pathogenic variant and remain asymptomatic.

Mitochondrial:

Mitochondrial biogenesis and function are regulated by both mitochondrial DNA (mtDNA) and nuclear DNA, indicating dual genetic control (Ng et al., 2021). Mutations in either DNA can lead to mitochondrial disorders, which are a significant category of inborn errors of metabolism. Whole-exome sequencing has revealed a variant in the mitochondrial *MT-ND1* gene in a patient diagnosed with West Syndrome, which subsequently progressed to Lennox–Gastaut Syndrome (Delmiro et al., 2013). In a recent study, a novel variant in the mitochondrial *MT-TF* gene was identified in a 14-year-old female patient diagnosed with epileptic encephalopathy (Itkis et al., 2019).

2.3 Disease mechanisms in childhood epilepsy and DEE

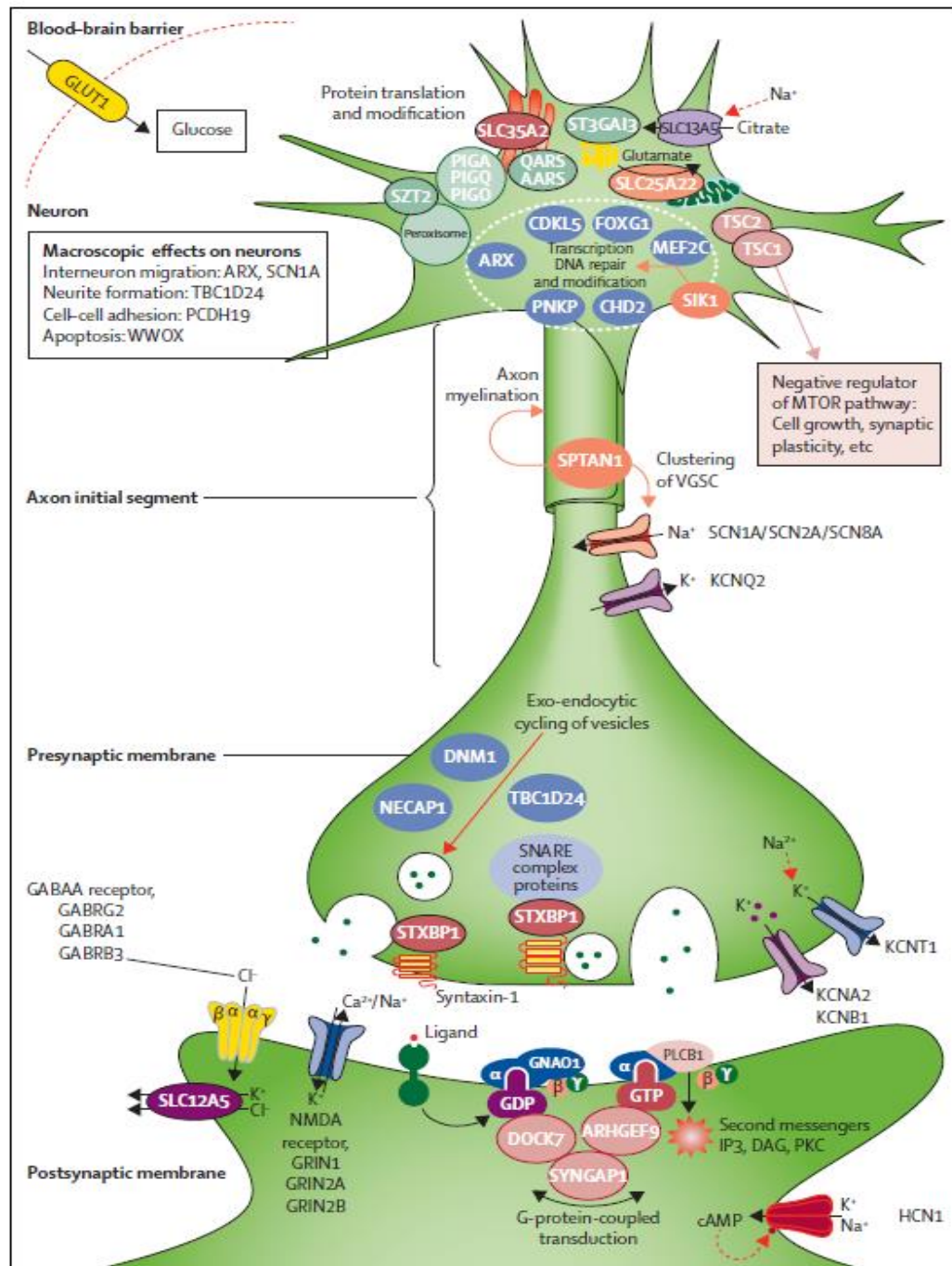


Figure 2.3: Pathways implicated in Childhood epilepsy and DEE (Adapted from McTague et al., 2016)

Epileptogenesis refers to the long-term process through which the functional architecture of a brain network is modified, leading to an increased susceptibility to seizures. This alteration results in a heightened likelihood of spontaneous and recurrent seizure activity. In genetically determined DEEs, the underlying mechanisms of epileptogenesis are frequently intertwined with the modified processes of brain development and neural connectivity (Guerrini et al., 2023).

Genetic influences can impact brain function across various levels, including ion channels, receptor activity, and synaptic connections. The identification of genes has revealed a wide array of disease mechanisms associated with severe epilepsies. These mechanisms encompass channelopathies, synaptic dysfunction, defects in transporters, transcriptional dysregulation, compromised DNA repair processes, chromatin remodeling errors, and metabolic abnormalities (**Figure 2.3**) Genes linked to DEE can be categorized into functional groups involved in various cellular processes including ion or transmitter or molecule transport, synaptic regulation, cell growth, division and proliferation, metabolism, intracellular signaling and trafficking, gene transcription, and protein synthesis and degradation.

2.3.1 Voltage-gated ion channels

Ion channels are integral membrane proteins that span the lipid bilayer of cellular membranes, creating a pore that allows for the movement of ions in and out of the cell by their electrochemical gradients. These channels demonstrate selective permeability to particular ions, with their activity regulated by distinct states: open, inactivated closed, and resting closed. Gating regulation occurs through various mechanisms, including transmembrane electric potential for voltage-gated channels and specific

extracellular ligands for ligand-gated channels. Other factors like intracellular second messengers, light, temperature, and mechanical stimuli may also influence this process. Most ion channels and pumps are heteromeric, featuring a pore-forming or catalytic subunit alongside accessory and regulatory subunits. Genetic mutations in ion channel genes are linked to various epilepsy syndromes. Genetic mutations affecting ion channel genes in the brain are implicated in a variety of rare neurodevelopmental disorders (NDDs). These include severe forms of epileptic encephalopathy observed in infancy and childhood, associated with genes such as *SCN1A*, *SCN2A*, *SCN8A*, *KCNQ2/KCNQ3*, *KCNA1*, *KCNA2*, *KCNT1*, *KCNB1*, and *CACNA1A*. Additionally, these mutations are linked to autism spectrum disorders (ASD) and psychomotor delays, with specific associations to *SCN1A*, *SCN2A*, *KCNJ10*, *KCNQ3*, and *CACNA1C*. Furthermore, certain ataxias accompanied by seizures are related to *KCNA1*, *KCND3*, and *CACNA1A*, while cerebellar atrophy and seizures have been connected to *KCNMA1*.

Voltage-gated sodium channel

Voltage-gated sodium channels (NaVs) are complex structures composed of a prominent alpha subunit that forms a central pore, accompanied by two smaller auxiliary beta subunits. Each alpha subunit has four homologous domains (I–IV), each with six transmembrane segments (S1–S6). Segments S5 and S6 form the pore, while S4 acts as a voltage sensor. NaVs are primarily found in the central and peripheral nervous systems, as well as in skeletal and cardiac muscles, playing a vital role in initiating and transmitting action potentials (Huang et al., 2017). Depolarization of the cell membrane causes a conformational change in the channel, allowing sodium ions

to enter and trigger action potentials in excitable cells (Bartolini et al., 2020). There are nine alpha subtypes of NaVs. Mutations in NaV1.1 (SCN1A), NaV1.2 (SCN2A), NaV1.3 (SCN3A), NaV1.6 (SCN8A), and NaV1.7 (SCN9A) are linked to epilepsy, while NaV1.4 (SCN4A), NaV1.5 (SCN5A), NaV1.8 (SCN10A), and NaV1.9 (SCN11A) show no such association (Menezes et al., 2020). Additionally, both alpha and beta subunits (SCN1B) are also implicated in epilepsy phenotypes.

The *SCN1A* gene encodes the alpha subunit of the voltage-gated sodium ion channel (**Figure 2.4**) and is linked to various epilepsy phenotypes, such as the Dravet syndrome, genetic epilepsy with febrile seizures plus (GEFS+), familial febrile seizures, focal epilepsies, and early infantile epileptic encephalopathy (Sadleir et al., 2017). A recent study in South India identified 29 *SCN1A* gene variants associated with Dravet syndrome, highlighted that children with truncating variants were more likely to have multiple seizure types, while those with missense variants had a greater chance of achieving seizure freedom (Krishna et al., 2024). view of the unfolded pore-forming protein shows domains DI-DIV and transmembrane segments S1-S6. Three dimensional representation of the NaV channel illustrates the positioning of the P-loop (ion permeation pathway) and Site 1 (binding site) within the cellular membrane (Cembella and Durán-Riveroll, 2021).

The *SCN2A* gene encodes the NaV1.2 sodium channel, which is predominantly expressed in the central nervous system, particularly within excitatory and glutamatergic neurons. To date, over 100 mutations have been identified in this gene, with approximately 300 patients having been analyzed (Reynolds et al., 2020). The primary disorders associated with mutations in *SCN2A* include West syndrome and

benign familial neonatal-infantile seizures. While mutations linked to epilepsy are distributed throughout the channel, certain regions, such as the ion selectivity filter, the voltage-sensing domain, the intracellular N-terminal, and the C-terminal domain, are notable hotspots. Given that NaV1.2 channels are found in excitatory neurons, gain-of-function mutations are implicated in epilepsy due to their role in promoting neuronal hyperexcitability.

The *SCN3A* gene is highly expressed in the brains of both neonates and adults, with similar levels in excitatory neurons and inhibitory interneurons. Point mutations in *SCN3A* have been found in individuals with focal epilepsy. Several *de novo* mutations in the *SCN3A* gene have been linked to severe neurological dysfunctions and cognitive deficits in infants. These mutations disrupt NaV1.3 channel function, impacting neuronal organization, migration, and proliferation during embryonic development. Key clinical manifestations include epileptic encephalopathy and polymicrogyria, the latter being a unique feature of *SCN3A* mutations not seen in other channelopathies (Inuzuka et al., 2020). Microduplication mutations linked to *SCN3A* have been associated with epilepsy. A study by Yoshitomi et al. identified three patients with focal seizures and infantile spasms who had a microduplication at the 2q24.3 chromosomal region (Yoshitomi et al., 2015).

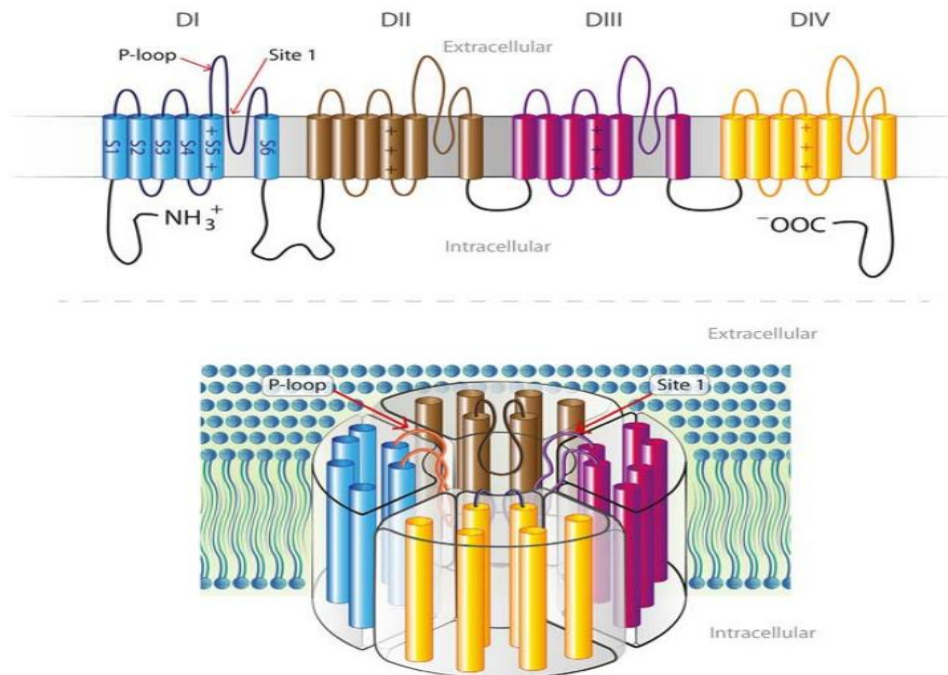


Figure 2.4: Representation of alpha subunit of voltage-gated sodium ion channel:

(Adapted from <https://sciencedirect.com/science/article/abs/pii/S187791821400104X>)

The *SCN8A* gene is responsible for encoding the alpha subunit of the type 8 voltage-gated sodium channel, known as NaV1.6, which is expressed from the prenatal stage throughout fetal development. This gene plays a crucial role in regulating neuronal excitability in excitable cells, including Purkinje cells located in the cerebellum, ganglion cells found in the retina, and hippocampal cells. Variations in the *SCN8A* gene are linked to early-infantile epileptic encephalopathy type 13, which is characterized by a diverse range of phenotypes in early-onset epilepsy, with seizures typically commencing prior to 18 months of age (Menezes et al., 2020). Affected individuals often experience intellectual disabilities, developmental delays, and movement disorders. Numerous exome sequencing studies have identified 11 *de novo* *SCN8A* mutations in epilepsy patients, with or without intellectual disability

(Ademuwagun et al., 2021). Additionally, another *de novo* mutation was found in a child with epileptic encephalopathy and congenital anomalies.

Voltage-gated Potassium channel

Voltage-gated potassium channels (Kv) play a crucial role in controlling and adjusting the resting membrane potential, as well as determining the threshold and duration of action potentials in excitable cells. Studies have shown that many K⁺ channel genes significantly influence the occurrence of DEE. Mutations in neuronal KCNQ channels are associated with hyperexcitability disorders, including neuropathic pain, benign familial neonatal seizures, and neonatal epileptic encephalopathy. Recent studies have identified variants in genes that encode voltage-gated potassium (Kv) channels, such as *KCNQ2*, *KCNQ3*, and *KCNB1*, as being linked to EIEE (Thiffault et al., 2015).

The Kv7 family of voltage-gated potassium channels, which includes five distinct members ranging from Kv7.1 to Kv7.5, is encoded by the *KCNQ* gene family (*KCNQ1–5*) (Nappi et al., 2020). *KCNQ2* gene variations are linked to neonatal-onset epileptic disorders with diverse phenotypes, ranging from benign familial neonatal seizures (BFNS) to severe neonatal-onset DEE. They are also associated with non-epileptic neurodevelopmental disorders on the autism spectrum. Mutations in the *KCNQ2* gene have also been found in neonates with pharmaco-resistant seizures, psychomotor retardation, a suppression-burst pattern on EEG, and distinct neuroradiological features, leading to the condition known as "KV 7.2 encephalopathy." These mutations are also seen in children with Ohtahara syndrome and early infantile epileptic encephalopathy, both severe forms of early-onset epileptic encephalopathy (Saitsu et al., 2012).

Variants in the *KCNQ3* gene are a rare cause of BFNS, with clinical features similar to *KCNQ2*-related BFNS. Both BFNS and BFIS are typically benign and self-resolving, but some individuals with *KCNQ3* mutations have shown varying degrees of intellectual disability (ID) alongside seizures, suggesting a potential link to more severe cases. Additionally, *de novo* pathogenic variants have been found in a few children with DEE, ID without epilepsy, cortical visual impairment, and ID with autism spectrum disorder (Nappi et al., 2020).

KCNQ5 is expressed in the brain, skeletal muscle and blood vessels. Recently, there have been reports of *de novo* heterozygous missense and truncating variants in the *KCNQ5* gene among individuals diagnosed with intellectual disability (ID), either as isolated cases or in conjunction with DEE. Kruger et al. recently found loss-of-function variants in *KCNQ5* linked to genetic generalized epilepsy (GGE) and mild to moderate intellectual disability (ID). This loss in KV7.5 channels may reduce the M-current, increasing neuronal excitability (Krüger et al., 2022).

Genetic variations in the *KCNT1* gene, which encodes a sodium-gated potassium channel, are associated with various epileptic conditions and neurodevelopmental disorders, including familial autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE), EIMFS, and developmental and epileptic encephalopathies (Bonardi et al., 2021).

KCNB1 encodes the Kv2.1 α subunit, a key component of a delayed rectifier potassium channel in neurons, essential for generating the primary delayed rectifier current and regulating neuronal excitability. Kv2.1 facilitates action potential repolarization and modulates neuronal activity. Mutations in *KCNB1* can cause EIEE, with *de novo*

missense variants linked to cognitive and motor impairments. These mutations disrupt ion selectivity and voltage-sensitive gating, increasing tonic inward cation conductance (Thiffault et al., 2015).

2.3.2 Ligand-gated ion channels

Ligand-gated ion channels, such as glycine receptors, γ -aminobutyric acid type A receptors, and nicotinic acetylcholine receptors, have been recognized for an extended period as possessing genetic variants linked to hyperekplexia and various types of epilepsy. In certain instances, missense variants either augment or diminish the receptor's inherent capacity to translate ligand binding into channel opening, as well as the effectiveness of receptor activation.

Genes associated with familial human epilepsy that encode ion channel subunits encompass the nicotinic acetylcholine receptor (nAChR) subunits *CHRNA2*, *CHRNA4*, and *CHRNA2*; the chloride channel subunit *CLCN2*; the GABAA receptor subunits *GABRA1*, *GABRA6*, *GABRB1*, *GABRB2*, *GABRB3*, *GABRG2*, and *GABRD*; potassium channel subunits *KCNA1*, *KCNQ2*, *KCNQ3*, *KCNJ11*, *KCNJ6*, and *KCNMA1*; sodium channel subunits *SCN1A*, *SCN1B*, and *SCN2A2*; as well as calcium channel subunits *CACNA1A*, *CACNA1H*, and *CACNB4*. Missense and nonsense mutations have been identified in the genes *GABRA1*, *GABRA2*, *GABRA5*, *GABRA6*, *GABRB1*, *GABRB2*, *GABRB3*, *GABRD*, and *GABRG2*, associated with monogenic autosomal dominant familial epilepsy and *de novo* epilepsy syndromes (Hernandez and Macdonald, 2019). Over 60 novel mutations in the *GABRA1*, *GABRA2*, *GABRA5*, *GABRB1*, *GABRB2*, *GABRB3*, and *GABRG2* genes, which encode GABA receptor subunits, are linked to various epileptic encephalopathies

(EEs) such as Ohtahara syndrome, infantile spasms, myoclonic astatic epilepsy, Lennox-Gastaut syndrome, early onset epileptic encephalopathy, and Dravet syndrome.

2.3.3 mTOR pathway

The mTOR signaling pathway is a key regulator of cellular metabolism, growth, proliferation, and survival, primarily through mTOR complex 1, which is inhibited by the antibiotic rapamycin. Mutations in genes that regulate mTOR complex 1 are associated with epilepsies and neurodevelopmental disorders. Germline and somatic mutations in genes of the mTOR signaling pathway contribute to epilepsies, malformations of cortical development (MCD), and neurodevelopmental disorders. Loss-of-function mutations in inhibitors like *TSC1*, *TSC2*, and *PTEN*, along with gain-of-function mutations in activators such as *PI3KCA* and *AKT3*, lead to mTORC1 hyperactivation. mTORopathies include a range of MCD, from severe brain abnormalities like megalencephaly to localized issues like focal cortical dysplasia (FCD) and cases with normal MRI. GATOR1 mutations are linked to non-lesional focal epilepsies, where subtle FCD may be undetectable on standard imaging. Additionally, mTOR-related MCD often coexists with intellectual disabilities and autism spectrum disorders (ASD) (Moloney et al., 2021).

Germline mutations in GATOR1 complex genes (*DEPDC5*, *NPRL2*, *NPRL3*), which repress mTORC1, are linked to various focal epilepsy syndromes, often with focal cortical dysplasia. Somatic mutations in the mechanistic/mammalian target of rapamycin (mTOR) signaling pathway, specifically in the genes *AKT1*, *AKT3*, *DEPDC5*, *MTOR*, *NPRL2*, *NPRL3*, *PIK3CA*, *PIK3R2*, *TSC1*, and *TSC2*, are now

recognized as the primary known contributors to focal cortical dysplasia type II (FCDII) and hemi-megalencephaly (HME) (Krochmalnek et al., 2023).

The pathways by which mTORopathies cause neuronal hyperexcitability and seizures are not fully understood. However, excessive mTORC1 activation likely contributes to epilepsy by disrupting neural circuit formation and modifying existing networks. It remains unclear whether seizures result directly from mTORC1 hyperactivation or as a secondary effect of structural cortical malformations.

2.3.4 Metabolic pathways

Inborn errors of metabolism (IEM) arise from genetic defects that affect the biochemical pathways regulating metabolism and neuronal activity. While IEM are classified as rare genetic disorders, the occurrence of epilepsy is frequently observed as a prominent symptom. The epilepsy phenotypes linked to IEM may frequently progress into developmental and epileptic encephalopathy. A recent study revealed pathogenic or likely pathogenic variants in metabolic genes in 39 of 89 patients (43.8%), linked to conditions such as progressive myoclonus epilepsies, developmental and epileptic encephalopathies, generalized epilepsy, early myoclonic encephalopathy, Lennox-Gastaut syndrome, and West syndrome (Jose et al., 2024).

Metabolic disorders can lead to seizures through various mechanisms. One such mechanism involves the accumulation of harmful metabolites, including ammonia. Hyperammonemia is associated with several metabolic conditions, particularly those affecting the urea cycle and organic acidemias. The accumulation of ammonia is neurotoxic, as it promotes the synthesis of glutamine, which in turn leads to astrocyte

swelling and cerebral edema. Metabolic disorders can disrupt neurotransmission and lead to seizures. In glycine encephalopathy, excess glycine overstimulates NMDA receptors, potentially triggering seizures. Energy deficits, particularly from hypoglycemia due to fatty acid oxidation defects or GLUT-1 deficiency, can also cause seizures, as the brain relies on glucose for energy. Mitochondrial disorders may reduce ATP production, contributing to seizure activity. Cofactor deficiencies, such as pyridoxal phosphate, are linked to specific epilepsies. Additionally, some metabolic disorders, like peroxisomal disorders causing polymicrogyria, can create structural brain abnormalities that increase seizure risk (Almannai et al., 2021).

Six hundred metabolic epilepsies have been identified, accounting for about 37% of all described inherited metabolic disorders and the predominant mode of inheritance for metabolic epilepsies is autosomal recessive (Tumiene et al., 2022).

2.3.5 Synaptic transmission

Synaptic transmission refers to the communication process between a presynaptic neuron and a postsynaptic neuron, along with the subsequent processing of the transmitted signal. This process is characterized by its complexity and stringent regulation, underscoring its significance in maintaining normal brain function and homeostasis. The continuity of synaptic transmission relies on a cyclical process involving the formation, release, and endocytosis of synaptic vesicles, which necessitates the involvement of various proteins, including dynamin, syndapin, synapsin, and synaptic vesicle protein 2A. Regulation of synaptic transmission occurs through a variety of mechanisms, encompassing presynaptic modulators that influence synaptic vesicle dynamics, postsynaptic receptor activity, and the modulation of

neurotransmission itself. Neurotransmitters released from the presynaptic neuron can interact with postsynaptic receptors, which include the inhibitory γ -aminobutyric acid (GABA) receptors and the excitatory glutamate receptors. Upon release, glutamate engages multiple postsynaptic receptors, such as α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), N-methyl-D-aspartate (NMDA), kainate, and metabotropic receptors. The activation of these receptors initiates downstream signaling pathways that produce a wide range of effects, which can be further modulated by various auxiliary regulatory subunits. Additionally, several neuropeptides, including neuropeptide Y, brain-derived neurotrophic factor (BDNF), somatostatin, ghrelin, and galanin, serve as regulators of various synaptic functions alongside traditional neurotransmitters. Dysregulation of synaptic transmission is implicated in the development of numerous neurological disorders, including epilepsy. The mechanisms involved in synaptic transmission serve as important targets for epilepsy treatments, encompassing a range of antiepileptic medications that are presently utilized in clinical settings, along with innovative therapeutic approaches that are currently being researched (Casillas-Espinosa et al., 2012).

Mutations of many proteins involved in the multi-step process of synaptic transmission cause epilepsy, which include the genes involved in a) presynaptic neurotransmitter release - *SV2A*, *SYNs*, *STX1B*, *SNP25B*, *STXBP1*, *CACNA1A*, *DNM-1*, *KCNA1*, *KCNA2* b) postsynaptic proteins - *CHRN4*, *GRIN2A*, *GRIN2B*, *CACNG2*, *SYNGAP1*, *PLPPR4*, *SLC12A5*, *GABRA1*, *GABRB3*, *GABRG2*, *IQSEC2* c) trans-synaptic communication - *NRXN1*, *DAG1*, *CNTNAP2*, *SRPX2*, *LGI-ADAM22-ADAM23*, *C1Q* and d) other proteins - *SLC1A2*, *VLGR1*, *PCDH19*.(Fukata and Fukata, 2017)

2.3.6 Chromatin remodeling and transcription regulation

Chromatin remodelers are crucial in modulating transcription, functioning as both activators and repressors by altering chromatin architecture. This alteration facilitates the accessibility of gene regulatory machinery or, alternatively, conceals regions where gene expression needs to be suppressed. The human nervous system comprises various neuronal cell types alongside supporting glial cells. All these distinct cell types originate from neural progenitor cells (NPCs), which undergo differentiation and migration to their designated locations throughout development. Epigenetic mechanisms significantly influence this process, thereby determining cell fate (Lamar and Carvill, 2018).

The chromodomain helicase DNA-binding (CHD) protein family consists of ATP-dependent chromatin remodelers that facilitate the restructuring of chromatin and the incorporation of histone variants essential for the regulation of gene expression. These CHD proteins are crucial in neurodevelopment, with pathogenic mutations identified in *CHD1*, *CHD2*, *CHD4*, *CHD7*, and *CHD8* linked to various neurological conditions, such as autism spectrum disorder (ASD), intellectual disability (ID), and DEE.

The *MECP2* and *ARX* genes play crucial roles in the normal development of the brain. Pathogenic sequence alterations in these genes are associated with Rett syndrome (RTT) and EIEE, respectively. Specifically, pathogenic variations within the *MECP2* coding region account for 90-95% of classical RTT cases and 40-60% of atypical RTT cases. *MECP2* was initially characterized as a nuclear protein that functions as a transcriptional repressor, specifically recognizing marks of DNA methylation (Gulmez Karaca et al., 2019).

2.4 Genetic testing in DEE

DEE show considerable genetic and phenotypic diversity, with various genetic testing methods available to investigate the growing list of epilepsy-related genes. An accurate genetic diagnosis of epilepsy is crucial for individuals and families, providing clinical and personal benefits. The advancements in understanding the etiology and fundamental pathophysiological mechanisms, as revealed by genetic research, have encouraged clinicians to embrace alternative viewpoints and diagnostic frameworks for precise phenotypic characterizations and the correlation between phenotype and genotype. In DEEs, early genetic testing is cost-effective and can reduce the need for invasive procedures. Identifying the specific causative gene can also help select anti-seizure medications for up to 76% of young children with epilepsy. Characteristics of genetic testing in DEE are illustrated in **Table 2.2**.

2.4.1 Next generation sequencing

The advent of NGS has resulted in a significant surge in the identification of genetic factors associated with both exceptionally rare diseases and prevalent yet diverse disorders. This technology can be utilized for comprehensive genome analysis or targeted examination of specific genomic regions, such as through genome or exome sequencing and gene panels. Advancements in sequencing technology, such as using fluorescent dyes instead of radiolabeling and capillary array electrophoresis instead of traditional gel methods, have automated the process and greatly increased sequencing capacity, enabling the analysis of thousands of base pairs in a single run, compared to just a few hundred previously. NGS is based on the "sequencing-by-synthesis" principle. This method involves monitoring nucleotide incorporation in real-time as

the DNA strand elongates, aided by the sequencing equipment. Importantly, improvements in sequencing capabilities have been matched by a substantial decrease in the costs of sequencing a human genome.

DEE exhibit significant genetic and phenotypic diversity. A wide array of NGS based genetic testing options is available (**Figure 2.5**), encompassing gene panels that may analyze a limited number of genes or extend to hundreds. WES examines approximately 20,000 genes, while WGS encompasses the entire three billion base pairs of the genome. Although WGS is primarily conducted in research environments, its integration into clinical practice is anticipated in the near future.

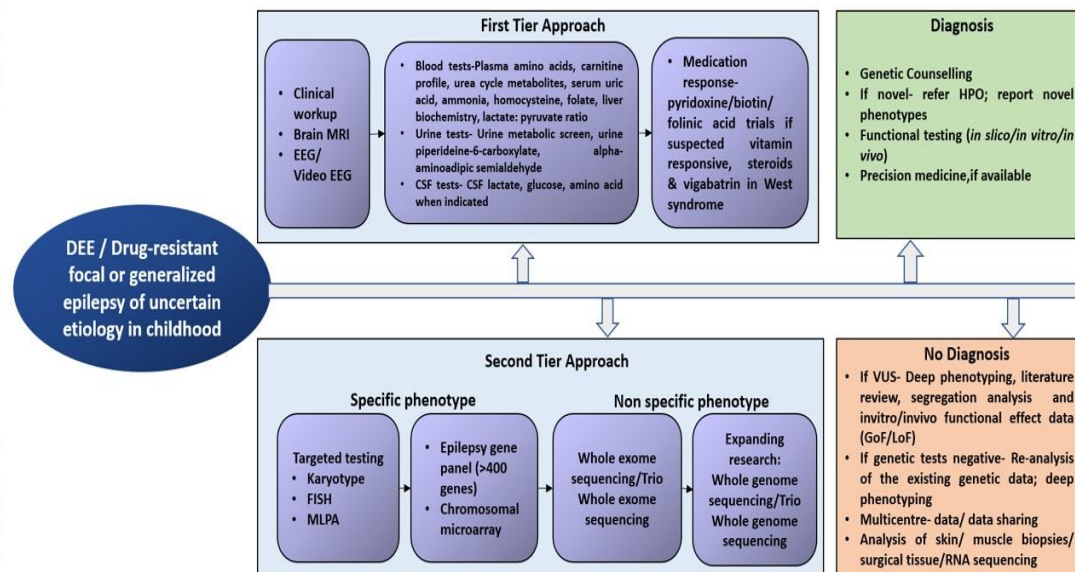


Figure 2.5: Diagnostic tests in refractory childhood epilepsies and DEE (Adapted from Fasaludeen et al., 2024)

Whole Exome Sequencing (WES)

WES involves the concurrent sequencing of the complete coding sequence along with adjacent intronic regions of the human genome, facilitating the identification of SNVs and CNVs. Typically, non-coding regions that are intra- and inter-genic are not included in the analysis, with the exception of splice sites located near exons. WES is regarded as the most effective method for genetic diagnosis in DEE, achieving a diagnostic yield of approximately 50%.

This yield is enhanced when employing trio strategies, which involve sequencing the exomes of both parents to facilitate the segregation of identified variants of interest. Additionally, bioinformatics algorithms are utilized for the identification of CNVs as well as for the detection of nucleotide expansions (Aledo-Serrano et al., 2021).

A recent study of patients with DEEs performed 82 WES, identifying 35 pathogenic variants, resulting in a 43% detection rate. Of these, 23 variants (66%) were *de novo*, with autosomal dominant inheritance being the most common (60%), followed by autosomal recessive (17%). The majority of mutations were missense (75%), followed by frameshift deletions (16%), duplications (5%), and splicing mutations (3%). The study supports WES as a primary method for genetic testing in DEEs (Vetri et al., 2024).

Schiffer et al. demonstrated the effectiveness of exome sequencing in the standard clinical management of pediatric patients with DEEs. The application of exome sequencing identified pathogenic or likely pathogenic variants in 30 of the 103 patients (29%), while variants of uncertain significance were found in 39 of the 103 patients, accounting for 38% (Scheffer et al., 2023). Trio Whole exome sequencing conducted

on a cohort of 221 pediatric Chinese patients diagnosed with epilepsy resulted in a diagnostic yield of 64.5%, with 55.2% of the identified variants being classified as novel (Jiang et al., 2021).

Table 2.2 : Comparison of Genetic tests in DEE

<i>Characteristics/ Genetic testing</i>	<i>Targeted Epilepsy panel</i>	<i>WES</i>	<i>WGS</i>	<i>CMA</i>	<i>MLPA</i>
<i>Cost</i>	Low	High	High	Low	Low
<i>Diagnostic rate</i>	19-48.5%	24-45%	34-48%	9-20%	~13.5%
<i>Time</i>	Rapid	Slow	Slow	Rapid	Rapid
<i>Advantages</i>	Useful when multiple genes contribute to a similar phenotype in different individuals. High diagnostic yield in well-selected individuals based on clinical characteristics.	The complete coding sequence is covered. Trio exome sequencing has the capability to identify de novo variants. It facilitates additional reanalysis opportunities. Furthermore, it may be utilized for the detection of CNVs.	Potential to detect mutations in non-coding regions and large structural variants.	Effective for identifying significant duplications, deletions, translocations, and CNVs.	Useful to detect deletions or duplications of one or more exons in a single gene with strong clinical suspicion.
<i>Disadvantages</i>	Difficult to detect large deletions/duplications and CNVs. Unable to identify novel genes.	Difficult to detect large CNV.	Longer analysis duration and significant data storage capacity is required, the sequencing depth is low, and the costs are significantly high.	Detection of small intragenic deletions or duplications (point mutations) is not feasible.	Ineffective for identifying point deletions or duplications.

A study using WES on 9,170 individuals with epilepsy and 8,436 controls of European ancestry found that people with epilepsy had a higher frequency of ultra-rare, harmful variants in constrained genes and genes linked to epilepsy, particularly among those with DEEs (Feng et al., 2019). A recent study on exome sequencing in patients with DEE found that WES can identify genetic causes in more individuals, even when targeted gene panels are inconclusive. Timely WES implementation can shorten the diagnostic journey for patients and families, especially for rare epilepsy forms. Therefore, WES is recommended as the new standard for diagnosing DEE, replacing targeted gene panel testing.

Whole Genome Sequencing (WGS)

WGS involves the comprehensive analysis of the entire human genome, which allows for the detection of SNVs and CNVs in both coding and non-coding regions, including intronic and intergenic areas. This capability enhances the effectiveness of genetic testing. Additionally, GS is expected to become increasingly advantageous for identifying specific variant types, such as repeat expansions and structural variants, which may be more readily detected. However, challenges remain in the interpretation of non-coding variants and the management of data storage. Typically, GS generates a greater number of variants of uncertain significance compared to exome sequencing (ES), and employing a trio approach can significantly aid in the interpretation of these variants (Krey et al., 2022).

WGS conducted on 30 thoroughly phenotyped patient trios with DEE, who remained undiagnosed following initial testing, revealed pathogenic variants in 19 of the 30 cases. WGS enhanced the diagnostic yield compared to WES, mainly by identifying complex structural variants (Palmer et al., 2021). 1. A comprehensive whole-genome analysis, which included whole-genome sequencing and extensive variant discovery methods, was conducted on a cohort of 14 subjects diagnosed with EIEE who had not received a diagnosis from previous genetic testing. This analysis successfully identified pathogenic or likely pathogenic mutations in all 14 individuals. The study revealed both *de novo* point mutations and insertions/deletions (INDELs), as well as *de novo* structural rearrangements in genes known to be associated with EIEE, in addition to mutations in genes that had not been previously linked to this condition.

This underscores the effectiveness of whole-genome analysis in expediting the clinical diagnosis of EIEE while also reducing associated costs (Ostrander et al., 2018).

Targeted Epilepsy panel

Panel sequencing targets genes linked to epilepsy by sequencing their coding regions and adjacent intronic areas. However, it only evaluates genes included in the panel, which can quickly become outdated due to rapid advancements in gene discovery. The diagnostic yield of targeted epilepsy panel in epilepsy varies between 19% and 48.5% (Fasaludeen et al., 2024). A recent study using large multigene panel and exome sequencing on 127 individuals suspected of monogenic epilepsy identified a definitive molecular diagnosis in 36% of cases (Krygier et al., 2024).

2.4.2 Chromosomal microarray

Chromosomal microarray analysis (CMA), a method employed to identify sub-microscopic chromosomal deletions/duplications or copy number variants, demonstrates a diagnostic yield ranging from 9% to 20% in cases of epilepsy (Burk et al., 2024). A total of 1.5-3% of prevalent epilepsies are linked to CNV (Niestroj et al., 2020). In the context of DEE, the diagnostic yield can rise to as high as 16%. NGS facilitates the simultaneous detection of both CNV and SNVs within a single assay, thereby rendering microarray analysis unnecessary in specific circumstances.

2.4.3 Karyotyping

Conventional karyotyping has been largely replaced by chromosomal microarray analysis (CMA) for the detection of CNVs. Nevertheless, karyotyping remains relevant for identifying significant structural chromosomal abnormalities, such as

translocations, inversions, and ring chromosomes, as well as numerical chromosomal anomalies like aneuploidy. However, such analyses are infrequently requested in the context of epilepsy. Patients with genetic syndromes, including Down syndrome (trisomy 21), are usually diagnosed before the emergence of epileptic symptoms. It is important to request karyotyping if there is a clinical suspicion of a ring chromosomal disorder, as CMA may not detect this condition effectively (Krey et al., 2022).

2.4.4 Multiplex-ligase dependent probe amplification

MLPA is a technique used to detect small deletions/duplications in a genomic region. The straightforward application of routine procedures for epilepsy diagnosis commenced with the introduction of MLPA for identifying CNVs in the *SCN1A* and *KCNQ2* genes. This technique is valuable for evaluating gene dosage in a particular gene when there is a significant clinical suspicion.

2.4.5 Sanger sequencing

Sanger sequencing has largely fallen out of favor in standard diagnostic procedures. Currently, it serves primarily as a confirmatory test to validate variants identified as pathogenic through NGS and is also employed for carrier testing in parents, siblings, or other family members who may be similarly affected.

2.5 Precision medicine

The field of precision medicine in epilepsy has garnered significant interest, particularly as advancements in gene discovery enhance our comprehension of the underlying biology of the disease. A number of targeted therapies are being developed, characterized by their complexity and customization to individual patients. Notably,

there have been infrequent successes in enhancing short-term outcomes for a select group of individuals with epilepsy. The potential for additional targeted, repurposed, or innovative treatments appears to be encouraging. Examples of precision medicine applied in the treatment of epilepsy are depicted in **Table 2.3**.

Table 2.3: Examples of precision medicine applied in the treatment of epilepsy (Striano and Minassian, 2020)

Gene	Epilepsy syndrome	Precision medicine
<i>SCN1A</i> (Voltage-gated sodium channel subunit)	Dravet syndrome	Avoid sodium channel blockers- eg. Carbamazepine, phenytoin
<i>SCN2A</i> (Voltage-gated sodium channel subunit)	EIEE 11	Use sodium channel blockers for GOF variants, but avoid them for LOF variants.
<i>SCN8A</i> (Voltage-gated sodium channel subunit)	EIEE 13, BFIS 5	Prioritize sodium channel blockers for GOF variants.
<i>KCNT1</i> (Sodium activated potassium channel)	EIEE 14, NFLE	Potassium channel activators (Quinidine for GOF variants)
<i>KCNQ2</i> (Voltage-gated potassium channel)	EIEE 7, BFNS1	Potassium channel activators like Retigabine and Ezogabine are for LOF variants, while sodium channel blockers such as Carbamazepine may also be effective
<i>KCNA2</i> (Voltage-gated potassium channel)	EIEE 32	4-aminopyridine (4-AP), an inhibitor of Kv1 channels FOR GOF variants.
<i>SLC2A1</i> (Glucose transporter)	GLUT1 deficiency	Ketogenic diet
<i>mTOR, TSC1, TSC2</i> (mTOR pathway effectors/regulators)	FCD type II	mTOR inhibitors (Everolimus)
<i>PRRT2</i> (Coregulator of synaptic transmission)	Familial infantile convulsions with paroxysmal choreoatetosis	Sodium channel blocker (carbamazepine)
<i>PNPO</i> (PNP oxidase)	PNP oxidase deficiency	Pyridoxal-5-phosphate
<i>ALDH7A1</i> (Aldehyde dehydrogenase)	Pyridoxin-dependent epilepsy	Pyridoxine (B6 vitamin)
<i>GRIN2A, GRIN2B</i> (NMDAR subunits)	Focal epilepsy with speech disorder, with/without mental retardation	NMDAR antagonists, such as memantine, along with dextromethorphan, for GOF variants
<i>ARG1</i> (Arginase)	LGS	Avoid Valproate and enzyme replacement
<i>BTD</i> (Biotinidase)	West syndrome	Biotin supplementation

2.6 Functional analysis

The ongoing advancements in technology are rendering the routine sequencing of human genomes more feasible, particularly within clinical environments. Nevertheless, the challenge of interpreting the clinical implications of genetic variants identified through sequencing continues to pose a significant obstacle to the advancement of precision medicine. Genome-wide association studies (GWAS) and expression quantitative trait locus (eQTL) analyses serve to connect genetic variants with diseases. Nonetheless, their application has predominantly focused on common variants, as these analyses necessitate precise assessments of allele frequency disparities between affected individuals and control groups. Rare variants within a gene can be compiled for gene-disease association testing (Lee et al., 2014). While these burden tests can identify genes that contain variants associated with disease, they often lack clarity in interpreting the specific individual rare variants. The identification of a rare or *de novo* genetic variant within a gene previously associated with an individual's phenotype led to the conclusion that the variant was causal. As the number of sequenced individuals continues to rise, it will be essential to implement more robust strategies to prevent erroneous interpretations, particularly in the context of Mendelian disorders (Starita et al., 2017). Aside from clearly pathogenic nonsense mutations and established splice-site variants, newly identified variants in actionable genes typically lack sufficient evidence for classification as either pathogenic or benign, resulting in their designation as VUSs. The presence of a VUS can lead to confusion for both patients and clinicians, as it introduces uncertainty and cannot serve as a reliable basis for diagnosis or treatment decisions (Richards et al., 2015). A range of strategies has been developed to address the challenges associated with VUSs,

including family segregation analysis, computational predictions of variant effects, data sharing initiatives, and functional assays. At present, computational prediction remains the sole method capable of delivering evidence for variant interpretation on the required scale. Functional data serve as one of the most compelling forms of evidence for determining whether a variant is pathogenic or benign; therefore, functional assays present a promising approach to address the challenges posed by VUS.

A thorough understanding of the molecular mechanisms underlying disease, along with the structural, functional, and biological interactions of the proteins implicated in its pathogenesis, is essential for the development of precision therapies. The growing prevalence of newly identified missense variants presents significant challenges in interpreting their pathogenicity. Not all mutations result in alterations to protein function, and those that do may exert varying effects on disease manifestation. Consequently, prior to conducting comprehensive analyses, it is challenging to categorize novel variants as either harmful or benign due to the uncertainty surrounding their functional and phenotypic implications. Therefore, it is essential to pursue additional research to confirm genetic diagnoses upon the discovery of a new missense variant. Ideally, *in vitro* assessments of these variants should be performed; however, practical constraints often hinder the experimental validation of a large number of mutations and their effects on protein function. An alternative and rigorous method for establishing pathogenicity involves the use of animal models to create homologous mutations that mirror the human phenotype. Nonetheless, similar to *in vitro* approaches, this method is characterized by its time-consuming, labor-intensive, and costly nature, rendering it impractical for the experimental evaluation of all newly

identified variants. To address these challenges and to facilitate rapid and efficient predictions regarding the functional impact of nonsynonymous variants on protein stability, structure, and function, a variety of computational tools have been developed.

The growing prevalence of newly identified missense variants presents significant challenges in interpreting their pathogenicity. Not all mutations result in alterations to protein function, and those that do may exert varying effects on disease manifestation. Consequently, prior to conducting comprehensive analyses, it is challenging to categorize novel variants as either harmful or benign due to the uncertainty surrounding their functional and phenotypic implications. Therefore, it is essential to pursue additional research to confirm genetic diagnoses upon the discovery of a new missense variant. Ideally, *in vitro* assessments of these variants should be performed; however, practical constraints often hinder the experimental validation of a large number of mutations and their effects on protein function. An alternative and rigorous method for establishing pathogenicity involves the use of animal models to create homologous mutations that mirror the human phenotype. Nonetheless, similar to *in vitro* approaches, this method is characterized by its time-consuming, labor-intensive, and expensive, rendering it impractical for the experimental evaluation of all newly identified variants. To address these challenges and to facilitate rapid and efficient predictions regarding the functional impact of nonsynonymous variants on protein stability, structure, and function, a variety of computational tools have been developed.

Bioinformatics is increasingly recognized as a vital component across various domains of biology, particularly in molecular biology, computer science, statistics, and genetics. These disciplines are essential for the analysis of gene and protein expression

as well as their regulation. Computational tools are instrumental in elucidating the implications of genomic variations, especially nonsynonymous single nucleotide polymorphisms (nsSNPs), by forecasting their effects on amino acid residue alterations in proteins. In essence, predictive tools are utilized to assess the potential structural and functional consequences of these genetic variants. To achieve a more accurate evaluation of the structural ramifications resulting from amino acid substitutions, molecular dynamics simulations can be employed, facilitating the examination of changes in numerous parameters, including the stability and flexibility of the protein (Al Mehdi et al., 2019).

3 MATERIALS AND METHODS

3.1 Subjects/study design

This hospital based prospective study was conducted at SCTIMST, Thiruvananthapuram, Kerala, India. All patients with complex pediatric epilepsy syndromes associated with/without developmental delay, particularly, refractory epilepsy of unknown cause including developmental epileptic encephalopathies who attended between 2019-2023 at the pediatric epilepsy clinic were included. Patients diagnosed with drug-resistant epilepsy, specifically those with DEE or anti-seizure medication (ASM) resistance of indeterminate or unknown origin, were categorized based on their electroclinical phenotype. This classification included individuals with or without accompanying developmental delays or regressions, provided that their families agreed to participate in genetic testing. Epilepsy syndromes were defined and classified as per ILAE criteria for diagnosis of electroclinical syndromes (Scheffer et al., 2017; Zuberi et al., 2022). Metabolic screening tests using plasma tandem mass spectroscopy for amino acids, organic acid and carnitine metabolites, ammonia, lactate:pyruvate ratios and standard biochemical tests such as blood glucose, blood gas analysis, urine ketones and specific tests like serum homocysteine, serum copper and ceruloplasmin levels, plasma and urine neopterin-biopterin levels and CSF glucose were done on a case to case basis.

Demographic details, family history, detailed clinical history, any comorbid condition were documented. A detailed pedigree analysis for each proband were carried out which provides information about the disorder and its mode of inheritance, course, and

treatment options. Video electroencephalography (VEEG) and multi modal magnetic resonance imaging (MRI) findings of the probands were reviewed. The domain specific, global developmental impairments (GDI) and regression of milestones were evaluated using revised Denver Developmental Scale (Denver II Technical Manual, 1990; Denver Developmental Materials, Inc.). Facial dysmorphisms, neurocutaneous markers and skeletal malformations were noted. Fundus examination carried out to detect optic atrophy, retinitis pigmentosa and macular cherry red spot were documented. Skin and muscle biopsy reports (light and electron microscopy) were reviewed if done, for evidence of inclusion bodies. The electrophysiological evaluation methods included nerve conduction studies (NCS), visual evoked potential (VEP) and somatosensory evoked potential (SSEP). The Institutional Ethics Committee clearance had been obtained for the study (SCT/IEC/1473/December 2019).

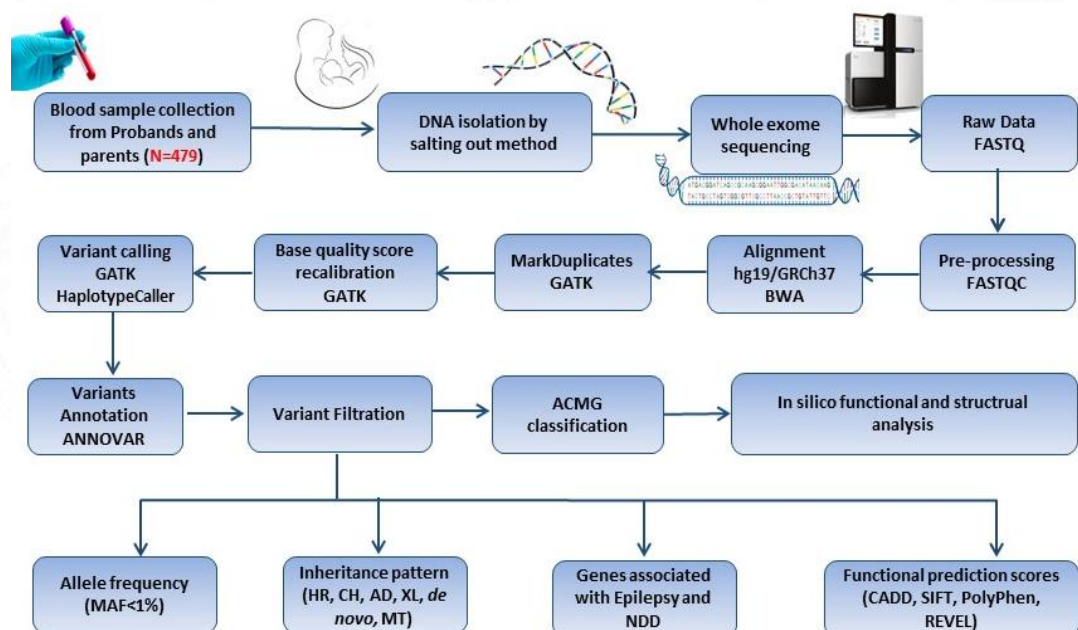


Figure 3.1: Workflow

Trios based study: The study was based on family trios. Variant identification was conducted on all sequenced first-degree relatives. This process facilitates more precise predictive assessments for all three family members, particularly the proband, as the variants are inherited from the parents. (with the exception of *de novo* mutations). WES in trios improves the capacity to assess the significance of variants found in regions with low coverage. Additionally, it allows for the identification of variants even at suboptimal coverage levels, thereby enhancing sensitivity while minimally affecting specificity. DEE are frequently attributed to rare or *de novo* mutations. Consequently, our trio-based study will adequately determine the significance of these rare and *de novo* mutations within the study population.

3.1.1 Inclusion criteria

All patients with complex pediatric epilepsy (<18 years age) with age of onset of epilepsy <12 years of age associated with/without developmental delay which includes refractory epilepsy of unknown cause or EE attended to between 2019-2022 at the pediatric epilepsy clinic conducted at SCTIMST, Trivandrum were included in this prospective cohort study as the core study arm.

The study subjects included two main categories:

Developmental Epileptic Encephalopathies (DEE): According to the 2017 International League Against Epilepsy (ILAE) classification, children with abnormal development due to an underlying cause, alongside an epileptic encephalopathy, are classified as having DEE. Numerous pathogenic gene variants play a role in developmental impairments, and the existence of epileptic encephalopathy worsens

the existing developmental difficulties, negatively impacting overall developmental outcomes.

This group included children whose interictal or ictal EEG abnormalities could result in clinically significant moderate to severe psychomotor delays or regressions that are not commensurate with the underlying epileptogenic condition. eg. malformation of cortical development (MCD). These children may demonstrate a tendency for progression over time, with considerably dependency for daily functioning. According to the 2017 International League Against Epilepsy (ILAE) classification, children with abnormal development due to an underlying cause, alongside an epileptic encephalopathy, are classified as having DEE. Numerous pathogenic gene variants independently contribute to developmental impairment, and the presence of epileptic encephalopathy exacerbates the pre-existing developmental challenges, thereby adversely affecting overall developmental outcomes. Thus the probands along with their parents and siblings were a part of our sample collection strategy.

Complex pediatric epilepsy of unknown cause/presumed genetic etiology (age of onset of epilepsy <12 years age): These group involved a diagnosis of Refractory epilepsy with or without developmental delay/regression; refractory to two or more trials of conventional antiepileptic drugs; recognised chromosomal aberrations. Children with drug-resistant generalized epilepsy syndromes including myoclonic atonic epilepsy, early onset refractory childhood absences, photosensitive occipital lobe epilepsies were also included.

To address the challenges posed by clinical heterogeneity, the constraints of clinical exome analysis, and the complexities associated with *de novo* variants across various

subtypes, we employed WES in trio studies. This approach provides enhanced insights and mitigates the limitations of statistical power within a clinically diverse array of disorders. Understanding population genetic parameters are important to resolve the issues of genetic heterogeneity in a clinically diverse group of disorders.

3.1.2 Exclusion criteria

- a) Acquired etiologies such as trauma, infections, immune mediated encephalitis
- b) Drug responsive conditions including typical self limiting childhood epilepsies, idiopathic generalized epilepsy, children with isolated developmental encephalopathy without epilepsy
- c) Denial of consent for genetic testing
- d) Proven chromosomal duplications/deletions, Trisomy or monosomy or established genomic imprinting defects.

3.1.3 Phenotyping

All participants in the study identified as probands underwent either an overnight or a 24-hour video electroencephalogram (EEG) along with multimodal imaging techniques that encompassed structural magnetic resonance imaging (MRI), diffusion tensor imaging (DTI), arterial spin labeling, and magnetic resonance spectroscopy (MR-Spectroscopy). Complex childhood epilepsy and DEE phenotypes were classified as per current International League Against Epilepsy (ILAE) criteria as Early myoclonic encephalopathy, Ohtahara syndrome, EIEE (< 6 months of age of onset with focal or multifocal epilepsy), West syndrome, Dravet syndrome, Myoclonic

atonic epilepsy, Lennox-Gastaut syndrome, Epilepsy of infancy with migrating focal seizures, Landau Kleffner syndrome/EE with spike and wave activation in sleep, Other focal/multifocal, Refractory generalized epilepsy and Progressive myoclonic epilepsy .

Developmental assessments were conducted by a neuropsychologist, a speech therapist, and an occupational therapist. Cognitive impairments were categorized according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as well as developmental and intelligence quotient scales. The initial metabolic evaluations performed on these subjects, as required by standard care guidelines to rule out metabolic disorders, included blood tandem mass spectrometry and urine tests for organic aciduria, which are routinely administered to these children. In cases where disorders such as neurotransmitter abnormalities or glucose transporter defects (GLUT-1 deficiency) were suspected, the findings from fasting cerebrospinal fluid (CSF) studies for glucose, neurotransmitter levels (including pterin levels), and amino acid concentrations, if conducted during the evaluation, were recorded as part of the endophenotyping process. CSF analyses were not performed separately for the purposes of this genetic study.

Demographic details, family history, detailed clinical history, any comorbid condition were recorded. Endophenotyping using clinical data including dysmorphism, developmental assessments, electrophysiological (EEG) and multi-modal MRI analysis were collected. All protocols and methodology were carried out with written informed consent of the parents of all proband children as per the established guidelines laid down by ICMR for resolving ethical concerns. Peripheral blood (5 ml)

samples of probands and their parents were collected in EDTA vials, for Deoxyribonucleic acid (DNA) isolation, from the study participants which included probands and their parents and siblings. Genomic DNA was isolated from peripheral lymphocytes using Salting out method of DNA extraction.

3.2 DNA isolation

The genomic DNA was isolated from lymphocytes obtained from anticoagulated blood. Modified salting out method was used for DNA extraction (Miller et al., 1988). Ten times volume 1X RBC Lysis buffer [155mM NH₄Cl, 12mM NaHCO₃, 0.1mM EDTA, pH-7.3] was added to 4 ml blood sample and kept for 10 minutes rotation at room temperature. Afterwards, it was centrifuged at 10,000 rpm for 10 minutes and the supernatant containing red blood cells was discarded. The WBC pellet was resuspended in equal volume of Tris-EDTA buffer (1M Tris, 0.5M EDTA, pH-8.0), to which 32 µl proteinase K (20mg/ml) and 100 µl of SDS (20 %) were added. This solution was then incubated at 55°C for 60 minutes followed by 15 minutes incubation at 65°C. During this step, the SDS ruptures the WBCs and the proteins in the cells is digested by the action of proteinase K.

1.52 ml of 4M NaCl was then added followed by equal volume of Chloroform: Isoamyl alcohol (24:1). The samples were rotated for 10 minutes and then centrifuged at 6000 rpm for 15 minutes at room temperature. The aqueous layer alone was collected and double the volume of absolute alcohol was added and mixed gently. The lump of DNA was collected into 1.5 ml fresh centrifuge tube. The DNA was washed twice in 70% alcohol. The pellet was then dried and dissolved in TE buffer and stored at -20°C. The quantity and quality of DNA was analyzed using Nanodrop. The ratio of absorbance

at 260 nm and 280 nm was used to estimate the purity of the DNA. A ratio between 1.7-1.9 was considered as good quality DNA without protein contamination. The concentration of 1µl DNA was quantified and samples were further diluted (1µg/10µl) for exome sequencing.

3.3 WES and Bioinformatics analysis

Genomic DNA samples extracted from peripheral blood of affected child and parents were enriched for sequence targets using Agilent SureSelect all exon V6 kit and sequenced to mean 80-100X coverage on a sequencing platform. The raw data obtained was evaluated for sufficient quality and aligned to human reference genome using BWA (Burrows-Wheeler Aligner) (Li and Durbin, 2010) and processed for duplicates removal and base recalibration. GATK HaplotypeCaller was used to call SNVs and INDELs (Van Der Auwera et al., 2013). Variant functional annotation integrates functional data from diverse sources to clarify the complex roles of genetic variants. ANNOVAR (Wang et al., 2010) was employed to annotate SNVs and insertions/deletions, which included assessing their functional implications on genes, deducing cytogenetic bands, providing functional importance scores, locating variants within conserved regions, and identifying variants documented in the 1000 Genomes Project and dbSNP. The annotation of the identified variants was performed against allele frequency data (1000 Genome Phase 3, ExAC, gnomAD, EVS, dbSNP, 1000 Japanese Genome & internal population databases), *in silico* prediction scores (SIFT, PolyPhen, MutationTaster, REVEL, CADD) and clinical relevance data (ClinVar, OMIM, GWAS, HGMD & HPO).

The variants which passed all filtering steps based on read depth (≥ 10) and genotype quality were considered for further analysis. The annotated trio WES data was filtered for allele frequency data ($MAF < 0.01$) in population databases and interrogated for seven different mode of inheritance: 1) homozygous recessive mutations in which two copies of the variant were inherited from heterozygous parents 2) autosomal dominant mutations where the variant was inherited to the proband from one affected heterozygous variant 3) germline denovo mutations in which the variant was present in the proband and absent in both parents 4) compound heterozygous mutations in the proband, for which two variants present in the same gene at different genomic positions, one variant transmitted from each heterozygous parent 5) X-linked recessive mutation in which the variant is present on the X-chromosome as hemizygous in the proband (male) inherited from heterozygous unaffected mother and 6) X-linked dominant mutations occurred on the X-chromosome as heterozygous in the proband (female) inherited from affected parents 7) Mitochondrial mutations can affect both males and females passed on by mother.

3.3.1 Functional prediction tools

The features of functional prediction tools used for the variant filtration and prioritization was depicted in **Table 3.1**.

Table 3.1: Functional prediction tools used for variant filtration and prioritization

<i>In silico</i> functional prediction tools	Features
CADD (Combined Annotation- Dependent Depletion) (https://cadd.gs.washington.edu/)	Uses 60 different annotations based on: Conservation (phastCons, phyloP, GERP) Epigenetic modification (H3K4Me1, H3K9ac, DNase-Seq) Functional prediction (amino acid change, TF motif disruption) Genetic context (GC content, CpG content, base transversion, InDel length) Annotations are obtained from the Ensembl Variant Effect Predictor (VEP) and include conservation metrics like phastCons, phyloP, and GERP. Additional data comes from UCSC Genome Browser tracks and epigenetic information from ENCODE and NIH Roadmap projects. Amino acid substitution scores for protein-coding sequences are evaluated using Grantham, SIFT, and PolyPhen2. The effects of InDels are determined by the VEP.
SIFT (Sorting Intolerant from the Tolerant) (https://sift.bii.a-star.edu.sg/)	Employs sequence homology to assess the likelihood that an amino acid substitution will influence protein function, thereby potentially modifying the phenotype. Utilizes the amino acid sequence as a basis for prediction. Based on the principles of evolutionary conservation. Considers the position at which amino acid change occurred and the type of amino acid change
PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/)	Predicting the possible impacts of an amino acid substitution on the structural and functional properties of a protein. Evaluates the detrimental impact of a missense mutation by utilizing SNP annotations, gene transcripts, protein sequence information, structural characteristics, and conservation profiles
MutationTaster (https://www.mutationtaster.org/)	Predictions based on the connection between protein structure and function, as well as evolutionary conservation. The analysis encompasses SNP evaluation, examination of splice sites, assessment of conservation, analysis of polyadenylation signals, investigation of the Kozak consensus sequence, and characterization of protein features, including the length of the protein. Utilizes protein characteristics sourced from SwissProt/UniProt, alongside DNA and protein sequences obtained from Ensembl,

	and genotype frequency data derived from HapMap. Limitation exists is the inability to effectively InDels that exceed 12 base pairs, as well as modifications that occur across an intron-exon boundary.
REVEL (Rare Exome Variant Ensembl Learner) (https://sites.google.com/site/revelgenomics/)	The Ensembl approach for assessing the pathogenicity of missense variants is grounded in the utilization of various individual tools, including MutPred, FATHMM, VEST, Polyphen, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP, SiPhy, phyloP, and phastCons. Differentiating pathogenic variants from infrequent neutral variants characterized by allele frequencies below 0.5%.

3.3.2 Variant prioritization and classification

Candidate variants including nonsynonymous, stop-gain, stop-loss, start-loss, frameshift and splice-site mutations which can alter the protein function were assessed with the functional prediction tools. The variant reporting was restricted to genes which are associated with epileptic encephalopathy, intellectual disability, developmental delay and autism spectrum disorder. Variants which are considered to be deleterious with a CADD score of greater than 20 and predicted to be deleterious by at least one prediction tools (SIFT/PolyPhen/MutationTaster/REVEL) (**Table 3.1**).

The selected gene list was compared to the OMIM database, specifically to see if the proband's clinical phenotype and inheritance pattern matched the OMIM disease phenotype and inheritance pattern. The qualifying mutation was also reviewed to see if it had previously been reported in a similarly affected patient by manually assessing the evidence from various clinical databases such as OMIM, ClinVar, HPO and HGMD. The selected variant was visually inspected using Integrative Genomics Viewer (IGV) followed by Sanger Validation.

According to the guidelines of American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP), the variants were classified in a five-class scheme (Pathogenic/Likely Pathogenic/Benign/Likely Benign/Uncertain significance) (Richards et al., 2015). A bioinformatics software tool called InterVar was used for classification of variants by the ACMG/AMP 2015 guideline (Li and Wang, 2017). wInterVar, a user-friendly web server of InterVar, was also used for automatically interpret the variant and to re-interpret the clinical significance by manually adjusting specific criteria.

3.4 ACMG classification criteria for SNVs and INDELs

The ACMG has previously established guidelines for the interpretation of sequence variants. Over the last ten years, advancements in sequencing technology, particularly the emergence of high-throughput next-generation sequencing, have significantly transformed the landscape of genetic testing. Clinical laboratories are now able to conduct a broader array of genetic tests, including genotyping, single-gene assessments, gene panels, exome and genome sequencing, transcriptomic analyses, and epigenetic evaluations for various genetic disorders. This increase in complexity has introduced new challenges in the interpretation of sequence data. In response to these developments, the ACMG formed a workgroup in 2013, which included representatives from the ACMG, the Association for Molecular Pathology (AMP), and the College of American Pathologists (CAP), to review and update the existing standards and guidelines for sequence variant interpretation. The workgroup comprised clinical laboratory directors and clinicians, and this report encapsulates their expert opinions, incorporating feedback from stakeholders of ACMG, AMP, and

CAP. The recommendations outlined in this report are primarily relevant to the diverse range of genetic tests utilized in clinical laboratories, such as genotyping, single-gene tests, panels, exomes, and genomes. It advocates for the adoption of standardized terminology, including ‘pathogenic’, ‘likely pathogenic’, ‘uncertain significance’, ‘likely benign’, and ‘benign’, to classify variants associated with Mendelian disorders. Furthermore, the report delineates a systematic approach for categorizing variants into these five classifications based on established criteria that utilize various types of evidence, including population data, computational analyses, functional studies, and segregation data (Richards et al., 2015).

Two distinct sets of criteria have been established: one for the classification of variants as Pathogenic or Likely Pathogenic, and another for categorizing variants as Benign or Likely Benign (**Figure 3.2**). The pathogenic criteria are assigned weights of very strong (PVS1), strong (PS1–4), moderate (PM1–6), or supporting (PP1–5), while the benign criteria are classified as standalone (BA1), strong (BS1–4), or supporting (BP1–6). These criteria are subsequently integrated according to specific scoring rules to determine a classification within a five-tier system. The application of these rules encompasses all relevant data pertaining to a variant, whether derived from the current case under examination or from rigorously validated data previously published.

		Benign		Pathogenic		
		Strong	Supporting	Supporting	Moderate	Strong
Population Data	MAF is too high for disorder <i>BA1/BS1</i> OR observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
Functional Data	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
Segregation Data	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
De novo Data				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
Allelic Data		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
Other Database		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
Other Data		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

Figure 3.2: Diagram that categorizes each criterion based on the type of evidence and the robustness of the criteria, distinguishing between benign assertions on the left side and pathogenic assertions on the right side (Adapted from Richards et al., 2015).

Variants can be categorized into the Pathogenic or Likely Pathogenic class based on several criteria (**Figure 3.3**). Frequency information obtained from diverse population databases such as 1000 Genomes, gnomAD, and the Exome Variant Server (EVS); *in silico* functional predictions derived from various disease-related databases including ClinVar, OMIM, and HPO; segregation analysis of affected family members confirmed through sequencing; well-established functional studies conducted in cell lines or animal models; *de novo* data analyzed through sequencing of parental samples; patient phenotype data corroborated with the gene's disease association found in multiple disease databases (ClinVar, OMIM, HPO); and allelic data assessed from the occurrence of variants (either *cis* or *trans*) on the chromosome.

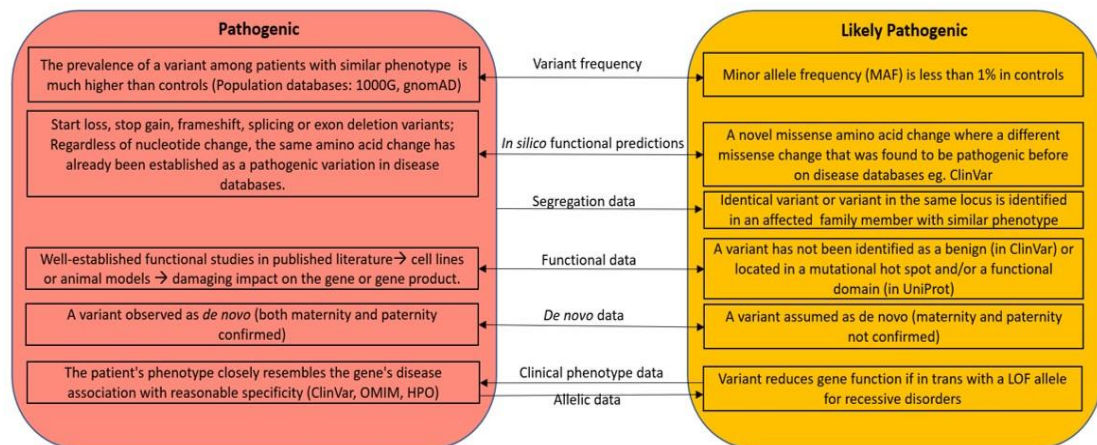


Figure 3.3: Illustration of variant categorization as either pathogenic or likely pathogenic, employing various tiers of evidence (Adapted from Fasaludeen et al., 2024).

3.5 Copy number variant calling

Copy-number variants (CNVs) represent a significant aspect of genetic diversity in humans. These variants may be harmless or contribute to the pathogenesis of diseases by causing dosage imbalances and interfering with genes and regulatory sequences. Therefore, precise detection and clinical characterization of CNVs are crucial.

The CoNIFER (copy number inference from exome reads) pipeline was utilized for the detection of CNVs (Krumm et al., 2012). CoNIFER employs exome sequencing data to identify CNVs and to determine the copy number of duplicated genes. Given that exome capture reactions are prone to significant and systematic biases across different sample batches, singular value decomposition (SVD) has been utilized to mitigate these biases in the exome data. CoNIFER facilitates the integration of exome sequences from various experimental runs by addressing batch effects. When used in conjunction with a short read aligner like mrsFAST, which is capable of aligning reads to multiple genomic locations, CoNIFER can effectively detect rare CNVs and

accurately estimate the copy number of duplicated genes, achieving precision for up to approximately eight copies with the current exome capture kits.

Exome sequencing data obtained from FASTQ files were segmented into non-overlapping 36-base pair units and subsequently aligned to designated target regions, permitting a maximum of two mismatches for each 36-base pair alignment. RPKM values were computed for each exon or targeted region, which were then converted into "ZRPKM" values utilizing the median and standard deviation of each exon across all samples. These ZRPKM values were subjected to singular value decomposition (SVD), during which the first 12 to 15 singular values were excluded. To mitigate the occurrence of false positives, a centrally weighted average of 15 exons was applied to the SVD-ZRPKM values, and a threshold of ± 1.5 SVD-ZRPKM was employed for the identification of CNVs.

The *classifycnv* tool was employed to evaluate the pathogenicity of copy CNVs by utilizing the classification criteria established by the ACMG in 2019 (Riggs et al., 2020). ClassifyCNV utilizes genomic coordinates along with the type of CNV as its input parameters, subsequently providing a clinical classification for each variant, a detailed breakdown of the classification score, and a compilation of genes that may be significant for the interpretation of the variant.

ClassifyCNV takes a BED file as its input and necessitates that the user specifies the genomic coordinates along with the type of copy number variation (CNV), whether it be a deletion or a duplication. Subsequently, the tool employs the criteria outlined in the ACMG scoring guidelines for assessing copy-number losses and gains to determine the clinical relevance of the CNVs. Points are allocated for each component

assessed within the rubric, and the clinical classification is derived from the cumulative points attributed to a particular CNV.

3.6 Sanger sequencing validation for pathogenic/likely pathogenic variants

Sanger sequencing represents a precise method for sequencing DNA that employs oligonucleotide primers to target specific regions of interest. The process initiates with the denaturation of double-stranded DNA, resulting in single-stranded DNA that is subsequently annealed to the oligonucleotide primers. This is followed by the elongation phase, which utilizes a mixture of deoxynucleotide triphosphates (dNTPs) that supply the necessary nucleotides—adenine (A), cytosine (C), thymine (T), and guanine (G)—to construct the new double-stranded DNA structure. Additionally, a limited quantity of chain-terminating dideoxynucleotide triphosphates (ddNTPs) corresponding to each nucleotide is incorporated into the reaction. The elongation process continues with dNTPs until a ddNTP is incorporated, leading to termination of the sequence at various lengths due to the equal probability of dNTPs and ddNTPs binding. Each ddNTP (ddATP, ddGTP, ddCTP, ddTTP) is tagged with a distinct fluorescent marker, which emits specific fluorescence upon incorporation into the growing DNA strand. The colors are standardized, with adenine represented by green, thymine by red, guanine by black, and cytosine by blue. An automated sequencing machine utilizes a laser to detect the emitted fluorescence, converting the intensity into a “peak” on a graph. In instances of heterozygous variants, two fluorescent signals of equal intensity will be observed at the corresponding loci, whereas in homozygous

variants, the original fluorescent signal is entirely replaced by that of the new base pair (Martin et al., 1985).

Primer designing

The primers of pathogenic/likely pathogenic variants identified were designed using the NCBI primers. The DNA sequences spanning the variant region for primer designing were obtained from NCBI. The length of primers ranged from 18-22bp. GC content of the primers designed was 50%. The melting temperature of forward and reverse primer, absence of primer dimers, loop formation etc were analyzed using the online tool for Oligocalc. Finally an *in silico* PCR was done using UCSC genome browser to assure the designed primers amplify the region containing the variant.

PCR amplification

The designed primers were obtained as lyophilized oligonucleotides from Sigma Genosys. The primer is suspended in the required amount of nuclease free water to make to 100uM. The suspended primer stock is stored at -20°C, and diluted to a concentration of 20ul for use in PCR. The constituents of the PCR reactions are as follows: Nuclease free water, 1X buffer, 200uM of deoxynucleotide mixture, 2pmol of primer and 0.5 units of Taq polymerase. 20ng of target DNA is added as sample in the reaction. All the components of the PCR reaction were mixed properly by vortexing. PCR was carried in Thermal cycler.

Cycle sequencing by BIG Dye Terminator

The amplification of DNA by PCR was verified by gel electrophoresis. Sequencing was done using Applied Biosystems Big Dye terminator v3.1 cycle sequencing kit. The sequencing reaction mix consists of nuclease free water, sequencing buffer, 10 microlitre of forward or reverse primer and sequencing reaction mix. The polymerase present in Big Dye terminator kit is Amplitaq Gold polymerase. The final volume of the reaction was made to 10 microlitre.

Clean up of sequencing product

After the sequencing PCR, the products were cleaned up to purify the sequenced DNA from other excess reagents. Sequencing clean up removes the remnants of post cycle sequencing reaction – excess dNTPs, primers, enzymes, unincorporated fluorescent dyes and residual salts. 2uL of 125uM EDTA, 10 microlitre of nuclease free water and 2uL of 3M sodium acetate was added to the sequencing product and vortexed well. The samples were incubated for 15 minutes followed by which centrifugation was done at 10,000 rpm for 20 minutes at 4°C. The supernatant was discarded and the pellet was washed with 70% alcohol. After 2 rounds of washing, the DNA pellet was air dried.

Capillary sequencing

Formamide was added to the air-dried samples and was subjected to PCR conditions for 10 minutes at 95°C for denaturation. The denatured samples were loaded into the capillary sequencer for analysis of the sequence. The ABI PRISM DNA analyzer used in the study is automated to handle 96 samples at a time. The long capillaries in the analyzer are filled in by POP7 polymer for resolving the sequences. DNA fragments move towards the end of capillary, with smaller fragments moving faster than longer

fragments. The fluorophores in the attached ddNTPs, a specific color fluorescence emission occurs. Thus a four color sequencing emission spectra is generate. DNA sequences upto 1000 bases can be read accurately using automated sequencer. The sequences obtained can be analyzed using the software sequence analyzer V.1.1.

3.7 In silico functional analysis

In order to evaluate the pathogenic potential of novel variants and VUSs concerning their impact on protein function and structural stability, computational analysis tools were employed. These tools focused on aspects such as stability, interaction patterns, and three-dimensional structural analysis of the proteins.

3.7.1 Protein network analysis

A protein engages in numerous interactions with various other proteins within the cellular environment, and these interactions play a crucial role in the protein's functionality and regulatory mechanisms. To identify the interactions of protein with other relevant proteins, STRING (Search Tool for the Retrieval of Interacting Genes) (<https://string-db.org/>) was applied. STRING utilizes various data types, including gene fusion, co-expression, functional associations, and experimental evidence, to predict the interacting partners of a given protein. When provided with a list of proteins, STRING is capable of identifying their neighboring interactors, which are the proteins that exhibit direct interactions with the specified proteins. Subsequently, STRING can construct a protein-protein interaction (PPI) network that encompasses all these proteins along with the interactions that occur among them.

3.7.2 Protein stability prediction

The stability of proteins is a crucial determinant of the functionality and activity of biological molecules. The free energy associated with protein unfolding plays a vital role in maintaining protein stability. So, it is important to precisely quantify the influence of mutation on protein stability by analysing the effect of mutation on free energy. The stability of protein variants was assessed using IMutant2.0 (<https://folding.biofold.org/i-mutant/i-mutant2.0.html>), and MuPro (<https://mupro.proteomics.ics.uci.edu/>). I-Mutant2.0 and MUpro have the capability to assess the impact of substitutions on protein stability by employing support vector machines (SVM). This analysis incorporates the protein sequence, the specific mutation, its position, and the residue of interest. The outputs generated predict alterations in stability by quantifying the change in free energy (DDG) and indicating the direction of change following a single point mutation in the protein. A positive score signifies an enhancement in protein stability, whereas a negative score indicates a destabilizing effect.

I-Mutant 3.0 is a tool based on Support Vector Machine (SVM) methodology that assesses the impact of single nucleotide polymorphisms (SNPs) on protein stability. It determines whether a mutation leads to stabilization or destabilization of the protein structure by calculating the change in free energy, utilizing predictions in conjunction with the energy-based FOLD-X tool. The scoring system spans from negative to positive values, with negative numbers indicating destabilization and positive numbers signifying stabilization.

Mupro serves as an online structural analysis resource designed to compute variations in protein stability resulting from various SNPs. A score approaching 0 indicates a state of stability that remains largely unchanged. Conversely, a score nearing -1 reflects a strong confidence in a decline in stability. In contrast, a score close to +1 signifies a high level of confidence in an improvement in stability.

3.7.3 Protein structural prediction

A mutation in a specific site, such as a conserved residue, can cause a change in the protein's structure and function, so, it is crucial to figure out how mutations affect protein structure to predict the pathogenicity of a variant. Missense3D (<http://missense3d.bc.ic.ac.uk/missense3d/>), a web-based tool to predict the structural changes introduced by an amino acid substitution, was used to predict the effect of mutation on protein structure.

3.7.4 Pathway analysis

A functional pathway refers to a series of biochemical processes that lead to a particular result or cellular function. This can involve the activation or inhibition of enzymes or metabolites, the modulation of signaling pathways, the repair of DNA integrity, or the coordination of immune responses. Examining pathways related to genes identified with deleterious variants provides an additional approach to identifying biological pathways enriched with variants associated with epilepsy. The Database for Annotation, Visualization and Integrated Discovery (DAVID) tool functional annotation tool was employed to conduct functional annotation and enrichment analysis of genes associated with deleterious variants. DAVID serves as a

widely utilized bioinformatics resource system that encompasses both a web server and web service dedicated to the functional annotation and enrichment analysis of gene lists. It features an extensive knowledgebase along with a variety of tools designed for functional analysis.

DAVID primarily offers standard batch annotation and gene Gene Ontology (GO) term enrichment analysis to emphasize the most pertinent GO terms linked to a specified gene list. The Functional Annotation Clustering feature of DAVID employs a Kappa statistic score to assess the relationships among annotation terms based on the extent of their co-associative genes. Additionally, it utilizes an innovative fuzzy clustering algorithm to organize similar, redundant, and diverse annotation content from various sources into cohesive annotation groups. This approach alleviates the challenge of associating similar redundant terms, thereby enhancing biological interpretation at a group level.

The identification of human phenotype ontology terms linked to genes with deleterious variants was conducted using WikiPathways (<https://www.wikipathways.org/>). WikiPathways serves as a repository for biological pathway models that have been gathered and curated by the research community.

3.7.5 Functional prediction for ion channel variants

funNCion (<https://funnc.shinyapps.io/shinyappweb/>)

Voltage-gated sodium channels (Navs) and calcium channels (Cavs) are essential for the initiation and propagation of action potentials in a diverse range of excitable cells and their associated physiological processes. When the membrane undergoes

depolarization, both Navs and Cavs are activated and subsequently inactivated within a matter of milliseconds, resulting in a brief influx of sodium or calcium ions into the cell. Genetic variants in the genes that encode these channel proteins, such as SCN1A, SCN2A, SCN8A, CACNA1E, and CACNA1A, have been linked to DEE.

funNCion, Functional variant prediction in Navs and Cavs ion channels, was used to predict the functional effect of variants in sodium and calcium channel genes. funNCion is a machine learning approach designed to forecast the functional implications of genetic missense variants within voltage-gated sodium and calcium channels. It assesses loss-of-function (LOF) or gain-of-function (GOF) effects associated with disease-related missense variants in these genes, achieving a receiver operating characteristic (ROC) score of 0.85 across 82 independent test variants (Heyne et al., 2020).

SCN Viewer (<https://scn-viewer.broadinstitute.org/>)

The SCN viewer, a web-based tool, was employed for the interpretation of sodium channel variants. This tool facilitates the examination of functional outcomes associated with pathogenic variants in genes belonging to the voltage-gated sodium channel family. It encompasses ten sodium channel genes: *SCN1A*, *SCN2A*, *SCN3A*, *SCN4A*, *SCN5A*, *SCN7A*, *SCN8A*, *SCN9A*, *SCN10A*, and *SCN11A*. The viewer presents a total of 369 variants that have undergone electrophysiological characterization, categorized into Gain of Function (GoF), Loss of Function (LoF), and Mixed effects. Functional data for each variant have been compiled from over 400 scholarly references. Upon alignment, it is often observed that the functional effects are comparable among analogous positions across various members of the gene family.

3.7.6 AlphaMissense

The assessment of the pathogenicity of missense variants was performed using AlphaMissense. AlphaMissense is an innovative artificial intelligence based methodology that combines several established strategies: (i) it uses weak labels from population frequency data for training, avoiding circularity by excluding human annotations; (ii) it employs unsupervised protein language modeling to capture amino acid distributions based on sequence context; and (iii) it incorporates structural context using a system derived from AlphaFold (Cheng et al., 2023). The AlphaMissense pathogenicity score of variants varies between zero and one, where elevated scores indicate a greater likelihood of pathogenicity.

3.8 Grouping of VUS into promising VUS/incidental VUS

VUS were grouped into Promising VUS/Incidental VUS according to *in silico* functional predictions, frequency data of variants, *de novo* data, clinical data and protein structure/stability data analysed with computational prediction tools (funNCion, SCN Viewer, iMutant, MUpro, Missense3D, etc) (**Figure 3.4**). Molecular dynamic simulation was used to re-classify the promising VUS.

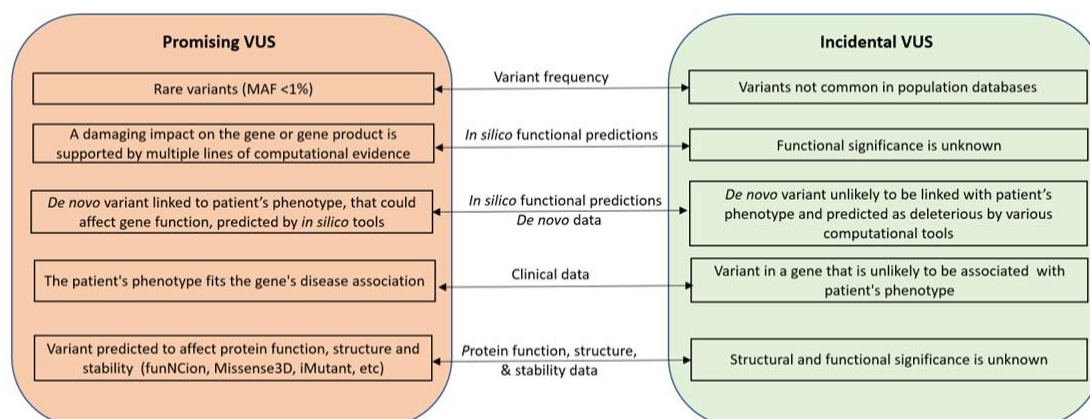


Figure 3.4: Grouping of VUS into Promising VUS/Incidental VUS evidence (Adapted from Fasaludeen et al., 2024).

3.9 Molecular dynamics simulation

AlphaFold predicted structure of the wild type protein was downloaded and quality assessment was done using Ramachandran plot. Mutant residue was introduced using Mutagenesis wizard tool of PyMOL molecular graphics system by selecting the most common and least strained backbone-dependent rotamers. A 100 nanosecond molecular dynamics simulation was carried out with periodic boundary conditions for wild type and mutant protein models using GROMACS 2022.4 (Pronk et al., 2013). The procedure involved energy minimization, followed by isothermal-isochoric (NVT) equilibration, and subsequent isobaric (NPT) equilibration of the system.

The force field utilized in the simulation was OPLS-AA. -The protein atoms were placed in a cubical box constructed using spc216 water model. Appropriate ions were utilized to achieve neutralization of the systems. Energy minimization was done using the steepest descent minimization algorithm for 5000 steps. For equilibrating the systems, isothermal-isochoric ensemble (NVT) was performed with a constant

temperature of 300K for 100ps and isothermal-isobaric (NPT) ensemble was performed at a constant temperature of 1 atm per 100 ps. All systems were soaked in a cubic box of water molecules using the gmx editconf module and gmx solvate module for solvation. All systems were equilibrated to a constant temperature (300K), by utilizing the NVT and NPT process for 100 ps. The final simulations were performed for 100ns. The trajectories obtained were applied to calculate Radius of gyration (Rg), Root mean square deviation (RMSD), and RMSF, using GROMACS utilities (gmx_gyrate, gmx_rms, gmx_rmsf).

3.9.1 Root mean square deviation

The RMSD diagrams illustrate the rate of spatial displacement and the dynamics of the protein model throughout the simulation, functioning as an indicator of protein stability in response to a mutation. A lower RMSD value during the simulation correlates with reduced protein stability. Throughout each simulation phase, the calculations for amino acid units were conducted for each structure, both with and without mutations, in relation to the initial (crystal) state of the simulation, focusing solely on the principal atoms of the molecule (Kufareva and Abagyan, 2011). gmx_rms command of Gromacs was used to calculate RMSD values from the trajectory.

3.9.2 Root mean square fluctuation

The examination of the dynamic behavior of Ca atoms within protein structures provides ample data to assess the influence of individual movements on the overall dynamics of the structure. To identify variations in motion and structural fluctuations, we employed Root Mean Square Fluctuation (RMSF) analysis. The gmx_rmsf

command in Gromacs was employed to compute the RMSF values derived from the trajectory data.

3.9.3 Radius of gyration

The assessment of parameters, including the overall dimensions and the molecular compaction level, was conducted through an analysis of the Radius of gyration (Rg). The `gmx_gyrate` command in Gromacs was utilized to calculate the Radius of gyration values obtained from the trajectory data.

3.9.4 Definition Protein Secondary Structure of Protein (DSSP) analysis

The `gmx_dssp` command of Gromacs processes the trajectory file to analyze the secondary structure at each time frame by invoking the DSSP analysis program (Gorelov et al., 2023). The analysis begins with a structural file that delineates the coordinates, alongside a topology file that provides details regarding the composition of protein residues. The topology is examined initially, focusing exclusively on the atoms constituting the main chain, specifically the C α , C, O, N, and H atoms, collectively referred to in Gromacs as “MainChain+H.” The coordinates of these atoms are recorded, and each residue is linked to its preceding and succeeding residues, where applicable. Subsequently, the molecular dynamics trajectory frames are scrutinized. Each frame is evaluated in isolation, during which the energy of hydrogen bonds between the designated protein residues is computed. The DSSP analysis was conducted to examine the secondary structural components, including alpha helices, beta sheets, and coils, throughout the entire duration of the simulation. Trajectories

were collected at different time intervals, and the resulting structures were visualized with the aid of xmgrace.

3.10 Multidisciplinary team meeting

A bi-weekly multidisciplinary team meeting has been established within our organization, comprising epileptologist, child developmental neurologists, molecular geneticist, and research fellows. The primary objective of these meetings was to analyze the patient's phenotype and review the results obtained. During these sessions, molecular and genetic data, along with clinical information, EEG recordings, and brain MRI scans, were presented. The genetic findings were systematically correlated with the patient's phenotype to enhance variant classification, facilitate further testing, and guide the management of the patient as well as genetic counseling. Following the analysis of genotype-phenotype correlation, the variants were categorized based on the ACMG classification criteria, and reports were subsequently provided to the patients.

3.11 Re-analysis of negative cases

Despite the wide array of diagnostic methods accessible, a minimum of 50% of patients with DEE fail to obtain a genetic diagnosis. For the re-analysis of negative cases, the following steps were performed:

- a) Deep phenotyping or precise and in-depth analysis of phenotypic abnormalities in which the individual aspects of the phenotype to establish the genotype-phenotype correlation
- b) Reverse phenotyping to test the novel genotype- disease associations thereby broadening phenotypic spectrum

- c) Re-analysis of sequencing data
- d) Literature review

The deep phenotyping process, which involved consultations with the clinician possessing extensive knowledge of individual patients, required the submission of approximately ten Human Phenotype Ontology terms for each patient. This methodology enabled a thorough and customized analysis, thereby enhancing the diagnostic capabilities considerably (Krey et al., 2022). A comprehensive examination of the patient's phenotype was conducted. Following this, the methodologies employed in prior tests were assessed.

It is crucial not only to perform high-quality genetic investigations but also to ensure the opportunity for reanalysis after a designated interval, ideally every 3 to 5 years. Exome sequencing that yields negative results at present may reveal positive findings in the future. Reverse phenotyping involves the enhancement of phenotypic characteristics through the analysis and interpretation of genetic information (Aledo-Serrano et al., 2021). For certain patients who have not received a diagnosis, conducting genetic reanalysis alongside more comprehensive or reverse phenotyping can serve as effective diagnostic methods to elucidate newly identified variants of uncertain significance.

Re-analysis of the generated sequencing data was conducted after a suitable time period for patients with ambiguous genetic diagnoses. Re-evaluating the data has demonstrated beneficial outcomes for individuals who had previously received negative test results.

In cases where patients present with significant VUS, we conducted a review of recent literature to identify publications that provide evidence of pathogenicity of these variants, such as data available in ClinVar or the Human Gene Mutation Database (HGMD). In addition, we examined benign variants, especially those that demonstrate a high prevalence in the gnomAD database.

4 RESULTS AND DISCUSSION

4.1 Genetic diagnostic yield

4.1.1 Samples and diagnosis

Between June 2019 and August 2023, the study enrolled 175 patients who had been diagnosed with complex pediatric epilepsy syndromes, specifically, refractory focal/generalized epilepsies of uncertain etiology and DEE. Whole exome sequencing was performed for a total of 479 individuals, which included 140 trios (proband-father and mother), 9 quads (two affected probands-father and mother), 6 duos (proband and single parent) and 11 proband only samples (**Figure 4.1**). In the case of the proband only and single parent samples (17/479), the parental samples were unattainable due to a lack of consent (6/17), the circumstance where one parent was abroad for work during the period of sample collection (8/17), the expiry of the parent (2/17), and the unavailability of the biological father (1/17).

The demographic details and clinical characteristics of 175 probands are provided in **Table 4.1**. Mean age at epilepsy onset was 23 months (range 0.1-144) and genetic diagnostics were conducted at a mean age of 94.3 months (range 2 to 216). Among the 175 patients enrolled in the study, 99 (56.6%) were male and 76 (43.4%) were female. Twenty one (12%) patients were identified as having consanguineous parentage. A positive family history of seizures was recorded for 82 (46.9%) patients, while MRI results were positive for 84 (48%) patients.

In this cohort of 175 patients with seizures, it was observed that 53.1% (93 out of 175) showed global developmental delay. Additionally, 20 patients (11.4% of the cohort)

showed delays in one of the following domains: motor (2.9%), social (1.1%), or language (7.4%). Furthermore, 12.6% (22/175) of the patients had Autism Spectrum Disorder (ASD). 53 (30.3%) were developmentally normal despite refractory epilepsy.

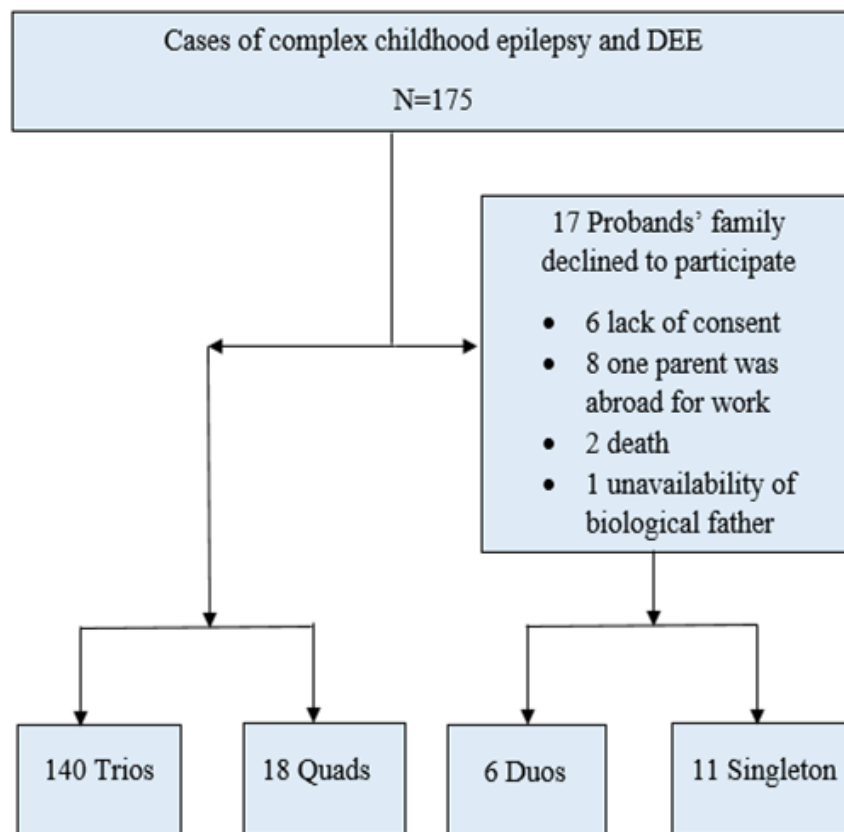


Figure 4.1: Study profile

Table 4.1 : Demographic and clinical characteristics of probands

Characteristics	Total, N=175
Sex, N (%)	
Male	99 (56.6)
Female	76 (43.4)
Age at seizure onset (months)	
Mean (SD)	23 (28.5)
Age of onset < 3 years	134 (76.6)
Age of onset > 3 years	41 (23.4)
Age at genetic testing was done (months)	
Mean (SD)	94.3 (59.8)
Developmental delay, N (%)	
Global	93 (53.1)
Motor	5 (2.9)
Language	13 (7.4)
Social	2 (1.1)
Motor and language	3 (1.7)
Motor and social	1 (0.57)
Language and social	5 (2.9)
Autism, N (%)	22 (12.6)
Seizure type, N (%)	
Focal	45 (25.7)
Generalized	39 (22.3)
Combined	91 (52)
Family history of seizures, N (%)	
Yes	82 (46.9)
No	93 (53.1)
Consanguinity, N (%)	
Yes	21 (12)
No	154 (88)
Brain MRI findings, N (%)	
Normal	91 (52)
Abnormal	84 (48)
Epilepsy phenotype, N (%)	
1. Ohtahara syndrome	1 (0.57)
2. Early Infantile epileptic encephalopathy (< 6months of age of onset with focal or multifocal epilepsy)	23 (13.1)
3. West syndrome	23 (13.1)
4. Dravet syndrome	10 (5.7)
5. Myoclonic atonic epilepsy	5 (2.8)
6. Lennox gastaut syndrome	12 (6.8)
7. Epilepsy of infancy migrating focal seizures	2 (1.1)
8. Landau kleffner syndrome/Continuous spike and wave during slow wave sleep	20 (11.4)
9. Other focal/multifocal epilepsy	45 (25.7)
10. Refractory generalized epilepsy	29 (16.5)
11. Progressive myoclonic epilepsy	5 (2.8)
Seizure in remission on last follow-up, N (%)	60 (34.3)

Epilepsy syndrome diagnoses at presentation were Ohtahara syndrome ($n = 1$), Early infantile epileptic encephalopathy of age of onset less than 6 months ($n = 23$), West syndrome ($n = 23$), Dravet syndrome ($n = 10$), Myoclonic atonic epilepsy ($n = 5$), Lennox-Gastaut syndrome ($n = 12$), Landau Kleffner syndrome or EE with spike and wave activation in sleep ($n = 20$), other focal/multifocal epilepsy ($n = 45$), refractory generalized epilepsy ($n = 29$), and progressive myoclonic epilepsy ($n = 5$).

4.1.2 Diagnostic yield

The diagnostic yield of WES in this cohort was determined by the number of P/LP variants identified. P/LP variants were identified in 71 out 175 patients (40.6%), 60 out of 175 (34.3%) had VUS, and 44 out of 175 (25.1%) had no potentially disease causing variants identified. Of the 71 diagnostic cases, 68 (95.8%) were single nucleotide variants or small insertions or deletions, and 3 (4.2%) were copy number variants (**Table 4.2**). A notable percentage of the disease-causing variants were identified in ion channel genes (40%), whereas neurometabolic genes represented 14% of the variants found (**Figure 4.2**). Most frequent disease causing variants identified were missense (73%) followed by non sense variants (10%), frameshift indels (7%), splice site mutations (4%) and copy number variants (4%) (**Figure 4.3**).

The study uncovered 19 novel variants that have not been reported in any population databases, such as 1000 Genomes, gnomAD, ExAC, and dbSNP (**Table 4.2**). Of these 19 novel variants, 14 were single nucleotide variants or small INDELs identified in *SCN8A*, *TUBA3E*, *NOVA2*, *CHRNA2*, *CDKL5*, *PCDH19*, *NHLRC1*, *GABRA1*, *STXBP1*, *GABRB2*, *ITPA*, *CLTC*, *SCN3A*, *SCN1A*, *SCN2A* and *NPRL2* genes, while one variant was a copy number variation on chromosome 13 (Chr13q32.3del).

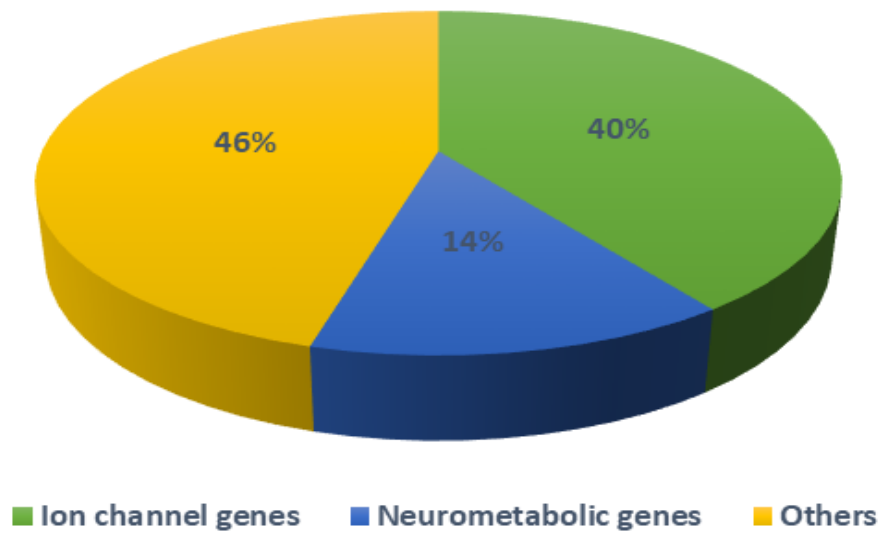


Figure 4.2: Type of disease causing variants

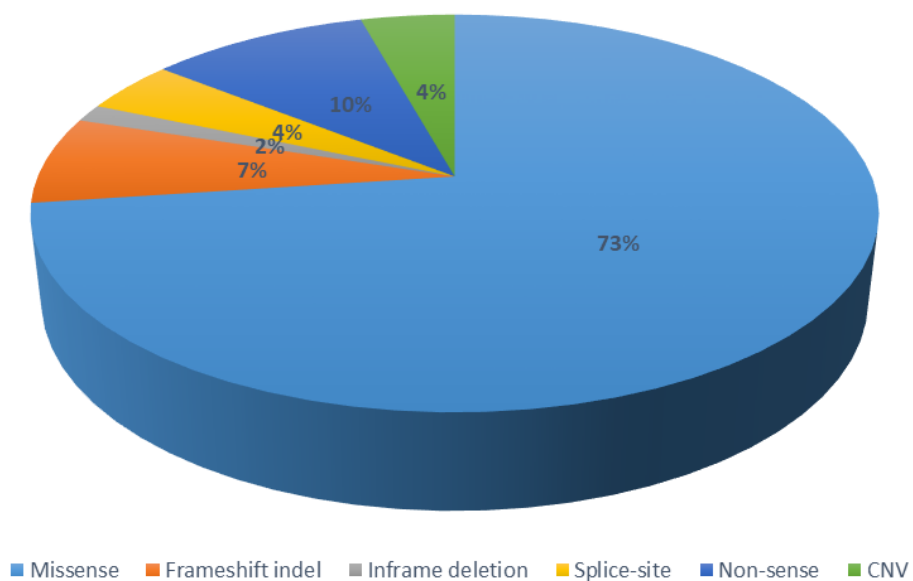


Figure 4.3: Types of variants in diagnostic cases

Table 4.2: List of pathogenic/likely pathogenic variants identified

Sl no	Gene	Variant description	Functional consequence	Allele frequency in population database	ClinVar accession	Functional prediction scores- SIFT/PolyPhen/ MutationTaster/ REVEL/CADD_score)	Inheritance pattern	ACMG class
1	CHRNA4	ENST00000370263:Exon 5:c.1105C>T:p.Arg369Trp	Missense	0.0032	NR	U/B/D/0.34/22.3	Het/De novo	Pathogenic
2	KMT2C	ENST00000262189:Exon 24:c.3758A>G:p.Glu1253Gly	Missense	NA	RCV000623369	U/-/D/0.61/27.2	Het/De novo	Likely pathogenic
3	SCN8A	NM_001330260.2:Exon 17:c.2936C>T:p.Ser979Phe	Missense	NA	NR	D/-/D/0.88/32	Het/Absent in mother	Pathogenic
4	SLC2A1	ENST00000426263:Exon 9:c.1134_1136del:p.Phe379del	In frame deletion	NA	RCV002927716	..	Het/De novo	Pathogenic
5	C12ORF57	ENST00000229281:Exon 1:c.1A>G:p.Met1?	Start loss	0.0062	RCV000760547	/-/D/-/36	Compound heterozygous	Pathogenic
6	C12ORF57	ENST00000229281:Exon 1:c.43C>T:p.Gln15*	Non sense	NA	NR	/-/D/-/60	Compound heterozygous	Pathogenic
7	DCX	DCX:ENST00000371993:Exon 3:c.557G>T:p.Arg186Leu	Missense	NA	RCV000145860	U/U/D/0.9/29.1	Het/De novo	Pathogenic
8	GNAO1	GNAO1:ENST00000262493:Exon 1:c.118G>C:p.Gly40Arg	Missense	NA	RCV000486165	D/-/D/0.98/27.1	Het/De novo	Pathogenic
9	PCDH19	ENST00000373034.4:Exon 1:c.593G>T:p.Arg198Leu	Missense	NA	RCV000519539	U/-/D/0.76/32	X-linked	Pathogenic
10	TUBA3E	ENST00000312988:Exon 4:c.392G>A:p.Gly131Glu	Missense	NA	NR	D/D/D/0.694/25.4	Het/De novo	Likely Pathogenic
11	Chr19p13.2 del	Chr19p13.2 deletion; 533.1KB	CNV	NA	RCV000141708.5	..	Het	Pathogenic
12	NOVA2	ENST00000263257:Exon 2:c.134G>C:p.Gly45Ala	Missense	NA	NR	D/D/D/0.527/25.7	Het/De novo	Likely Pathogenic
13	SCN1A	ENST00000423058:Exon 29:c.5339T>C:p.Met1780Thr	Missense	NA	RCV001854242	D/D/D/0.99/26.3	Het/De novo	Pathogenic
14	TSC2	ENST00000219476:Exon 9:c.826_827del:p.Met276Valfs*61	Frameshift deletion	NA	RCV000180518	..	Het/De novo	Pathogenic
15	KCNT1	ENST00000371757:Exon 15:c.1421G>A:p.Arg474His	Missense	NA	RCV000032795	U/D/D/0.77/33	Het/De novo	Pathogenic
16	GABRG2	ENST00000356592:Exon 7:c.917C>T:p.Ser306Phe	Missense	NA	RCV002047455	D/D/D/33/0.826	Het/De novo	Pathogenic
17	Chr13q32.3del	Chromosome 13q32.3deletion;19.7KB	CNV	NA	NR	..	De novo	Pathogenic
18	SLC2A1	ENST00000426263.3:Exon 4:c.458G>A:p.Arg153His	Missense	NA	RCV000178275	D/-/D/0.9/33	De novo	Likely pathogenic
19	PTPN11	ENST00000351677.2:Exon 13:c.1493G>T:p.Arg498Leu	Missense	NA	RCV000212896	D/U/D/0.94/33	De novo	Pathogenic
20	CHRNA2	ENST00000407991:Exon 6:c.859_860del:p.Phe287Leufs*29	Frameshift deletion	NA	NR	/-/D/-/31	De novo	Pathogenic
21	SCN1A	ENST00000423058:Exon 29:c.5293T>C:p.Phe1765Leu	Missense	NA	RCV001237218	U/-/D/0.98/29.8	Het/De novo	Pathogenic
22	CDKL5	ENST00000379996:Exon 8:c.517G>C:p.Alal173Pro	Missense	NA	NR	U/-/D/0.8/31	Het/De novo	Pathogenic
23	ARG1	ENST00000368087:Exon 8:c.923G>A:p.Arg308Gln	Missense	0.008	RCV000665488	U/D/D/0.83/29.3	Hom/Homozygous recessive	Pathogenic
24	TBL1XR1	ENST00000457928.2:Exon 11:c.974G>A:p.Cys323Tyr	Missense	NA	RCV000656402	D/-/D/0.91/33	Het/De novo	Likely pathogenic
25	PCDH19	ENST00000373034:Exon 1:c.815A>C:p.Gln272Pro	Missense	NA	NR	U/-/D/0.38/24.5	Het/De novo	Pathogenic
26	CYFIP2	ENST00000377576:Exon 11:c.1072C>T:p.Arg358Cys	Missense	NA	RCV001997354	U/-/D/0.47/33	Het/De novo	Pathogenic
27	CLN5	ENST00000377453:Exon 2:c.230G>A:p.Cys77Tyr	Missense	NA	RCV001039257	/D/D/0.95/29	Hom/Homozygous recessive	Pathogenic
28	SCN8A	ENST00000354534:Exon 22:c.3967G>A:p.Alal1323Thr	Missense	NA	RCV001253356	D/-/D/0.88/33	Het/De novo	Pathogenic
29	AKT3	ENST00000263826:Exon 14:c.1393C>T:p.Arg465Trp	Missense	NA	RCV000033035	U/-/D/0.4/31	Het/De novo	Pathogenic
30	NHLRC1	ENST00000340650:Exon 1:c.138C>G:p.Cys461Trp	Missense	NA	NR	U/-/D/0.92/24.5	Hom/Homozygous recessive	Pathogenic
31	GABRA1	ENST00000393943:Exon 9:c.934G>C:p.Alal12Pro	Missense	NA	NR	U/-/D/0.93/33	Het/De novo	Pathogenic
32	SCN1A	ENST00000423058:Exon 14:c.1724del:p.Phe575Serfs*48	Frameshift deletion	NA	RCV001972659	..	Het/De novo	Pathogenic
33	BTD	ENST00000383778:Exon 4:c.1503del:p.Arg502Glyfs*10	Frameshift deletion	0.0033	RCV000668694	..	Hom/Homozygous recessive	Pathogenic
34	TSC1	ENST00000298552:Exon 15:c.1960C>T:p.Gln654*	Non-sense	NA	RCV000698745	..	Het/De novo	Pathogenic
35	STXBP1	ENST00000373299:Exon 7:c.434A>C:p.Tyr145Ser	Missense	NA	NR	U/-/D/0.75/24	Het/De novo	Pathogenic

Sl no	Gene	Variant description	Functional consequence	Allele frequency in population database	ClinVar accession	Functional prediction scores- SIFT/PolyPhen/ MutationTaster/ REVEL/CADD_score)	Inheritance pattern	ACMG class
36	GABRB2	NM_001371727.1:Exon 8:c.949T>G;p.Phe317Val	Missense	NA	NR	U/-/D/0.91/29.4	Het/ De novo	Pathogenic
37	ITPA	ENST00000380113:Exon 8:c.511G>T;p.Glu171*	Non-sense	NA	NR	..	Hom/Homozygous recessive	Pathogenic
38	ALG3	ENST00000397676:Exon 1:c.28C>T;p.Arg10Trp	Missense	0.0062	NR	U/U/D/0.33/25.6	Het/ Compound heterozygous	Likely pathogenic
39	ALG3	ENST00000397676:Exon 1:c.16C>T;p.Arg6Trp	Missense	0.01	RCV001964864	U/D/D/0.52/27.4	Het/ Compound heterozygous	Likely pathogenic
40	SETD1B	NM_001353345.2:Exon 2:c.22dup;p.His8Profs*30	Frameshift insertion	0.02	RCV003388163	-/-/29.8	Het/ De novo	Pathogenic
41	SETD1A	ENST00000262519:Exon 7:c.1370G>A;p.Arg457Gln	Missense	0.01	RCV002652795	U/B/B/0.31/21.3	Het/ Autosomal dominant	Likely pathogenic
42	CLTC	ENST00000269122.3:Exon 29:c.4605+2T>C	splice site	NA	NR	/-/D/-/33	Het/De novo	Likely pathogenic
43	KCNT1	ENST00000371757:Exon 31:c.3691G>A;p.Asp1231Asn	Missense	NA	RCV000766050	U/U/D/0.24	Het/ De novo	Pathogenic
44	SCN3A	ENST00000283254:Exon 23:c.4054A>C;p.Met1352Leu	Missense	NA	NR	D/D/D/0.917/22.8	Het/ De novo	Pathogenic
45	SZT2	NM_001365999.1:Exon 2:c.137C>T;p.Thr46Ile	Missense	0.01	RCV001840730	B/B/B/0.03/18.69	Compound heterozygous	Likely pathogenic
46	SZT2	NM_001365999.1:Exon 55:c.7708G>A;p.Gly2570Arg	Missense	NA	RCV000869508	B/U/B/0.32/23.7	Compound heterozygous	Likely pathogenic
47	SCN1A	ENST00000423058:Exon 18:c.2836C>G;p.Arg946Gly	Missense	NA	NR	D/-/D/0.99/32	Het/De novo	Pathogenic
48	SCN2A	ENST00000375437:Exon 22:c.4231G>A;p.Gly1411Arg	Missense	NA	NR	D/D/D/0.97/26.5	Het/ De novo	Pathogenic
49	SCN1A	ENST00000423058:Exon 24:c.4168G>A;p.Val1390Met	Missense	NA	RCV000059411	D/D/D/0.74/23.1	Het/ De novo	Pathogenic
50	NR2F1	ENST00000327111:Exon 3:c.1117C>T;p.Arg373*	Non-sense	NA	RCV000985203	-/-/D/-/45	Het/ AD	Pathogenic
51	SCN1A	ENST00000423058:Exon 14:c.1738C>T;p.Arg580*	Non-sense	NA	RCV000180844	..	Het/ De novo	Pathogenic
52	KCNT1	ENST00000371757:Exon 25:c.2896G>A;p.Ala966Thr	Missense	NA	RCV000697406	U/-/D/0.86/23.9	Het/ AD	Pathogenic
53	UFSP2	ENST00000264689:Exon 5:c.344T>A;p.Val115Glu	Missense	0.012	RCV002281647	U/U/D/0.33/25.7	Hom/AR	Pathogenic
54	SCN1A	ENST00000423058:Exon 12:c.1171-1G>A	splice site	NA	RCV003159277	..	Het/ De novo	Pathogenic
55	SCN1A	ENST00000423058:Exon 15:c.2134C>T;p.Arg712*	Non-sense	NA	RCV000188886	..	Het/ De novo	Pathogenic
56	GNAO1	ENST00000262493:Exon 6:c.626G>A;p.Arg209His	Missense	NA	RCV000490633	D/-/D/0.91/34	Het/ De novo	Pathogenic
57	WDR45	ENST00000376372:Exon 9:c.827+1G>A	splice site	NA	RCV000578469	..	Het/ De novo	Pathogenic
58	SCN1A	ENST00000423058:Exon 29:c.5436G>C;p.Trp1812Cys	Missense	NA	RCV000658882	D/-/D/0.94/33	Het/ AD	Pathogenic
59	Chr13q14	Chr13q14.2-q14.3DEL	CNV	NA	RCV000136624.5	..	De novo	Likely pathogenic
60	ALDH18A1	ENST00000371224:Exon 17:c.2159T>C;p.Phe720Ser	Missense	NA	RCV001199068	D/-/D/0.91/31	Hom/AR	Pathogenic
61	CHD8	ENST00000399982:Exon 4:c.1261G>A;p.Val421Ile AND	Missense	0.0004	RCV003552414	U/U/D/0.48/26.9	Het/ De novo	Likely pathogenic
62	SCN1A	ENST00000423058:Exon 18:c.2799C>A;p.His933Gln	Missense	NA	RCV000479119	D/-/D/0.93/25.1	Het/ De novo	Pathogenic
63	KCNT1	ENST00000371757:Exon 20:c.2280C>G;p.Ile760Met	Missense	NA	RCV000413294	U/D/D/0.66/17.1	Het/ De novo	Pathogenic
64	MT-ATP6	ENST00000361899.2:Exon 1:c.256G>A;p.Gly86*	Non-sense	NA	RCV000854321	..	Het/ De novo	Pathogenic
65	SHQ1	ENST00000325599:Exon 7:c.850T>C;p.Tyr284His	Missense	0.0048	NR	D/D/D/0.8/26.7	Hom/AR	Pathogenic
66	SETD1B	ENST00000542440:Exon 6:c.931C>T;p.Arg311Trp	Missense	0.005	RCV003136602	D/D/D/0.67/23.4	Het/ AD	Likely Pathogenic
67	KCTD7	ENST00000275532:Exon 3:c.458G>A;p.Arg153His	Missense	0.0039	RCV000680019	U/D/D/0.64/28.7	Compound heterozygous	Pathogenic
68	KCTD7	ENST00000275532:Exon 4:c.533C>T;p.Ala178Val	Missense	0.0021	RCV000677304	D/D/D/0.51/27.1	Compound heterozygous	Pathogenic
69	SCN2A	ENST00000375437:Exon 5:c.585T>A;p.Asp195Glu	Missense	NA	NR	D/-/D/0.94/24.4	Het/De novo	Pathogenic
70	NPRL2	ENST00000232501:Exon 7:c.689A>G;p.Tyr230Cys	Missense	NA	NR	U/-/D/0.65/26.5	Het/De novo	Pathogenic
71	ABCA2	ENST00000341511.6:Exon 14:c.1826T>C;p.Ile609Thr	Missense	NA	NR	U/-/D/0.77/22.7	Hom/Homozygous recessive	Likely Pathogenic

The identification of pathogenic or likely pathogenic variants was more prevalent in trios (43%) than in duos (16.7%) and singletons (18.2%), whereas the occurrence of VUSs was greater in duos (50%) and singletons (45.5%) compared to trios (32.3%) (**Figure 4.4**).

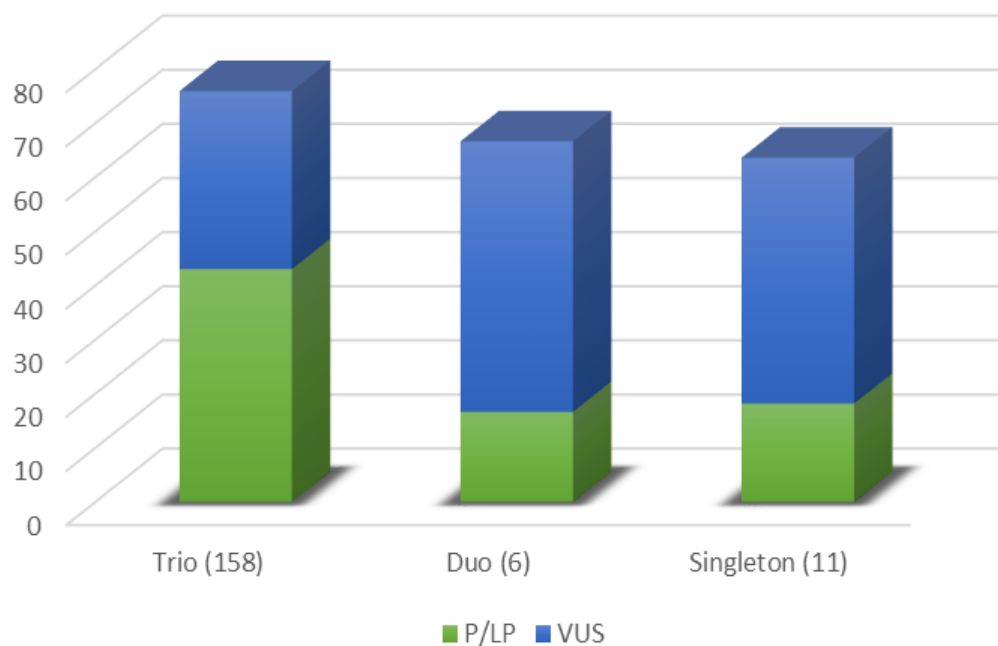


Figure 4.4: Diagnostic yield comparison between trios, duos and singletons

The detection rate of P/ LP variants was influenced by several factors. The comparison of diagnostic yield and clinical characteristics including Epilepsy syndromes, inheritance pattern, age of onset of epilepsy, family history, MRI, type of seizures, etc. are summarised in **Table 4.3**.

The most common mode of inheritance identified was *de novo* , accounting for 45 out of 71 cases (63%). This was followed by autosomal recessive at 16% (11/71), inherited autosomal dominant also at 10% (7/71), compound heterozygous at 5.6% (4/71),

autosomal dominant with unknown inheritance at 4% (3/71), and x-linked inheritance at 1.4% (1/71) (**Figure 4.5**).

Diagnostic yield varied between epilepsy syndromes (**Figure 4.6**). 1 out of 1 (100%) for Ohtahara syndrome, 13 of 23 (56.5%) for EIEE less than 6 months of age of onset, 9 of 23 (39.1%) for West syndrome, 9 of 10 (90%) for Dravet syndrome (p-value = 0.001; **Table 4.3**), 4 of 12 (33.3%) for Lennox-Gastaut Syndrome, 2 of 2 (100%) for Epilepsy of infancy with migrating focal seizures, 7 of 20 (35%) for Landau Kleffner syndrome/EE with spike and wave activation in sleep, 10 of 45 (22.2%) for focal/multifocal epilepsy, 11 of 29 (37.9%) for refractory generalized epilepsy, 2 of 5 (40%) for myoclonic atonic epilepsy, and 3 of 5 (60%) for progressive myoclonic epilepsy.

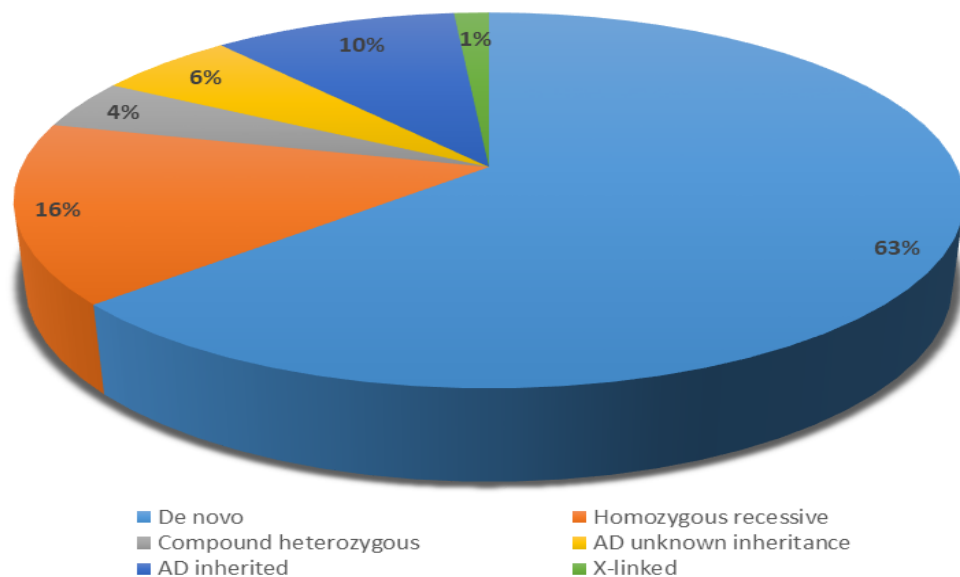


Figure 4.5: Inheritance pattern of P/LP variants

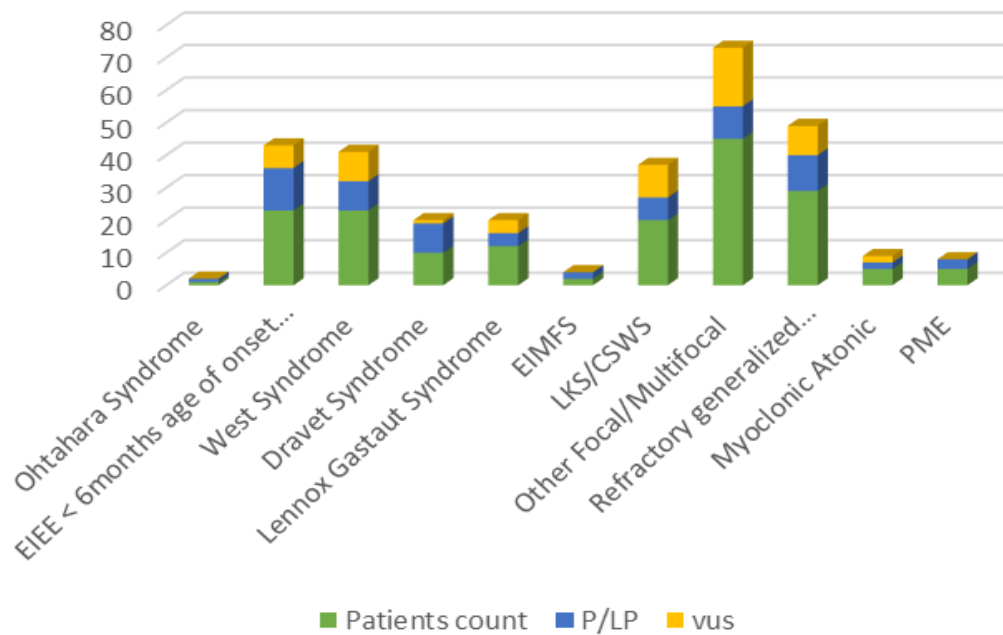


Figure 4.6: Genetic diagnosis of epilepsy syndromes

Pathogenic variants were more likely to be found in children with earlier seizure onset. A significant difference in diagnostic yield was found between patients with an age of onset of epilepsy under 3 years and those whose onset occurred after 3 years, with a p-value of 0.041 from Fisher's exact test (**Table 4.3**). The genetic aetiology was identified in 60 out of 134 (44.8%) probands with seizure onset under the age of 3 years compared with 11 out of 41 (26.8%) with seizure onset over the age of 3 years (**Figure 4.7**). The analysis of correlation between the age of onset and diagnostic yield revealed a negative relationship, suggesting that an increase in age of onset is associated with a decline in diagnostic yield (**Figure 4.8**)

Table 4.3: Clinical characteristics and diagnostic yield

Clinical characteristics	Genetic diagnosis, <i>n</i> =71	No genetic diagnosis, <i>n</i> =104	P-value (Fisher's exact/two-tailed Chi square)
Sex, N (%)			0.109
Male	35 (35.4)	64 (64.6)	..
Female	36 (47.4)	40 (52.6)	..
Age at seizure onset in months, N (%)			0.041
Age of onset < 3 years	60 (44.8)	74 (55.2)	..
Age of onset > 3 years	11 (26.8)	30 (73.2)	..
Global developmental delay, N (%)	44 (47.3)	49 (52.7)	0.053
Autism, N (%)	8 (36.4)	14 (63.6)	0.667
Seizure type			
Focal	10 (22.2)	35 (77.8)	0.004
Generalized	16 (41)	24 (59)	0.933
Combined	45 (49.5)	46 (50.5)	0.013
Family history, N (%)			0.593
Yes	35 (42.7)	47 (57.3)	..
No	36 (38.7)	57 (61.3)	..
Brain MRI findings			0.368
Normal	34 (37.4)	57 (62.6)	..
Abnormal	37 (44)	47 (56)	..
Epilepsy phenotype, N (%)			
1. Ohtahara syndrome	1 (100)	0	0.406
2. Early Infantile epileptic encephalopathy (< 6months of age of onset with focal or multifocal epilepsy)	13 (56.5)	10 (43.5)	0.113
3. West syndrome	9 (39.1)	14 (60.9)	0.88
4. Dravet syndrome	9 (90)	1 (10)	0.001
5. Myoclonic atonic epilepsy	2 (40)	3 (60)	1
6. Lennox gastaut syndrome	4 (33.3)	8 (66.7)	0.764
7. Epilepsy of infancy migrating focal seizures	2 (100)	0	0.163
8. Landau kleffner syndrome/Continuous spike and wave during slow wave sleep	7 (35)	13 (65)	0.59
9. Other focal/multifocal epilepsy	10 (22.2)	35 (77.8)	0.008
10. Refractory generalized epilepsy	11 (37.9)	18 (62.1)	0.751
11. Progressive myoclonic epilepsy	3 (60)	2 (40)	0.397
Seizure in remission on last follow-up	21 (35)	39 (65)	0.278

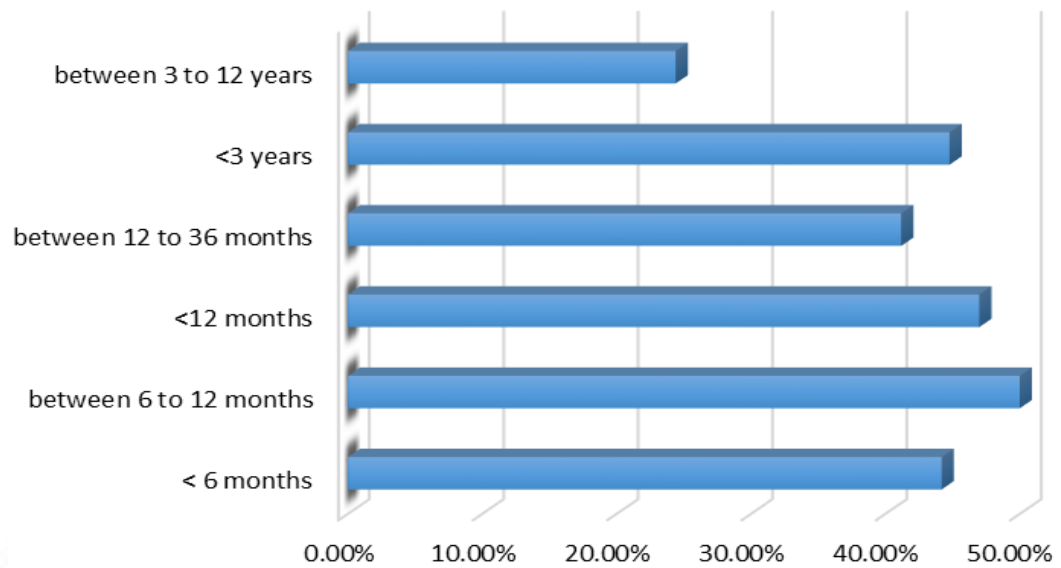


Figure 4.7 : Comparison of diagnostic yield (in %) and age of onset of epilepsy

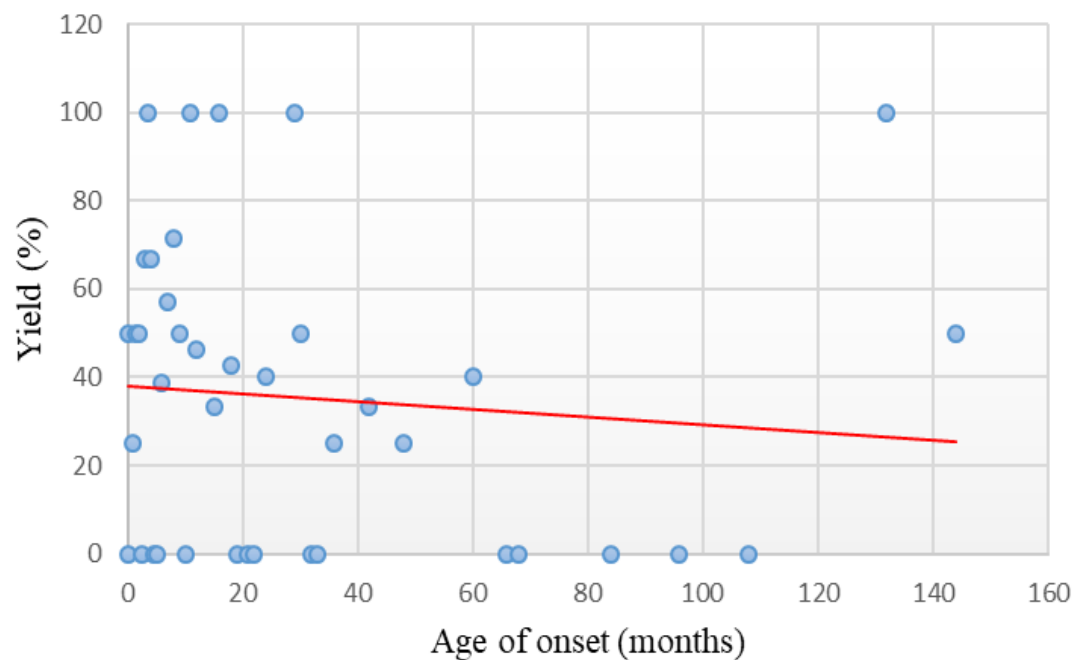


Figure 4.8: Correlation analysis between age of onset and diagnostic yield

Causative genes exhibit a variety of cellular functions including transcription regulation (*KMT2C* (*n*=1), *TBL1XR1* (*n*=1), *SETD1B* (*n*=3), *SETD1A* (*n*=1), *NR2F1* (*n*=1), *CHD8* (*n*=1)), transporters (*SLC2A1* (*n*=2), *ABCA2* (*n*=1)), ion channel subunits (*CHRNA4* (*n*=1), *SCN8A* (*n*=2), *SCN1A* (*n*=13), *GABRG2* (*n*=1), *CHRNA2* (*n*=1), *GABRA1* (*n*=1), *GABRB2* (*n*=1), *KCNT1* (*n*=4), *SCN3A* (*n*=1), *KCTD7* (*n*=1), *SCN2A* (*n*=2)), mechanistic target of rapamycin (mTOR) pathway (*TSC1* (*n*=1), *TSC2* (*n*=1), *NPRL2* (*n*=1)), metabolic pathways (*ARG1* (*n*=1), *BTD* (*n*=1), *STXBP1* (*n*=1), *ITPA* (*n*=1), *ALG3* (*n*=1), *ALDH18A1* (*n*=1), *MT-ATP6* (*n*=1)), G-protein transduction (*GNAO1* (*n*=2)) and others (*C12ORF57* (*n*=1), *DCX* (*n*=1), *TUBA3E* (*n*=1), *NOVA2* (*n*=1), *PTPN11*, *CDKL5* (*n*=1), *PCDH19* (*n*=2), *CYFIP2* (*n*=1), *CLN5* (*n*=1), *AKT3* (*n*=1), *NHLRC1* (*n*=2), *CLTC* (*n*=1), *SZT2* (*n*=1), *UFSP2* (*n*=2), *WDR45* (*n*=1), *SHQ1* (*n*=1)). (Figure 4.9)

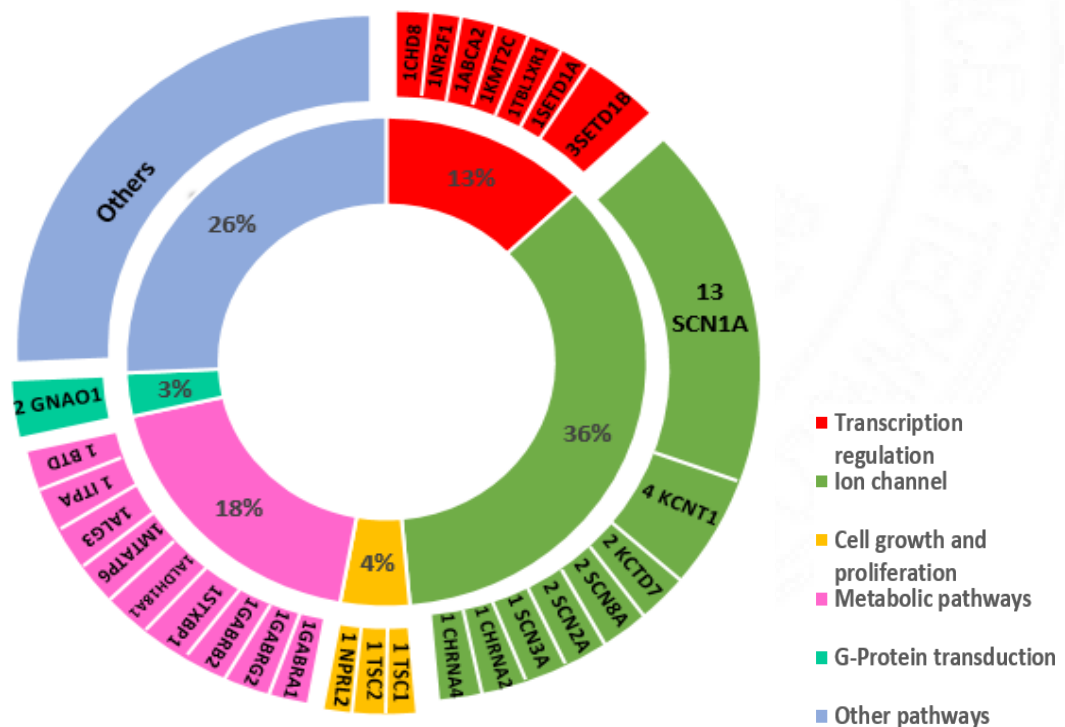


Figure 4.9 : Functional mechanisms of pathogenic variants identified in our cohort

4.1.3 Epilepsy syndromes and associated genes

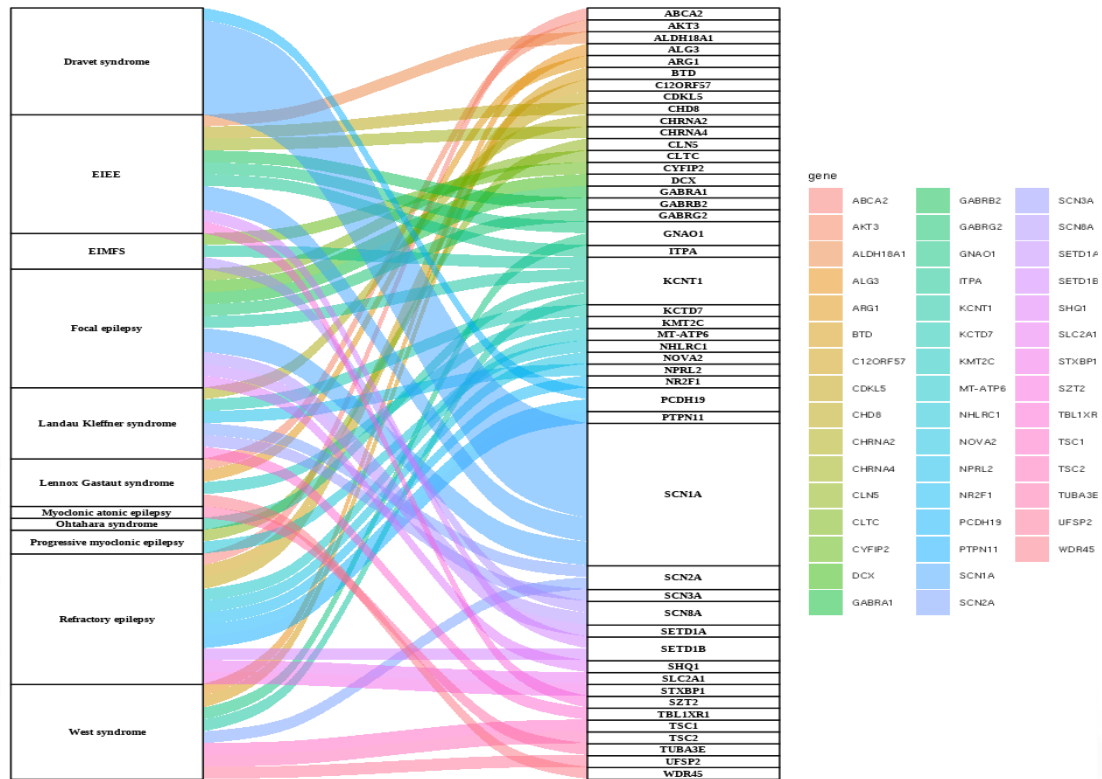


Figure 4.10 : Epilepsy syndromes and associated disease causing genes

The genes associated with various types of epilepsy syndromes are illustrated in the **Figure 4.10**. A significant number of *SCN1A* variants has been linked to Dravet syndrome and EIEE followed by *KCNT1* variants associated with EIMFS, focal epilepsy, Ohtahara syndrome and West syndrome. Multiple genes are involved in developmental and epileptic encephalopathies including West syndrome, early infantile epileptic encephalopathy, Lennox-Gastaut syndrome, and Landau-Kleffner syndrome.

4.1.4 Copy number variants

Causative CNVs were identified in 3 patients (4.22%, 3/71) (**Table 4.4**). These include deletions in chromosome 13 (Chr13q14.3 and Chr13q32.3) and chromosome 19 (Chr19p13.2). Chromosome 13q deletion was detected in a patient diagnosed with West syndrome, as well as in another patient suffering from myoclonic atonic epilepsy. Additionally, a deletion of chromosome 19 was identified in a patient with Landau-Kleffner syndrome, characterized by continuous spike and wave activity during slow-wave sleep.

Table 4.4: List of P/LP copy number variants identified

Sl No.	CNV	Position	Genes	Diagnosis	Inheritance	ACMG classification
1	Chr 13q14.3 deletion	DEL:CHR13:50121284-52333905; 2.21 MB	RCBTB1, ARL11, EBPL, KPNA3, SPRYD7, TRIM13, KCNRG, DLEU7, RNASEH2B, C13orf42, FAM124A, SERPINE3, INTS6, WDFY2	West syndrome	Het/ unknown	Pathogenic
2	Chr 13q32.3 deletion	DEL:chr13-100617645-100637318; 19.7kb	ZIC5, ZIC2	Myoclonic atonic epilepsy	De novo	Likely pathogenic
3	Chr 19p13.2 deletion	Chr19:10597370 - 1113042711,244,471;977.4kb	KEAP1, S1PR5, SLC4A2, CARM1, YIPF2, SMARCA4	LKS/CSWS	De novo	Pathogenic

4.1.5 Precision medicine

Among the 71 patients identified with causative variants, precision medicine was applied in 25 (35.2%) cases, summarized in **Table 4.5**.

Table 4.5: Precision medicine applied

Sl. No	Gene	No. of patients	Diagnosis	Precision medicine applied
1	<i>SCN1A</i>	12	EIEE, Focal epilepsy, <u>Dravet syndrome</u>	Avoid sodium channel blockers, <u>Stiripentol</u> , CBD oil, <u>Fenfluramin</u>
2	<i>SCN2A</i>	2	LKS/CSWS, West syndrome	Avoid/Provide Sodium channel blockers according to prognosis
3	<i>SCN8A</i>	2	EIEE, focal/multifocal epilepsy	Sodium channel blockers
4	<i>KCNT1</i>	4	OS, focal/multifocal epilepsy, West syndrome, EIMFS	Potassium channel openers
5	<i>TSC1</i>	1	West syndrome	<u>Everolimus</u>
6	<i>TSC2</i>	1	West syndrome	<u>Everolimus</u>
7	<i>ARG1</i>	1	LGS	Avoid Valproate and enzyme replacement
8	<i>BTBD</i>	1	West syndrome	Biotin supplementation
9	<i>SLC2A1</i>	1	Myoclonic epilepsy	Ketogenic Diet

4.1.6 Sanger sequencing validation

Disease causing variants identified were validated using Sanger sequencing. The representative electropherogram images of a proband identified with a *de novo* variant in *CHRNA4* (c.1105C>T) are illustrated in **Figure 4.11**. The variant was identified as heterozygous in child and homozygous for the wild type allele in parents.

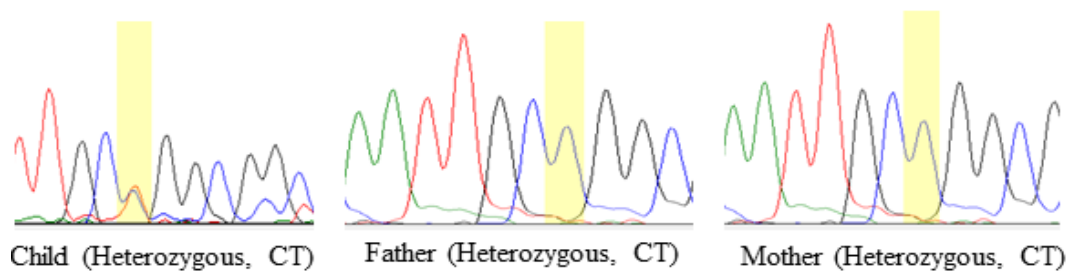


Figure 4.11: Representative Sanger sequencing validation image of a *de novo* variant CHRNA4 (c.1105C>T)

4.1.7 IGV visualization

The IGV visualization image of a *de novo* variant in WDR45 (c.827+1G>A) gene is illustrated in **Figure 4.12**. Child was heterozygous for the mutant allele, CT and parents were homozygous for the wild type allele, CC.

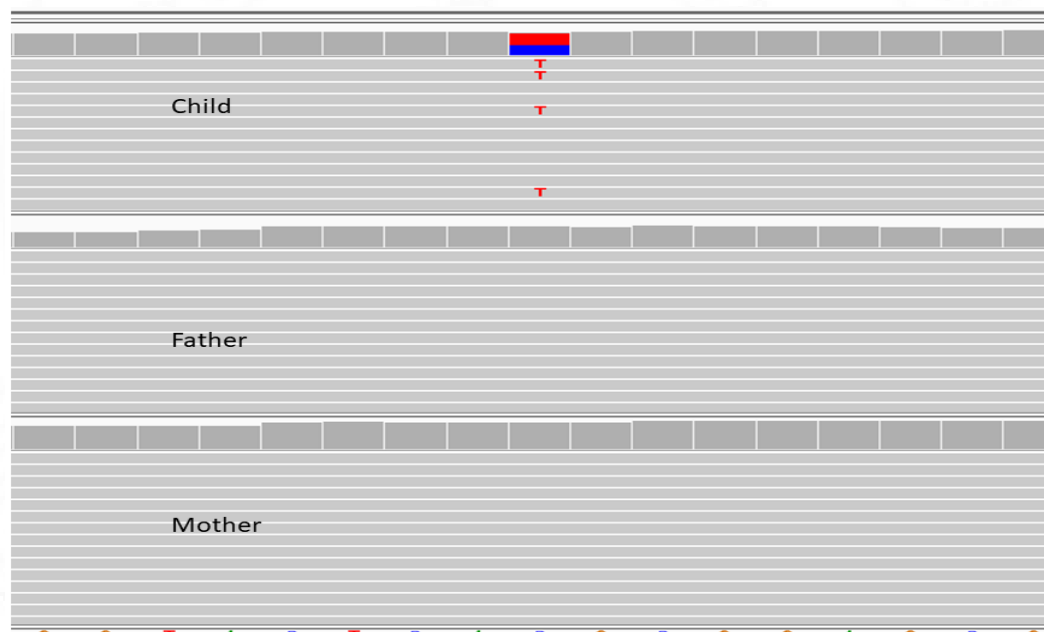


Figure 4.12 : Representative IGV visualization image of a *de novo* variant in WDR45 (c.827+1G>A)

4.2 Variants of uncertain significance

Table 4.6 : List of VUS

Gene	Variant description	Functional predictions- SIFT/Polyphen/MutationTaster/CADD/Revel	Inheritance	Diagnosis
WASHC4	ENST00000332180.5Exon 29:c.2935G>A:p.Asp979Asn	T/B/D/22.4/0.266	Compound het	West syndrome
WASHC4	ENST00000332180.5Exon 31:c.3188A>C:p.Gln1063Pro	D/D/D/29.1/0.761	Compound het	West syndrome
CACNA1E	ENST00000367573.2Exon 20:c.2732G>A:p.Arg911Gln	U/U/D/22.4/0.36	AD	West syndrome
GCH1	NM_000161.3Exon 5:c.610G>A:p.Val204Ile	D/D/D/24.5/0.669	AD	Focal/multifocal epilepsy
KMT2E	ENST00000311117.3Exon 27:c.4952C>G:p.Ser1651Cys	T/B/D/22.6/0.292	AD	Focal/multifocal epilepsy
CPT2	ENST00000371486.3Exon 4:c.692G>A:p.Arg231Gln	T/D/N/20.9/0.297	AD	Focal/multifocal epilepsy
DNM1L	ENST00000549701.1Exon 17:c.1849C>T:p.Pro617Ser	T/B/D/23.3/0.359	AD	Refractory generalized epilepsy
USP9X	ENST00000378308.2Exon 16:c.2219T>G:p.Phe740Cys	D/D/D/28.5/0.773	X-linked	LGS
PLXNA3	ENST00000369682.3Exon 2:c.203C>T:p.Thr68Met	D/-/D/24.8/0.409	Compound het	Focal/multifocal epilepsy
PLXNA3	ENST00000369682.3Exon 3:c.856G>A:p.Gly286Ser	T/-/D/22.5/0.328	Compound het	Focal/multifocal epilepsy
NDUFS2	NM_001377299.1Exon 13:c.1324C>T:p.His442Tyr	D/D/D/24.9/0.702	De novo	Focal/multifocal epilepsy
RNF112	ENST00000461366.1Exon 2:c.59G>T:p.Arg20Ile	D/D/D/24.7	De novo	Focal/multifocal epilepsy
SYNE2	ENST00000358025.3Exon 6:c.323T>C:p.Leu108Pro	D/D/D/24.4/0.353	De novo	Focal/multifocal epilepsy
STXBP1	ENST00000373299.1Exon 5:c.281C>T:p.Pro94Leu	AD	AD	LKS/CSWS
TSPYL2	ENST00000375442.4Exon 1:c.130C>G:p.Gln44Glu	D/D/N/23/0.091	X-linked	West syndrome
UBE3A	NM_130839.5Exon 8:c.1905C>G:p.Tyr635*	/./D/40/.	De novo	Focal/multifocal epilepsy
ADGRV1	ENST00000405460.2Exon 77:c.16544T>C:p.Leu5515Ser	T/D/D/23.4/0.127	Compound het	Myoclonic epilepsy
ADGRV1	ENST00000405460.2Exon 77:c.16544T>C:p.Leu5515Ser	T/B/N/8.67/0.04	Compound het	Myoclonic epilepsy
POLA1	NM_001330360.2Exon 20:c.2137C>A:p.His713Asn	T/B/D/21.4/0.076	X-linked	West syndrome
BRD4	NM_001379291.1Exon 8:c.1445C>T:p.Pro482Leu	D/D/D/25.8/0.071	De novo	Myoclonic atonic epilepsy
KCNT1	ENST00000371757.2Exon 31:c.3671C>T:p.Ser1224Phe	D/-/N/24.9/0.069	AD, inheritance unknown	EIEE
BSN	ENST00000296452.4Exon 5:c.4630C>T:p.Arg1544Cys	D/D/D/25/0.174	AD, CH	EIEE
EFHC1	ENST00000371068.5Exon 4:c.661C>T:p.Arg221Cys	D/D/D/33/0.369	Compound het	EIEE
EFHC1	ENST00000371068.5Exon 11:c.1892A>G:p.Tyr631Cys	D/D/D/27.2/0.413	Compound het	EIEE
PLXNB1	ENST00000296440.6Exon 37:c.6238G>A:p.Gly2080Arg	D/D/D/33/0.247	AD	EIEE
ACTL6B	ENST00000160382.5Exon 2:c.27T>G:p.Asp9Glu	D/D/D/25.3/0.726	De novo	LKS/CSWS
MT-CYTB	ENST00000361789.2Exon 1:c.899T>C:p.Ile300Thr	..	De novo	Focal/multifocal epilepsy
SETD1B	NM_001353345.2Exon 15:c.5585G>A:p.Arg1862His	D/-/D/23.8/0.721	AD, inheritance unknown	Refractory generalized epilepsy
Chr13q12.11 Dup	CNV	..	AD, inheritance unknown	West syndrome
CACNA1A	ENST00000360228.5Exon 43:c.6250A>G:p.Met2084Val	B/-/B/17.7/0.18	AD	Focal/multifocal epilepsy
BRWD3	ENST00000373275.4Exon 38:c.4362C>A:p.Ser1454Arg	D/D/D/25.2/0.232	X-linked	Myoclonic atonic epilepsy
Chr 22q11.23	CNV	..	De novo	Refractory generalized epilepsy
SCN9A	NM_001365536.1Exon 16:c.2647G>T:p.Val883Phe	D/-/D-32/0.924	AD	Refractory generalized epilepsy
Chr 4p16.3 dup	CNV	..	De novo	Refractory generalized epilepsy
ATRX	ENST00000373344.5Exon 9:c.3359A>T:p.Asp1120Val	D/B/D/22.3/0.303	X-linked	EIEE
AP2M1	ENST00000292807.5Exon 3:c.327T>G:p.Tyr109*	/-/D/35/-	De novo	EIEE
WDR62	ENST00000401500.2Exon 28:c.3407G>A:p.Arg1136Gln	T/B/N/-/0.005	Compound het	LKS/CSWS
WDR62	ENST00000401500.2Exon 22:c.2565G>T:p.Lys855Asn	T/D/N/23.3/0.115	Compound het	LKS/CSWS
SETD1B	NM_001353345.2Exon 12:c.4243C>A:p.Pro1415Thr	D/-/D/23/0.228	AD	Focal/multifocal epilepsy

Gene	Variant description	Functional predictions- SIFT/Polyphen/MutationTaster/CADD/Revel	Inheritance	Diagnosis
USP9X	ENST00000378308.2Exon 23:c.3409C>A.p.Pro1137Thr	D/D/D/24.6/0.277	X-Linked	Focal/multifocal epilepsy
TCF20	NM_001378418.1Exon 2:c.4810A>G.p.Ile1604Val	D/D/D/22.9/0.201	AD	Focal/multifocal epilepsy
HIVEP2	ENST00000367603.2Exon 5:c.2269C>T.p.Arg757Cys	D/D/D/28.1/0.145	AD	LKS/CSWS
AP2M1	ENST00000292807Exon 3:c.327T>G.p.Tyr109*	/-/D/-/35	De novo	EIEE
GAS2L2	NM_139285.4Exon 3:c.730A>G.p.Ile244Val	D/D/D/20.4/0.182	De novo	Focal/multifocal epilepsy
PPP3CA	ENST00000394854.3Exon 2:c.83G>A.p.Arg28Gln	D/D/D/24.2/0.215	AD, inheritance unknown	LKS/CSWS
MED13L	ENST00000281928.3Exon 7:c.833T>C.p.Met278Thr	D/D/D/23.8/0.86	AD, inheritance unknown	LKS/CSWS
ARFGEF1	ENST00000262215.3Exon 37:c.5220T>G.p.Asp1740Glu	T/D/D/20.9/0.443	De novo	LKS/CSWS
Chr 17q12dup	CNV	..	De novo	Focal/multifocal epilepsy
DNM1L	ENST00000549701.1Exon 2:c.163A>G.p.Thr55Ala	D/B/D/22.3/0.631	AD	West syndrome
SPTBN1	ENST00000356805.4Exon 33:c.6460G>A.p.Ala2154Thr	T/B/D/20.9/0.107	AD	West syndrome
SYNGAP1	NM_006772.3Exon 19:c.3956C>G.p.Ala1319Gly	D/D/D/22.8/0.128	De novo	LGS
SLC4A10	ENST00000446997.1Exon 1:c.43C>T.p.Pro15Ser	T/B/D/22.4/0.272	AD	LGS
Chr 8p23.1 del	CNV	..	Inheritance unknown	Refractory generalized epilepsy
CHRNA7	ENST00000306901.3Exon 1:c.5G>A.p.Arg2His	D/B/D/23.4/0.175	De novo	Focal/multifocal epilepsy
Chr 2q11.2dup	CNV	..	De novo	Focal/multifocal epilepsy
STXBP1	ENST00000373299.1Exon 18:c.1702+1843A>C	D/B/D/23/0.26	De novo	Focal/multifocal epilepsy
GABBR2	ENST00000259455.2Exon 17:c.2503A>C.p.Asn835His	D/D/D/26/0.461	De novo	Focal/multifocal epilepsy
STARD9	ENST00000290607.7Exon 23:c.7216G>A.p.Ala2406Thr	D/-/N/16.7/0.045	Compound het	Refractory generalized epilepsy
STARD9	ENST00000290607.7Exon 8:c.572A>G.p.His191Arg	D/-/N/23.1/0.340	Compound het	Refractory generalized epilepsy
UBR5	ENST00000520539.1Exon 20:c.2530C>G.p.Leu844Val	T/D/D/24/0.326	AD, CH	Refractory generalized epilepsy
CIC	NM_001386298.1Exon 11:c.4394C>A.p.Thr1465Asn	D/B/N/21.8/0.025	AD	EIEE
MBD5	ENST00000407073.1Exon 12:c.3539A>G.p.Asp1180Gly	D/D/D/23.5/0.223	AD	EIEE
BSN	ENST00000296452.4Exon 6:c.10658A>G.p.Tyr3553Cys	D/D/D/29.7/0.492	Compound het	Focal/multifocal epilepsy
BSN	ENST00000296452.4Exon 6:c.9563G>T.p.Gly3188Val	D/B/D/21.6/0.084	Compound het	Focal/multifocal epilepsy
Chr 22q13.1dup	CNV	..	De novo	Refractory generalized epilepsy
TET3	ENST00000409262.3Exon 12:c.5354C>T.p.Thr1785Met	/-/D/29.8/0.539	AD	West syndrome
MBD5	NM_001378120.1Exon 9:c.3232G>T.p.Gly1078Trp	..	AD	West syndrome
RYR3	NM_001036.6Exon 52:c.7862A>T.p.Tyr2621Phe	T/B/D/21.4/0.365	AD	West syndrome
CHRNA2	ENST00000407991.1Exon 6:c.859_860del:p.Phe287Leufs*29	/-/31/-	AD	LKS/CSWS
GATAD2B	ENST00000368655.4Exon 2:c.68A>C.p.Glu23Ala	T/D/D/24.5/0.202	AD, inheritance unknown	West syndrome
CAMK2B	ENST00000395749.2Exon 19:c.1399G>A.p.Glu467Lys	T/B/D/23.1/0.075	AD	LKS/CSWS
DNM1L	ENST00000549701.1Exon 15:c.1654_1659del:p.Ser552_Ala553del	/-/20.9/-	AD	Focal/multifocal epilepsy
PLXNA3	ENST00000369682.3Exon 2:c.409C>T.p.Arg137Cys	T/-/D/33/0.313	X-linked, inheritance unknown	LGS
TWNK	ENST00000311916.2Exon 2:c.1357C>T.p.Arg453Trp	D/D/D/31/0.756	AD	LKS/CSWS
DLL1	ENST00000366756.3Exon 7:c.950G>A.p.Arg317Gln	T/P/D/21.3/0.225	AD	LKS/CSWS
KCNQ2	ENST00000359125.2Exon 14:c.1559A>G.p.Asp520Gly	D/P/D/23.5/0.791	AD	LKS/CSWS
KCNQ5	ENST00000370398.1Exon 14:c.2775C>A.p.Ser925Arg	D/B/D/23.5/0.494	AD	Dravet syndrome
PAK3	ENST00000372007.5Exon 10:c.715C>T.p.Arg239Trp	D/D/D/32/0.566	X-linked	Focal/multifocal epilepsy
KMT2D	ENST00000301067.7Exon 25:c.5477G>T.p.Gly1826Val	D/D/D/23.7/0.379	AD, inheritance unknown	EIEE
TCF20	NM_001378418.1Exon 2:c.2483C>T.p.Ala828Val	D/D/D/23.5/0.252	AD, inheritance unknown	EIEE
BSN	ENST00000296452.4Exon 2:c.440A>G.p.Asp147Gly	D/B/D/23.3/0.087	Compound het	LGS
BSN	ENST00000296452.4Exon 7:c.11284C>A.p.Pro3762Thr	D/B/N/21.1/0.066	Compound het	LGS
MIB1	ENST00000261537.6Exon 11:c.1576C>T.p.Arg526*	/-/39/-	AD	Refractory generalized epilepsy
RYR3	NM_001036.6Exon 14:c.1471C>T.p.Arg491Cys	D/D/D/35/0.865	AD	Refractory generalized epilepsy

Out of 175, 60 patients were identified with variants of uncertain significance. The variant description, inheritance pattern and associated phenotypes are provided in **Table 4.6**. The predominant inheritance pattern observed for VUSs was autosomal dominant (38.3%) followed by *de novo* variants (26.7%). The majority of the identified VUSs were linked to focal or multifocal epilepsy (30%), with LKS/CSWS (16.7%) being the next most common association.

To the best of our knowledge, this is the first study from India in a large cohort of epileptic encephalopathies studying the genetic diagnostic yield of WES in Trios. Exome sequencing is revolutionizing the diagnostic process for human diseases, particularly in the context of severe epilepsies referred to as DEEs. The feasibility of whole exome sequencing in trios affected by developmental and epileptic encephalopathy is demonstrated, showcasing a high diagnostic yield and significant clinical impact. Our results advocate for the implementation of trio-WES as a means to enable timely aetiological diagnosis, which can guide immediate and targeted interventions.

We achieved a diagnostic yield of 40.6% and 34.3% of the patients had VUS, and 25.1% had no potentially disease causing variants identified. 95.8% of the disease causing variants identified were single nucleotide variants or small insertions or deletions, and 4.2% were copy number variants. The diagnostic yield we observed falls within the expected range reported by similar studies (Balciuniene et al., 2019; Scheffer et al., 2023; Sun et al., 2021).

The diagnostic yield for DEEs is subject to significant variability, influenced by a range of factors including epilepsy syndromes, inheritance pattern, age of onset of

epilepsy, family history, MRI, and type of seizures. A comprehensive understanding of these factors is essential for optimizing recruitment strategies and study designs aimed at enhancing diagnostic yield. Our findings indicate a marginally higher diagnostic yield in trio studies (43%) compared to singleton studies (18.2%), which is consistent with earlier studies highlighting the advantages of utilizing trio designs (The 100,000 Genomes Project Pilot Investigators, 2021; Wright et al., 2022).

Similar to previous studies, genes associated with ion channels (40%), particularly sodium and potassium ion channels, represented a significant portion of the identified pathogenic and likely pathogenic variants (Sun et al., 2021). The most common variants were missense mutations (73%), followed by nonsense mutations (10%), frameshift insertions and deletions (7%), splice site alterations (4%), and copy number variations (4%). The predominant factor contributing to DEE in our cohort was identified as *de novo* point mutations, accounting for 45 out of 71 resolved cases. In contrast, among 71 resolved cases, the remaining cases were inherited mutations in 26, autosomal dominant with unknown inheritance in 3 and *de novo* copy number variations in 3 out of 71 resolved cases. Notably, *de novo* missense variants represented a greater percentage of individuals with DEE in our study. This is consistent with the previous large cohort study in DEE using whole genome sequencing (Hamdan et al., 2017b) and whole exome sequencing (Jiang et al., 2021; Sedlackova et al., 2024; Vetri et al., 2024).

No statistically significant differences were observed in brain MRI findings, family history, age, and sex when comparing subjects with identified variants to those with unidentified variants. A significant difference in diagnostic yield was observed

between patients with an age of onset of epilepsy under 3 years and those whose onset occurred after 3 years. The reduction in the diagnostic yield of genetic testing as the age of epilepsy onset increases is consistent with findings reported in prior literature (Cavirani et al., 2024; Hunter et al., 2020). Statistical significance was noted among patients experiencing focal seizures (p value- 0.004) as well as those with combined focal and generalized seizures (p value- 0.013) (**Table 4.3**). In patients diagnosed with focal epilepsy, a weak significance was noted (p value-0.008), indicating that the diagnostic yield was lower for this group. The diagnostic yield of next-generation sequencing, whether through gene panels or whole exome sequencing, in individuals with focal epilepsy ranged from 16% to 29% (Perucca, 2018). Conversely, a substantial diagnostic yield was attained for patients with Dravet syndrome (p value- 0.001).

The integration of CNV analysis from exome sequencing enhances the thoroughness of testing and improves cost-effectiveness by eliminating the need for chromosomal microarray (CMA) testing. The exome-based CNV algorithm utilized in this study detected pathogenic variants in 3 out of 71 positive cases (4.2%), contributing an additional 1.7% (3 out of 175) to the overall diagnostic yield. These findings are in agreement with the recent findings (Balciuniene et al., 2019; Scheffer et al., 2023).

Following a thorough diagnostic assessment and a re-examination of clinical and genetic reports, as well as the classification of variants, it was found that the identified genetic variations exhibited an uncertain clinical significance in 34.3% of the cases. This percentage is similar that reported in a recently published pediatric cohort, where 36 out 103 the tested patients had variant of uncertain significance identified through

WES analysis (Scheffer et al., 2023). In a very recent study involving a pediatric cohort, it was found that 16.4% of the patients tested exhibited at least one variant of uncertain significance identified through chromosomal microarray analysis, while 41.9% were detected using next-generation sequencing of a panel focused on epilepsy-related genes (Burk et al., 2024).

Epileptic encephalopathies exhibit significant genetic and phenotypic diversity. Distinct epilepsy phenotypes, including West syndrome, Lennox-Gastaut syndrome, and Dravet syndrome, may arise from mutations in different genes, while alterations in a specific gene can lead to a range of phenotypic expressions (Helbig and Abou Tayoun, 2016). Understanding the correlations between genotype and phenotype in relation to individual genes and their associated age-related epileptic encephalopathies is crucial for elucidating the specific cellular mechanisms that contribute to various epilepsy syndromes (Syrbe, 2022). Factors contributing to phenotypic pleiotropy may encompass the nature and timing of mutations, which can include somatic mosaicism or genomic rearrangements; the specific localization of mutations within the protein; mechanisms of loss- or gain-of-function resulting from these mutations; epigenetic influences; and the presence of modifier genes (Ko et al., 2018; Kobow et al., 2020). Nevertheless, individuals with different pathogenic genes often exhibit certain shared characteristics, such as the timing of symptom onset and the classification of epilepsy syndromes, as previously noted. This phenomenon is particularly pronounced in specific instances, such as Dravet syndrome, where a significant majority of patients (88.9% of those genetically diagnosed) had *SCN1A* identified as the causative gene.

In our study group, the genes linked to DEE and epilepsy were implicated in several

fundamental biological mechanisms, including ion transport, mTOR signalling, modulation of synaptic processes, transport and metabolism of small molecules, and the regulation of gene transcription and translation.

The disease causing variants identified in ion transport genes include *SCN1A*, *SCN2A*, *SCN8A*, *SCN3A*, *KCNT1*, *KCTD7*, *CHRNA4*, *CHRNA2*, *GABRG2*, *GABRB2*, and *GABRA1*. Ion channels play a vital role in the generation and regulation of excitability within the nervous system and have been associated with the occurrence of human epilepsy. Each ion channel contributes uniquely to the processes of generating and modulating neuronal excitability, which may correlate with specific epilepsy phenotypes or a spectrum of phenotypes characterized by distinct mechanisms. Notably, the *SCN1A* gene has been established as linked to epilepsy through its association with the specific phenotype spectrum of febrile seizure-related epilepsies. Numerous genes exhibit varied funotypes (functional alterations), among which a specific correlation between funotype and phenotype identified in *SCN1A*. The differential expression of *SCN1A* in excitatory pyramidal neurons compared to inhibitory interneurons plays a crucial role in the development of epilepsy. Furthermore, the varied and relatively localized distribution of inhibitory interneurons accounts for the prevalence of partial seizures (Wei et al., 2017).

Mutations in the *KCNT1* gene are associated with a range of focal epilepsies accompanied by intellectual disabilities. While the exact mechanisms remain elusive, all mutations in *KCNT1* that have been functionally analyzed exhibit a pronounced gain-of-function (GOF) effect on the properties of the ion channel (Villa and Combi, 2016). Alterations in the function of ion channel genes can lead to heightened neuronal

excitability, as seen in gain-of-function mutations in NMDAR, or to reduced inhibitory activity, as observed in loss-of-function (LOF) mutations in GABAA receptors within the central nervous system. Conversely, the pathogenic mechanisms linked to certain genes involved in the development of epilepsy, such as SCN1A, KCNC1, and CACNA2D2, are considerably more complex.

A total of six hundred metabolic epilepsies have been identified, accounting for about 37% of all recognized inherited metabolic diseases (IMD). The presentation of seizures in metabolic epilepsies can differ widely, and most of these conditions are marked by their complexity and the involvement of multiple systems (Tumiene et al., 2022). A significant genetic diagnostic yield (43.8%) was observed in genetic testing aimed at identifying metabolic causes of DEE (Jose et al., 2024). This finding was particularly evident in the electroclinical phenotypes associated with progressive myoclonus epilepsies, as well as in unclassified DEE phenotypes characterized by focal or multifocal seizures. In the present study, pathogenic or likely pathogenic variants were identified in genes associated with transport and metabolism. These include *STXBP1* (involved in metabolic cell signalling), *ITPA* (linked to metabolism of heterocyclic compounds), *ALG3* (linked to complex molecule and organelle metabolism), *BTD* and *ALDH18A1* (involved in co-factor and mineral metabolism), *ARG1* and *MT-ATP6* (associated with intermediary metabolism of nutrients), and *SLC2A1* (associated with intermediary metabolism of nutrients).

The mechanistic target of rapamycin (mTOR) signaling pathway functions as a fundamental regulator of cellular metabolism, growth, proliferation, and survival. Central to this pathway is the mTOR complex 1 (mTORC1), which is inhibited by

rapamycin, a macrolide antibiotic derived from the bacterium *Streptomyces hygroscopicus*. Mutations in genes that encode upstream regulators of mTORC1 are implicated in various epilepsies and neurodevelopmental disorders. Tuberous sclerosis complex (TSC) is a multisystem condition resulting from mutations in the mTOR regulators TSC1 or TSC2, characterized by significant neurological symptoms such as epilepsy, focal cortical dysplasia, and neuropsychiatric disorders (Cho, 2011). Focal cortical dysplasia type II is associated with somatic mutations in mTOR pathway activators, including MTOR, AKT3, PIK3CA, and RHEB, and is a leading cause of drug-resistant epilepsy. *De novo* pathogenic variants in *TSC1* and *TSC2* genes were identified in two patients; both were diagnosed with West syndrome accompanied by focal cortical dysplasia. The genes *DEPDC5*, *NPRL2*, and *NPRL3* encode subunits of the GTPase-activating protein (GAP) activity towards the Rags 1 complex (GATOR1), which serves as the primary amino acid-sensing regulator of mTORC1. Germline pathogenic variants in GATOR1 genes are linked to non-lesional focal epilepsies and those associated with malformations of cortical development (Moloney et al., 2021).

In the last two decades, genetic factors have been identified in over half of epilepsy cases, with single gene mutations in ion channels or neurotransmitter receptors linked to many inherited forms. Genetic testing options have expanded, allowing for whole exome and whole genome sequencing, which has improved our understanding of the molecular mechanisms of epilepsy and identified potential targets for precision medicine in specific syndromes like Dravet syndrome and glucose transporter 1 deficiency (Striano and Minassian, 2020). However, precision medicine has mainly focused on seizure management, leaving other aspects, such as neurodevelopmental

and neuropsychiatric comorbidities, largely unaddressed. In our cohort, precision medicine was applied to 35.2% of patients identified with a disease causing variant and 25% are currently in seizure remission.

4.3 Re-analysis of negative cases

Upon conducting a re-analysis of patients for whom no specific disease-causing variants had been detected, the application of deep phenotyping, reverse phenotyping, sequencing data analysis, and literature reviews led to the identification of pathogenic or likely pathogenic variants in four patients, while VUSs were found in three patients. The P/LP variants were identified in *PCDH19* (c.593G>T), *PTPN11* (c.1493G>T), *TBL1XR1* (c.974G>A) and *CLTC* (c.4605+2T>C) and VUSs were identified in genes *CACNA1E* (c.2732G>A), and *CACNA1A* (c.6250A>G), and *AP2M1* (c.327T>G).

The reanalysis of genomic data has identified the underlying cause in an additional four patients. These results underscore the significance of reanalysis for individuals with DEEs who initially test negative on molecular assessments. Previous studies have emphasized the utility of reanalysis for patients with DEEs of indeterminate origin (Rochtus et al., 2020; Scheffer et al., 2023). Our findings indicate that, despite recent negative exome sequencing results, consistent and prompt reanalysis can be productive, often leading to the identification of novel DEE-associated genes. Furthermore, the reanalysis of single nucleotide variants within exome data is expected to enhance diagnostic yield over time.

4.4 *In silico* functional analysis of disease causing variants

Missense variants have the potential to disrupt protein structure, primarily by influencing aspects such as protein folding, interactions between proteins, solubility, or the overall stability of protein molecules. Although no single computational tool can accurately predict the functional consequences of causative variants, functional analysis was effectively performed using various complementary tools. The structural consequences of these mutational alterations were investigated using *in silico* methods, which rely on three-dimensional structural data, the interactions of proteins, protein stability assessments, and molecular dynamics simulations. Consequently, the initial *in silico* analysis of deleterious variants, followed by the identification of a limited set of predicted deleterious mutations for subsequent experimental validation, is a crucial component of contemporary genetic and genomic research.

4.4.1 Protein interaction network

The functional interaction pattern of causative genes was predicted using STRING database (**Figure 4.13**) (Colored node indicates the query protein and first shell of interactors and edges represents protein-protein associations). Most of the deleterious variants were enriched in ion channel genes; particularly variants in voltage gated sodium channel. Strong functional interaction pattern was observed in ion channel genes such as *SCN1A*, *SCN2A*, *SCN3A*, *SCN8A*, *GABRA1*, *GABRG2*, *CHRNA4*, *CHRNA2*, *GABRB2* and *KCNT1*.

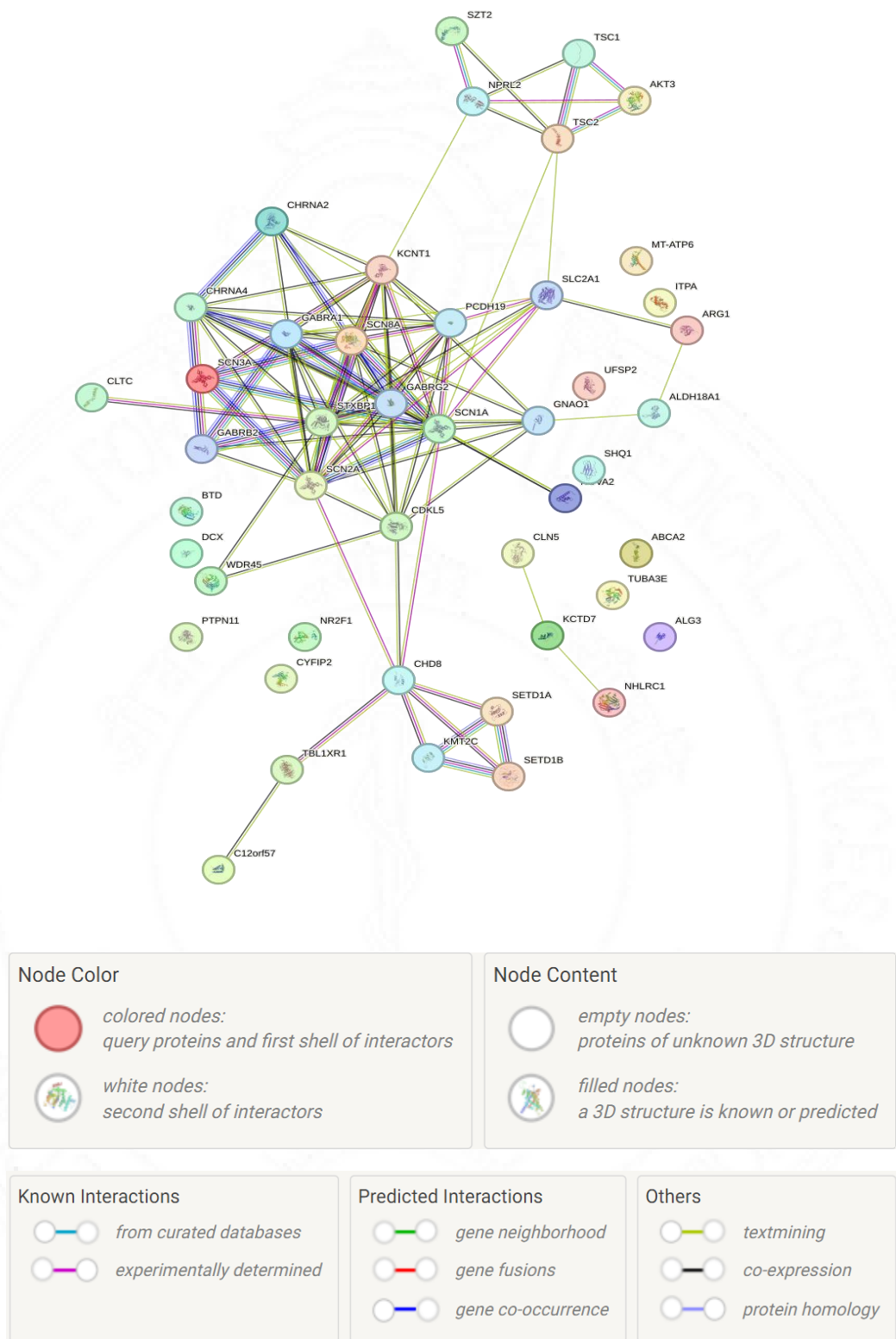


Figure 4.13: Gene-gene interaction of causative genes predicted by STRING

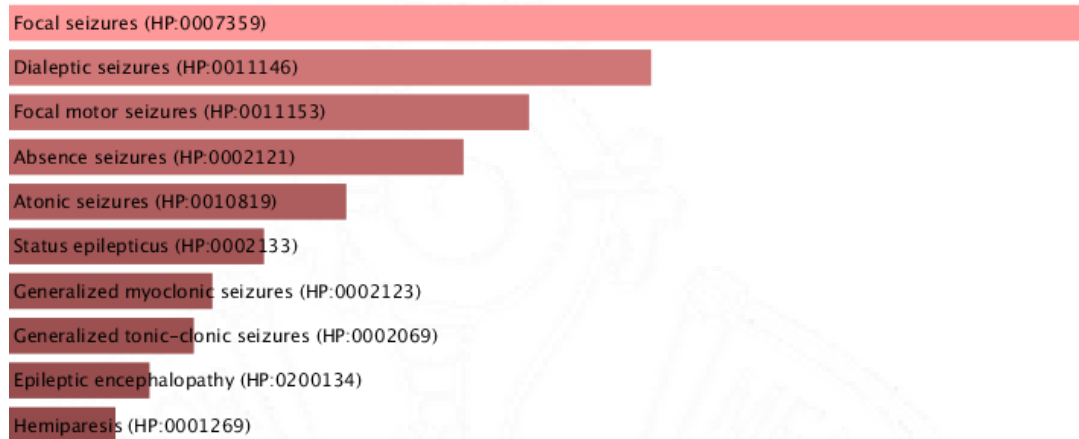
The functional interaction was also observed in mTOR pathway genes (*TSC1*, *TSC2*, and *NPRL2*) and transcription regulation genes (*SETD1A*, *SETD1B*, *TBL1XR1*, and *KMT2C*). A recent study on genotype-phenotype correlation in developmental epileptic encephalopathy with spike-wave activation in sleep found a genetic cause in 34% of cases, highlighting the significant co-expression of associated genes in the brain and the importance of channelopathies and transcriptional regulators (Viswanathan et al., 2024).

Numerous aspects of atypical neuronal activity, such as DNA repair mechanisms, transcriptional control, axonal myelination, transport of metabolites and ions, and peroxisomal activities, alongside channelopathies and synaptic impairments, are associated with childhood epileptic encephalopathies (McTague et al., 2016). Ion channels are integral membrane proteins that form pores within the cell membrane. They are crucial for generating action potentials and sustaining homeostasis by controlling the flow of ions across the membrane, thereby regulating ionic movement between cells and influencing cell volume. Given the importance of these functions in neuronal excitability, ion channels are likely to be significant contributors to the development of epilepsy (Wei et al., 2017). The relationship between ion channel genes and epilepsy may offer valuable insights into the underlying mechanisms of this condition.

4.4.2 Human phenotype ontology enrichment analysis

The enrichment analysis of causative genes in relation to the human phenotype ontology was performed using Wikipathways (**Figure 4.14**).

a)



b)

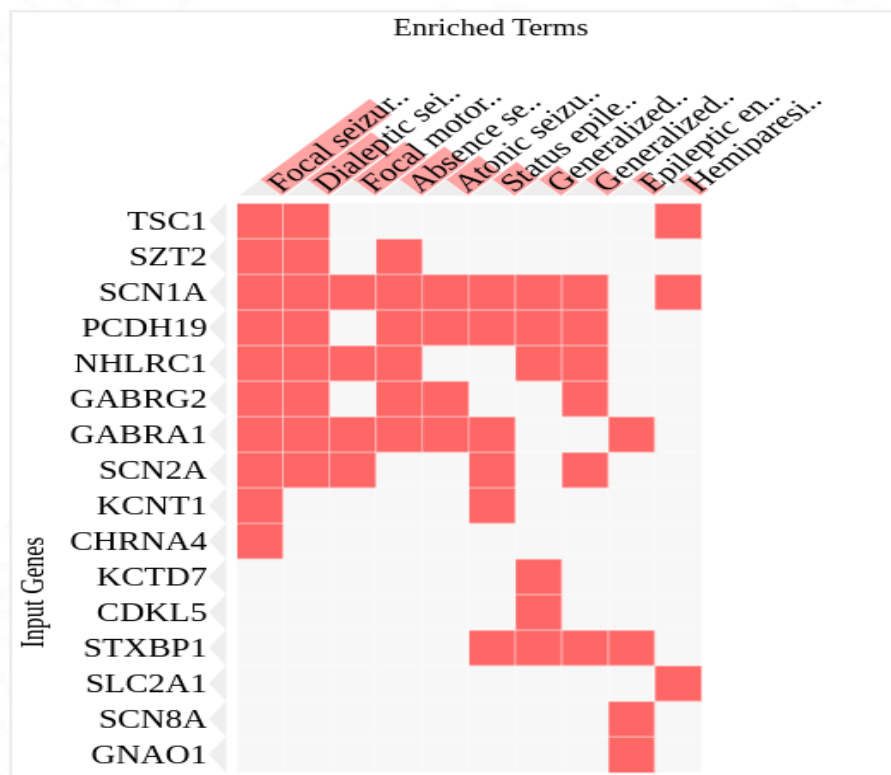


Figure 4.14 : Phenotype enrichment analysis a) HPO enrichment analysis using wikipathways b) Phenotype enrichment analysis of causative genes

The results revealed that a notable number of disease-associated genes were enriched in the context of focal seizures, particularly *TSC1*, *SZT2*, *SCN1A*, *PCDH19*, *NHLRC1*, *GABRG2*, *GABRA1*, *SCN2A*, *KCNT1*, and *CHRNA4*.

The genes *GABRA1*, *STXBP1*, *SCN8A* and *GNAO1* have been linked to Epileptic encephalopathy, while a notable enrichment of the voltage-gated sodium channel gene, *SCN1A*, was observed in relation to various phenotype terms, including focal seizures, dialeptic seizures, focal motor seizures, absence seizures, atonic seizures, status epilepticus, generalized myoclonic seizures, and generalized tonic-clonic seizures.

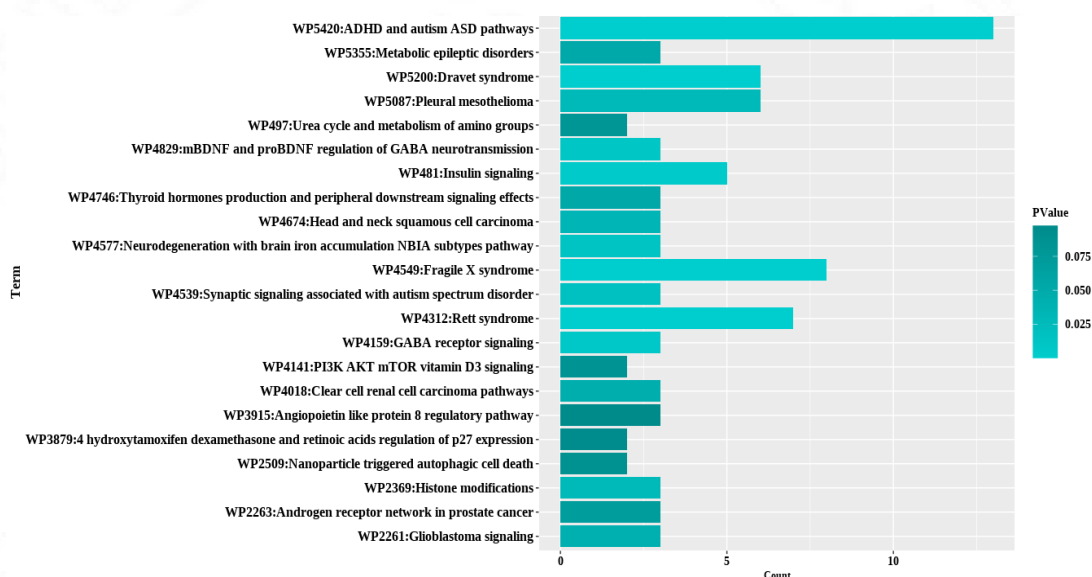


Figure 4.15 : Human phenotype enrichment analysis of causative genes using David functional database

In the context of rare diseases, next-generation sequencing often identifies potential candidate variants across multiple genes. Establishing accurate diagnoses necessitates a thorough analysis of the implications of these variants, which must be integrated with clinical observations (Groza et al., 2015). Research conducted by various groups has demonstrated that employing the Human Phenotype Ontology (HPO) to evaluate

phenotypic data significantly enhances the capacity of NGS methodologies to pinpoint candidate genes associated with diseases and facilitate clinical diagnoses (Soden et al., 2014; Zemojtel et al., 2014). In the present study, a notable enrichment (P-value < 0.05) of genes implicated in diseases related to ADHD and autism pathways, as well as Dravet syndrome, Rett syndrome, Fragile X syndrome, and metabolic epileptic disorders, was observed (Figure 4.15).

4.4.3 Functional annotation and gene ontology analysis

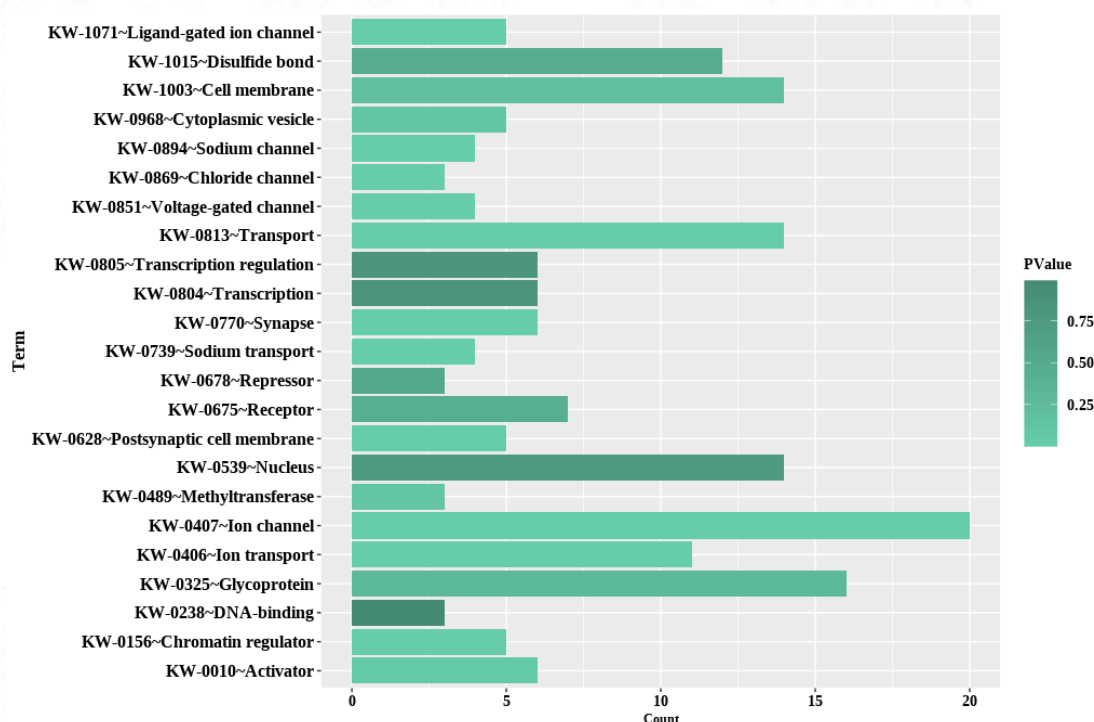


Figure 4.16: Functional annotation of causative genes

The Gene ontology (GO) enrichment analysis of disease causing genes were performed using DAVID database, which provides a systematic and comprehensive biological function annotation of genes. The functional annotation showed that most of the disease causing genes were enriched in ion channel, cell membrane and transport (Figure 4.16).

The GO biological process (BP) analysis revealed that the causative genes were mainly enriched in signal transduction (**Figure 4.17**) and significant enrichment was observed in negative regulation of TORC1 signalling, membrane depolarization during action potential, cardiac muscle cell action potential involved in contraction, nervous system process, sodium ion transmembrane transport, neuronal action potential, brain development, central nervous system development, and sodium ion transport.

The GO molecular function (MF) analysis revealed that the causative genes were primarily enriched in extracellular ligand gated mono atomic ion channel activity, and protein kinase binding, and significant enrichment was observed in voltage gated sodium channel activity, transmitter gated mono atomic ion channel activity involved in regulation of postsynaptic membrane potential, GABA-gated chloride ion channel activity, histone H3K4 methyl transferase activity, GABA-A receptor activity and chloride channel activity (**Figure 4.18**).

The GO cellular component (CC) analysis revealed that the disease causing genes were mainly enriched in membrane and nucleoplasm. Significant enrichment was also noted in perinuclear region of cytoplasm, neuron projection, axon, synapse, GABA-A receptor complex, transmembrane transporter complex, postsynaptic specialization membrane, and GABA-ergic synapse (**Figure 4.19**).

In a recent investigation focused on elucidating the clinical and genetic traits of individuals with epilepsy within a neurodevelopmental disorder (NDD) cohort, the gene ontology analysis of 129 epilepsy-associated genes demonstrated a systematic interrelationship among these genes. This analysis highlighted functional associations

with various components, including ion channels, neuronal cells, and neurotransmitters (Ko et al., 2023). Notable interactions were identified with cation channel complexes, voltage-gated ion channel activity, neuronal cell body membranes, glutamate-gated calcium ion channel activity, and processes related to associative learning. This observation was consistent with our results.

Numerous biological pathways may play a role in the pathogenesis of developmental and epileptic encephalopathies. The most common contributors are channelopathies and synaptic dysfunctions (Wei et al., 2017). Additionally, disruptions in various cellular functions, including DNA repair mechanisms, transcriptional regulation, axonal myelination, transport of metabolites and ions, and peroxisomal activity, may also contribute to the development of DEEs. Various cellular compartments, such as the soma, dendrites, axon, and synapses, could be significant in this context (McTague et al., 2016).

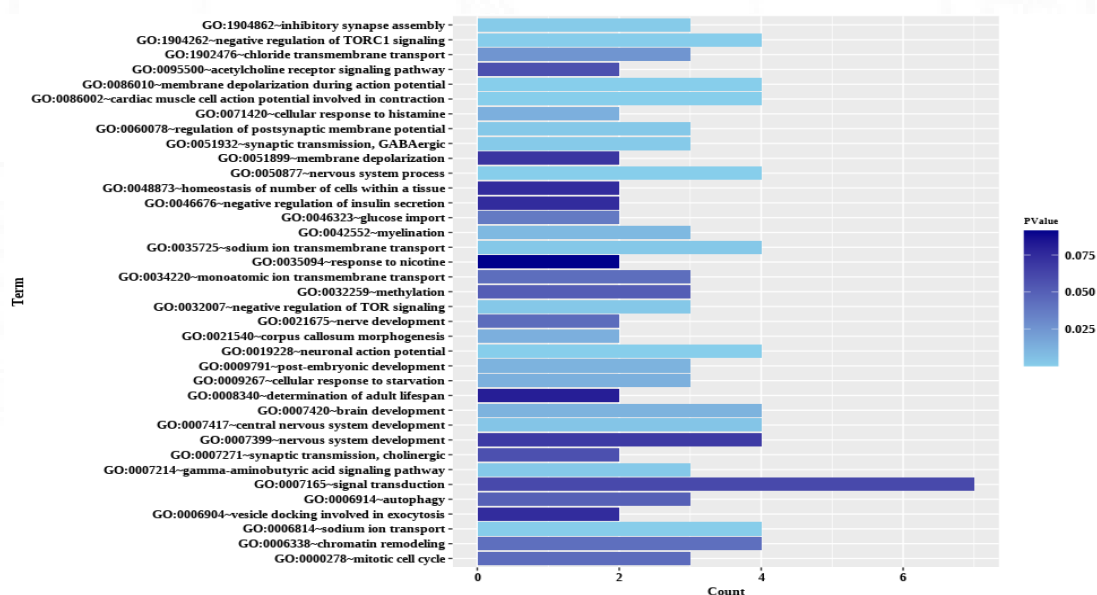


Figure 4.17: Gene ontology analysis (Biological process)

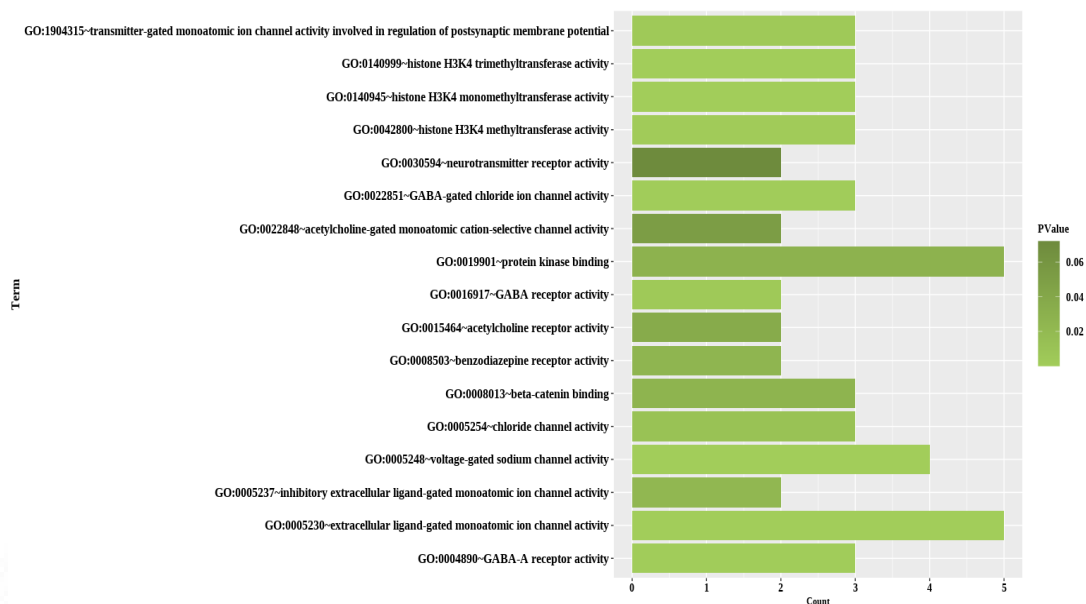


Figure 4.18: Gene ontology analysis (Molecular function)

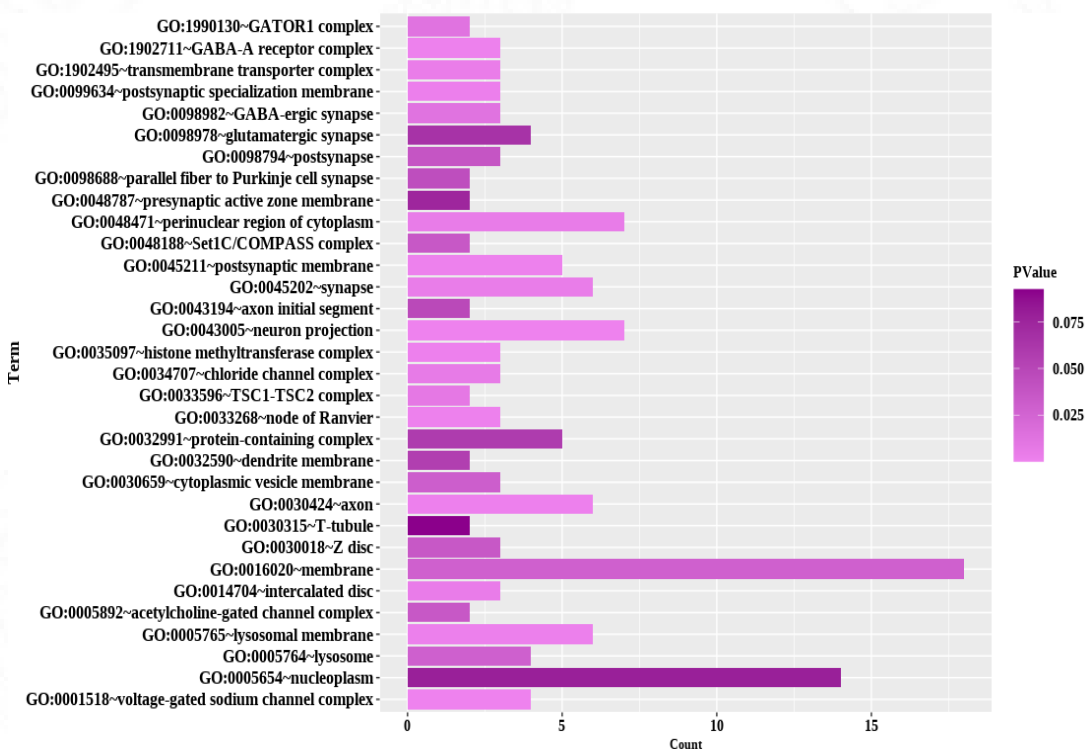


Figure 4.19: Gene ontology analysis (Cellular component)

4.4.4 Functional prediction of voltage-gated sodium ion channel variants

To date, numerous genes associated with epilepsy have been identified, with approximately 25% of these genes encoding ion channel proteins, which include both voltage-gated and ligand-gated ion channels. Abnormalities in ion channel function can influence neural excitability and brain development, potentially contributing significantly to the pathogenesis of epilepsy (Hahn and Neubauer, 2009). In the current study, it was found that 37% of the variants responsible for the disease are located in genes that encode ion channels, thereby enhancing the potential for targeted genetic testing and personalized therapeutic approaches for individual epilepsy patients. Among these, *SCN1A* ($n=13$) stands out as the most pertinent epilepsy gene, exhibiting strong correlations between genotype, functional changes, and clinical manifestations. Loss of function (LOF) of *SCN1A*, primarily due to protein truncation variants, deletions, and missense variants in the pore region, is associated with severe clinical outcomes, such as Dravet syndrome.

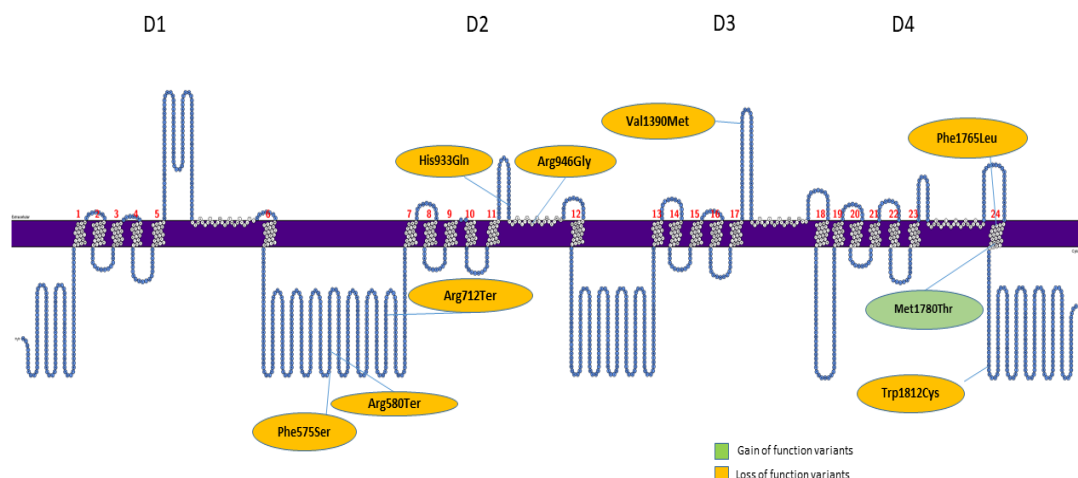
In this study, six missense variants in the *SCN1A* gene were identified, with four located in the pore region and two situated within the cytoplasmic loop (**Figure 4.20**). Additionally, there were two truncating variants and one splice site variant found in the cytoplasmic loop. Eight of these variants were predicted as loss-of-function variants through the *in silico* functional prediction tool, funNCion (**Table 4.7**). A prior investigation revealed that the ratio of missense to truncating variants in the *SCN1A* gene associated with Dravet syndrome was comparable. Notably, a significant trend was observed indicating that children with truncating variants were more likely to

experience multiple seizure types, whereas those with missense variants had a higher likelihood of achieving seizure freedom. Furthermore, all patients possessing truncating variants exhibited epilepsy that was resistant to antiepileptic medications (Krishna et al., 2024).

In a cohort of four individuals diagnosed with Dravet syndrome, the identified variants included truncation mutations (2 out of 4), missense variants in the pore region (S5-S6, 1 out of 4), and in the voltage sensor region (S4, 1 out of 4) (Chengyan et al., 2024). These individuals experienced frequent seizures despite treatment with valproate (VPA), clonazepam (CZP), or topiramate (TPM). Recently, fenfluramine and stiripentol have emerged as additional first-line treatment options for Dravet syndrome (Wirrell et al., 2022).

Table 4.7: *In silico* functional prediction of causative variants in *SCN1A* gene

Sl No	Variant	Inheritance	Functional prediction (funNCion)	Functional prediction (SCN Viewer)	Position on protein structure
1	Exon 29 c.5339T>C p.Met1780Thr	De novo	GoF	Mixed effect	D4/S24; Pore region
2	Exon 14 c.1724del p.Phe575Serfs*48	De novo	LoF	-	D1/S6-D2/S7; Cytoplasmic loop
3	Exon 15 c.2134C>T p.Arg712*	De novo	Truncating	-	D1/S6-D2/S7; Cytoplasmic loop
4	Exon 14 c.1738C>T p.Arg580*	De novo	Truncating	-	D1/S6-D2/S7; Cytoplasmic loop
5	Exon 18 c.2836C>G p.Arg946Gly	De novo	LoF	R946C, R946H (LoF)	D2/S11-D2/S12; Pore loop
6	Exon 18 c.2799C>A p.His933Gln	De novo	probability to be loss-of-function	-	D2/S11-D2/S12; Pore loop
7	Exon 24 c.4168G>A p.Val1390Met	De novo	probability to be loss-of-function	-	D3/S17-S18; Pore loop
8	Exon 12 c.1171-1G>A	Assumed De novo	Splicing	-	D4/S24 Cytoplasmic loop
9	Exon 29 c.5436G>C p.Trp1812Cys	Autosomal Dominant	LoF	-	D4/S24 Cytoplasmic loop
10	Exon 29 c.5293T>C p.Phe1765Leu	De novo	LoF	LoF	D4/S24; Pore region



Protter - visualize proteoforms Omasits et al., Bioinformatics. 2013 Nov 21.

Figure 4.20: Causative variants in SCN1A labeled in primary structure

Table 4.8 : Functional prediction of voltage-gated sodium channel variants

Sl No	Gene	Variant	Inheritance	Functional prediction (funNCion)	Functional prediction (SCN Viewer)
1	SCN2A	SCN2A:Exon 22 c.4231G>A p.Gly1411Arg	De novo	LoF	LoF
2	SCN2A	SCN2A:Exon 5 c.585T>A p.Asp195Glu	De novo	LoF	-
3	SCN3A	SCN3A:Exon 23 c.4054A>C p.Met1352Leu	De novo	probability to be loss-of-function	-
4	SCN8A	SCN8A: Exon 17 c.2936C>T p.Ser979Phe	Assumed De novo	GoF	GoF
5	SCN8A	SCN8A: Exon 22 c.3967G>A p.Ala1323Thr	De novo	GoF	GoF

The *in silico* functional prediction analysis of variants in other voltage-gated sodium channel genes are summarised in **Table 4.8**. NaV1.2 is encoded by the *SCN2A* gene, which is situated on chromosome 2q24.3. This gene is predominantly expressed in the central nervous system, particularly within excitatory and glutamatergic neurons. In contrast, the NaV1.1 channel exhibits a high expression level in GABAergic interneurons. NaV1.2 channels are predominantly found in excitatory neurons, which

links gain-of-function (GoF) mutations to epilepsy due to their role in promoting neuronal hyperexcitability. Conversely, loss-of-function (LoF) mutations are associated with autism and intellectual disability phenotypes. It is noteworthy that certain studies have also established a connection between loss of function and epilepsy (Lossin et al., 2012). In the present study, we identified two LoF variants in the *SCN2A* gene among two patients diagnosed with West syndrome and LKS/CSWS.

SCN3A is primarily expressed in the central nervous system during the stages of embryonic and neonatal development, exhibiting very low levels or occasionally being undetectable in individuals after birth. *De novo* mutations in the *SCN3A* gene are linked to severe neurological dysfunctions and cognitive deficits in infants. These mutations can disrupt NaV1.3 channel function and impact neuronal organization, migration, and proliferation during embryonic development, leading to conditions such as epileptic encephalopathy and polymicrogyria (Menezes et al., 2020). A *de novo* variant in the *SCN3A* gene was detected in our cohort, which is predicted to be a possible LoF variant.

SCN8A mutations can be GoF or LoF alterations, necessitating distinct therapeutic strategies and targets. Notably, even among individuals harboring identical *SCN8A* mutations, variability in drug response can occur. Interestingly, a significant proportion of *SCN8A*-related epilepsies demonstrate a favorable response to channel blockers. Mutations in the *SCN8A* gene are associated with early-infantile epileptic encephalopathy type 13, characterized by early-onset epilepsy with seizures starting before 18 months (Menezes et al., 2020). In our cohort two patients were identified with pathogenic GoF variants in *SCN8A* with age of onset of seizures before 18

months. This finding aligns with the earlier reported cases in the literature. Affected individuals may have intellectual disabilities, developmental delays, and movement disorders, along with occasional co-occurring conditions like autism spectrum disorders, severe juvenile osteoporosis, bradyarrhythmia, cortical visual impairment, and gastrointestinal issues.

4.4.5 *In silico* functional prediction of a novel variant in *STXBP1* gene

Trio-WES conducted on a 6.5-year-old female patient diagnosed with epileptic encephalopathy identified a previously unreported variant in the *STXBP1*(c.A434C;p.Y145S) gene. The variant was not identified in her parents. Mutations in the *STXBP1* gene have been linked to a range of neurodevelopmental disorders collectively known as *STXBP1* encephalopathy. These disorders encompass intellectual disabilities and severe early-onset epileptic encephalopathies (EOEEs), including conditions such as Dravet syndrome, Ohtahara syndrome, Lennox–Gastaut syndrome, and West syndrome (Suo et al., 2021). *De novo* mutations in *STXBP1* were identified in association with EOEE, mental retardation and nosyndromic epilepsy (Hamdan et al., 2009; Li et al., 2018). The functional *in silico* predictions of the variant was performed using protein interaction network, structure, stability and molecular dynamics simulation.

The *STXBP1* gene is situated on the long arm of chromosome 9 at the locus 34.11 and consists of 20 exons (Hamdan et al., 2009). It encodes the Syntaxin1a binding protein, a member of the SEC1 protein family. The *Stxbp1* protein comprises 603 amino acids organized into three distinct domains. The first domain spans residues 4 to 134 and features a five-stranded parallel beta-sheet, which is flanked by five alpha-helices. The

second domain, encompassing residues 135 to 245 and 480 to 592, exhibits a six-stranded beta-sheet structure, characterized by five parallel strands and one antiparallel strand. The third domain, which includes residues 246 to 479, is further divided into two subdomains: 3a and 3b. STXBP1 is predominantly expressed in the brain (**Figure 4.21**) (<https://gtexportal.org/home/gene/STXBP1>) and is believed to play a crucial role in the process of synaptic vesicle exocytosis.

Mutations in the *STXBP1* gene result in the production of a dysfunctional protein that cannot interact with syntaxin1a. This inability renders syntaxin1a inactive, preventing it from binding to Synaptobrevin and synaptosomal-associated protein 25 (SNAP25). The resultant lack of a functional SNARE complex inhibits the attachment of synaptic vesicles to the presynaptic membrane, thereby obstructing the release of inhibitory neurotransmitters. This disruption leads to neuronal hyperexcitability, which can ultimately precipitate convulsive seizures.

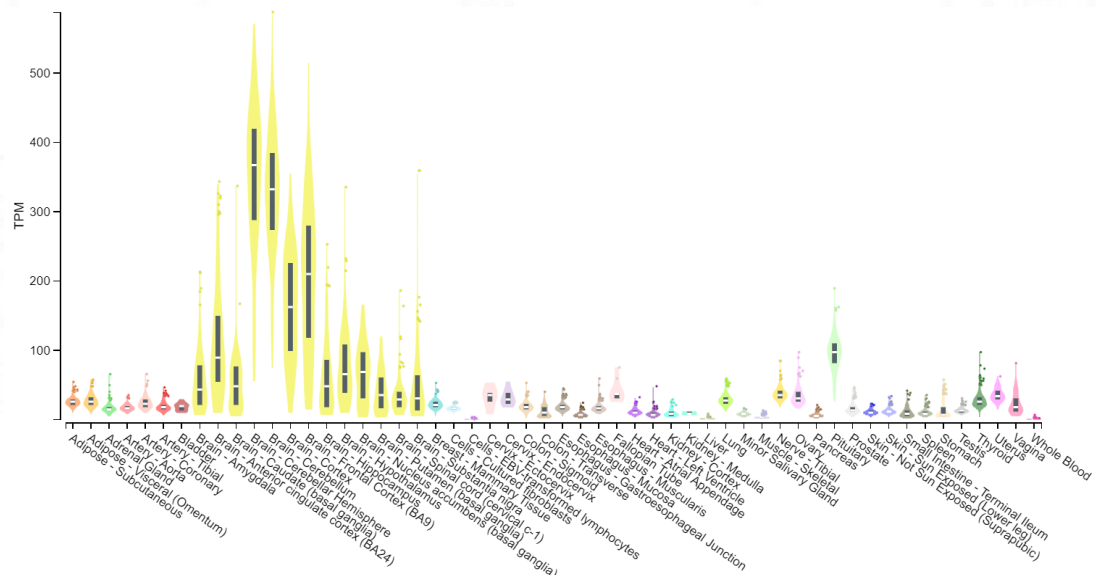


Figure 4.21: Gene expression of STXBP1

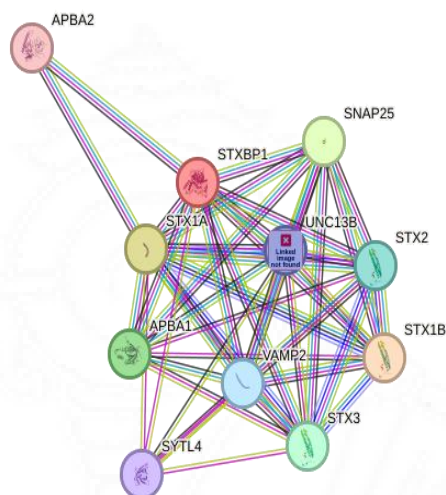


Figure 4.22: Protein interaction network of STXBP1 protein

The STRING tool was employed to identify the closely interacting proteins of STXBP1. The findings indicated that STXBP1 exhibits significant interactions with the proteins SNAP25, STX1A, APBA1, APBA2, UNC13B, VAMP2, SYTL4, STX3, STX1B and STX2 (**Figure 4.22**).

I-Mutant and MUpro tools were used to check the stability of the STXBP1 protein for the Serine amino acid substitution at the position 145. The I-Mutant2.0 revealed a DDG score of -1.77, while MUpro provided a score of -1.56. These results indicate that the variant is likely to reduce the stability of the protein, which may consequently lead to a more significant detrimental effect on the STXBP1 protein.

Molecular dynamics simulation of STXBP1 wild type and mutant protein

The AlphaFold predicted structure of STXBP1 was obtained and subsequently evaluated for quality through the use of a Ramachandran plot. The Missense3D tool was employed to identify the structural alterations resulting from the substitution of

amino acid Serine at position 145. According to the predictions made by the Missense3D tool, the c.A434C:p.Y145S variant has the potential to induce structural damage to the protein by modifying its cavity. Variations in cavity volume may have implications for the stability of the protein (Ittisoponpisan et al., 2019).

The assessment of how pathogenic variants affect the protein structure of STXBP1 was conducted through molecular dynamics simulations utilizing GROMACS. The trajectories from the molecular dynamics simulations for both the wild type and mutant proteins were analyzed over a duration of 100 nano seconds, employing RMSD, RMSF, and Rg.

During the initial phase of the dynamics simulation, the RMSD) values for both the wild type and mutant type remained indistinguishable over the first 10 nanoseconds (**Figure 4.23**). Nevertheless, subsequent analysis of the trajectory revealed a marked increase in the RMSD value for the mutant protein in comparison to the wild type. This suggests that the Y145S variant may exert a considerable detrimental effect on the stability of the protein.

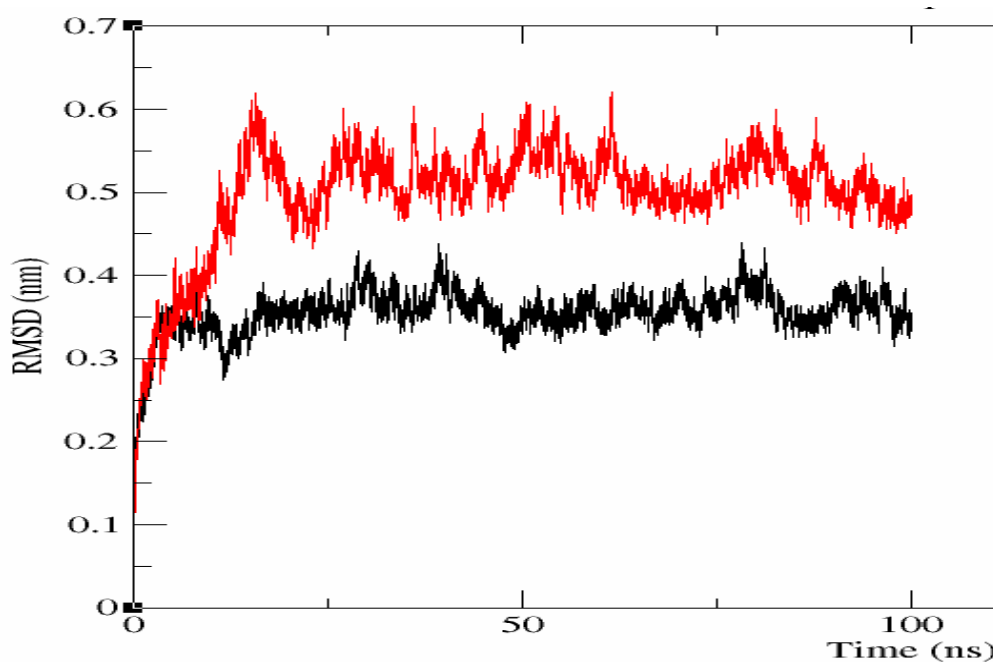


Figure 4.23 : RMSD analysis of STXBP1 wild type (black) and mutant protein (red)

A previous study utilizing molecular dynamics simulations of non-synonymous SNPs in STXBP1 associated with early infantile epileptic encephalopathy revealed that the R190W and L256P variants lead to a decrease in protein stability (Al Mehdi et al., 2019). Several studies have indicated that this reduction in stability can result in increased degradation, aggregation, and misfolding of the protein (Singh et al., 2010).

The assessment of parameters, including the overall dimensions and the molecular compaction level, was conducted through a radius of gyration analysis. The R_g represents the mass-weighted root mean square distance of a collection of atoms from their collective center of mass. This metric offers insights into the overall dimensions of a protein.

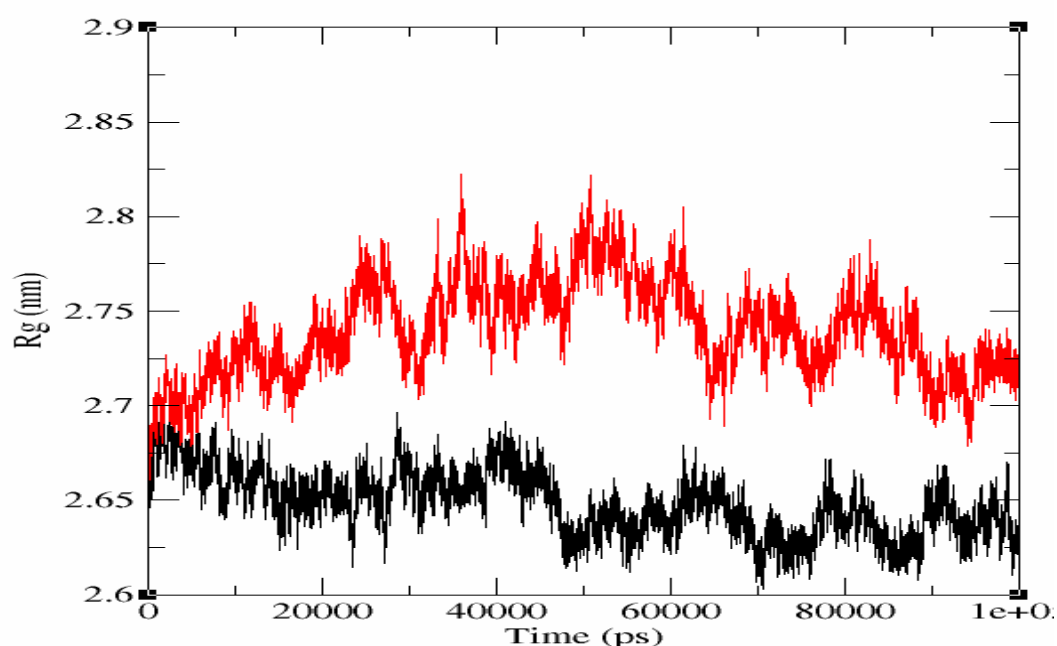


Figure 4.24: Changes in the radius of gyration of the wild type structure (black) and mutant structure (red)

The R_g data is founded on the principle that increased compactness of a protein correlates with enhanced stability (Lobanov et al., 2008). Consequently, R_g serves as an effective metric for examining the properties of proteins during MD simulations conducted over a duration of 100 nanoseconds, during which the radius of gyration is measured (**Figure 4.24**). Notably, significant fluctuations were observed in both the native and mutant forms during the time interval of 0 to 100 ns. The R_g analysis revealed that protein with the variant Y145S had a lower level of compaction than the wild type protein.

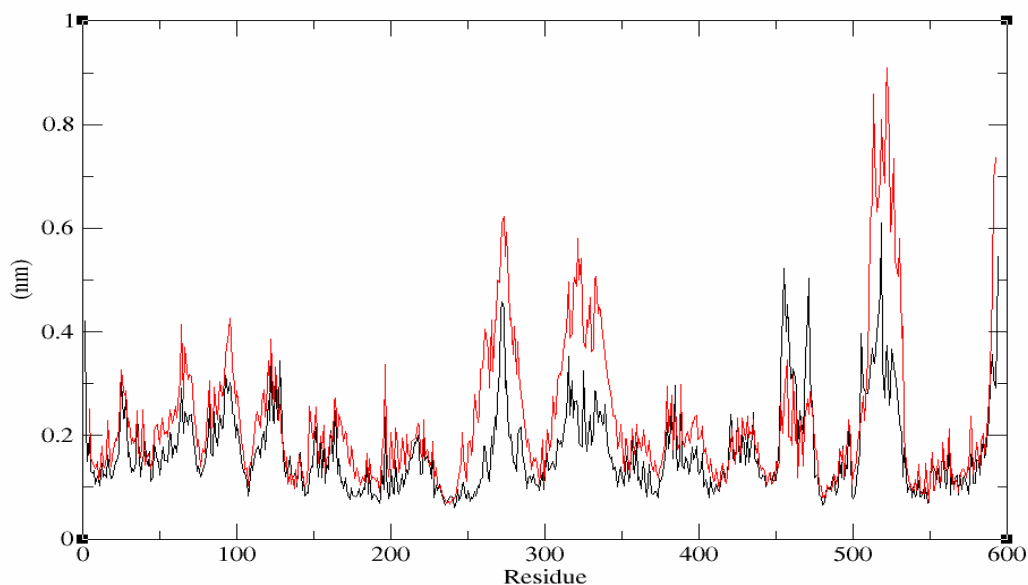


Figure 4.25: RMSF analysis of the wild type (black) and mutant protein (red) during the 100ns MD simulation

The analysis of RMSF indicated that the Y145S variant affect flexibility of the protein at various levels (**Figure 4.25**). Elevated peaks of fluctuations were noted around the 500-520 position, attaining a maximum RMSF value of 0.9 nm. The RMSF values of the mutant protein showed that the Y145S variant exhibited higher values compared to the native protein, thereby enhancing its flexibility. Flexibility is an essential characteristic of protein functionality, enabling proteins to respond to environmental fluctuations and chemical alterations. Molecular flexibility governs numerous functions, including enzymatic catalysis and the regulation of protein activity. Consequently, any disruption that impacts protein flexibility may adversely affect their function, potentially leading to the progression of diseases (Al Mehdi et al., 2019).

An analysis of secondary structure assignments, utilizing the DSSP, was conducted to examine the secondary structural components, including alpha helices, beta sheets, and

coils, throughout the entire duration of the simulation. Trajectories were collected at different time intervals, and the resulting structures were visualized with the aid of xmgrace (**Figure 4.26**).

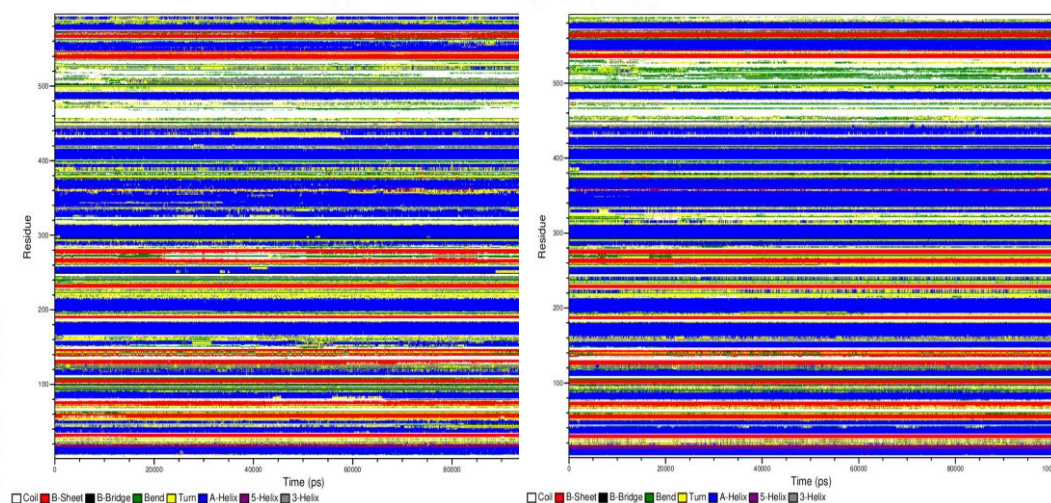


Figure 4.26 : Definition Secondary Structure of Protein analysis of wild-type (left) and mutant structure (right) at different time period

In the wild-type protein, the residues between position 140-160 appeared as a combination of beta sheets, turns, coils, and alpha helices, along with some traces of turn and bend conformations. Conversely, in the mutant structure, this same region displayed beta sheets, turns, coils, and alpha helices, but lacked any evidence of turn and bend conformations. This observation suggests that alterations in the secondary structural elements of the protein could lead to a decrease in its stability (Tanwar and George Priya Doss, 2018).

4.4.6 *In silico* functional prediction of a homozygous recessive variant in *UFSP2* gene

WES analysis revealed a pathogenic homozygous recessive variant in exon 5 of the *UFSP2* (c.344T>A;p.Val115Glu) gene in two siblings from a third-degree consanguineous family. The proband was a 5-year-old male who experienced the onset of epilepsy at the age of 3 months. He received a diagnosis of West syndrome. Additionally, his elder sister, a 14-year-old female, had her epilepsy onset at 7 months of age and is currently in remission from West syndrome. Parents were heterozygous for this variant.

UFM1-specific peptidase 2 (UFSP2), which participates in the UFMylation pathway of protein modification and are implicated in Developmental epileptic encephalopathy-106. UFMylation represents a form of post-translational modification of proteins that bears resemblance to ubiquitination, as both processes involve an E1-E2-E3 reaction cascade. The initiation of UFMylation occurs through the cleavage of the ubiquitin-like peptide known as ubiquitin-fold modifier-1 (UFM1), which reveals a glycine residue on UFM1, thus enabling its conjugation. Following this cleavage, UFM1 undergoes adenylation and is subsequently conjugated to the E1 enzyme, Ubiquitin-like modifier activating enzyme (UBA5). The UFM1 moiety is then transferred to the E2 enzyme, UFM1-conjugase 1 (UFC1). The UFM1-conjugated UFC1, along with a target protein for UFMylation, is directed to the endoplasmic reticulum (ER) membrane by the noncatalytic RING-type E3 enzyme, UFM1-ligase 1 (UFL1) (Liu et al., 2009). The subsequent steps, which remain partially elucidated,

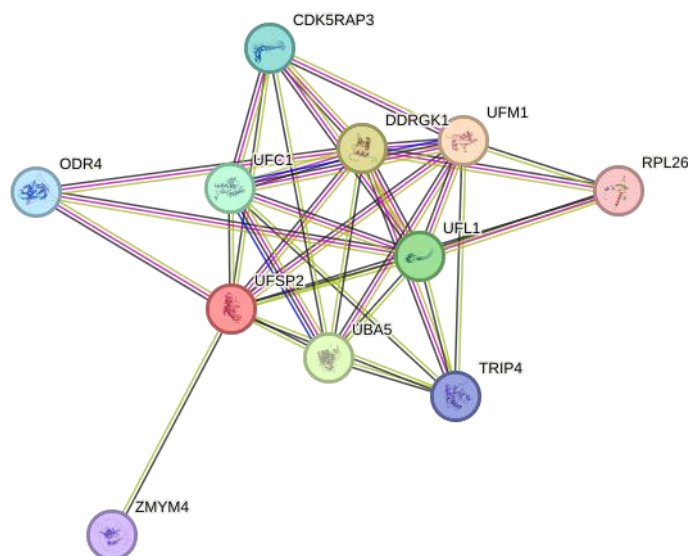


Figure 4.28: Protein interaction network of UFSP2 protein

Molecular dynamics simulation of UFSP2 wild type and mutant (V115E) protein

AlphaFold predicted structure was downloaded for the structural prediction analysis (<https://www.uniprot.org/uniprotkb/Q9NUQ7/entry#structure>) of UFSP2 (AF-Q9NUQ7-F1). The quality assessment of the structure was done using the Ramachandran plot. Mutant residue was introduced using the Mutagenesis wizard tool of PyMOL molecular graphics system by selecting the most common and least strained backbone-dependent rotamers.

A 100 nanosecond molecular dynamics simulation was carried out with periodic boundary conditions for wild type and mutant protein models using GROMACS 2022.4. The procedure involved energy minimization, followed by isothermal-isochoric (NVT) equilibration, and subsequent isobaric (NPT) equilibration of the system. Molecular dynamics simulation was performed for 100ns. The RMSD, root-mean-square-fluctuation (RMSF), and radius of gyration (Rg), were calculated.

For the protein with the variant Val115Glu, its RMSD trajectory showed a higher value compared to the wild type protein during the simulation suggesting that Val115Glu variant decreased the stability of the protein. The variant V115E may cause decrease in function of UFSP2 (**Figure 4.29**).

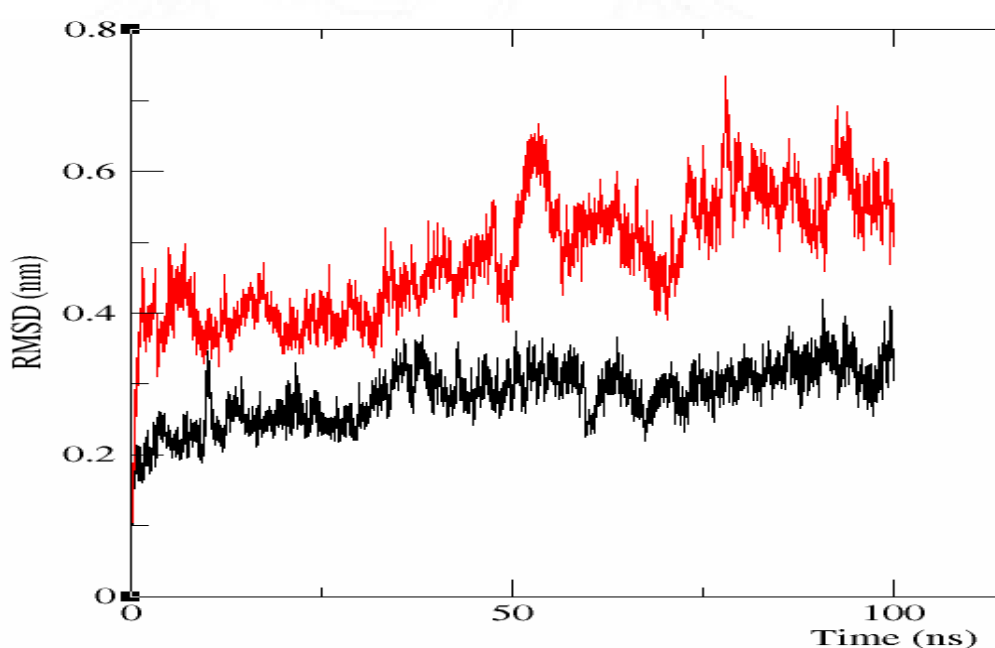


Figure 4.29: RMSD analysis of wild type (black) and mutant type (red) protein structures of UFSP2 during 100 nanoseconds MD simulation

The RMSF analysis showed that there is a significant change in the compaction level of mutant protein when compared to the wild type protein (**Figure 4.31**). The mutant protein showed higher RMSF value than the native protein and influences the flexibility by its increase.

For the mutant protein, Rg values were higher than the native protein values during the first 20ns and between 60ns to 80ns. After this period, the two trajectories showed

no much difference upto 40ns. Between 40ns to 60ns and 80ns to 100ns, the Rg values were lower than the native protein (**Figure 4.30**).

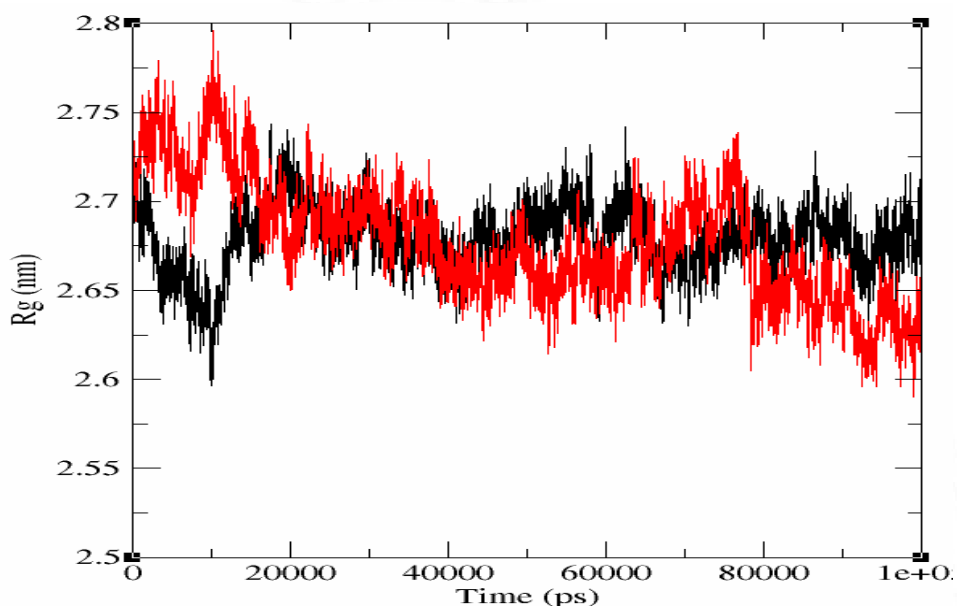


Figure 4.30 : Rg analysis of wild type (black) and mutant type (red) UFSP2 protein

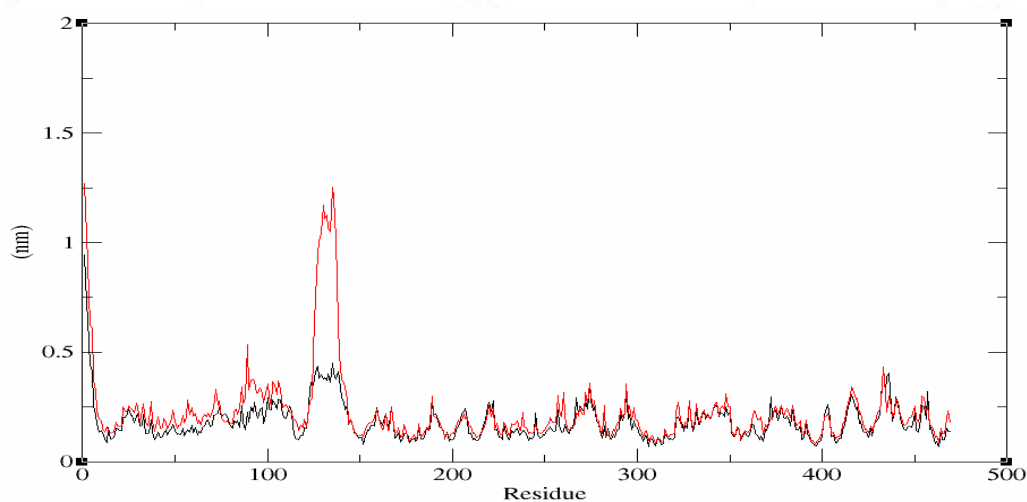


Figure 4.31 : RMSF analysis wild type (black) and mutant type (red) UFSP2 protein

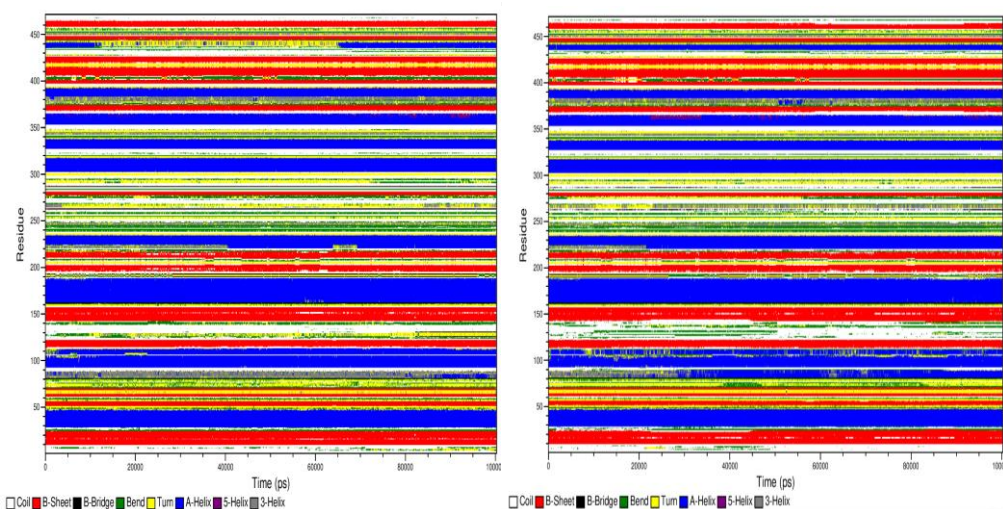


Figure 4.32: DSSP analysis of wild-type (left) and mutant structure (right) at different time period

In the wild-type protein, DSSP analysis showed that the segment spanning positions 100 to 120 exhibited a combination of alpha helices and beta sheets, with some indications of turns noted during the initial 20 nanoseconds of simulation (**Figure 4.32**). In contrast, the mutant structure revealed a similar arrangement of alpha helices and beta sheets in this region; however, turns were consistently observed throughout the molecular dynamics simulation. This observation suggests that alterations in the secondary structural elements of the protein could lead to a decrease in its stability.

Substantial evidence derived from human genetic studies suggests that UFMylation plays a crucial role in brain development, as numerous elements of this pathway exhibit mutations associated with autosomal recessive disorders impacting the central nervous system. Variants in UFM1 are linked to a condition characterized by profound intellectual disability, refractory epilepsy, microcephaly, and poor growth. Previous study demonstrated that immunoblotting of fibroblasts derived from patients indicated

a decrease in UFSP2 levels alongside an increase in UFMylated targets. This suggests that the variant may hinder the process of de-UFMylation rather than affecting UFMylation itself. When patient-derived fibroblasts were reconstituted with wild-type UFSP2, there was a notable reduction in UFMylation marks. Structural analysis of UFSP2 revealed that the variants associated with skeletal disorders are localized within the catalytic domain, while the V115 variant is situated in an N-terminal domain that may play a role in substrate binding (Ni et al., 2021).

4.5 Functional and structural analysis of VUS with potential phenotype match

The VUSs were classified into two categories: incidental VUS and promising VUS (VUS-D). This classification was determined by assessing whether the patient's phenotype aligns with the disease association of the gene, as well as through *in silico* functional predictions. Among the 60 VUSs identified, 10 were designated as promising VUS. These variants were subsequently classified into either incidental VUS or P/LP variants based on assessments of pathogenicity prediction, as well as evaluations of protein stability and structure (**Table 4.9**).

Table 4.9: *In silico* functional prediction of promising VUS

VUSD	Stability prediction: Mupro	Structural prediction: Missense3D	Functional prediction: fuNCion	Pathogenicity prediction	Franklin database classification	Re-classification
GABBR2:Exon 17 c.2503A>C p.Asn835His;	Decreasing	No PDB structure available	–	0.0754; Benign	VUS	Incidental VUS
RNF112:Exon 2 c.59G>T p.Arg20Ile	Decreasing	No PDB structure available	–	0.2332; Benign	VUS	Incidental VUS
BSN: Exon 6 c.10658A>G p.Tyr3553Cys	Decreasing	No PDB structure available	–	0.1149; Benign	VUS	Incidental VUS
BSN:Exon 6 c.9563G>T p.Gly3188Val	Decreasing	No PDB structure available	–	0.0766; Benign	VUS	Incidental VUS
SCN3A:Exon 16 c.2410A>G p.Thr804Ala	Decreasing	No PDB structure available	No functional effect	0.2078; Benign	VUS	Incidental VUS
TCF20:Exon 2 c.2483C>T p.Ala828Val	Increasing	No PDB structure available	–	0.078; Benign	VUS	Incidental VUS
SHQ1:Exon 7 c.850T>C p.Tyr284His	Decreasing	No PDB structure available	–	0.4931; Ambiguous	VUS	Pathogenic
SETD1B:Exon 6 c.931C>T p.Arg311Trp	Decreasing	No PDB structure available	–	0.654; Pathogenic	VUS	??
PDK2:c.704C>T :p.Pro235Leu	Increasing	No structural damage was detected	–	0.1173; Benign	VUS	Incidental VUS
CACNA1E:c.2732G>A: p.Arg911Gln	Decreasing	No PDB structure available	No functional effect	0.103; Benign	VUS	Incidental VUS

Out of the ten promising variants of uncertain significance (VUS-D) identified, eight were classified as incidental VUS after conducting *in silico* functional analysis. The variant SETD1B (c.931C>T;p.Arg311Trp) was predicted as a pathogenic variant (pathogenicity score- 0.654) by AlphaMissense pathogenicity prediction tool. This variant was identified in two siblings who presented with focal epilepsy, macrocephaly, and global developmental delay. After conducting deep phenotyping, the SETD1B (c.931C>T;p.Arg311Trp) variant was reclassified as a pathogenic variant.

The SHQ1 (c.850T>C:p.Tyr284His) variant was initially classified as uncertain by Alaphamissense. Consequently, additional molecular dynamic simulation analyses were conducted for this variant, leading to its reclassification as a likely pathogenic variant.

4.5.1 *In silico* functional analysis of SHQ1 wild type and mutant protein

A female patient who developed epilepsy at 8 months old exhibited unprovoked gelastic seizures and reflex bathing epilepsy. Magnetic resonance imaging (MRI) revealed a focal cortical dysplasia (FCD) in the right middle frontal gyrus. She was later diagnosed with static encephalopathy along with dyskinesia. Trio-WES detected a homozygous recessive VUS in the *SHQ1* (c.850T>C | p.Tyr284His) gene. The variant was carried by the parents.

SHQ1 plays a crucial role in stabilizing the formation of H/ACA ribonucleoproteins (RNPs) by interacting with NAP57 during the initial stages of their assembly, thereby facilitating the biogenesis of H/ACA RNPs. Consequently, the absence of SHQ1 results in the degradation of the RNP assembly. Furthermore, mutations in genes that encode the components of the H/ACA RNP complex, including DKC1, can have profound implications for neurological development (Sleiman et al., 2022).

Biallelic variants in the *SHQ1* gene are associated with two distinct disorders: Dystonia-35, which manifests in childhood, and Neurodevelopmental Disorder with Seizures and Dystonia (NEDDS). The clinical presentation of NEDDS is characterized by profound intellectual disability, motor and speech impairments, behavioral issues,

epilepsy, extrapyramidal symptoms, and neuroimaging findings that may include cerebellar hypoplasia.

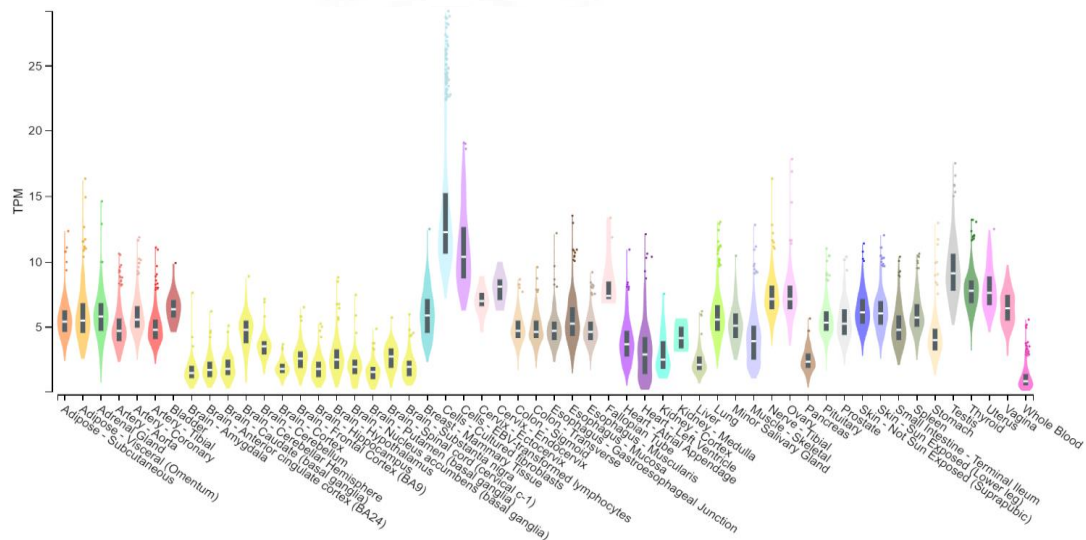


Figure 4.33: Gene expression analysis of SHQ1

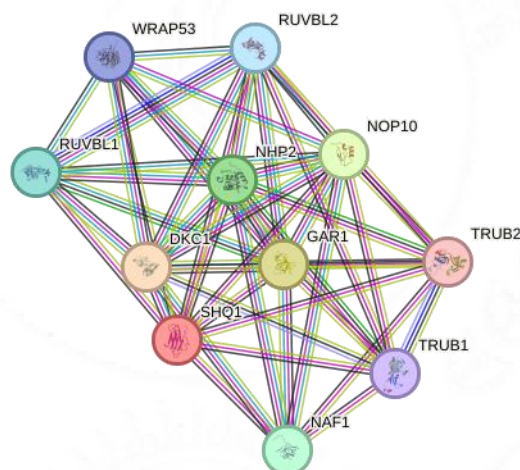


Figure 4.34: Protein interaction network analysis of SHQ1 protein

SHQ1 protein is expressed in most of the tissues (**Figure 4.33**). Protein interactome analysis revealed several potential genetic interactors, including GAR1, NAF1, TRUB2, DKC1, NOP10, NHP2, and RUVBL2 (**Figure 4.34**). The pathways

associated with the prioritized genes encompass ribosome biogenesis in eukaryotes, RNA polymerase activity, RNA transport mechanisms, spliceosome function, ribosomal assembly, the cytosolic DNA-sensing pathway, DNA replication processes, mismatch repair, base excision repair, nucleotide excision repair, and the action of basal transcription factors. These genes play a significant role in vital biological processes that are fundamental to human physiology. Consequently, alterations in the *SHQ1* gene may affect its structural integrity, modify these biological pathways, and ultimately lead to the manifestation of particular phenotypes (Gowda et al., 2024).

Computational analyses of the protein-protein interaction (PPI) network provide additional evidence corroborating earlier findings that the SHQ1 protein engages with H/ACA-specific proteins, including NAF1, GAR1, DKC1, RRP9, and snoRNA, all of which are essential constituents of H/ACA ribonucleoproteins (RNPs). These findings underscore the significant relationship between SHQ1 and H/ACA RNPs, suggesting its vital function in the correct folding and stabilization of the complex, thereby maintaining its functional integrity.

Molecular dynamic simulation of SHQ1 (c.850T>C | p.Tyr284His)

AlphaFold predicted structure was downloaded for the structural prediction analysis of SHQ1. (AF-Q6PI26-F1-v4: <https://alphafold.ebi.ac.uk/entry/Q6PI26>). The quality assessment of the structure was done using the Ramachandran plot. Mutant residue was introduced using the Mutagenesis wizard tool of PyMOL molecular graphics system by selecting the most common and least strained backbone-dependent rotamers.

A 100 nanosecond molecular dynamics simulation was carried out with periodic boundary conditions for wild type and mutant protein models using GROMACS 2022.4. The procedure involved energy minimization, followed by isothermal-isochoric (NVT) equilibration, and subsequent isobaric (NPT) equilibration of the system. Molecular dynamics simulation was performed for 100ns. The RMSD, RMSF, and Rg, were calculated.

The protein variant p.Tyr284His exhibited a significant alteration in its RMSD trajectory relative to the wild type protein throughout the simulation, indicating that this variant could potentially influence the protein's stability. Furthermore, the p.Tyr284His variant may lead to a reduction in the functionality of SHQ1 (**Figure 4.35**).

The RMSF analysis showed that there is a significant change in the flexibility of mutant protein when compared to the wild type protein (**Figure 4.36**). The RMSF value of the mutant type displayed more flexible residues at residue number 450 to 475 as compared to the wild type. The peak at residues 140, 375, and 500 is more defined in wild type indicating more movement.

The Rg data for proteins is based on the principle that increased compactness correlates with enhanced stability (Lobanov et al., 2008). Consequently, Rg serves as an effective metric for examining the properties relevant to the simulation of protein structures. The Rg analysis of SHQ1: p.Tyr284His variant showed that there is a significant change in the compaction level of mutant protein when compared to the wild type protein (**Figure 4.37**). The two trajectories of wild type and mutant type protein showed no much difference upto initial 20ns. After 20 nanoseconds, the two

trajectories exhibited a notable divergence that persisted throughout the 100 nanoseconds of simulation. This observation suggests that the mutation influences the compactness of the protein, thereby impacting its stability.

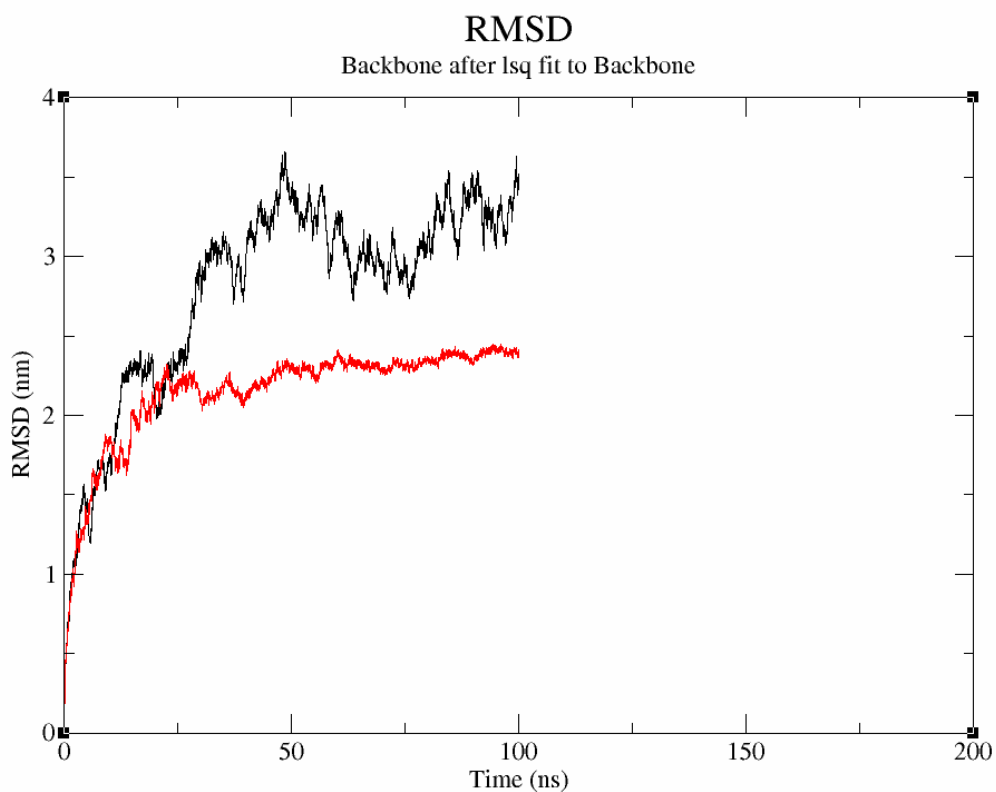


Figure 4.35: RMSD analysis of SHQ1 wild type (black) and mutant type (red) protein

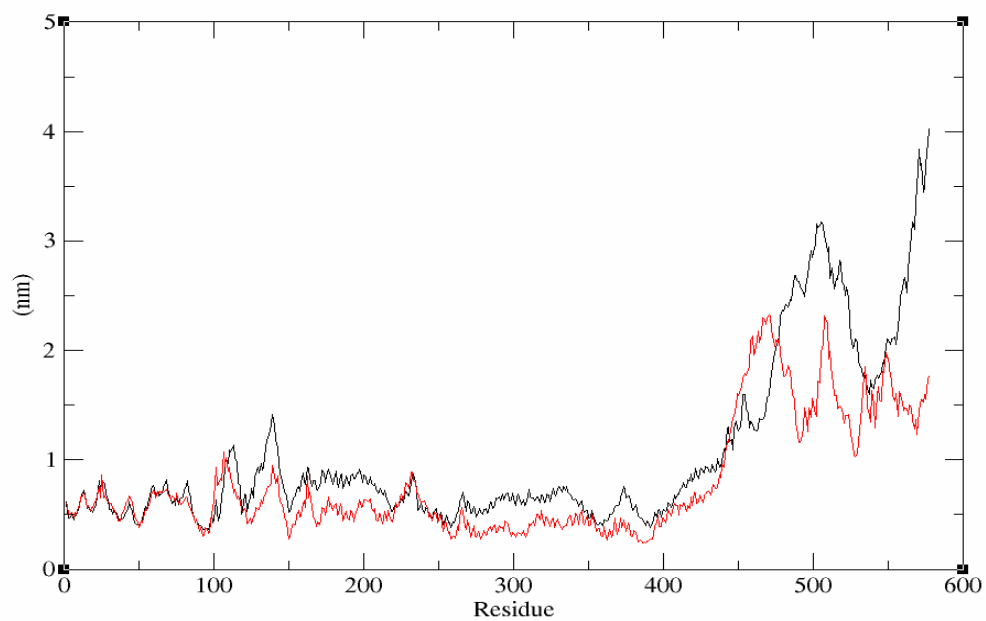


Figure 4.36: RMSF analysis SHQ1 wild type (black) and mutant type (red) protein

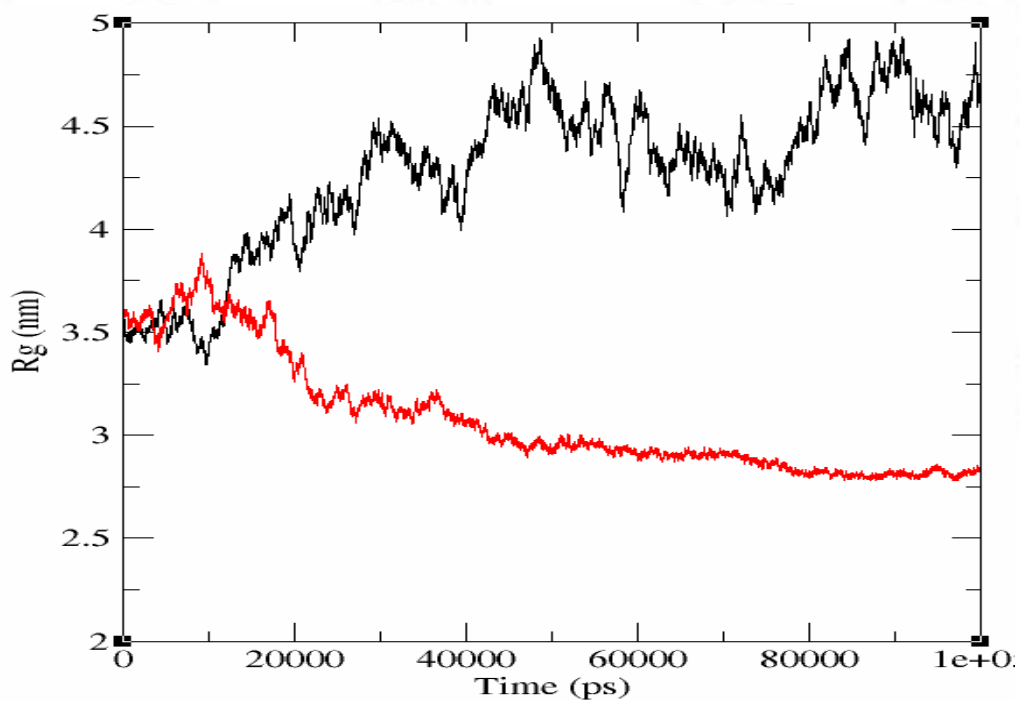


Figure 4.37: Rg analysis of SHQ1 wild type (black) and mutant type (red) protein

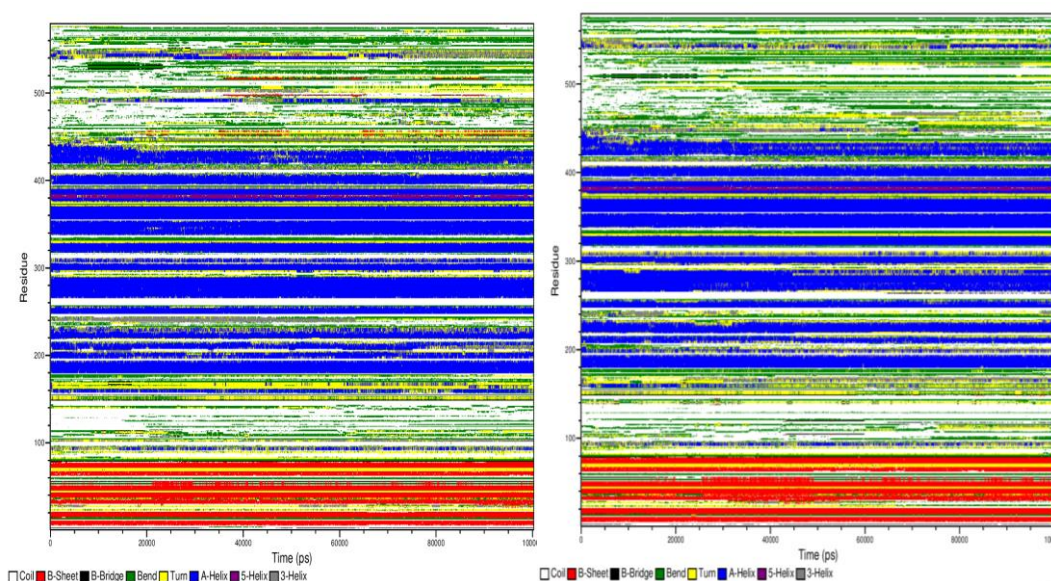


Figure 4.38: Time evolution of secondary structural elements (DSSP) of wild type (left) and mutant type protein (right)

Time evolution analysis of the secondary structural elements of the protein (DSSP) showed the secondary structural elements as a function of time (**Figure 4.38**). In the wild-type protein, the segment spanning positions 260 to 290 exhibited a combination of coils and alpha helices. The mutant structure revealed a similar arrangement of coils and alpha helices during the initial 45 nano seconds in this region. After that, turns were consistently observed throughout the molecular dynamics simulation. This observation suggests that alterations in the secondary structural elements of the protein could lead to a decrease in its stability.

SHQ1 has demonstrated interactions with proteins that regulate the cell cycle, suggesting its possible involvement in the modulation of cell cycle checkpoints, cell division, or signaling pathways associated with the cell cycle. These observations underscore the multifunctional characteristics of the SHQ1 protein, which go beyond

its role in the biogenesis of H/ACA RNPs. Its connections to telomere maintenance, mRNA splicing, and cell cycle regulation imply that SHQ1 is integral to the coordination of diverse cellular functions (Gowda et al., 2024). A previous study demonstrated the functional impact of mutations (R335C and A426V) in the *SHQ1* gene, elucidating the essential role of SHQ1 in facilitating the activity of the pseudouridine synthase dyskerin. Furthermore, it highlighted that insufficient functionality of SHQ1 leads to adverse effects on the synthesis of H/ACA snoRNPs and ribosomes (Alidou-D'Anjou et al., 2023). Recent studies link compound heterozygous variants in the *SHQ1* gene to early-onset dystonia and other neurological issues. When expressed in yeast cells, these mutated SHQ1 variants disrupt the assembly of H/ACA small nucleolar ribonucleoproteins (snoRNPs) and ribosomes (Sleiman et al., 2022). In our study, the *in silico* functional predictions, which encompassed molecular dynamics simulations, suggested that the Tyr284His mutation could potentially influence protein functionality. Additional investigations are required to elucidate the molecular mechanisms that regulate these interactions and their implications for cellular homeostasis, development, and disease.

5 SUMMARY AND CONCLUSIONS

- To the best of our knowledge, this is the first study in India to evaluate a large group of individuals with epileptic encephalopathies, and other refractory epilepsy of uncertain etiology, specifically assessing the genetic diagnostic effectiveness of whole exome sequencing in trios. Trio-based whole exome sequencing serves as an essential tool for the genetic diagnosis of DEEs. Timely genetic testing can facilitate the initiation of suitable therapeutic interventions and yield precise molecular diagnoses.
- In our study, pathogenic variants were detected in 40.6% of the patients, while variants of uncertain significance were found in 34.3% of the cases. Precision medicine was utilized for 35.2% of the patients who were found to have a causative variant through genetic testing.
- Majority of the pathogenic variants consisted of single nucleotide variants or small insertions and deletions. A significant proportion of these pathogenic variants were located within ion channel genes. The most commonly identified pathogenic variants were missense mutations, followed by nonsense mutations, frameshift insertions and deletions, splice site alterations, and copy number variations.
- The study identified 19 novel variants that have not been reported in any population databases. The disease causing variants were confirmed using Sanger sequencing and analyzed by visualizing them in the Integrative Genomics Viewer tool.

- The diagnostic yield was primarily affected by the age at which seizures began, the specific types of seizures experienced, and the various epilepsy syndrome classifications. The predominant mode of variants identified were *de novo* autosomal dominant, which constituted 63.4% of cases, and the diagnostic yield was notably high for individuals diagnosed with Dravet syndrome. A significant diagnostic yield was observed in patients who experienced the onset of epilepsy before the age of 3 years.
- The functional analysis of disease-related variants was performed *in silico* by utilizing structural data, examining protein interactions, assessing protein stability, and executing molecular dynamics simulations. A considerable proportion of the disease-associated variants in DEE were linked to biological pathways involved in the transport of ions and metabolites, as well as transcriptional regulation. A significant functional interaction pattern was identified among ion channel genes, subsequently followed by genes associated with the mTOR pathway and those involved in transcriptional regulation.
- The analysis of causative genes through the human phenotype ontology indicated a significant enrichment of the voltage-gated sodium channel gene, *SCN1A*. This enrichment was associated with a range of phenotype terms, such as focal seizures, dialeptic seizures, focal motor seizures, absence seizures, atonic seizures, status epilepticus, generalized myoclonic seizures, and generalized tonic-clonic seizures.
- The analysis of functional annotation revealed that a significant proportion of the genes associated with DEE were concentrated in biological pathways

related to ion channels, cell membranes, and the transport of ions and metabolites. The *in silico* functional analysis of voltage-gated sodium channel variants revealed that 78.6% of these variants result in a loss of function effect on the protein.

- The *in silico* functional analysis and deep phenotyping of VUS-D resulted in the reclassification of two variants as pathogenic or likely pathogenic, thereby enhancing the genetic diagnostic yield of the investigation.
- The reanalysis of patients who had not previously shown any identifiable disease-causing variants, combined with the implementation of deep phenotyping, reverse phenotyping, analysis of sequencing data, and comprehensive literature reviews, resulted in the discovery of pathogenic or likely pathogenic variants in four individuals.
- The utilization of trio whole exome sequencing for the genetic evaluation of patients experiencing severe early-onset DEE represents a significant diagnostic resource for healthcare professionals, particularly to identify *de novo* variants in dominant-inherited genes. This approach not only aids in the clinical management of these patients but also has important implications for the counseling of their families.

Strengths of the study

- To the best of our knowledge, this prospective study is the first in India and one of the rare studies across the world to investigate a large group of individuals diagnosed with refractory epilepsies, particularly DEE, specifically focusing on the effectiveness of trio-based whole exome sequencing for

genetic diagnostics. Trio-based whole exome sequencing is an essential tool for diagnosing DEEs.

- Apart from prognostication, timely genetic testing can facilitate the prompt adoption of suitable therapeutic approaches and result in accurate molecular diagnoses. The study reported an overall diagnostic yield of 40.6%, with precision medicine applied to 35.2% of the participants identified with pathogenic variants.
- The present study identified 19 novel variants that have not been reported in any population databases. The data regarding these novel variants can be utilized to develop a chip for the genetic diagnosis of epileptic encephalopathy.
- An integrated pipeline for the analysis of whole exome sequencing data was established, utilizing raw data and subsequently validated against the Franklin Database.
- An integrated pipeline was developed for the *in silico* structural and functional assessment of novel variants and VUSs, which will be valuable for centres without facilities for *in vitro* and *in vivo* testing, although validation studies are required.

Limitations of the study

- The identification of VUSs is a common characteristic of all genomic testing, reflecting the incomplete understanding of the genome and its associations with various diseases. A notable aspect of this uncertainty pertains to the identification of variants that are not deemed clinically diagnostic; this was observed in 34.3% of our cohort, indicating that further time or additional investigations may be necessary to clarify their implications. *In vivo* testing

facilities are not available however we attempted to circumvent this by computational modeling in selected variants.

- Technical challenges were associated with the detection of substantial CNVs and mitochondrial variants from WES data.
- Testing was based on informed consent of family and hence consecutive patients could not be tested in our referral centre wherein a sampling bias is inherent.

Future directions

The application of genetic testing has emerged as a crucial instrument in the diagnosis and treatment of epilepsy, facilitating tailored therapeutic approaches and guidance for patients and their families. The discovery of epigenetic factors associated with epilepsy, including DNA methylation, histone alterations, and non-coding RNAs, presents novel opportunities for therapeutic strategies. By influencing the activity of particular genes or modifying the roles of synaptic proteins, therapies aimed at epigenetic mechanisms may contribute to seizure prevention and enhance the overall prognosis for individuals affected by epilepsy.

Gene therapy is emerging as a promising personalized treatment for rare epilepsy types linked to specific genetic mutations, enabling previously unfeasible targeted interventions. To maximize precision medicine's potential in epilepsy, collaboration between scientists and clinical researchers is crucial. Advanced technologies like artificial intelligence (AI) enhance the accuracy of genetic diagnostics, treatment development, and patient management, leading to improved outcomes. Future research

should focus on identifying genes and mutations associated with epilepsy, driven by advancements in genomic technologies.

Investigating the structural and functional aspects of defective genes may facilitate the development of therapies utilizing existing or repurposed pharmaceuticals. Future advancements in epilepsy genetics are expected to incorporate additional genomic methodologies, such as transcriptomic and metabolomic analyses, alongside the broader implementation of whole genome sequencing. Furthermore, novel therapeutic strategies could emerge from drug screening conducted in both *in vivo* and *in vitro* models, guided by the genetic diagnosis of epilepsy.

Recent developments in cellular reprogramming techniques have enabled the generation of nearly any cell type from induced Pluripotent Stem Cells (iPSCs). iPSCs serve as a compelling model for studying neurological disorders, particularly due to the scarcity of accessible live human tissue for culture purposes. The findings indicate that neurons derived from iPSCs specific to various epilepsy syndromes are valuable for simulating epileptic-like hyperactivity, thereby providing a foundation for the evaluation of novel antiepileptic treatments.

The identification of novel genetic variants improves our understanding of genotype-phenotype correlations and supports the creation of personalized treatment strategies. Recent advancements in drug discovery have led to the emergence of promising targeted therapies, highlighting the importance of each genetic variant as a potential therapeutic target in personalized medicine. The combination of NGS technology, functional characterisation of variants, and personalised therapies is set to enhance the treatment of epilepsy.

Integrating genetic and epigenetic knowledge into clinical practice has the potential to revolutionize epilepsy treatment, paving the way for precision medicine approaches that seek to enhance patient outcomes and overall quality of life. This important translational impact allows for the tailoring of treatments to the specific genetic and epigenetic profiles of individual patients, thereby improving the effectiveness of therapeutic interventions and the well-being of those living with epilepsy. The future of genetic research in advancing both the diagnosis and management of epilepsy looks promising, indicating that the field may soon witness the development of more innovative and effective treatment options. Nevertheless, it is essential to continue research to achieve a thorough understanding of the genetic basis of epilepsy and to create more targeted therapeutic strategies.

6 REFERENCES

- Ademuwagun IA, Rotimi SO, Syrbe S, et al. (2021) Voltage Gated Sodium Channel Genes in Epilepsy: Mutations, Functional Studies, and Treatment Dimensions. *Frontiers in Neurology* 12: 600050.
- Al Mehdi K, Fouad B, Zouhair E, et al. (2019) Molecular Modelling and Dynamics Study of nsSNP in STXBP1 Gene in Early Infantile Epileptic Encephalopathy Disease. *BioMed Research International* 2019: 1–14.
- Aledo-Serrano A, Sánchez-Alcudia R, Toledano R, et al. (2021) Developmental and epileptic encephalopathies after negative or inconclusive genetic testing: what is next? *Journal of Translational Genetics and Genomics*. Epub ahead of print 2021. DOI: 10.20517/jtgg.2021.40.
- Alidou-D’Anjou I, Patel A, Sleiman S, et al. (2023) Human SHQ1 variants R335C and A426V lead to severe ribosome biogenesis defects when expressed in yeast. *Frontiers in Genetics* 14: 1240416.
- Almannai M, Al Mahmoud RA, Mekki M, et al. (2021) Metabolic Seizures. *Frontiers in Neurology* 12: 640371.
- Amudhan S, Gururaj G and Satishchandra P (2015) Epilepsy in India I: Epidemiology and public health. *Annals of Indian Academy of Neurology* 18(3): 263.
- Balciuniene J, DeChene ET, Akgumus G, et al. (2019) Use of a Dynamic Genetic Testing Approach for Childhood-Onset Epilepsy. *JAMA Network Open* 2(4): e192129.
- Balestrini S, Arzimanoglou A, Blümcke I, et al. (2021) The aetiologies of epilepsy. *Epileptic Disorders* 23(1): 1–16.
- Bartolini E, Campostrini R, Kiferle L, et al. (2020) Epilepsy and brain channelopathies from infancy to adulthood. *Neurological Sciences* 41(4): 749–761.
- Beal JC, Cherian K and Moshe SL (2012) Early-Onset Epileptic Encephalopathies: Ohtahara Syndrome and Early Myoclonic Encephalopathy. *Pediatric Neurology* 47(5): 317–323.
- Bonardi CM, Heyne HO, Fiannacca M, et al. (2021) *KCNT1* -related epilepsies and epileptic encephalopathies: phenotypic and mutational spectrum. *Brain* 144(12): 3635–3650.
- Burk KC, Kaneko M, Quindipan C, et al. (2024) Diagnostic Yield of Epilepsy-Genes Sequencing and Chromosomal Microarray in Pediatric Epilepsy. *Pediatric Neurology* 150: 50–56.
- Burke W, Parens E, Chung WK, et al. (2022) The Challenge of Genetic Variants of Uncertain Clinical Significance: A Narrative Review. *Annals of Internal Medicine* 175(7): 994–1000.

- Camfield PR (2011) Definition and natural history of Lennox-Gastaut syndrome. *Epilepsia* 52(s5): 3–9.
- Caraballo RH, Chamorro N, Darra F, et al. (2013) Epilepsy With Myoclonic Atonic Seizures: An Electroclinical Study of 69 Patients. *Pediatric Neurology* 48(5): 355–362.
- Casillas-Espinosa PM, Powell KL and O'Brien TJ (2012) Regulators of synaptic transmission: Roles in the pathogenesis and treatment of epilepsy. *Epilepsia* 53(s9): 41–58.
- Cavirani B, Spagnoli C, Caraffi SG, et al. (2024) Genetic Epilepsies and Developmental Epileptic Encephalopathies with Early Onset: A Multicenter Study. *International Journal of Molecular Sciences* 25(2): 1248.
- Cembella AD and Durán-Riveroll LM (2021) Marine guanidinium neurotoxins: Biogenic origins and interactions, biosynthesis and pharmacology. In: *Advances in Neurotoxicology*. Elsevier, pp. 1–47. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2468748021000205> (accessed 14 August 2024).
- Cheng J, Novati G, Pan J, et al. (2023) Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 381(6664): eadg7492.
- Chengyan L, Chupeng X, You W, et al. (2024) Identification of genetic causes in children with unexplained epilepsy based on trio-whole exome sequencing. *Clinical Genetics* 106(2): 140–149.
- Cho C-H (2011) Frontier of Epilepsy Research - mTOR signaling pathway. *Experimental and Molecular Medicine* 43(5): 231.
- Claes L, Del-Favero J, Ceulemans B, et al. (2001) *De novo* Mutations in the Sodium-Channel Gene SCN1A Cause Severe Myoclonic Epilepsy of Infancy. *The American Journal of Human Genetics* 68(6): 1327–1332.
- Coppola G (2009) Malignant migrating partial seizures in infancy: An epilepsy syndrome of unknown etiology. *Epilepsia* 50(s5): 49–51.
- Delmiro A, Rivera H, García-Silva MT, et al. (2013) Whole-Exome Sequencing Identifies a Variant of the Mitochondrial *MT-ND1* Gene Associated with Epileptic Encephalopathy: West Syndrome Evolving to Lennox-Gastaut Syndrome. *Human Mutation* 34(12): 1623–1627.
- Depienne C, Trouillard O, Gourfinkel-An I, et al. (2010) Mechanisms for variable expressivity of inherited SCN1A mutations causing Dravet syndrome. *Journal of Medical Genetics* 47(6): 404–410.
- D’Gama AM, Mulhern S, Sheidley BR, et al. (2023) Evaluation of the feasibility, diagnostic yield, and clinical utility of rapid genome sequencing in infantile epilepsy (Gene-STEPS): an international, multicentre, pilot cohort study. *The Lancet Neurology* 22(9): 812–825.
- Divi PK, Mukesh N, Mohapatra S, et al. (2023) *In silico* analysis of the altered protein

structure and functions caused by existing synaptic gene mutations in Indian population having neurological disorders. Available at: <http://biorxiv.org/lookup/doi/10.1101/2023.11.02.565337> (accessed 8 August 2024).

- Dyment DA, Tétreault M, Beaulieu CL, et al. (2015) Whole-exome sequencing broadens the phenotypic spectrum of rare pediatric epilepsy: a retrospective study. *Clinical Genetics* 88(1): 34–40.
- Ellis CA, Petrovski S and Berkovic SF (2020) Epilepsy genetics: clinical impacts and biological insights. *The Lancet Neurology* 19(1): 93–100.
- Epi4K Consortium and Epilepsy Phenome/Genome Project (2013) *De novo* mutations in epileptic encephalopathies. *Nature* 501(7466): 217–221.
- Esmaeeli Nieh S and Sherr EH (2014) Epileptic Encephalopathies: New Genes and New Pathways. *Neurotherapeutics* 11(4): 796–806.
- Fasaludeen A, McTague A, Jose M, et al. (2024) Genetic variant interpretation for the neurologist – A pragmatic approach in the next-generation sequencing era in childhood epilepsy. *Epilepsy Research* 201: 107341.
- Feng Y-CA, Howrigan DP, Abbott LE, et al. (2019) Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. *The American Journal of Human Genetics* 105(2): 267–282.
- Fisher RS, Boas WVE, Blume W, et al. (2005) Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* 46(4): 470–472.
- Fukata Y and Fukata M (2017) Epilepsy and synaptic proteins. *Current Opinion in Neurobiology* 45: 1–8.
- Gorelov SV, Titov AI, Tolicheva OA, et al. (2023) DSSP in Gromacs: tool for defining secondary structures of proteins in trajectories. Available at: <http://biorxiv.org/lookup/doi/10.1101/2023.08.21.554196> (accessed 30 August 2024).
- Gowda VK, Srinivasan VM, Srivastava S, et al. (2024) *In silico* characterization and identification of compound heterozygous variants in H/ACA Ribonucleoprotein Assembly Factor (SHQ1) from Indian population. *Journal of Family Medicine and Primary Care* 13(1): 208–220.
- Groza T, Köhler S, Moldenhauer D, et al. (2015) The Human Phenotype Ontology: Semantic Unification of Common and Rare Disease. *The American Journal of Human Genetics* 97(1): 111–124.
- Guerrini R, Conti V, Mantegazza M, et al. (2023) Developmental and epileptic encephalopathies: from genetic heterogeneity to phenotypic continuum. *Physiological Reviews* 103(1): 433–513.
- Gulmez Karaca K, Brito DVC and Oliveira AMM (2019) MeCP2: A Critical Regulator of Chromatin in Neurodevelopment and Adult Brain Function.

International Journal of Molecular Sciences 20(18): 4577.

Hahn A and Neubauer BA (2009) Sodium and potassium channel dysfunctions in rare and common idiopathic epilepsy syndromes. *Brain and Development* 31(7): 515–520.

Hamdan FF, Piton A, Gauthier J, et al. (2009) *De novo* STXBP1 mutations in mental retardation and nonsyndromic epilepsy. *Annals of Neurology* 65(6): 748–753.

Hamdan FF, Myers CT, Cossette P, et al. (2017a) High Rate of Recurrent *De novo* Mutations in Developmental and Epileptic Encephalopathies. *The American Journal of Human Genetics* 101(5): 664–685.

Hamdan FF, Myers CT, Cossette P, et al. (2017b) High Rate of Recurrent *De novo* Mutations in Developmental and Epileptic Encephalopathies. *The American Journal of Human Genetics* 101(5): 664–685.

Hancock EC and Cross JH (2013) Treatment of Lennox-Gastaut syndrome. *Cochrane Database of Systematic Reviews* Cochrane Epilepsy Group (ed.). Epub ahead of print 28 February 2013. DOI: 10.1002/14651858.CD003277.pub3.

Happ HC and Carvill GL (2020) A 2020 View on the Genetics of Developmental and Epileptic Encephalopathies. *Epilepsy Currents* 20(2): 90–96.

Helbig I and Abou Tayoun AN (2016) Understanding Genotypes and Phenotypes in Epileptic Encephalopathies. *Molecular Syndromology* 7(4): 172–181.

Helbig KL, Farwell Hagman KD, Shinde DN, et al. (2016) Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genetics in Medicine* 18(9): 898–905.

Hernandez CC and Macdonald RL (2019) A structural look at GABAA receptor mutations linked to epilepsy syndromes. *Brain Research* 1714: 234–247.

Heyne HO, Baez-Nieto D, Iqbal S, et al. (2020) Predicting functional effects of missense variants in voltage-gated sodium and calcium channels. *Science Translational Medicine* 12(556): eaay6848.

Hollingsworth SA and Dror RO (2018) Molecular Dynamics Simulation for All. *Neuron* 99(6): 1129–1143.

Hu Y, Shan Y, Du Q, et al. (2021) Gender and Socioeconomic Disparities in Global Burden of Epilepsy: An Analysis of Time Trends From 1990 to 2017. *Frontiers in Neurology* 12: 643450.

Huang W, Liu M, Yan SF, et al. (2017) Structure-based assessment of disease-related mutations in human voltage-gated sodium channels. *Protein & Cell* 8(6): 401–438.

Hunter MB, Yoong M, Sumpter RE, et al. (2020) Incidence of early-onset epilepsy: A prospective population-based study. *Seizure* 75: 49–54.

Inuzuka LM, Macedo-Souza LI, Della-Ripa B, et al. (2020) Neurodevelopmental

- disorder associated with *de novo* SCN3A pathogenic variants: two new cases and review of the literature. *Brain and Development* 42(2): 211–216.
- Itkis Y, Krylova T, Pechatnikova NL, et al. (2019) A novel variant m.641A>T in the mitochondrial MT-TF gene is associated with epileptic encephalopathy in adolescent. *Mitochondrion* 47: 10–17.
- Ittisoponpisan S, Islam SA, Khanna T, et al. (2019) Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? *Journal of Molecular Biology* 431(11): 2197–2212.
- Jain P, Sharma S and Tripathi M (2013) Diagnosis and management of epileptic encephalopathies in children. *Epilepsy Research and Treatment* 2013: 501981.
- Jiang T, Gao J, Jiang L, et al. (2021) Application of Trio-Whole Exome Sequencing in Genetic Diagnosis and Therapy in Chinese Children With Epilepsy. *Frontiers in Molecular Neuroscience* 14: 699574.
- Jose M, Poulouse P, Sundaram S, et al. (2022) Utility of clinical exome sequencing in progressive myoclonus epilepsy syndromes: An exploratory analysis. *Clinical Genetics* 101(2): 270–271.
- Jose M, Fasaludeen A, Pavuluri H, et al. (2024) Metabolic causes of pediatric developmental & epileptic encephalopathies (DEE)- genetic variant analysis in a south Indian cohort. *Seizure: European Journal of Epilepsy* 115: 20–27.
- Kingsmore SF, Nofsinger R and Ellsworth K (2024) Rapid genomic sequencing for genetic disease diagnosis and therapy in intensive care units: a review. *npj Genomic Medicine* 9(1): 17.
- Kjeldsen MJ, Corey LA, Christensen K, et al. (2003) Epileptic seizures and syndromes in twins: the importance of genetic factors. *Epilepsy Research* 55(1–2): 137–146.
- Ko A, Youn SE, Kim SH, et al. (2018) Targeted gene panel and genotype-phenotype correlation in children with developmental and epileptic encephalopathy. *Epilepsy Research* 141: 48–55.
- Ko YJ, Kim SY, Lee S, et al. (2023) Epilepsy phenotype and gene ontology analysis of the 129 genes in a large neurodevelopmental disorders cohort. *Frontiers in Neurology* 14: 1218706.
- Kobow K, Reid CA, Van Vliet EA, et al. (2020) Epigenetics explained: a topic “primer” for the epilepsy community by the ILAE Genetics/Epigenetics Task Force. *Epileptic Disorders* 22(2): 127–141.
- Kothur K, Holman K, Farnsworth E, et al. (2018) Diagnostic yield of targeted massively parallel sequencing in children with epileptic encephalopathy. *Seizure* 59: 132–140.
- Krey I, Platzer K, Esterhuizen A, et al. (2022) Current practice in diagnostic genetic testing of the epilepsies. *Epileptic Disorders* 24(5): 765–786.

- Krishna S, Fasaludeen A, Jose M, et al. (2024) Impact of variant subtype on electro-clinical phenotype of Dravet syndrome- a South Indian cohort study. *Seizure: European Journal of Epilepsy* 115: 81–86.
- Krochmalnek E, Accogli A, St-Onge J, et al. (2023) mTOR Pathway Somatic Pathogenic Variants in Focal Malformations of Cortical Development: Novel Variants, Topographic Mapping, and Clinical Outcomes. *Neurology Genetics* 9(6): e200103.
- Krüger J, Schubert J, Kegele J, et al. (2022) Loss-of-function variants in the KCNQ5 gene are implicated in genetic generalized epilepsies. *eBioMedicine* 84: 104244.
- Krumm N, Sudmant PH, Ko A, et al. (2012) Copy number variation detection and genotyping from exome sequence data. *Genome Research* 22(8): 1525–1532.
- Krygier M, Pietruszka M, Zawadzka M, et al. (2024) Next-generation sequencing testing in children with epilepsy reveals novel clinical, diagnostic and therapeutic implications. *Frontiers in Genetics* 14: 1300952.
- Kufareva I and Abagyan R (2011) Methods of Protein Structure Comparison. In: Orry AJW and Abagyan R (eds) *Homology Modeling*. Methods in Molecular Biology. Totowa, NJ: Humana Press, pp. 231–257. Available at: https://link.springer.com/10.1007/978-1-61779-588-6_10 (accessed 30 August 2024).
- Lamar K-MJ and Carvill GL (2018) Chromatin Remodeling Proteins in Epilepsy: Lessons From CHD2-Associated Epilepsy. *Frontiers in Molecular Neuroscience* 11: 208.
- Lee S, Abecasis GR, Boehnke M, et al. (2014) Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics* 95(1): 5–23.
- Li H and Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26(5): 589–595.
- Li Q and Wang K (2017) InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *The American Journal of Human Genetics* 100(2): 267–280.
- Li T, Cheng M, Wang J, et al. (2018) *De novo* mutations of *STXBP1* in Chinese children with early onset epileptic encephalopathy. *Genes, Brain and Behavior* 17(8): e12492.
- Liu G, Forouhar F, Eletsy A, et al. (2009) NMR and X-RAY structures of human E2-like ubiquitin-fold modifier conjugating enzyme 1 (UFC1) reveal structural and functional conservation in the metazoan UFM1-UBA5-UFC1 ubiquitination pathway. *Journal of Structural and Functional Genomics* 10(2): 127–136.
- Lobanov MYu, Bogatyreva NS and Galzitskaya OV (2008) Radius of gyration as an indicator of protein structure compactness. *Molecular Biology* 42(4): 623–628.

- Loddenkemper T, Fernández IS and Peters JM (2011) Continuous Spike and Waves During Sleep and Electrical Status Epilepticus in Sleep. *Journal of Clinical Neurophysiology* 28(2): 154–164.
- Lossin C, Shi X, Rogawski MA, et al. (2012) Compromised function in the Nav1.2 Dravet syndrome mutation R1312T. *Neurobiology of Disease* 47(3): 378–384.
- Martin WJ, Warmington JR, Galinski BR, et al. (1985) Automation of DNA Sequencing: A System to Perform the Sanger Dideoxysequencing Reactions. *Bio/Technology* 3(10): 911–915.
- Matricardi S, De Liso P, Freri E, et al. (2020) Neonatal developmental and epileptic encephalopathy due to autosomal recessive variants in *SLC13A5* gene. *Epilepsia* 61(11): 2474–2485.
- McTague A, Appleton R, Avula S, et al. (2013) Migrating partial seizures of infancy: expansion of the electroclinical, radiological and pathological disease spectrum. *Brain* 136(5): 1578–1591.
- McTague A, Howell KB, Cross JH, et al. (2016) The genetic landscape of the epileptic encephalopathies of infancy and childhood. *The Lancet Neurology* 15(3): 304–316.
- Menezes LFS, Sabiá Júnior EF, Tibery DV, et al. (2020) Epilepsy-Related Voltage-Gated Sodium Channelopathies: A Review. *Frontiers in Pharmacology* 11: 1276.
- Mercimek-Mahmutoglu S, Patel J, Cordeiro D, et al. (2015) Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. *Epilepsia* 56(5): 707–716.
- Miller SA, Dykes DD and Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* 16(3): 1215–1215.
- Mitta N, Menon RN, McTague A, et al. (2020) Genotype-phenotype correlates of infantile-onset developmental & epileptic encephalopathy syndromes in South India: A single centre experience. *Epilepsy Research* 166: 106398.
- Møller RS, Hammer TB, Rubboli G, et al. (2019) From next-generation sequencing to targeted treatment of non-acquired epilepsies. *Expert Review of Molecular Diagnostics* 19(3): 217–228.
- Moloney PB, Cavalleri GL and Delanty N (2021) Epilepsy in the mTORopathies: opportunities for precision medicine. *Brain Communications* 3(4): fcab222.
- Nappi P, Miceli F, Soldovieri MV, et al. (2020) Epileptic channelopathies caused by neuronal Kv7 (KCNQ) channel dysfunction. *Pflügers Archiv - European Journal of Physiology* 472(7): 881–898.
- Ng YS, Bindoff LA, Gorman GS, et al. (2021) Mitochondrial disease in adults: recent advances and future promise. *The Lancet Neurology* 20(7): 573–584.
- Ni M, Afroze B, Xing C, et al. (2021) A pathogenic UFSP2 variant in an autosomal

recessive form of pediatric neurodevelopmental anomalies and epilepsy. *Genetics in Medicine* 23(5): 900–908.

Niestroj L-M, Perez-Palma E, Howrigan DP, et al. (2020) Epilepsy subtype-specific copy number burden observed in a genome-wide study of 17 458 subjects. *Brain* 143(7): 2106–2118.

Ostrander BEP, Butterfield RJ, Pedersen BS, et al. (2018) Whole-genome analysis for effective clinical diagnosis and gene discovery in early infantile epileptic encephalopathy. *npj Genomic Medicine* 3(1): 22.

Palmer EE, Schofield D, Shrestha R, et al. (2018) Integrating exome sequencing into a diagnostic pathway for epileptic encephalopathy: Evidence of clinical utility and cost effectiveness. *Molecular Genetics & Genomic Medicine* 6(2): 186–199.

Palmer EE, Sachdev R, Macintosh R, et al. (2021) Diagnostic Yield of Whole Genome Sequencing After Nondiagnostic Exome Sequencing or Gene Panel in Developmental and Epileptic Encephalopathies. *Neurology* 96(13).

Papuc SM, Abela L, Steindl K, et al. (2019) The role of recessive inheritance in early-onset epileptic encephalopathies: a combined whole-exome sequencing and copy number study. *European Journal of Human Genetics* 27(3): 408–421.

Peljto AL, Barker-Cummings C, Vasoli VM, et al. (2014) Familial risk of epilepsy: a population-based study. *Brain* 137(3): 795–805.

Perucca P (2018) Genetics of Focal Epilepsies: What do we know and where are we Heading? *Epilepsy Currents* 18(6): 356–362.

Perucca P, Bahlo M and Berkovic SF (2020) The Genetics of Epilepsy. *Annual Review of Genomics and Human Genetics* 21(1): 205–230.

Poke G, Stanley J, Scheffer IE, et al. (2023) Epidemiology of Developmental and Epileptic Encephalopathy and of Intellectual Disability and Epilepsy in Children. *Neurology* 100(13).

Pronk S, Páll S, Schulz R, et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29(7): 845–854.

Rastin C, Schenkel LC and Sadikovic B (2023) Complexity in Genetic Epilepsies: A Comprehensive Review. *International Journal of Molecular Sciences* 24(19): 14606.

Reynolds C, King MD and Gorman KM (2020) The phenotypic spectrum of SCN2A-related epilepsy. *European Journal of Paediatric Neurology* 24: 117–122.

Richards S, Aziz N, Bale S, et al. (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine* 17(5): 405–424.

- Riggs ER, Andersen EF, Cherry AM, et al. (2020) Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genetics in Medicine* 22(2): 245–257.
- Rochtus A, Olson HE, Smith L, et al. (2020) Genetic diagnoses in epilepsy: The impact of dynamic exome analysis in a pediatric cohort. *Epilepsia* 61(2): 249–258.
- Sadleir LG, Mountier EI, Gill D, et al. (2017) Not all *SCN1A* epileptic encephalopathies are Dravet syndrome: Early profound Thr226Met phenotype. *Neurology* 89(10): 1035–1042.
- Saitsu H, Kato M, Koide A, et al. (2012) Whole exome sequencing identifies *KCNQ2* mutations in Ohtahara syndrome. *Annals of Neurology* 72(2): 298–300.
- Sánchez Fernández I, Loddenkemper T, Gáinza-Lein M, et al. (2019) Diagnostic yield of genetic tests in epilepsy: A meta-analysis and cost-effectiveness study. *Neurology* 92(5).
- Sands TT and Gelinas JN (2024) Epilepsy and Encephalopathy. *Pediatric Neurology* 150: 24–31.
- Scheffer IE and Liao J (2020) Deciphering the concepts behind “Epileptic encephalopathy” and “Developmental and epileptic encephalopathy”. *European Journal of Paediatric Neurology* 24: 11–14.
- Scheffer IE, Berkovic S, Capovilla G, et al. (2017) ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 58(4): 512–521.
- Scheffer IE, Bennett CA, Gill D, et al. (2023) Exome sequencing for patients with developmental and epileptic encephalopathies in clinical practice. *Developmental Medicine & Child Neurology* 65(1): 50–57.
- Sedlackova L, Sterbova K, Vlckova M, et al. (2024) Yield of exome sequencing in patients with developmental and epileptic encephalopathies and inconclusive targeted gene panel. *European Journal of Paediatric Neurology* 48: 17–29.
- Seifi M and Walter MA (2018) Accurate prediction of functional, structural, and stability changes in PITX2 mutations using *in silico* bioinformatics algorithms. *PLOS ONE* Cai T (ed.) 13(4): e0195971.
- Singh SM, Kongari N, Cabello-Villegas J, et al. (2010) Missense mutations in dystrophin that trigger muscular dystrophy decrease protein stability and lead to cross- β aggregates. *Proceedings of the National Academy of Sciences* 107(34): 15069–15074.
- Singhi P and Gupta A (2021) Epilepsy in Children—Important Facets. *Indian Journal of Pediatrics* 88(10): 991–992.
- Sisodiya SM (2021) Precision medicine and therapies of the future. *Epilepsia* 62(S2).

- Sleiman S, Marshall AE, Dong X, et al. (2022) Compound heterozygous variants in *SHQ1* are associated with a spectrum of neurological features, including early-onset dystonia. *Human Molecular Genetics* 31(4): 614–624.
- Soden SE, Saunders CJ, Willig LK, et al. (2014) Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Science Translational Medicine* 6(265).
- Starita LM, Ahituv N, Dunham MJ, et al. (2017) Variant Interpretation: Functional Assays to the Rescue. *The American Journal of Human Genetics* 101(3): 315–325.
- Striano P and Minassian BA (2020) From Genetic Testing to Precision Medicine in Epilepsy. *Neurotherapeutics* 17(2): 609–615.
- Sun D, Liu Yan, Cai W, et al. (2021) Detection of Disease-Causing SNVs/Indels and CNVs in Single Test Based on Whole Exome Sequencing: A Retrospective Case Study in Epileptic Encephalopathies. *Frontiers in Pediatrics* 9: 635703.
- Suo G, Cao X, Zheng Y, et al. (2021) A *de novo* nonsense mutation of STXBP1 causes early-onset epileptic encephalopathy. *Epilepsy & Behavior* 123: 108245.
- Syrbe S (2022) Developmental and epileptic encephalopathies – therapeutic consequences of genetic testing. *Medizinische Genetik* 34(3): 215–224.
- Tanwar H and George Priya Doss C (2018) An Integrated Computational Framework to Assess the Mutational Landscape of α -L-Iduronidase *IDUA* Gene. *Journal of Cellular Biochemistry* 119(1): 555–565.
- Thakran S, Guin D, Singh Pooja, et al. (2020) Genetic Landscape of Common Epilepsies: Advancing towards Precision in Treatment. *International Journal of Molecular Sciences* 21(20): 7784.
- The 100,000 Genomes Project Pilot Investigators (2021) 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care — Preliminary Report. *New England Journal of Medicine* 385(20): 1868–1880.
- The International League Against Epilepsy Consortium on Complex Epilepsies, Abou-Khalil B, Auce P, et al. (2018) Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nature Communications* 9(1): 5269.
- Thiffault I, Speca DJ, Austin DC, et al. (2015) A novel epileptic encephalopathy mutation in *KCNB1* disrupts Kv2.1 ion selectivity, expression, and localization. *Journal of General Physiology* 146(5): 399–410.
- Tumiene B, Ferreira CR and Van Karnebeek CDM (2022) 2022 Overview of Metabolic Epilepsies. *Genes* 13(3): 508.
- Vadlamudi L, Milne RL, Lawrence K, et al. (2014) Genetics of epilepsy: The testimony of twins in the molecular era. *Neurology* 83(12): 1042–1048.

- Van Der Auwera GA, Carneiro MO, Hartl C, et al. (2013) From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics* 43(1).
- Veeramah KR, Johnstone L, Karafet TM, et al. (2013) Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia* 54(7): 1270–1281.
- Vetri L, Calì F, Saccone S, et al. (2024) Whole Exome Sequencing as a First-Line Molecular Genetic Test in Developmental and Epileptic Encephalopathies. *International Journal of Molecular Sciences* 25(2): 1146.
- Villa C and Combi R (2016) Potassium Channels and Human Epileptic Phenotypes: An Updated Overview. *Frontiers in Cellular Neuroscience* 10.
- Viswanathan S, Oliver KL, Regan BM, et al. (2024) Solving the Etiology of Developmental and Epileptic Encephalopathy with Spike–Wave Activation in Sleep (D/EE-SWAS). *Annals of Neurology*: ana.27041.
- Wang K, Li M and Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* 38(16): e164–e164.
- Wei F, Yan L-M, Su T, et al. (2017) Ion Channel Genes and Epilepsy: Functional Alteration, Pathogenic Potential, and Mechanism of Epilepsy. *Neuroscience Bulletin* 33(4): 455–477.
- Wheless JW, Fulton SP and Mudigoudar BD (2020) Dravet Syndrome: A Review of Current Management. *Pediatric Neurology* 107: 28–40.
- Wirrell E, Farrell K and Whiting S (2005) The Epileptic Encephalopathies of Infancy and Childhood. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques* 32(4): 409–418.
- Wirrell EC, Hood V, Knupp KG, et al. (2022) International consensus on diagnosis and management of Dravet syndrome. *Epilepsia* 63(7): 1761–1777.
- Wright CF, Campbell P, Eberhardt RY, et al. (2022) Optimising diagnostic yield in highly penetrant genomic disease. Available at: <http://medrxiv.org/lookup/doi/10.1101/2022.07.25.22278008> (accessed 25 September 2024).
- Yoshitomi S, Takahashi Y, Ishizuka M, et al. (2015) Three patients manifesting early infantile epileptic spasms associated with 2q24.3 microduplications. *Brain and Development* 37(9): 874–879.
- Zemojtel T, Köhler S, Mackenroth L, et al. (2014) Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Science Translational Medicine* 6(252).
- Zuberi SM, Wirrell E, Yozawitz E, et al. (2022) 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* 63(6): 1349–1397.

ANNEXURES

List of publications from Thesis

1. **Alfiya, F.**, Jose, M., Chandrasekharan, S. V., Sundaram, S., Urulangodi, M., Thomas, B., Radhakrishnan, A., Banerjee, M., & Menon, R. N. (2022). *C12orf57* pathogenic variants: a unique cause of developmental encephalopathy in a south Indian child. *Journal of genetics*, 101, 30.
2. **F, Alfiya.**, Jose, M., Rudrabhatla, P. K., Chandrasekharan, S. V., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. (2023). Visual-sensitive epilepsy in GLUT-1 deficiency syndrome: Expanding the phenotype. *Epileptic disorders : international epilepsy journal with videotape*, 25(2), 265–268. <https://doi.org/10.1002/epd2.20008>
3. Krishna, S.*, **Fasaludeen, A*.**, Jose, M.*, Banerjee, M., Sundaram, S., Radhakrishnan, A., & Menon, R. N. (2024). Impact of variant subtype on electro-clinical phenotype of Dravet syndrome- a South Indian cohort study. *Seizure*, 115, 81–86. Advance online publication. <https://doi.org/10.1016/j.seizure.2024.01.004>
4. Jose, M*., **Fasaludeen, A*.**, Pavuluri, H., Rudrabhatla, P. K., Chandrasekharan, S. V., Jose, J., Banerjee, M., Sundaram, S., Radhakrishnan, A., & Menon, R. N. (2024). Metabolic causes of pediatric developmental & epileptic encephalopathies (DEE)- genetic variant analysis in a south Indian cohort. *Seizure*, 115, 20–27. Advance online publication. <https://doi.org/10.1016/j.seizure.2023.12.017>
5. Harini Pavuluri*, Manna Jose*, **Alfiya F***, Soumya Sundaram, Ashalatha Radhakrishnan, Moinak Banerjee, Ramshekhar N Menon (2023). Arginase deficiency- an unheralded cause of developmental epileptic encephalopathy. *Epileptic Disorders* (2023) doi: 10.1002/epd2.20081.
6. Jose, M*., **Fasaludeen, A*.**, Pavuluri, H., Rudrabhatla, P. K., Chandrasekharan, S. V., Jose, J., Banerjee, M., Sundaram, S., Radhakrishnan, A., & Menon, R. N. (2024). Challenges in genetic testing for metabolic causes of developmental epileptic encephalopathy- relevance of genotype-phenotype correlations. *Seizure*, 117, 307–308. <https://doi.org/10.1016/j.seizure.2024.02.012>
7. **Fasaludeen, A.**, McTague, A., Jose, M., Banerjee, M., Sundaram, S., Madhusoodanan, U. K., Radhakrishnan, A., & Menon, R. N. (2024). Genetic variant interpretation for the neurologist - A pragmatic approach in the next-generation sequencing era in childhood epilepsy. *Epilepsy research*, 201, 107341. <https://doi.org/10.1016/j.eplepsyres.2024.107341>

Curriculum Vitae

ALFIYA FASALUDEEN

PhD Scholar, R Madhavan Nayar Center for Comprehensive Epilepsy Care,
Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology,
Thiruvananthapuram, Kerala, India.

Email: alfiyafazal31@gmail.com; Contact No: 00919567905436

Experience

- **Senior Research Fellow** 06.2020 to 08.2023

Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram

- Performed whole exome sequencing analysis of 175 patients diagnosed with developmental epileptic encephalopathy.
- Processed raw NGS data (fastq) generated from Illumina sequencing platform.
- Conducted variant calling from BAM files using GATK-HaplotypeCaller tool and annotation using ANNOVAR tool.
- Utilized Integrated Genomic Viewer (IGV) tool for visualization of genomic data.
- Classified variants according to ACMG classification criteria using InterVar and Franklin databases.
- Performed molecular dynamics simulation of variants using GROMACS software.
- Published 6 case studies in international journals.

- **Junior Research Fellow** 10.2018 to 05.2020

Human Molecular Genetics Lab, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram

- Assessed the genetic and epigenetic landscapes of methylation data across ethnicities
- Utilized different R-packages (RnBeads, minfi, missMethyl, DMRCate, etc.) for the analysis of Illumina Infinium 450K methylation array data

Education

- **PhD** 01.03.2021 to Present

Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram

Title: Genotype-phenotype Correlations in Complex Childhood Epilepsies and Developmental Epileptic Encephalopathies in an Indian Cohort

- Conducted whole exome sequencing analysis of 159 trios with probands being diagnosed to have developmental epileptic encephalopathy
- Performed *in silico* functional analysis of pathogenic variants
- Conducted molecular dynamics simulation of variants of uncertain significance to predict the pathogenicity
- Provided guidance to over 20 Undergraduate/Graduate students in the completion of their dissertations.
- Published 14 articles including case studies.

- **Master of Philosophy in Bioinformatics** 2015-2016

Department of Computational Biology and Bioinformatics, University of Kerala, Thiruvananthapuram

- Grade: A
- Dissertation: "Analysis of Protein annotation methods"

- **Master of Technology in Technology Management** 2012-2014

Department of Futures Studies, University of Kerala

- Class: First class with Distinction

- Dissertation: Prediction of divergence of Dengue virus serotypes by comparative analysis of nucleotide and amino acid sequences
- **Bachelor of Technology in Biotechnology** 2008-2012
Sree Buddha College of Engineering, University of Kerala
 - CGPA: 7.0; First class
 - Dissertation: Antimicrobial activity of microorganism in sea water
 - GATE-2011 qualified
- **Secondary Education**
 - **Class 12 |2008|**
Board of Higher secondary education, Govt. of Kerala;
Major: Mathematics, Physics, Chemistry, Biology – 89%
 - **Class 10 |2006|**
Secondary School Leaving Certificate, Govt. of Kerala;
General- 93%

Areas of Interest

- Clinical genetics; Clinical Bioinformatics; Genomics; Real world data; Next generation sequencing; Molecular Biology; Modelling and simulation

Key Skills

- Expertise in NGS data analysis
- Professional level expertise in bash scripts and R Programming languages
- Intermediate level knowledge in Perl, BioPerl, Python and C++
- Bioinformatics Tools and databases: FastQC, Samtools, BWA, Bcftools, VCFtools, GATK, IGV, VARSCAN, ANNOVAR, XHMM CNV caller, CoNIFER, GROMACS, Databases (Franklin database, InterVar, UCSC genome browser, Ensembl genome browser, ClinVar, OMIM, dbSNP, etc)
- Molecular Biology Techniques: DNA isolation from human blood, Primer designing, Polymerase Chain Reaction, Gel electrophoresis, Sanger Sequencing

Publications

1. **Fasaludeen, A.**, McTague, A., Jose, M., Banerjee, M., Sundaram, S., Madhusoodanan, U. K., Radhakrishnan, A., & Menon, R. N. (2024). Genetic variant interpretation for the neurologist - A pragmatic approach in the next-generation sequencing era in childhood epilepsy. *Epilepsy research*, 201, 107341. <https://doi.org/10.1016/j.eplepsyres.2024.107341>
2. **Fasaludeen A.**, Jose, M., Rudrabhatla, P. K., Chandrasekharan, S. V., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. (2023). Visual-sensitive epilepsy in GLUT-1 deficiency syndrome: Expanding the phenotype. *Epileptic disorders: international epilepsy journal with videotape*, 25(2), 265–268. <https://doi.org/10.1002/epd2.20008>
3. **Alfiya, F.**, Jose, M., Chandrasekharan, S. V., Sundaram, S., Urulangodi, M., Thomas, B., Radhakrishnan, A., Banerjee, M., & Menon, R. N. (2022). *C12orf57* pathogenic variants: a unique cause of developmental encephalopathy in a south Indian child. *Journal of genetics*, 101, 30.
4. Krishna, S.*, **Fasaludeen, A*.**, Jose, M.*, Banerjee, M., Sundaram, S., Radhakrishnan, A., & Menon, R. N. (2024). Impact of variant subtype on electro-clinical phenotype of Dravet syndrome- a South Indian cohort study. *Seizure*, 115, 81–86. Advance online publication. <https://doi.org/10.1016/j.seizure.2024.01.004>
5. Jose, M.*, **Fasaludeen, A*.**, Pavuluri, H., Rudrabhatla, P. K., Chandrasekharan, S. V., Jose, J., Banerjee, M., Sundaram, S., Radhakrishnan, A., & Menon, R. N. (2024).

Metabolic causes of pediatric developmental & epileptic encephalopathies (DEE)- genetic variant analysis in a south Indian cohort. *Seizure*, 115, 20–27. Advance online publication. <https://doi.org/10.1016/j.seizure.2023.12.017>

6. Saranya B. Gomathy, Ajit Valaparambil Karthika, Balaswamy Reddy, Adarsh Anilkumar, **Alfiya Fasaludeen**, Soumya Sundaram, Bejoy Thomas, Ramshekhar N. Menon. (2024). De Barsy syndrome due to ALDH18A1 mutation – Expanding the spectrum of a rare neurocutaneous syndrome, *Human Gene*, 40, 201284, <https://doi.org/10.1016/j.humgen.2024.201284>.
7. Pavuluri, H., F, **Alfiya.**, Menon, R. N., Nair, S. S., & Sundaram, S. (2022). Pyruvate Dehydrogenase Complex Deficiency Due to PDHA1 Mutation-A Rare Treatable Cause for Episodic Ataxia in Children. *Indian journal of pediatrics*, 89(5), 519. <https://doi.org/10.1007/s12098-021-04068-x>
8. Pavuluri, H., Jose, M., **Fasaludeen, A.**, Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. (2023). Arginase deficiency-An unheralded cause of developmental epileptic encephalopathy. *Epileptic disorders : international epilepsy journal with videotape*, 25(4), 556–561. <https://doi.org/10.1002/epd2.20081>
9. Rudrabhatla, P. K., Divya, K. P., **Fasaludeen, A.**, Menon, R. N., Cherian, A., Urulangodi, M., & Sundaram, S. (2023). Generalized Stiffness in Hereditary Hyperekplexia Responsive to Trihexyphenidyl: A Novel Finding. *Clinical pediatrics*, 99228231203300. Advance online publication. <https://doi.org/10.1177/00099228231203300>
10. Valaparambil, K. A., **Fasaludeen, A.**, Priya, L., Menon, R. N., Menon, R., & Sundaram, S. (2023). Clinical Utility of Proband Only Clinical Exome Sequencing in Neurodevelopmental Disorders. *Indian journal of pediatrics*, 10.1007/s12098-023-04916-y. Advance online publication. <https://doi.org/10.1007/s12098-023-04916-y>

Awards

- DST-SERB International Travel Scheme award for the platform presentation "Comparing the diagnostic yield and clinical utility of Trio-Whole exome sequencing and Clinical exome sequencing in children with Developmental epileptic encephalopathy" at International Child Neurology Congress 2024 (ICNC2024), Cape Town, South Africa, May 6-10, 2024.
- Second prize for the award paper presentation titled “Diagnostic yield of Trio-Whole exome sequencing and Clinical exome sequencing in children with Developmental and epileptic encephalopathy” at Epilepsy conference, ECON23, Marriot hotel, Jaipur, Rajasthan, 21st to 23rd July, 2023.
- Qualified Graduate Aptitude Test in Engineering, Govt. of India, 2011.

International Conference Presentations

1. “Functional and structural characterization of variants of uncertain significance identified from whole exome sequencing of developmental and epileptic encephalopathies in an Indian cohort: an *In Silico* approach”; Digital poster presentation; 35th International Epilepsy Congress, 2-6 September, 2023, Dublin, Ireland
2. “Metabolic basis of pediatric developmental and epileptic encephalopathies (DEE)- genetic variant analysis in a South Indian cohort”; Digital poster presentation; 35th International Epilepsy Congress, 2-6 September, 2023, Dublin, Ireland
3. “Comparative analysis of Whole Exome Sequencing and Global Screening Array yield in evaluation of childhood developmental epileptic encephalopathies”; Digital poster presentation; American Epilepsy Society Annual Meeting (AES 2021), Dec 2-9; 2021.

National Conference Presentations

Oral Presentations

1. **Alfiya, F.**, Jose, M., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. Denovo mutations in Developmental and Epileptic Encephalopathies (DEE) in an Indian cohort, Epilepsy Conference ECON2022, August 2022
2. **Alfiya, F.**, Jose, M., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. Genetic Diagnostic Yield of Childhood Developmental Epileptic Encephalopathies: A Comparative Study of Whole Exome Sequencing and Global Screening Array, RMNC silver jubilee conference, August 2022
3. **Alfiya, F.**, Jose, M., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. Denovo mutation identified in SCN1A gene associated with Dravet syndrome using whole exome sequencing at the 8th International Symposium on Computational Biology and Bioinformatics, Bioinformatica Indica, January 6th -8th 2022.
4. “Diagnostic yield of Trio-Whole exome sequencing and Clinical exome sequencing in children with Developmental and epileptic encephalopathy” at Epilepsy conference, ECON23, Marriot hotel, Jaipur, Rajasthan, 21st to 23rd July, 2023.
5. “Pathogenic Metabolic gene variations in Developmental and epileptic encephalopathy” at the National Seminar: Epigenetics meets Metabolomics, Institute for Communicative and Cognitive Neurosciences (ICCON), Shoranur, Kerala, June 2-3, 2023.

Poster presentations

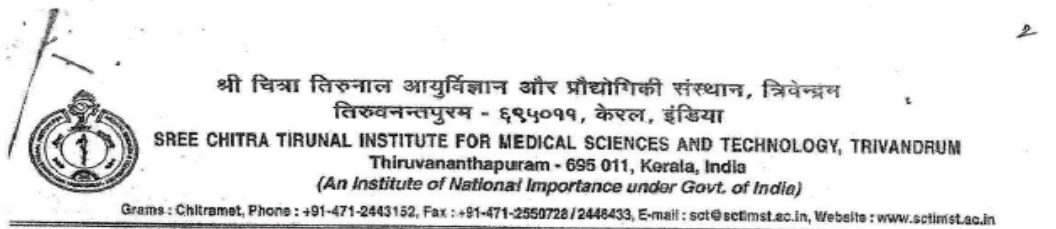
1. **Alfiya, F.**, Jose, M., Sundaram, S., Radhakrishnan, A., Banerjee, M., & Menon, R. N. A novel ITPA variant associated with Developmental and Epileptic Encephalopathy (DEE), ECON, August 2022
2. **Alfiya, F.**, Jose, M., Banerjee, M., & Menon, R. N. Comparing Yield of Whole Exome Sequencing and Global Screening Array for Genetic Diagnosis of Childhood Developmental Epileptic Encephalopathies” Science day presentations, SCTIMST, February 2022

References:

1. **Dr. Ramshekhar N Menon:** Professor, Dept. of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India. E-mail: rsnmenon@yahoo.com
2. **Dr. Moinak Banerjee:** Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Trivandrum, Kerala, India. E-mail: moinak@gmail.com
3. **Dr. Achuthsankar S Nair:** Professor, Department of Computational Biology and Bioinformatics, Trivandrum, Kerala, India. E-mail: achuthsankar@keralauniversity.ac.in

Appendices

APPENDIX A – ETHICS COMMITTEE APPROVAL



Institutional Ethics Committee (IEC Regn No. ECR/189/Inst/KL/2013/RR-16)

SCT/IEC/1473/DECEMBER-2019

24.12.2019

Dr. Ramshekhar N Menon
Additional Professor
Department of Neurology
SCTIMST, Thiruvananthapuram

Dear Dr. Ramshekhar N Menon,

The Institutional Ethics Committee reviewed and discussed your application to conduct the study entitled "GENETICS OF COMPLEX PEDIATRIC EPILEPSY SYNDROMES: ELECTROCLINICO-IMAGING BASED GENOTYPE- PHENOTYPE CORRELATIONS IN AN INDIAN COHORT (IEC/1473)" on 21st December, 2019.

The following documents were reviewed:

1. Covering Letter addressed to the Chairman, IEC, SCTIMST dated 27.11.2019 with checklist
2. Response to comments
3. TAC Approval Letter
4. IEC Application Form
5. Project Proposal
6. Subject Information and Informed Consent Form in English and Malayalam
7. Memorandum of Understanding
8. Approval from funding agency (ICMR)
9. Declaration Form
10. CV of Principal Investigator and Co-Principal Investigators

Page 1 of 2

The following members of the Ethics Committee were present at the meeting held on 21st December, 2019 at G. Parthasarathi Board Room, AMCHSS, SCTIMST

SL. No.	Member Name	Highest Degree	Gender	Scientific /Non Scientific	Affiliation with Institution(s)
1.	Dr. R V G Menon	M Tech. PhD	Male	Lay Person (Chairman)	No
2.	Dr. Kala Kesavan. P	MBBS, MD	Female	Basic Medical Scientist	No
3.	Dr. K.R.S. Krishnan	M.E., Ph.D.	Male	Medical Technology	Yes
4.	Dr. Anand Kumar A	MD, DM	Male	Clinician	No
5.	Dr. Rema M. N	MD	Female	Basic Medical Scientist	No
6.	Dr. Hankrishna Varma PR	Ph.D (Materials Science)	Male	Medical Technology	Yes
7.	Dr. V. Raman Kutty	M D, M Phil, M P H	Male	Health Sciences Expert/Clinician	Yes
8.	Dr. P. Manickam	BSMS, MSc (Epid), PhD	Male	Health Science Expert/ Social Scientist	No
9.	Dr. Aneesh V Pillai	BA, LLB (Hons.), LL.M, Ph. D, SET (Law)	Male	Legal Expert	No
10.	Dr. Hankrishnan S	MD, DM (Cardiology) DNB (Cardiology)	Male	Clinician	Yes
11.	Dr. Mala Ramanathan	PhD	Female	Social Scientist (Member Secretary)	Yes

IEC Decision

The IEC approved the conduct of the study in the present form.

Remarks:

The Institutional Ethics Committee expects to be informed about the progress of the study, any SAE occurring in the course of the study, any changes in the protocol and patient information/informed consent and asks to be provided a copy of the final report.

There was no member of the study team who participated in voting / decision making process. The ethics committee is organized and operated according to the requirements of Good Clinical Practice and the requirements of the Indian Council of Medical Research (ICMR).

Sincerely,



Mala Ramanathan
Member Secretary, IEC

APPENDIX B – SUPPLEMENTARY TABLES

Supplementary Table 1: Demographic details of the patients

Patient ID	Gender (1-Male;2-Female)	Age (months)	Age-onset of epilepsy (months)	Consanguinity (1= Yes;0= No)	Trio-1; Quad-2; Duo-3;Proband only-4	DEE Type: 1-EME;2-OS;3- EIEE<6 months of age of onset with focal or multifocal epilepsy;4-WS;5-DS;6-MAE;7-LGS;8-EIMFS;9-LKS/CSWS;10-other myoclonic epilepsy; 11-other focal/multifocal epilepsy;12-refractory generalized epilepsy;13-PME	Family history (0-No;1-Yes)	MRI Brain(0-Normal;1-Positive)	Chief complaints(0-Developmental delay;1-ASD;2-seizures)
EP1	1	65	1.5	0	1	4	1	0	0,1,2
EP2	2	68	1.5	0	1	3	0	0	0,2
EP3	1	108	36	0	1	7	1	1	0,2
EP4	2	23	0.1	0	3	3	1	1	0,2
EP5	2	108	108	0	4	11	0	1	2
EP6	1	12	4	0	1	4	1	0	0,2
EP7	1	64	12	0	1	11	1	0	0,2
EP8	2	47	24	0	1	10	0	1	0,1,2
EP9	1	100	18	0	1	10	0	0	0,2
EP10	2	93	42	0	3	7	1	1	2
EP11	1	50	2	0	1	10	1	1	0,2
EP12	1	64	36	0	1	11	1	1	0,2
EP13	2	127	6	0	1	11	0	0	0,2
EP14	1	126	2	0	1	11	1	0	0,2
EP15	1	77	1	0	1	4	0	1	0,2
EP16	2	8	2	0	1	11	1	1	0,2
EP17	1	37	33	0	1	6	0	1	2
EP18	2	45	9	0	3	5	1	0	1,2
EP19	2	216	18	0	1	11	0	1	2
EP20	1	107	30	0	1	9	0	1	1,2
EP21	2	53	6	0	1	10	0	1	0,2
EP22	1	110	36	0	1	6	1	0	2
EP23	1	50	12	0	1	9	0	0	0,2
EP24	1	31	15	0	1	10	0	1	0,2
EP25	2	12	7	0	1	5	0	0	2
EP26	1	37	9	0	1	11	0	1	2
EP27	2	64	3.5	0	1	4	0	1	0,2
EP28	2	30	5	0	1	4	0	0	0,2
EP29	1	9	3	0	1	2	0	0	0,2
EP30	1	52	22	0	1	4	0	0	0,2
EP31	1	34	0.2	1	1	11	1	1	0,2
EP32	1	66	9	0	1	13	0	1	0,2
EP33	2	30	15	0	1	10	1	0	0,2
EP34	1	120	24	0	2	11	1	0	0,2
EP35	2	35	5	0	2	11	1	1	0,2
EP36	1	19	4	0	1	11	0	1	0,2
EP37	2	144	60	0	1	11	1	0	0,2
EP38	1	147	6	0	1	12	0	0	0,2
EP39	1	34	30	0	1	4	1	0	0,1,2
EP40	2	66	36	1	1	6	0	0	2
EP41	2	59	48	0	1	11	1	1	2
EP42	1	61	36	0	1	9	1	1	2
EP43	1	91	29	0	1	9	1	0	2
EP44	1	168	8	0	4	10	0	0	0,2
EP45	1	81	9	0	1	5	1	1	2
EP46	2	72	18	0	1	10	0	0	0,2
EP47	2	64	30	1	1	7	1	1	0,2
EP48	1	162	42	1	1	10	1	1	0,2
EP49	2	77	48	0	1	9	0	0	0,1,2
EP50	1	178	6	0	3	3	0	0	0,2

Patient ID	Gender (1-Male;2-Female)	Age (months)	Age-onset of epilepsy (months)	Consanguinity (1= Yes;0= No)	Trio-1; Quad-2; Duo-3;Proband only-4	DEE Type: 1-EME;2-OS;3- EIEE<6 months of age of onset with focal or multifocal epilepsy;4-WS;5-DS;6-MAE;7-LGS;8-EIMFS;9-LKS/CSWS;10-other myoclonic epilepsy; 11-other focal/multifocal epilepsy;12-refractory generalized epilepsy;13-PME	Family history (0-No;1-Yes)	MRI Brain(0-Normal;1-Positive)	Chief complaints(0-Developmental delay;1-ASD;2-seizures)
EP51	1	68	4	0	1	3	0	0	2
EP52	2	14	7	0	1	12	1	0	2
EP53	2	150	12	0	1	11	0	1	0,2
EP54	1	75	19	0	1	4	0	0	0,2
EP55	2	66	60	1	1	7	0	0	0,1,2
EP56	2	216	12	0	1	11	0	0	0,2
EP57	1	83	6	1	1	3	1	1	0,2
EP58	2	74	42	0	1	9	1	1	2
EP59	2	50	30	0	1	8	0	1	0,2
EP60	1	62	18	0	1	12	0	0	0,2
EP61	1	75	68	0	1	11	1	1	0,2
EP62	1	111	6	1	1	4	0	1	0,1,2
EP63	1	85	60	0	1	13	0	1	0,2
EP64	2	216	8	0	1	11	0	0	0,2
EP65	1	216	5	0	3	3	1	0	0,2
EP66	2	207	84	0	3	12	0	0	2
EP67	2	90	18	0	1	7	0	1	0,2
EP68	2	43	66	0	4	4	0	0	0,2
EP69	1	42	30	0	1	11	1	1	2
EP70	1	133	15	0	1	6	1	0	0,2
EP71	2	175	12	1	2	13	1	1	0,2
EP72	1	153	132	1	2	13	1	0	2
EP73	1	68	1	0	1	3	1	1	0,2
EP74	1	157	48	0	1	12	1	0	0,1,2
EP75	1	75	6	0	1	12	1	1	0,2
EP76	2	92	1	0	1	7	1	1	0,2
EP77	1	18	5	0	1	3	1	1	0,2
EP78	1	41	4.5	0	1	3	1	1	0,1,2
EP79	2	27	18	0	1	11	0	0	0,1,2
EP80	1	187	42	1	1	9	1	0	0,2
EP81	2	82	3	0	1	5	1	1	0,2
EP82	2	88	30	0	1	4	1	1	0,1,2
EP83	1	207	204	0	1	11	1	1	2
EP84	1	104	84	0	1	11	0	0	2
EP85	1	19	8	0	1	4	1	1	2
EP86	1	129	36	0	1	9	1	0	0,2
EP87	1	170	6	0	1	4	1	1	0,2
EP88	1	141	7	0	1	11	0	0	0,2
EP89	2	73	11	0	1	12	0	1	0,2
EP90	1	108	3	0	4	9	0	1	0,2
EP91	1	136	96	0	1	9	1	0	2
EP92	2	38	4	0	1	3	0	1	2
EP93	2	198	48	0	1	9	0	1	0,1,2
EP94	1	5	4	1	1	3	1	1	0,2
EP95	1	22	6	0	1	4	0	1	0,2
EP96	2	134	12	0	1	11	0	0	0,1,2
EP97	2	41	6	0	1	4	0	0	0,2
EP98	2	216	12	0	1	7	0	0	2
EP99	1	196	12	0	1	12	0	0	0,1,2
EP100	2	133	36	0	1	11	0	1	2

Patient ID	Gender (1-Male;2-Female)	Age (months)	Age-onset of epilepsy (months)	Consanguinity (1= Yes;0= No)	Trio-1; Quad-2; Duo-3;Proband only-4	DEE Type: 1-EME;2-OS;3- EIEE<6 months of age of onset with focal or multifocal epilepsy;4-WS;5-DS;6-MAE;7-LGS;8-EIMFS;9-LKS/CSWS;10-other myoclonic epilepsy; 11-other focal/multifocal epilepsy;12-refractory generalized epilepsy;13-PME	Family history (0-No;1-Yes)	MRI Brain(0-Normal;1-Positive)	Chief complaints(0-Developmental delay;1-ASD;2-seizures)
EP101	1	124	9	0	1	11	0	0	0,2
EP102	2	163	39	0	1	12	0	0	0,2
EP103	1	185	168	0	1	11	1	1	2
EP104	2	216	84	0	1	11	1	0	0,2
EP105	1	114	6	0	1	3	1	0	0,2
EP106	2	132	7	0	1	11	1	1	0,2
EP107	2	102	42	0	1	11	0	1	0,2
EP108	1	101	18	0	1	6	0	1	0,2
EP109	2	56	30	0	1	11	1	0	0,2
EP110	1	45	12	0	1	11	1	1	0,1,2
EP111	1	68	18	0	1	11	1	1	0,2
EP112	2	189	24	0	1	9	1	0	0,2
EP113	1	120	12	0	1	3	1	0	0,2
EP114	2	76	36	0	1	12	1	1	0,2
EP115	1	216	204	0	1	12	1	1	0,2
EP116	1	165	3	0	4	12	1	1	0,1,2
EP117	2	42	1	0	1	3	0	0	0,2
EP118	2	119	48	1	1	12	1	0	0,2
EP119	1	67	12	1	2	12	1	0	0,2
EP120	2	75	48	0	2	11	1	1	0,2
EP121	2	19	4	1	1	3	0	0	0,2
EP122	1	38	33	0	1	11	0	0	0,2
EP123	1	4	0.1	0	1	12	0	1	0,2
EP124	1	72	18	0	1	9	1	1	0,2
EP125	2	162	6	0	1	12	0	0	0,2
EP126	1	74	18	0	1	11	0	1	0,2
EP127	1	184	60	0	1	7	0	0	0,2
EP128	2	52	3	0	1	4	0	1	0,2
EP129	1	103	0.1	0	1	4	0	0	0,2
EP130	2	157	84	0	1	9	0	0	2
EP131	2	35	32	0	1	13	1	0	2
EP132	2	63	0.1	0	4	4	1	1	0,1,2
EP133	2	176	6	0	1	5	1	1	0,2
EP134	1	48	16	0	1	10	1	0	0,2
EP135	1	52	6	1	1	9	1	0	0,2
EP136	1	154	24	1	2	9	1	0	0,2
EP137	1	42	8	0	2	5	1	0	0,2
EP138	1	168	108	0	4	11	0	1	2
EP139	2	36	18	0	1	4	1	0	0,1,2
EP140	1	170	5	0	1	11	1	0	0,2
EP141	2	16	10	0	2	11	1	1	0,2
EP142	1	12	5	1	4	7	0	1	0,2
EP143	2	48	3	0	2	4	1	1	0,1,2
EP144	1	46	3	1	1	4	1	0	0,2
EP145	1	154	48	1	2	7	1	1	0,1,2
EP146	1	136	18	0	1	9	1	0	0,2
EP147	2	155	7	1	1	4	1	1	0,1,2
EP148	1	112	6	0	2	5	0	1	0,2
EP149	1	134	12	0	4	11	0	1	0,2
EP150	1	81	21	0	1	9	1	0	0,2

Patient ID	Gender (1-Male;2-Female)	Age (months)	Age-onset of epilepsy (months)	Consanguinity (1= Yes;0= No)	Trio-1; Quad-2; Duo-3;Proband only-4	DEE Type: 1-EME;2-OS;3- EIEE<6 months of age of onset with focal or multifocal epilepsy;4-WS;5-DS;6-MAE;7-LGS;8-EIMFS;9-LKS/CSWS;10-other myoclonic epilepsy; 11-other focal/multifocal epilepsy;12-refractory generalized epilepsy;13-PME	Family history (0-No;1-Yes)	MRI Brain(0-Normal;1-Positive)	Chief complaints(0-Developmental delay;1-ASD;2-seizures)
EP151	1	143	6	0	1	3	1	0	0,2
EP152	2	121	8	0	2	7	0	0	0,2
EP153	2	145	3	0	1	3	1	0	0,2
EP154	1	94	9	0	2	11	1	0	0,2
EP155	2	64	2.5	0	2	5	0	0	0,2
EP156	2	22	0.1	1	4	3	1	0	0,2
EP157	1	8	2	0	1	3	1	1	0,2
EP158	2	18	6	0	1	3	1	1	0,2
EP159	2	90	7	0	1	5	0	1	0,2
EP160	1	21	8	0	1	11	1	1	0,2
EP161	1	2	0.1	0	1	8	1	1	0,2
EP162	2	94	48	0	1	10	0	0	2
EP163	1	81	0.1	0	1	10	0	1	0,2
EP164	1	96	60	0	1	11	0	0	0,2
EP165	2	194	8	0	1	11	1	0	0,2
EP166	2	33	7	0	4	3	1	0	0,2
EP167	1	216	3	0	1	5	1	0	0,2
EP168	2	216	0.1	0	1	3	1	0	0,2
EP169	1	216	18	0	2	3	1	1	0,2
EP170	1	180	6	0	2	7	0	0	0,2
EP171	2	45	30	0	1	9	0	0	2
EP172	1	34	12	0	1	4	1	1	0,2
EP173	1	160	18	0	1	11	1	0	0,2
EP174	1	63	42	0	1	9	1	0	0,2
EP175	2	72	24	0	2	12	1	0	2

Supplementary Table 2: List of primers used for Sanger sequencing

Sl No	Gene/variant	Forward primer	Reverse primer
1	C12ORF57	CATACGCAGCAGTGTTACAGC	CTTAGAAGGACGGTACCCGG
2	DCX/rs587783563	GCCTTAGGGGAAAGCTATGTC	ATGTGGAGGAAGAGTCCGTC
3	SCN1A/rs121917952	GCACCCATTCTCAACAGTAAG	GAGTTTGTGGTTGTGGCAG
4	ARG1/rs377280518	GTTGTAGGGCTACTCTCAG	CCAAGAGGGAATTTGTAGAG
5	SLC2A1	GAGAATGCTGGGCCAGCAC	GCCCAATGCTGGATACACAAC
6	KCNT1/rs397515404	CAGGATGGACAATGGGGAGG	GACGCACCAGCAAACCTTGAC
7	TSC2/rs137853977	CAGACAATGGCCTGTTGAG	GACAGAGGAAGGAGAGGAC
8	SCN8A	CAGGCTGATGCATAGGAA	GGGATTTGTGCTAGCTGC
9	CNDL5	AGCATGGGATGACTTCGA	CAGCTGAGACCCCTAAAC
10	NHLRC1	GGGTACCTAATGTCTTGG	GTACACAATCCATCCATG
11	ITPA	CCACTGCTTCTGTCAAC	ACTCCCACCAAGTACAGC
12	CLN5/rs267606738	ACAGTGGGATCAGGGAAC	TCTTCTCACCTTCCAGGA
13	GABRB2	GTTTGGTCACTCCTCTTG	GGGATCAAACGTTCACTG
14	BTD/rs1163419871	CAATGGCCTCTGCTGTTA	GCTTGAGACCAATCGCAT
15	ALG3	TGGCGCTCAGGTAAGGGATA	AAGACTGGCTACTCGGATCG
16	KCTD7	ATAAGTCCTTGTACCCGAC	CTCTCACTCCGCTCAAATCG
17	SCN3A	GCTATGTTCCCTCGCTGTC	CATTCAAACGAAGAATCATCAGG
18	TSC1	GAAAGCATTGTAGTGTCTC	AGCCTTTACTCCCATAGACC
19	CHRNA4	CACCGTGTGGTCTTCTACC	TACTGGATGCTCCGAGACCG
20	SETD1B	CCGGAAGGGTTATTCTCTCG	GGCGGTTGCTGGACATCT
21	SETD1A/ rs543794172	CTGCATCTTCAGCCTCCAC	CTGTAGGTGCCACATCCTC
22	STXBP1	GTTACCTTCTCAGCAAGTCC	TCCTGATCTTCTTGAGACAAC
23	GABRA1	ATATCCCCCTCAATGCTGCC	CCATCCAATACCACAGGCCA
24	PCDH19	AGCTTTGGCGTGCAGACTTA	CCGTCCGTCCACCAGAATAG
25	AKT3	ACAATGGTGGGCTCATGACT	CAGACTTTGGATGCTGTGG
26	SCN2A	TGTGCCAGGTTGTTGTAAATGC	CTGCCCCCTCCCTGACTACTT
27	SZT2/ rs563782968	TAGAAGAGGGTGAGGTGGGG	ATCTCTGGCTCTGAACCCCT
28	SZT2/ rs139147641	TAGAAGAGGGTGAGGTGGGG	GGTCCAACTCAATGACAAA
29	SCN1A	GTTAAGGTCATGGTGCCAGT	TCTCAAGGTTGCCGTTCTGT
30	GABRG2	AGGACGTCAGAAAGCACAG	CCTTGCGATAGTTTGTGAG
31	SETD1B	GGAATGACCCTCAAGGTTC	TGGGGTCTGATACGGAGAG
32	KMT2C	GGTAACATGGGAAAAAGGAGA	CTGGCACTCCTGCATAAAC
33	EP4SCN8A	GTATGAAAGGATGGCACTTC	GCTGCTAACATCCTCTGTG
34	SHQ1	CAGTGTGCTTTATTTAGCAC	CTTGCCAATGTGAAGTGTC
35	EP157/KCNT1	TTAAGGTTCCGGACCTGC	GCTCCAACCTGTTCTCAAG
36	CHD8	GTACCAGATGTTTGCTCTGC	AATAAGTGATCAGTCTCCC
37	GABBR2	CTGAGCTTGCTTGTCTGCTTC	CAGGACACTGGGAATGTGG
38	MT-ATP6	CTGAGAACCAAAATGAACG	GGCTTGATTAAGGCGAC
39	NPRL2	CCACTTGAAGGTGATTGAGCA	CCACTACCTTGCTTGGTCAC
40	KCNT1/rs146193090	GACTCAGCTCATCCTGGA	TGTAAAGAGAAGTGGACTGC
41	SCN1A/rs1064795736	ATGGTCCCAGCACATCCTTT	GACGGTTAGGGCAGATCAGA
42	SCN1A/rs121917986	TTGTTTGGTCTGGTCACTAC	CAATTCTGACTTCATCTGGC
43	NOVA2	GTCATCAGGAGTCCCACCAT	GCAAGTGTGTGGAACATGGC
44	EP133/NR2F1	TGTGCCTCTCTCTCCAGCT	TTTCCTCGCCAACGTGCT
45	GNAO1/rs797044878	ATGGGTGGCTCACTTCTG	TGTTGCCTGTACATTCTTGC
46	WDR45	CTTCCTTGTTGACTCTTGGC	GGTAAAAGGTCAGAGGCTGC
47	ALDH18A1	GGTCATGTTGTAAAGTCCC	GAGTGTGTAGCGTAATC
48	EP144/SCN1A	AGAGTGATGCCAAGGATTG	CTCACTTGAAAAATACCCCA
49	EP139/KCNT1	ATTTCGGGATGGCAAGAG	AGCACCACATGACCCACA
50	EP148/SCN1A	GATCCGCACGCTGCTCTTT	AAGAAGATTAGCCCCACCTTTG
51	UFSP2	ACCTATGCGTGGCTTTAACG	GCATGCTACCCTACTTCTC
52	GNAO1/rs886041715	AGTCTCCGCTGCTGGAAT	GCAATGTGGAGTAAGCAATG
53	SCN1A/rs794726736	TGTTTACCACGTCCATTGCC	GAATCATTCTCAAGGTTGCC
54	SCN1A/rs794726730	TGCCTTGATCACTTGTAGGAC	GATGAGGCTGAAAAGGGC

Received: 26 August 2022 | Accepted: 27 October 2022
DOI: 10.1002/epd2.20908

CASE VIGNETTE

Epileptic
Disorders

Visual-sensitive epilepsy in GLUT-1 deficiency syndrome: Expanding the phenotype

Alfiya Fasaludeen¹ | Manna Jose¹ | Pavan Kumar Rudrabhatla¹ |
Soumya V. Chandrasekharan¹ | Soumya Sundaram¹ | Ashalatha Radhakrishnan¹ |
Moinak Banerjee² | Ramshekhar N. Menon¹

¹Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences & Technology (SCTIMST), Thiruvananthapuram, Kerala, India

²Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala, India

Correspondence

Ramshekhar N. Menon, Department of Neurology, SCTIMST, Thiruvananthapuram, Kerala 695011, India.
Email: rsnmenon@yahoo.com

Funding information

Indian Council of Medical Research, Grant/Award Number: 2018-2670

Keywords: GLUT-1 deficiency syndrome, hypoglycorrhachia, next generation sequencing, reflex photosensitive epilepsy

Glucose, the primary fuel for brain energy metabolism, is transported by glucose transporter type 1 (GLUT-1) across the blood–brain barrier. Mutations in solute carrier family 2 member 1 (*SLC2A1*) gene result in defective transfer of glucose and a spectrum of neurological manifestations together called GLUT-1 deficiency syndrome (GLUT-1DS). The association between GLUT-1DS and intractable infantile epilepsy responsive to ketogenic diet (KD) has been clearly elucidated.^{1–3} Low fasting cerebrospinal fluid (CSF) glucose levels are demonstrated in most patients and provide a diagnostic clue, thereby driving early genetic confirmation and initiation of KD. Often, patients with GLUT-1DS are treated with antiepileptic medication (ASM) in the early stages which fails to address the underlying metabolic disturbance during early brain development. Studies show that low brain glucose supply early in life can lead to profound diminution of cerebral microvasculature.⁴ However, reflex phenotypes, such as photosensitive epilepsy or eyelid myoclonia with/without absence among the genetic generalized epilepsies (GGE) that have thalamo-cortical generators, have not been reported. Time delay to genetic confirmation and initiation of the ameliorative KD has not been addressed critically. We give an account of the clinical and genetic odyssey of a child with a novel visual-sensitive epilepsy phenotype associated with GLUT-1DS who was evaluated using whole-exome sequencing (WES) trio analysis, by virtue of variants of uncertain significance (VUS), detected using a limited

epilepsy gene panel. We also compared this phenotype with those from other patients with GLUT-1DS in our prospectively maintained database.

A 5-year-old girl, born of a nonconsanguineous parentage, presented with recurrent seizures and developmental delay with no prior antecedents or perinatal insults. From 6 months of age, the child was noted to have a tendency to look at bright objects, resulting in multiple brief episodes of eye blinks, after which she would cry holding her head on a daily to weekly basis. She was also noted to have moderate developmental delay, especially in language, and was prone to temper tantrums, reduced eye contact, and interest in rotating objects which worsened from 3 years of age. These paroxysmal events were refractory to ASMs including sodium valproate, ethosuximide, and clobazam. From 4 years of age, the seizure burden (severity, frequency, and self-induction) and behavioral issues worsened. On examination, she had no facial dysmorphism, microcephaly, neurocutaneous markers, or focal neurological deficits.

The developmental assessment was performed using the Revised Denver Developmental Scale (Denver II Technical Manual, 1990; Denver Developmental Materials, Inc.), Vineland Social Maturity Scale (VSMS), and Childhood Autism Rating Scale (CARS). She was found to have global delay; developmental quotients for the different domains were social—27, fine motor—42, language—30, and gross motor—21. Receptive Expressive Emergent Language Scale



RESEARCH NOTE

C12orf57 pathogenic variants: a unique cause of developmental encephalopathy in a south Indian child

F. ALFIYA¹, MANNA JOSE¹, SOUMYA V. CHANDRASEKHARAN¹, SOUMYA SUNDARAM¹, MADHUSOODANAN URULANGODI², BEJOY THOMAS³, ASHALATHA RADHAKRISHNAN¹, MOINAK BANERJEE⁴ and RAMSHEKHAR N. MENON^{1*}

¹Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology (SCTIMST), Thiruvananthapuram 695 011, India

²Department of Biochemistry, Sree Chitra Tirunal Institute for Medical Sciences and Technology (SCTIMST), Thiruvananthapuram 695 011, India

³Department of Imaging Sciences, Sree Chitra Tirunal Institute for Medical Sciences and Technology (SCTIMST), Thiruvananthapuram 695 011, India

⁴Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram 695 014, India

*For correspondence. E-mail: rsnmenon@yahoo.com.

Received 5 February 2022; revised 17 March 2022; accepted 1 April 2022

Abstract. Open reading frame variants which lack stop codons such as *C12orf57* variants are known to cause Temtamy syndrome, an extremely rare disorder characterized by intellectual disability, seizures, facial dysmorphism and agenesis of corpus callosum. *C12orf57* was initially reported to be required for human corpus callosum development. We report the first child who is of Indian origin with developmental and epileptic encephalopathy (DEE) with a unique phenotypic evolution as focal onset reflex seizures. We performed whole exome sequencing of genomic DNA isolated from peripheral blood samples of proband and his parents. Two pathogenic compound heterozygous variants, a start loss variant (Chr12:7053285:c.1A>G) and a premature stop gain variant (Chr12:7053327:c.43C>T), involving the *C12orf57* gene were identified in the proband. Our case report which details genotyping in this rare syndromic developmental encephalopathy, with no prior cases reported from India, expands the ethnic spectrum of patients.

Keywords. developmental encephalopathy; agenesis of corpus callosum; epilepsy; *C12orf57*; compound heterozygous.

Introduction

Temtamy syndrome is a rare disorder, which is caused by mutations involving open reading frame 57 (*C12orf57*) in chromosome 12. It was first described in 1991 and is phenotypically characterized by intellectual disability (ID), epilepsy, craniofacial dysmorphism, agenesis of the corpus callosum, and optic coloboma (Temtamy *et al.* 1991, 1996). No prior reports from the Indian subcontinent exist in literature focussing on potential mutations in this syndrome that lead to epileptogenesis and callosal anomalies. Pathogenic variants were first reported in 2013 in four siblings of a Saudi Arabian family with a deleterious nucleotide change (c.1A>G; p.Met1Val) and eight pathogenic variants have been reported till date (Salih *et al.* 2013; Alrakaf *et al.* 2018;

Wang *et al.* 2020) of which seven are of Middle Eastern and one of Asian (China) ethnicity. Homozygous and compound heterozygous mutations in *C12orf57* have been described (Platezer *et al.* 2014). We report the first Indian male with genetically proven Temtamy syndrome who as diagnosed using whole exome sequencing in trios to have compound heterozygous variants involving *C12orf57*. Informed consent for publication was obtained from the parents.

Case presentation

A male, born of south Indian nonconsanguineous parentage, presented to us in early childhood with global developmental delay and no history of perinatal insults. There was a history

Published online: 14 June 2022



Contents lists available at ScienceDirect

Seizure: European Journal of Epilepsy

journal homepage: www.elsevier.com/locate/seizure

Impact of variant subtype on electro-clinical phenotype of Dravet syndrome- a South Indian cohort study

S Krishna^{1,a}, Alfiya Fasaludeen^{1,a}, Manna Jose^{1,a}, Moinak Banerjee^b, Soumya Sundaram^c,
Ashalatha Radhakrishnan^a, Ramshekhar N Menon^{a,c,*}

^a R Madhavan Nayar Centre for Comprehensive Epilepsy Care, Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India

^b Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Trivandrum, Kerala, India

^c Comprehensive Care Centre for Neurodevelopmental Disorders, Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India

ARTICLE INFO

Keywords:

Dravet syndrome
SCN1A
Genetics
Pathogenic variants

ABSTRACT

Objective: We aimed to compare the electroclinical correlates of truncating and missense variants of *SCN1A* variants in children with Dravet syndrome (DS) and to determine phenotypic features in relation to variants identified and seizure outcomes.

Methods: A single center prospective study was carried out on a South Indian cohort. Patients below 18 years of age who met the clinical criteria for DS who had undergone genetic testing and completed a minimum of one year follow up were included. We compared the differences in clinical profile, seizure outcome, developmental characteristics and anti-seizure medication (ASM) responsiveness profiles between patients with missense and truncating variants.

Results: Out of a total of 3967 children with drug-resistant epilepsy during the period 2015–2021, 49 patients who fulfilled the inclusion criteria were studied. Thirty-seven had positive genetic tests, out of which 29 were *SCN1A* variants and 9 were other novel variants. The proportion of missense (14; 48.3%) and truncating *SCN1A* variants (15; 51.7%) was similar. A significant trend for developing multiple seizure types was noted among children with truncating variants ($p = 0.035$) and seizure freedom was more likely among children with missense variants ($p = 0.042$). All patients with truncating variants had ASM resistant epilepsy ($p = 0.020$). Developmental outcomes did not differ between the variant subtypes.

Conclusion: Our results show that children harbouring missense variants demonstrated a significantly lower propensity for multiple seizure subtypes and a higher proportion with seizure freedom. However developmental implications appear to be independent of variant subtype.

1. Introduction

The prevalence of DS is found to be 1 in 20,000 to 1 in 40,000 of the general population, with a male-to-female ratio of 2:1¹. Variants in the *SCN1A* gene which codes for the alpha subunit of the voltage-gated sodium channel 1.1 composed of 26 coding exons, are found in more than 80% of patients with DS, which includes severe myoclonic epilepsy of infancy (SMEI) and severe myoclonic epilepsy borderland (SMEB) or Dravet borderline phenotype [1–5]. The latter patients however have a better outcome compared to DS patients. Phenotypes of patients with

pathogenic truncating and missense *SCN1A* variants can range from the severe DS to mild genetic epilepsy with febrile seizures plus (GEFS+) [6]. Deletions, duplications and intronic splice site changes in *SCN1A* gene have been implicated in DS. Majority of the variants follow de novo inheritance [7]. Patients present with recurrent and prolonged seizures, usually febrile, hemi-clonic or generalised tonic clonic, often resulting in status epilepticus [2]. Cognitive outcome is poor, with a high risk of psychomotor impairment. Dravet syndrome is usually sporadic, however familial cases have been occasionally reported [8]. Patients with DS have an increased risk of mortality which might be due to status

* Corresponding author at: Dept of Neurology SCTIMST, Thiruvananthapuram, Kerala 695011, India.
E-mail address: rsamenon@yahoo.com (R.N. Menon).

¹ First, second and third authors have made equal contributions to the manuscript

<https://doi.org/10.1016/j.seizure.2024.01.004>

Received 26 September 2023; Received in revised form 21 December 2023; Accepted 9 January 2024

Available online 14 January 2024

1059-1311/© 2024 British Epilepsy Association. Published by Elsevier Ltd. All rights reserved.



Contents lists available at ScienceDirect

Seizure: European Journal of Epilepsy

journal homepage: www.elsevier.com/locate/seizure

Metabolic causes of pediatric developmental & epileptic encephalopathies (DEE)- genetic variant analysis in a south Indian cohort

Manna Jose^{a,1}, Alfiya Fasaludeen^{a,1}, Harini Pavuluri^a, Pavan Kumar Rudrabhatla^a, Soumya V. Chandrasekharan^a, Jithu Jose^a, Moinak Banerjee^b, Soumya Sundaram^a, Ashalatha Radhakrishnan^a, Ramshekhar N. Menon^{a,*}

^a Department of Neurology, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India

^b Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Trivandrum, Kerala, India

ARTICLE INFO

Keywords:

Developmental and epileptic encephalopathy

Metabolic genes

Variants

Next generation sequencing

Biochemical tests

ABSTRACT

Purpose: Drug-resistant epilepsy is seen in patients with inborn errors of metabolism and metabolic dysfunction in neurons is crucial to brain disorders associated with psychomotor impairment. Diagnostic rates of metabolic causes of developmental and epileptic encephalopathy (DEE) using next generation sequencing have been rarely studied in literature.

Methods: A prospective hospital study was carried out in 384 children with DEE, who underwent genetic testing. Metabolic disorders were evaluated with biochemical blood/urine assays and when required CSF estimations performed.

Results: A total of 154 pathogenic/likely pathogenic variants in 384 children were identified. Out of 384 children, 89 were clinically suspected to have probable or possible metabolic disorders. Pathogenic/likely pathogenic variants in metabolic genes were identified in 39 out of 89 (43.8 %) and promising VUS in 28 (31.4 %). These included variants for progressive myoclonus epilepsies (21; 53.8 %), DEE with focal/multifocal seizures (8; 20.5 %), generalized epilepsy (5; 12.8 %), early myoclonic encephalopathy (2; 5.1 %), LGS (1; 2.6 %) and West syndrome (2; 5.1 %).

Conclusion: Our cohort demonstrates for the first time from the Indian subcontinent that identification of metabolic variants can guide investigations and has therapeutic implications in patients with variable DEE phenotypes. A high utility is noted with regard to diagnosis and prognostication, given the low yield of available biochemical tests, indicating cost-effectiveness of this approach.

1. Introduction

Inborn errors of metabolism (IEM) are caused by defects in genes involved in biochemical pathways which control metabolism and neuronal function. Although IEM are rare genetic diseases, epilepsy is a common presenting feature. Early identification and diagnosis is important because some are treatable with variable developmental prognosis [1,2]. Often the epilepsy phenotypes associated with IEM can potentially evolve as developmental and epileptic encephalopathy (DEE). The severity of epilepsy may range from infrequent seizures to anti-seizure medication (ASM) resistant phenotypes including status epilepticus. DEE refers to a group of severe epilepsies characterized both by often ASM resistant seizures as well as encephalopathy in the form of

significant developmental delay, stagnation or regression of attained developmental skills [3]. An ever-growing list of genetic causes of DEE exists and other etiologies associated with DEE include structural (e.g. tuberous sclerosis complex, hypothalamic hamartomas, hemimegalencephaly), metabolic (e.g. pyridoxine or biotinidase deficiency, GLUT1 deficiency) and immune disorders (e.g. Rasmussen syndrome). The number of treatable metabolic epilepsies have increased considerably due to available precision-medicine modalities, making early diagnosis a prime necessity. [4].

The most recently adopted classification of metabolic disorders, the International Classification of Inherited Metabolic Disorders (ICIMD), classifies IEM according to the current understanding of genes responsible for these disorders and includes more than 1500 IEM [5]. The

* Corresponding author at: Dept of Neurology, SCTIMST, Thiruvananthapuram, Kerala, India, 695011.

E-mail address: rsnmenon@yahoo.com (R.N. Menon).

¹ First and second authors have made equal contributions to the manuscript.

<https://doi.org/10.1016/j.seizure.2023.12.017>

Received 30 March 2023; Received in revised form 12 December 2023; Accepted 25 December 2023

Available online 3 January 2024

1059-1311/© 2023 British Epilepsy Association. Published by Elsevier Ltd. All rights reserved.



Genetic variant interpretation for the neurologist – A pragmatic approach in the next-generation sequencing era in childhood epilepsy

Alfiya Fasaludeen^a, Amy McTague^{b,c}, Manna Jose^a, Moinak Banerjee^d, Soumya Sundaram^a, U.K. Madhusoodanan^c, Ashalatha Radhakrishnan^a, Ramshekhar N. Menon^{a,*}

^a Dept of Neurology, Sree Chitra Tirunal Institute for Medical Sciences & Technology (SCTIMST), Thiruvananthapuram, Kerala, India

^b Developmental Neurosciences, UCL Great Ormond Street Institute of Child Health, London, United Kingdom

^c Department of Neurology, Great Ormond Street Hospital, London, United Kingdom

^d Human Molecular Genetics Laboratory, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala, India

^{*} Department of Biochemistry, Sree Chitra Tirunal Institute for Medical Sciences & Technology (SCTIMST), Thiruvananthapuram, Kerala, India

ARTICLE INFO

Keywords:

Genetic testing
Next generation sequencing
Childhood epilepsy
Developmental epileptic encephalopathy
Variant interpretation

ABSTRACT

Genetic advances over the past decade have enhanced our understanding of the genetic landscape of childhood epilepsy. However a major challenge for clinicians has been understanding the rationale and systematic approach towards interpretation of the clinical significance of variant(s) detected in their patients. As the clinical paradigm evolves from gene panels to whole exome or whole genome testing including rapid genome sequencing, the number of patients tested and variants identified per patient will only increase. Each step in the process of variant interpretation has limitations and there is no single criterion which enables the clinician to draw reliable conclusions on a causal relationship between the variant and disease without robust clinical phenotyping. Although many automated online analysis software tools are available, these carry a risk of misinterpretation. This guideline provides a pragmatic, real-world approach to variant interpretation for the child neurologist. The focus will be on ascertaining aspects such as variant frequency, subtype, inheritance pattern, structural and functional consequence with regard to genotype-phenotype correlations, while refraining from mere interpretation of the classification provided in a genetic test report. It will not replace the expert advice of colleagues in clinical genetics, however as genomic investigations become a first-line test for epilepsy, it is vital that neurologists and epileptologists are equipped to navigate this landscape.

1. Introduction

Advances in genetic testing have revolutionized our understanding of childhood epilepsy. The availability of comprehensive epilepsy gene panels and recently whole exome (WES) and genome sequencing (WGS) for clinical use has enabled etiological diagnosis with therapeutic and prognostic utility (Krey et al., 2022). With reduction in sequencing costs worldwide, exomic/genomic data has now entered everyday clinical practice of epilepsy. This especially applies to the developmental-epileptic encephalopathies (DEE) and refractory epilepsies of early childhood, often associated with developmental delay, regression and non-diagnostic neuroimaging/metabolic tests. (Merckem-Mahmutoglu et al., 2015; McTague et al., 2016; Mitta et al., 2020; Koh et al., 2023). The relevance of genetic testing in the clinical evaluation and management of DEE has grown recently due to reports that

genetic variations are involved in many electroclinical phenotypes (McKnight et al., 2022; Chang et al., 2023). With next generation sequencing (NGS) techniques, a diagnostic yield of ~30–50% could be achieved for several complex childhood epilepsies (Mitta et al., 2020; Stöberg et al., 2020; Horák et al., 2022; Boonsimma et al., 2023). Precision medicine is possible for a small number of pathogenic variants, the functional consequence of which can be modulated by certain existing or novel therapies (McTague et al., 2016). A study about the current impact of genetic diagnosis in epilepsy showed that 32% of patients had treatment modifications including precision medicine and/or treatment changes implemented as a consequence of their genetic results, indicating the need for further development in strategies for precision medicine in genetic epilepsies (Balestrini et al., 2021). The clinical utility, diagnostic yield and potential cost-effectiveness of this approach (with regard to avoiding needless investigations and guiding

* Correspondence to: Dept of Neurology, SCTIMST, Thiruvananthapuram, Kerala 695011, India.
E-mail address: rsnmenon@yahoo.com (R.N. Menon).

<https://doi.org/10.1016/j.epilepsyres.2024.107341>

Received 27 November 2023; Received in revised form 14 February 2024; Accepted 29 February 2024

Available online 2 March 2024

0920-1211/© 2024 Elsevier B.V. All rights reserved.

APPENDIX D – PLAGIARISM CHECK REPORT

Alfiya PhD Thesis

ORIGINALITY REPORT

9%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|--|------------------------|
| 1 | dspace.sctimst.ac.in
<small>Internet</small> | 1854 words — 6% |
| <hr/> | | |
| 2 | S. Harikrishnan, P.P. Mohanan, V.K. Chopra, Roy Ambuj et al. "Cardiological society of India position statement on COVID-19 and heart failure", Indian Heart Journal, 2020
<small>Crossref</small> | 1081 words — 3% |

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES < 3%

EXCLUDE MATCHES < 9 WORDS