GENETIC STUDIES ON FAMILIAL GENETIC GENERALIZED EPILEPSY: A WHOLE EXOME SEQUENCING APPROACH

CHAN CHUNG KIN

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR
2021

GENETIC STUDIES ON FAMILIAL GENETIC GENERALIZED EPILEPSY: A WHOLE EXOME SEQUENCING APPROACH

CHAN CHUNG KIN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2021

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **CHAN CHUNG KIN**

Matric No: **17028263/2**

Name of Degree: **MASTER OF SCIENCE (EXCEPT MATHEMATICS AND SCIENCE PHILOSOPHY)**

Title of Dissertation ("this Work"):
GENETIC STUDIES ON FAMILIAL GENETIC GENERALIZED EPILEPSY: A WHOLE EXOME SEQUENCING APPROACH

Field of Study:
GENETICS AND MOLECULAR BIOLOGY

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 6/9/2021

Subscribed and solemnly declared before,

Witness's Signature Date: 6/9/2021

Name:
Designation:

GENETIC STUDIES ON FAMILIAL GENETIC GENERALIZED EPILEPSY: A WHOLE EXOME SEQUENCING APPROACH

ABSTRACT

Genetic generalized epilepsy (GGE) is a form of epilepsy that is potentially caused by genetic factors. It accounts for 15-20% of all epilepsy worldwide and 18.7% of the epilepsy cases in Malaysia. Unlike focal epilepsy, patients with generalized epilepsy mainly rely on antiepileptic drugs to achieve seizure control. As a result, understanding of the molecular pathway behind generalized epilepsy would be beneficial as it can aid the treatment decision and the design of newer antiepileptic drugs. The pathway analyses in GGE were conducted on European population and these results may not be applied to Malaysian or Asian due to genetic variation. Hence, this project was conducted to identify the genetic factors and plausible biological processes associated with GGE in Malaysia. The project started with whole exome sequencing on 12 unrelated Malaysian probands (4 Malays, 3 Chinese and 5 Indians) and 2 unrelated Malaysian Chinese proband-parent trios, all the 14 probands were diagnosed with GGE and having family history of epilepsy. The exome sequencing data was then analysed with GATK and the variants were annotated with wANNOVAR. Following variant filtration, a 3-step functional characterization analysis encompassing gene prioritization, protein-protein interaction analysis and DAVID enrichment was conducted. The analysis found the GO terms related with sodium and calcium ion transport, such as GO:0035725, GO:0006814, GO:0070588 and GO:0070509, were highly associated with GGE among Malaysians. The active role of ion channels in epilepsy makes them an ideal therapeutic target for antiepileptic drugs in seizure suppression. From this prospective, the protein-protein interaction analysis has suggested the potential of dystrophin (DMD) to be used as therapeutic target due to its

regulatory role on sodium transporters and calcium ion channels. Besides, the functional characterization analysis has also intimated the potential risk of tyrosine kinases NTRK1, NTRK2 and ERBB4 in the pathogenesis of GGE. Additionally, segregation analysis was conducted on 2 families to study the pathogenicity of the *SCN1A* c.5753C>T and *ERBB4* c.1972A>T variants identified from this cohort. The results showed that the *SCN1A* variant was likely pathogenic but the *ERBB4* variant was expected to induce epilepsy via an interaction with gamma-aminobutyric acid type A (GABA-A) receptor gene *GABRA1* c.448G>A variant. A disease model illustrating the combined effect between the mutant *ERBB4* and *GABRA1* in GGE pathogenesis was proposed, this hypothesized disease model may provide a new insight on the impact of ERBB4 in the disturbance of inhibitory postsynaptic current (iPSC) and excitation-inhibition (E/I) balance. While this project is just a preliminary study on the genetics behind GGE among Malaysians, this study has demonstrated that ion channels are not the only cause of GGE, tyrosine kinases may also contribute to GGE by manipulating numerous biological pathways like the expression of GABA-A receptors. Moreover, the identification of DMD as a potential therapeutic target warrants further investigation as the findings may contribute to the design of new antiepileptic drugs in future.

Keywords: Genetic generalized epilepsy, whole exome sequencing, functional characterization, segregation analysis

KAJIAN GENETIK EPILEPSI MENYELURUH GENETIK FAMILIAL: SATU KAEDAH PENJUJUKAN SELURUH EKSOM

ABSTRAK

Epilepsi menyeluruh genetik (GGE) merupakan sejenis epilepsi yang boleh disebabkan oleh faktor genetik. Ia merangkumi 15-20% kes epilepsi seluruh dunia dan 18.7% kes epilepsi di Malaysia. Berbeza daripada epilepsi fokal, pesakit epilepsi menyeluruh bergantung kepada ubat antiepileptik untuk mengawal serangan sawan. Oleh itu, pemahaman terhadap laluan molekul epilepsi menyeluruh boleh mendatangkan manfaat kerana dapat membantu dalam penentuan kaedah rawatan serta perekaan ubat antiepileptik baru. Analisis terhadap laluan molekul GGE pernah dibuat dalam kalangan populasi Eropah tetapi hasil daripada kajian-kajian ini tidak semestinya menggambarkan laluan molekul GGE dalam kalangan orang Malaysia atau Asia atas sebab variasi genetik. Justeru, projek ini dijalankan untuk mencari faktor genetik serta proses biologi yang berkaitan dengan GGE di Malaysia. Projek ini dimulakan dengan penjujukan seluruh eksom pada 12 individu Malaysia (4 Melayu, 3 Cina dan 5 India) tidak bersaudara serta 2 trio proban-ibu-bapa Cina Malaysia tidak bersaudara, kesemua 14 proban ini telah didiagnos menghidapi penyakit GGE dan mempunyai sejarah keluarga epilepsi. Data penjujukan eksom ini adalah dianalisis dengan GATK dan varian-varian adalah dianotasi dengan wANNOVAR. Selepas penapisan varian, analisis pencirian fungsi tiga langkah yang merangkumi pengutamaan gen, analisis interaksi protein-protein serta pengayaan fungsi telah dijalankan. Keputusan analisis ini menunjukkan bahawa istilah GO yang bersangkutan dengan pengangkutan ion natrium dan kalsium, seperti GO:0035725, GO:0006814, GO:0070588 dan GO:0070509, adalah berkaitan dengan GGE dalam kalangan orang Malaysia. Peranan aktif saluran ion dalam epilepsi menjadikan mereka

sebagai sasaran terapi ubat antiepileptik yang ideal dalam pengurangan serangan sawan. Dari segi perspektif ini, analisis interaksi protein-protein menyarankan potensi dystrophin (DMD) sebagai sasaran terapi lantaran peranannya dalam regulasi pengangkut natrium dan saluran ion kalsium. Selain itu, analisis pencirian fungsi juga mencadangkan kinase tirosina NTRK1, NTRK2 dan ERBB4 sebagai risiko yang berpotensi menimbulkan patogenesis GGE. Sementara itu, analisis pengasingan telah dijalankan dalam 2 keluarga untuk mengkaji kepatogenan varian *SCN1A* c.5753C>T dan *ERBB4* c.1972A>T yang dikenal pasti daripada kohort ini. Keputusan analisis pengasingan menunjukkan bahawa varian *SCN1A* ini adalah berkemungkinan patogenik tetapi varian *ERBB4* ini diduga menimbul epilepsi melalui interaksi dengan varian gen reseptor asid aminobutirik-gamma jenis A (GABA-A) *GABRA1* c.448G>A. Satu model penyakit yang menyarankan patogenesis GGE malalui gabungan efek mutan *ERBB4* dan *GABRA1* telah dikemukakan, model penyakit ini boleh membawa fahaman yang baru terhadap kesan ERBB4 dalam gangguan arus postsinaptik penginhibitan (iPSC) serta keseimbangan ujaan-rencatan (E/I). Meskipun projek ini hanya kajian awalan tentang genetik GGE dalam kalangan orang Malaysia, kajian ini telah menunjukkan bahawa saluran ion bukan penyebab tunggal GGE, kinase tirosina juga boleh mengakibatkan GGE dengan memanipulasi laluan-laluan biologi seperti ekspresi reseptor GABA-A. Sementelahan, pengenalpastian DMD sebagai sasaran terapi berpotensi mewajarkan penyelidikan yang lebih lanjut terhadap DMD kerana penemuan ini mungkin dapat menyumbang kepada perekaan ubat antiepileptik baru pada masa hadapan.

Kata kunci: Epilepsi menyeluruh genetik, penjujukan seluruh eksom, pencirian fungsi, analisis pengasingan

ACKNOWLEDGEMENTS

First of all, I would like to take this opportunity to express my greatest gratitude to my supervisors, Professor Dr Lim Kheng Seang and Associate Professor Dr Ng Ching Ching, for all the guidance and advice over the entire duration of this study. The guidance and advice have enlightened me with plenty of knowledge covering from experiment planning to data analysis and finally manuscript writing. This learning process is a fruitful experience and I have learned a great deal along the journey including the soft skills in communication with patients as well as the thinking and analytical skills of a researcher.

Besides, I wanted to present my vote of thanks to my colleagues Amy Khor and Chin Yoon Ming for sharing the knowledge and experience on laboratory works and bioinformatic analyses, not to mention the medical staffs from University Malaya Medical Centre for the technical support given during blood sample collection, your kind assistance has smoothened the progress of my study. I would also like to grab this chance to express my thankfulness to the patients and family members for your involvement in this study, this project would not be completed without the participation from all of you.

Last but not least, I wish to extend my gratefulness to my parents Mr Chan Hock Ming and Madam Thien Nyuk Len for the continuous spiritual and financial support throughout the journey of my study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

1KGP	:	1000 Genomes Project
ACMG	:	American College of Medical Genetics and Genomics
ACS	:	American Chemical Society
AED	:	antiepileptic drug
BDNF	:	brain derived neurotrophic factor
bp	:	base pair
BWA	:	Burrows-Wheeler Alignment
CADD	:	Combined Annotation Dependent Depletion
CAE	:	childhood absence epilepsy
CBZ	:	carbamazepine
cDNA	:	complementary DNA
CDX	:	Chinese Dai in Xishuangbanna, China
DAVID	:	Database for Annotation, Visualization and Integrated Discovery
DEE	:	developmental and epileptic encephalopathy
DMD	:	dystrophin
DNA	:	deoxyribonucleic acid
E/I	:	excitation-inhibition
EAS	:	East Asian
EDTA	:	ethylenediaminetetraacetic acid
EEG	:	electroencephalography
ERBB4	:	erb-b2 receptor tyrosine kinase 4
ERK	:	extracellular signal-regulated kinase
FS	:	febrile seizures
GABA	:	gamma-aminobutyric acid

GABA-A	:	gamma-aminobutyric acid type A
GABRA1	:	gamma-aminobutyric acid type A receptor subunit alpha1
GABRB3	:	gamma-aminobutyric acid type A receptor subunit beta3
GABRG2	:	gamma-aminobutyric acid type A receptor subunit gamma2
GATK	:	Genome Analysis Toolkit
GEFS+	:	generalized epilepsy with febrile seizure plus
GGE	:	genetic generalized epilepsy
gnomAD	:	Genome Aggregation Database
GO	:	Gene Ontology
GSW	:	generalized spike and wave
GTC	:	generalized tonic-clonic seizures
IGE	:	idiopathic generalized epilepsy
ILAE	:	International League Against Epilepsy
iPSC	:	inhibitory postsynaptic current
JAE	:	juvenile absence epilepsy
JME	:	juvenile myoclonic epilepsy
KEGG	:	Kyoto Encyclopedia of Genes and Genomes
KHV	:	Kinh in Ho Chi Minh City, Vietnam
LTG	:	lamotrigine
LVT	:	levetiracetam
MAF	:	minor allele frequency
MRI	:	magnetic resonance imaging
N/A	:	not applicable
NGS	:	next generation sequencing
NRG	:	neuregulin
PCR	:	polymerase chain reaction

PI3K	:	phosphatidylinositol 3-kinase
Polyphen-2	:	Polymorphism Phenotyping V2
PPI	:	Protein-Protein Interaction
PPR	:	photoparoxysmal response
PSW	:	polyspike-waves
SAS	:	South Asian
SCN1A	:	sodium voltage-gated channel alpha subunit 1
SCN2A	:	sodium voltage-gated channel alpha subunit 2
SIFT	:	Sorting Intolerant from Tolerant
STRING	:	Search Tool for Retrieval of Interacting Genes/Proteins
SUDEP	:	sudden unexpected death in epilepsy
TBE	:	tris-borate-EDTA
TLE	:	temporal lobe epilepsy
TPX	:	topiramate
TrkB	:	tropomyosin-related kinase B
UMMC	:	University Malaya Medical Centre
VPA	:	sodium valproate
VUS	:	variants of uncertain significance
WES	:	whole exome sequencing
WGS	:	whole genome sequencing

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background

Genetic generalized epilepsy (GGE), which was formerly known as idiopathic generalized epilepsy, is a subtype of generalized epilepsy that is potentially induced by genetic factors (Scheffer et al., 2017). GGE constitutes 15-20% of all epilepsies worldwide and 18.7% of the epilepsy cases in Malaysia (Jallon & Latour, 2005; Lim et al., 2017). GGE can be divided into several subtypes, which includes childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), and generalized tonic-clonic seizures alone (Scheffer et al., 2017). Among the subtypes, JME is the most common form of GGE, it is responsible for 5-10% of all epilepsies and about 18% of GGE cases (Camfield et al., 2013).

Epilepsy has brought huge psychosocial impacts on the patients as well as their caregivers (Mula & Sander, 2016). Even in Malaysia, patients with uncontrolled seizures may suffer from unemployment (Lim et al., 2013; Wo et al., 2015), and their caregivers can experience additional burden due to higher seizure frequency among patients with uncontrolled epilepsy (Lai et al., 2019). Hence, it is crucial to keep the epilepsy under control and subsequently reduce the seizure frequency.

Unlike focal epilepsy in which the seizures can be treated with surgical resection, surgical treatments for generalized epilepsy mainly rely on disconnection procedures such as corpus callosotomy (Jette et al., 2014). However, corpus callosotomy might not guarantee a seizure-free as some patients can develop focal seizures after the surgery (Jenssen et al., 2006). In this case, achieving seizure control through antiepileptic drugs will be a more effective option and this is the aspect in which genetic knowledge is contributing to epilepsy treatment.

Antiepileptic drugs reduce seizures by decreasing neuronal excitation signals or by increasing neuronal inhibition signals (Santulli et al., 2016). This is achieved by targeting

specific cellular components in the neurons. For instance, lamotrigine, carbamazepine and lacosamide tend to block the voltage-gated sodium channels; bromide and topiramate enhance the effect of inhibitory neurotransmitter gamma-aminobutyric acid; and levetiracetam binds to synaptic vesicle proteins and inhibits neurotransmitter release at the synapse (Kobayashi et al., 2020; Moavero et al., 2017; Stefanović et al., 2018). Understanding of the genetic mechanism of epilepsy and pharmacogenomics of antiepileptic drugs is very beneficial in deciding drug treatment, as the efficacy of each drug can be influenced by genetic mutation (Arroyo et al., 2002; Balestrini & Sisodiya, 2018). For example, bromide has higher efficacy than carbamazepine in treating epilepsy patients with *SCN1A* mutation, suggesting that drugs targeting the sodium ion channels are less effective compared to drugs targeting the sodium channel antagonists like gamma-aminobutyric acid in treating *SCN1A*-positive patients (Psarropoulou, 2010; Shi et al., 2016).

Over the years, many studies have been conducted to identify the genetic factors associated with epilepsy. According to Wang et al. (2017a), a total of 977 epilepsy-associated genes have been identified in 2017. Despite there is evidence showing cellular ion channels and signal transduction molecules are closely related with epilepsy, the exact mechanism of epilepsy remains unknown (Song & Deng, 2018; Wang et al., 2017a). Furthermore, increase in the number of epilepsy-associated genes and the complex inheritance patterns suggest the polygenic nature in most epilepsy with genetic basis (Dhiman, 2017). For epilepsy that is polygenic, its molecular mechanism is very complex as it can involve the interaction between multiple genes, proteins, and even genetic variants.

In order to investigate the interaction between genes, some researchers apply network biology in their studies, with the aim to identify the key pathways or genes that can be used as therapeutic target (Bezhentsev et al., 2018; Jia et al., 2011; Korotkov et al., 2017).

Nevertheless, only two pathway analysis studies have been conducted on GGE, and both of them are conducted on patients with European origin (Epi25 Collaborative, 2019; Ozdemir et al., 2019). Even though GGE normally shows a good prognosis, 12-20% of the patients are still suffering from drug refractoriness (Gesche et al., 2020; Gomez-Ibañez et al., 2017; Vorderwülbecke et al., 2017). Therefore, studies on the genetic mechanism and the search for new therapeutic targets for GGE are still required.

Malaysia is a southeast Asian country with population size of 32.73 million composed of multiple ethnicities, with Malays being the largest ethnic group followed by Chinese and Indians (Department of Statistics Malaysia, 2020). Despite there is no large-scale study showing the genetic diversity between the ethnic groups in Malaysia yet, a study in Singapore, a country with similar ethnic composition as Malaysia, has demonstrated the genetic differences between Malays, Chinese and Indians (Wu et al., 2019). According to Wu et al. (2019), the Chinese and Indians in southeast Asia are having genetic origin from East Asian and South Asian respectively. Meanwhile, the genetic composition of Malays is more specialized and does not overlap with East Asian or South Asian, the only established populations to date with the closest genetic relation to Malays are Kinh (Vietnamese) and Chinese Dai in Xishuangbanna (Wu et al., 2019). Due to different genetic origin, Malaysians are having a different genetic composition compared to Europeans and Africans. As a result, the genetic findings from Europeans and Africans may not absolutely apply to Malaysians because of genetic variation.

1.2 Problem Statement

Despite there is an increasing number of genetic studies conducted on epilepsy worldwide, the genetic mechanism underlying the disease remains unknown. Since antiepileptic drug is a preferred treatment option for GGE, understanding of the genetic mechanism behind GGE via network biology will be very beneficial as it may help in

making drug treatment decision. Nevertheless, network biology has not been conducted on GGE in Asians yet. Despite such studies have been conducted on Europeans, the results may not reflect on Malaysians due to genetic differences. Hence, this study is conducted to study the key genes and the biological processes associated with GGE among Malaysians.

1.3 Scope of Research

This project focuses on the genetic basis of GGE among 14 Malaysian probands with positive family history. As its name suggests, GGE indicates possible genetic predisposition in the pathogenesis of the disease. Since genetic mutations are inheritable, it is hypothesized that patients with genetic diseases are likely to have family history. Therefore, patients with family history of GGE are focused in this project, with the assumption that these patients are having an increased likelihood that their epilepsy is caused by genetic factors.

1.4 Objectives

The objectives of this project include the following:

1. To identify the genes and variants which are potentially associated with GGE among Malaysians.
2. To investigate the biological processes that are potentially associated with GGE, which may be used as therapeutic target in future.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction to Epilepsy

2.1.1 Background

Epilepsy is one of the neurological disorders that is induced by the abnormal generation of seizures. International League Against Epilepsy (ILAE) defines epilepsy as follow, “Epilepsy is a disease of the brain defined by any of the following conditions: (1) At least two unprovoked (or reflex) seizures occurring >24 h apart; (2) One unprovoked (or reflex) seizure and a probability of further seizures similar to the general recurrence risk (at least 60%) after two unprovoked seizures, occurring over the next 10 years; (3) Diagnosis of an epilepsy syndrome” (Fisher et al., 2014). Meanwhile, seizure is defined by ILAE as “a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain” (Fisher et al., 2017).

It is reported that lifetime prevalence of epilepsy is having a pooled estimate of 7.60 per 1,000 persons worldwide (Fiest et al., 2017). Nevertheless, the distribution of epilepsy is uneven, the prevalence of epilepsy is generally higher in rural areas due to increased exposure to risk factors such as head trauma, perinatal injuries and central nervous system infections (Espinosa-Jovel et al., 2018). In Asia, the prevalence of epilepsy is estimated to be 6 per 1,000 persons (Trinka et al., 2019), Table 2.1 shows the lifetime prevalence of epilepsy among Asian countries between 2000 and 2019.

Table 2.1: Lifetime prevalence of epilepsy among Asian countries between 2000 and 2019.

Region/Country	Lifetime prevalence (per 1,000 persons)	Reference
Iran	18.00	(Mohammadi et al., 2006)
Laos	7.70	(Tran et al., 2006)
Malaysia	7.80	(Fong et al., 2021)
Patiala, India	10.30	(Hara et al., 2015)
Taiwan	5.85	(Chen et al., 2012)
Thailand	7.20	(Asawavichienjinda et al., 2002)
Yueyang, China	4.50	(Pi et al., 2012)

Epilepsy has been shown to bring huge physical impacts on patients. Systematic analysis revealed that epilepsy ranked the fifth highest in global age-standardized disability-adjusted life-years rate, after stroke, migraine, Alzheimer's disease, and meningitis (Feigin et al., 2019). People with epilepsy are sometimes present with comorbidities such as psychiatric disorders, migraine, heart diseases and intellectual disabilities (Keezer et al., 2016). In China, 26.2% and 9.3% of the epilepsy patients are having psychiatric and neurological disorders respectively (Zhu et al., 2020). Meanwhile, 23.0% of the epilepsy children in India are exhibiting the symptoms of attention deficit hyperactivity disorder (Choudhary et al., 2018); and in Malaysia, 38.0% of the epilepsy patients are diagnosed with mental comorbidities like neurodevelopmental and psychiatric disorders (Lai et al., 2019). Besides psychiatric and cardio-cerebrovascular comorbidities, epilepsy can also lead to premature death (Mbizvo et al., 2019). In the context of epilepsy-related death, sudden unexpected death in epilepsy (SUDEP) is the leading cause of mortality (Chahal et al., 2020), accounting for 76.0% of all epilepsy-related death among children in United Kingdom (Abdel-Mannan & Sutcliffe, 2020). Other causes of death among epilepsy patients include drowning, motor accidents and suicide (Devinsky et al., 2016).

People with epilepsy are often suffering from poor quality of life (Yogarajah & Mula, 2019), they are generally having lower educational achievement, marriage rate and

employment rate compared to people without epilepsy (Jennum et al., 2021; Kaur et al., 2019; Wo et al., 2015). The poor life quality is mainly associated with the comorbidities and public stigmatization against the patients (Alsaadi et al., 2017; Beghi, 2019; Mond et al., 2019; Trinka et al., 2019; Welton et al., 2020). Furthermore, the quality of life worsens in patients with poor seizure control, as they are prone to comorbidities, such as psychiatric symptoms, that have further elevated stigmatization on people with epilepsy (Blixen et al., 2020; Chatterjee et al., 2020; Deleo et al., 2020). Apart from health-related comorbidities and stigmatization, the lifestyle of people with epilepsy is also disturbed. For instance, people with uncontrolled seizures are restricted from driving in some countries (Beran et al., 2020; Inoue et al., 2004; Winston & Jaiser, 2012). Besides, they might face issues in obtaining fitness certificates for certain sport activities (Capovilla et al., 2016).

2.1.2 Types of Epilepsy and Its Symptoms

According to ILAE, epilepsy can be classified into four main categories, namely generalized epilepsy, focal epilepsy, combined generalized and focal epilepsy, and unknown epilepsy type for the cases that cannot be determined whether it is generalized or focal (Scheffer et al., 2017). The classification of epilepsy is mainly based on seizure characteristics. Generalized epilepsy comprised of multiple subtypes, including but not limited to childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), and generalized tonic-clonic seizures alone (Scheffer et al., 2017). Meanwhile, focal epilepsy can be divided into temporal lobe epilepsy (TLE), occipital lobe epilepsy, parietal lobe epilepsy, and frontal lobe epilepsy, based on the brain region in which the epileptic spike is initiated (Scheffer et al., 2017).

Seizures in focal epilepsy are generally originated from one side of the brain and are often restricted to one cerebral hemisphere; in contrast, generalized epilepsy involves

seizures that spread rapidly across both sides of the brain (Berg et al., 2010; Fisher et al., 2017). Due to differences in seizure dynamics, the phenotypic symptoms exhibited among epilepsy patients may vary from mild to severe (Borlot et al., 2019b; Chatron et al., 2018; Chen et al., 2020a; Cvetkovska et al., 2018; Fang et al., 2019; Marini et al., 2018; Zhang et al., 2017). In order to ease the diagnosis of epilepsy, an online diagnostic manual was published by ILAE, a clinician may start the diagnosis with the classification of seizure types, followed by epilepsy types and finally epilepsy syndromes; besides, findings from electroencephalography (EEG) may aid the diagnosis of epilepsy in some cases (Scheffer et al., 2017). Table 2.2 shows some common epilepsy syndromes and their corresponding clinical characteristics in accordance with ILAE diagnostic manual (International League Against Epilepsy, 2020).

Table 2.2: Common epilepsy syndromes and their corresponding clinical characteristics according to ILAE diagnostic manual (International League Against Epilepsy, 2020), a patient might not manifest all the clinical characteristics in a given epilepsy syndrome.

Epilepsy type	Epilepsy syndrome	Clinical characteristic	EEG finding
Focal epilepsy	Temporal lobe epilepsy	• Behavioural arrest • Impaired awareness • Automatisms • Sensory seizures • Emotional seizures • Cognitive seizures	• Spike and wave or sharp slow waves in the anterior temporal lobe
Focal epilepsy	Frontal lobe epilepsy	• Focal hyperkinetic seizures • Focal bilateral motor seizures • Prominent vocalization • Bizarre behaviour • Urinary incontinence • Head and eye deviation	• Midline or bi-frontal discharge • Normal EEG results in some cases
Focal epilepsy	Occipital lobe epilepsy	• Focal sensory visual seizures	• Spike and wave or sharp slow waves • Normal EEG results in some cases
Focal epilepsy	Parietal lobe epilepsy	• Paraesthesia • Disorientation • Visual hallucinations • Visual vertigo • Receptive language impairment • Ipsilateral or contralateral rotatory body movements	• Spike and wave or sharp slow waves in the posterior parietal lobe • Normal EEG results in some cases
Generalized epilepsy	Juvenile myoclonic epilepsy	• Myoclonic seizures • Generalized tonic-clonic seizures	• Generalized spike and wave • Polyspike and wave
Generalized epilepsy	Juvenile absence epilepsy	• Absence seizures • Generalized tonic-clonic seizures	• Generalized spike and wave • Fragments of polyspike and wave

Table 2.2, continued.

Epilepsy type	Epilepsy syndrome	Clinical characteristic	EEG finding
Generalized epilepsy	Childhood absence epilepsy	• Absence seizures	• Generalized spike and wave
Generalized epilepsy	Generalized tonic-clonic seizures alone	• Generalized tonic-clonic seizures	• Generalized spike and wave • Polyspike and wave

Generalized epilepsy	Generalized epilepsy with febrile seizure plus	• Febrile seizures and afebrile seizures that can be either
		  • Generalized seizures, such tonic-clonic, atonic, myoclonic, myoclonic-atonic or absence seizures, or
		  • Focal seizures
			• Generalized spike and wave

Abbreviation: EEG, Electroencephalography.

2.1.3 Causes of Epilepsy

Epilepsy can be arisen from multiple causative factors, ranging from genetic to environmental factors such as exposure to pesticides (Guekht et al., 2017; Requena et al., 2018; Sokka et al., 2017). ILAE divides the epilepsy causes into 6 aetiological categories, which are structural, genetic, infectious, metabolic, immune, and unknown (Scheffer et al., 2017). It is worth noting that the aetiologies are not mutually exclusive and an epileptic syndrome can be classified into more than one aetiology (Falco-Walter et al., 2018). This condition is best exemplified by tuberous-sclerosis-related epilepsy, which can be classified under both structural and genetic aetiology (Scheffer et al., 2017). The new classification aids in patient management and epilepsy research as it may shed some light on why a seizure is developed (Pack, 2019).

Structural aetiology refers to epilepsy causes that involve the changes or abnormalities in brain structure, examples of structural aetiology include brain tumour, stroke, trauma, tuberous sclerosis and abnormalities in brain development (Scheffer et al., 2017; Sokka et al., 2017). It is a common aetiology for epilepsy onset in the elderly. In western China, it is found that 62.4% of the epilepsy onset in elderly are due to structural aetiology, of which stroke is the major cause and has constituted 48.7% of the cases, followed by trauma (17.5%) and brain tumour (9.7%) (Guo et al., 2018). Similar findings are reported by studies conducted in India, Niger and Saudi Arabia, stroke is the main cause of epilepsy onset among elderly (Assadeck et al., 2019; Shariff & AlKhamis, 2017; Verma & Kumar, 2017). While stroke is a common cause of epilepsy in the elderly, it is not the major cause of epilepsy among teenagers and young adults (Dahl-Hansen et al., 2019). Instead, their epilepsy onset largely results from perinatal insults, neoplasia and trauma (Dahl-Hansen et al., 2019).

Epilepsy with genetic aetiology suggests the involvement of possible or known genetic factors in the seizure onset (Scheffer et al., 2017). This aetiology is highly prevalent in

childhood epilepsy and is responsible for 30-60% of the cases (Aaberg et al., 2017; Farghaly et al., 2018; Sokka et al., 2017; Veri et al., 2018). Seizures with genetic aetiology often begin in young ages, analysis conducted by Aaberg et al. (2017) showed that 77% of the cases are experiencing seizure onset before 5 years old. Nevertheless, genetic seizure onset can also occur in adulthood, a multi-centre study involving 243 adult patients revealed that 14.7% of the adult-onset epilepsy are due to genetic mutation (Kang et al., 2019). Genetic aetiology is often related with structural and metabolic aetiologies (Scheffer et al., 2017), the roles of genetics in epilepsy are further discussed in Section 2.2.

Infectious, metabolic and immune are new aetiological categories officially announced by ILAE in 2017 (Scheffer et al., 2017). In comparison with structural and genetic aetiologies, the occurrence of infectious, metabolic and immune aetiologies in epilepsy is less frequent (Aaberg et al., 2017; Dahl-Hansen et al., 2019). In spite of this, infectious aetiology, which implies the epilepsy onset due to an infection, is still a common cause of epilepsy in low-middle-income countries (Scheffer et al., 2017; Thurman et al., 2018). In these countries, the frequency of infectious aetiology can range from 7.0% to 23.2% (Caprara et al., 2020; Espinosa Jovel et al., 2016; Guekht et al., 2017). The high prevalence of epilepsy with infectious aetiology is related to the increased exposure to central nervous system infectious agents, such as *Plasmodium falciparum*, *Onchocerca volvulus*, and *Taenia solium* (Kind et al., 2017; Lenaerts et al., 2018; Mital et al., 2020; Mmbando et al., 2018). Besides, an occasional viral outbreak, such as Zika virus epidemic in Brazil, can also contribute to the increased risk of epilepsy (Carvalho et al., 2020).

Epilepsy cases with metabolic aetiology are also known as “metabolic epilepsy” because they mainly result from metabolic disorders (Scheffer et al., 2017; Sharma & Prasad, 2017). Metabolic epilepsy is largely found in patients who developed seizure during infancy, it is reported that about 3.6% of the neonatal seizure is caused by

metabolic disorder (Vučetić Tadić et al., 2020). Metabolic disorders are normally related to genetic factors (Scheffer et al., 2017; Tumienė et al., 2018), examples of genes associated with metabolic disorders include *SLC2A1*, *MTHFS* and *ADSL* (Banerjee et al., 2021; Ivanova et al., 2018; Rodan et al., 2018).

Immune aetiology refers to epilepsy cases caused by immune disorders (Scheffer et al., 2017). Examples of immune epilepsy include anti-glutamic acid decarboxylase 65 antibody-associated epilepsy, febrile infection-related epilepsy syndrome and Rasmussen's encephalitis (Caputo et al., 2018; Daif et al., 2018; Steinman, 2018). Immune epilepsies are mostly related with neural-specific antibodies and neuroinflammation (Bosco et al., 2020; Dubey et al., 2017; Husari & Dubey, 2019; Koh, 2018). The cause of immune epilepsy is complicated, as neuroinflammation can also be triggered by both genetic and environmental factors like traumatic brain injury (Chaves et al., 2020; Leal et al., 2018; Sharma et al., 2019).

Despite many studies have been conducted to study the aetiologies of epilepsy, there are epilepsy cases where its causal factors remain unidentified, ILAE groups these cases under the unknown aetiology (Scheffer et al., 2017). One good example of unknown epilepsy is the rare case of isolated myoclonus reported in United States, where the diagnosis evaluation conducted on the 72-year-old male patient failed to illicit a clear underlying aetiology (Kapoor & Kinsella, 2018). This case suggests the complexity and heterogeneity of the epilepsy causes; a lot of studies are still needed in order to elucidate the aetiologies of epilepsy.

2.2 Genetics and Epilepsy

2.2.1 Brief History of Genetic Discoveries in Epilepsy

The history of genetics in epilepsy can be dated back to 1889 when Russell (1889) pinpointed possible heritability predisposition in epilepsy among family members. In

1911, Davenport & Weeks (1911) published the first study on epilepsy heritability by performing analysis of multiple family pedigrees using Mendelian method, they proposed the frequency of having affected children among epileptic parents. However, this work was criticized by Schuster (1912) and he pointed out that some of the conclusions made by Davenport & Weeks (1911) were bias and were not derived from the results. Since then, not much research has been made on epilepsy genetics until the 1930s, when various studies including Rosanoff et al. (1934) have consistently observed that the proportion of having both affected co-twins was higher in monozygotic twins compared to dizygotic twins. 13 years later, Lennox (1947) concluded the involvement of genetic factors in epilepsy from his analysis of 2,130 probands and 12,119 relatives. Henceforth, genetic research in epilepsy started to blossom and more studies had been conducted on this topic. For instance, Kimball & Hersh (1955) studied on sibships and found 36% chance that the children will be epileptic if one of the parents is affected, Sarlin et al. (1960) reported an autosomal recessive inheritance in a family with myoclonic epilepsy, Bray et al. (1964) implicated the presence of genetic factors in familial temporal lobe epilepsy by analysing the incidence of epilepsy in the families, Tsuboi & Christian (1973) studied on 319 patients and concluded 27.3% prevalence of genetic predisposition in impulsive petit mal (juvenile myoclonic epilepsy), and Heijbel et al. (1975) analysed the pedigrees from 19 probands and deduced that an autosomal dominant gene is responsible for the benign epilepsy of childhood with centrotemporal spikes. In 1985, ILAE officially recognized genetics as one of the causes of epilepsy and the term “idiopathic” was used to describe the epilepsy cases which are suspected to be induced by genetic factors (Commission on Classification and Terminology of the International League Against Epilepsy, 1985).

2.2.2 "Idiopathic Generalized Epilepsy" versus "Genetic Generalized Epilepsy"

As mentioned in the last sentence in Section 2.2.1, ILAE recognized genetics as one of the causing factors of epilepsy in 1985 and the term "idiopathic" was used to describe the epilepsy cases with this aetiology. In 2010, ILAE introduced a new aetiological classification and the term "idiopathic" was replaced with "genetic", owing to the reason that "idiopathic" indicates the possible involvement of genetic factor in the pathogenesis of epilepsy; besides, "idiopathic" may also mean that the epilepsy is benign or drug-responsive which may not apply to all cases, the use of "genetic" will resolve this ambiguity as it gives a more direct meaning on the epilepsy aetiology (Berg, 2011; Berg et al., 2010; Berg & Millichap, 2013; Berg & Scheffer, 2011). Following the new classification, "idiopathic generalized epilepsy (IGE)" was renamed as "genetic generalized epilepsy (GGE)" (Berg et al., 2010). The classification change has brought about an uproar among the clinicians and researchers. Shinnar (2010) opposed the idea of choosing genetics as a major aetiology as there are only a few genetic cases that are being reported in epilepsy; furthermore, genetic abnormalities were not identified in most cases that are classified under the "genetic" category. However, Beghi (2011) and Moshé (2011) supported the use of "genetic" as the change will improve clinical practice such as prognosis and treatment decision. Meanwhile, Panayiotopoulos (2012) proposed that "genetic" should be considered as a new aetiology rather than replacing the existing "idiopathic". Due to strong argument, the term "GGE" was not used in all epilepsy-genetic studies. For example, DiFrancesco et al. (2011), Heinzen et al. (2012) and Lachance-Touchette et al. (2011) were still using "IGE" in their research topic. In 2013, van Campen et al. (2013) performed a test to investigate the agreement between the old and new aetiological classification, despite they did not comment specifically on the agreement between "idiopathic" and "genetic", they gave a general overview that the new aetiological classification is comparable with the old classification. Nevertheless, French

(2014) still disagreed with the use of "GGE" and flagged up the issue that clinicians may confuse it with other genetically related generalized syndromes like Dravet syndrome. In 2016, ILAE reviewed the debates over "idiopathic" and "genetic", they pinpointed that the disagreement was mainly arose from the points that the underlying genetic factors were not identified and the lack of family history in "GGE"; while it is true that causal genetic factor remains unidentified in most cases, having genetic aetiology is not equivalent to having family history as genetic mutations can arise *de novo* (Scheffer et al., 2016). In 2017, ILAE came up with another new aetiological classification, and stated that the term "GGE" should be used for the IGE cases with evidence suggesting the involvement of genetic factors in disease pathogenesis; meanwhile, the use of "IGE" is still accepted for CAE, JAE, JME, and generalized tonic-clonic seizures alone (Scheffer et al., 2017). For this project, the term "GGE" is used to describe the cases from our cohort as at least one potential pathogenic genetic variant has been identified from all the study subjects (Section 4.4).

2.2.3 Previous Genetic Discoveries on Genetic Generalized Epilepsy and Its Impact

Despite GGE indicates the potential involvement of genetic factors in epilepsy pathogenesis, it is an interesting fact that the first epilepsy-associated gene was not reported from GGE. In fact, the first epilepsy-associated gene is the *CHRNA4* identified from a family with benign neonatal familial convulsions (now known as benign familial neonatal seizures) (Beck et al., 1994). Meanwhile, the first gene reported in generalized epilepsy with febrile seizures plus (GEFS+) is *SCN1B* (Steinlein, 2008; Wallace et al., 1998), this may also be the first gene identified in GGE. As the technology advances, more GGE-associated genes are being discovered (Table 2.3).

Table 2.3: Example of genes reported to be potentially associated with GGE, the gene symbols are in accordance with current Human Genome Organization Gene Nomenclature Committee (Braschi et al., 2019).

Gene symbol	GGE syndrome	Reference
BRD2	Juvenile myoclonic epilepsy	(Pal et al., 2003)
CACNA1H	Childhood absence epilepsy, Juvenile absence epilepsy, Juvenile myoclonic epilepsy	(Heron et al., 2007)
CACNB4	Genetic generalized epilepsy	(Escayg et al., 2000; Escayg et al., 1998)
CLCN2	Juvenile myoclonic epilepsy, Generalized tonic-clonic seizure alone	(Saint-Martin et al., 2009)
CPA6	Juvenile myoclonic epilepsy	(Sapio et al., 2015)
EFHC1	Juvenile myoclonic epilepsy	(Suzuki et al., 2004)
GABRA1	Juvenile myoclonic epilepsy, Generalized epilepsy with febrile seizures plus, Generalized tonic-clonic seizure alone	(Cossette et al., 2002; Johannesen et al., 2016)
GABRA6	Childhood absence epilepsy	(Hernandez et al., 2011)
GABRG2	Childhood absence epilepsy, Generalized epilepsy with febrile seizures plus	(Baulac et al., 2001; Kananura et al., 2002)
HCN1	Childhood absence epilepsy, Genetic generalized epilepsy, Generalized epilepsy with febrile seizures plus	(Bonzanni et al., 2018; Marini et al., 2018)
KCNN3	Juvenile myoclonic epilepsy	(Vijai et al., 2005)
MAST4	Childhood absence epilepsy, Genetic generalized epilepsy, Juvenile myoclonic epilepsy	(Landoulsi et al., 2018)
NIPA2	Childhood absence epilepsy	(Xie et al., 2014)
SCN1A	Generalized epilepsy with febrile seizures plus	(Abou-Khalil et al., 2001)
SCN1B	Generalized epilepsy with febrile seizure plus	(Wallace et al., 1998)
SCN2A	Generalized epilepsy with febrile seizures plus	(Baulac et al., 1999; Sugawara et al., 2001)
SLC2A1	Childhood absence epilepsy	(Soto-Insuga et al., 2019)

Abbreviation: GGE, Genetic generalized epilepsy.

The discovery of new epilepsy genes in turn promoted the idea to investigate the genetic mechanism behind epilepsy (Stafstrom & Tempel, 2000). Since early discovered epilepsy genes, such as *SCN1A*, *SCN1B*, *CACNB4*, *GABRA1*, *GABRG2* and *CHRNA4*, were mainly involved in ion channels, epilepsy was once classified as channelopathy

disease (Hirose et al., 2002a; Hirose et al., 2002b; Mulley et al., 2003), and it was hypothesized to be a result of excitation-inhibition imbalance (Stafstrom, 2014). Due to active involvement of ion channels in the pathogenesis of epilepsy, the ion channels are selected as the main therapeutic target in many antiepileptic drugs (Kobayashi et al., 2020; Moavero et al., 2017), even the latest third generation antiepileptic drugs are still designed to influence the activity of ion channels to achieve seizure suppression (Stefanović et al., 2018).

Nevertheless, channelopathy fails to explain all causes of genetic epilepsy as there are epilepsy-associated genes that do not encode ion channels (Mulley et al., 2003; Ptácek & Fu, 2003), examples of such genes in GGE are the *BRD2* and *ICK* which respectively encode the transcription regulator and Ser/Thr kinase (Gilsoul et al., 2019). Both *BRD2* and *ICK* have been shown to influence the development of central nervous system in mouse models (Gilsoul et al., 2019). The discovery of non-ion-channel-coding epilepsy genes has provided some new insights into epilepsy genetics, it is proposed that epilepsy can be induced not only by ion channel genes, but also by genes that are involved in brain development (Steinlein, 2004).

Besides, there are also studies suggesting that GGE might be polygenic (Koeleman, 2018; Thakran et al., 2020). Nevertheless, no polygenic disease model has been proposed for GGE until 2018 (Mullen et al., 2018), when the segregation analysis conducted by Landoulsi et al. (2018) showed the inheritance of two GGE-associated genes in some GGE families. Meanwhile, Lee et al. (2018) performed a gene-panel screening on 57 Korean GGE patients and found that each patient is averagely carrying 2.8 potential GGE-inducing variants. This study provides indirect evidence that GGE might be polygenic. If GGE is polygenic, then the pathogenesis of GGE is probably coordinated by a complex molecular mechanism. In order to understand the genetic mechanism behind GGE, some

studies utilize network biology techniques to investigate the biological pathways associated with GGE, these studies are described in Section 2.5.3.

2.3 Whole Exome Sequencing

2.3.1 Impact and Application of Whole Exome Sequencing in Epilepsy

As illustrated by Boycott et al. (2013), one of the greatest impacts of whole exome sequencing (WES) on human disease research is speeding up the discovery of disease-related genes. In addition to inherited mutations, WES facilitates the discovery of *de novo* mutations (Boycott et al., 2013; Koboldt et al., 2013). *De novo* mutations, which can arise from mutagenesis in the parents during oogenesis and spermatogenesis, have been shown to be associated with a great deal of human diseases including epilepsy (Acuna-Hidalgo et al., 2016; Pranckėnienė et al., 2018; Veeramah et al., 2013). Many new disease-related genes, including epilepsy genes, are identified through *de novo* mutations (Acuna-Hidalgo et al., 2016), examples of such epilepsy genes include *KCNA2*, *SCN8A* and *SNAP25* (Pena & Coimbra, 2015; Rohena et al., 2013; Syrbe et al., 2015; Veeramah et al., 2012). Apart from *de novo* mutations, the use of WES has been recently expanded to detect somatic mutations in epilepsy (Koh & Lee, 2018; Stosser et al., 2018), and current studies have identified a few somatic mutations which are associated with neurodevelopmental disorders including focal cortical dysplasia (Heinzen, 2020; Iffland & Crino, 2019). Interestingly, there is a study demonstrating the use of WES for detecting copy number variation in epilepsy (Tsuchida et al., 2018). Despite there is limitation in the exome-sequencing-based copy number variant detection at current stage (Marchuk et al., 2018; Zhao et al., 2020a), it does demonstrate the potential of WES in future copy number variant research.

The use of WES is not limited to research activities but also in clinical diagnosis and management (Jackson et al., 2018; Tetreault et al., 2015). This is best exemplified by Wei

et al. (2019) in diagnosis of glucose transporter deficiency type 1 syndrome in which the drug-resistant patients achieved seizure control after changing the treatment to ketogenic diet following the detection of *SLC2A1* mutation. Another example showing the practicality of WES in clinical diagnosis is demonstrated by Osman et al. (2020), where the patient is diagnosed as having pyridoxine-dependent epilepsy after a *ALDH7A1* mutation has been identified by WES. Meanwhile, Peng et al. (2019) showed the improvement of seizure control in some paediatric epilepsy patients following the use of gene panels and WES-guided targeted therapies. Despite the use of WES seems promising in clinical application, it does show some drawbacks. One of the biggest disadvantages of WES is the identification of variants of uncertain significance (VUS) (Timmermans et al., 2017), which can be a challenge in genetic counselling and clinical decision making (Christiaans et al., 2019; Hoelz et al., 2020; van Lint et al., 2019). It is advised that the identified VUS to be reviewed by a multidisciplinary team before any genetic pre-test and post-test counselling (Christiaans et al., 2019).

2.3.2 Advantages and Limitations of Whole Exome Sequencing over Gene Panels and Whole Genome Sequencing

Gene panel (also known as targeted sequencing), whole exome sequencing (WES) and whole genome sequencing (WGS) are the three popular next generation sequencing (NGS) approaches in which most researchers will consider in genetic research (Sun et al., 2015). Each approach has their own advantages and disadvantages. Among the three approaches, gene panel costs the least compared to WES and WGS as gene panel covers the smallest sequencing region by targeting only specific genes; in contrast, WES targets all the exons of every gene and WGS targets the entire genome (Dunn et al., 2018; Liu et al., 2015). However, the identification of causative variants by gene panel will only success if the gene is included in the panel (Sun et al., 2015), increasing the number of genes in the

panel may increase the diagnostic yield rate of the panel (Wu et al., 2020). In contrast, WES and WGS are able to identify novel disease-associated genes in addition to disease-causing variants in reported disease genes (Barozzi et al., 2019; Lionel et al., 2018; Liu et al., 2015; Sun et al., 2015). The cost of WES is significantly cheaper than WGS, it is estimated that a single WES test can cost from $555 to $5,169 and a single WGS test can cost from $1,906 to $24,810 (Schwarze et al., 2018). The lower cost of WES is largely achieved by reduced sequencing cost as it targets only the exonic regions which constitute about 1.5% of the genome (Falardeau et al., 2017). In comparison with WGS, WES remains as a popular technique in genetic diagnosis as it is postulated that 85% of disease-causing mutations are located within the exons (Liu et al., 2015; Tetreault et al., 2015). Despite this, WGS offers several advantages over WES. Firstly, WGS can identify the mutations in the intronic regions which will potentially be missed by WES (Falardeau et al., 2017). Even though intron does not encode any protein, it may still play crucial role in molecular processes such as improving translational efficiency and alternative splicing (Vaz-Drago et al., 2017). Mutations in intronic regions, such as those in the *Alu* element may associate with human diseases (Kim et al., 2016; Vaz-Drago et al., 2017). Besides, WGS can detect copy number variants with higher accuracy compared to WES (Hehir-Kwa et al., 2015; Tetreault et al., 2015). In comparison with gene panel and WES, WGS provides the most comprehensive data, the major drawbacks limiting the use of WGS are the high sequencing cost and the lack of bioinformatics tools (Dunn et al., 2018; Han & Lee, 2020). However, as the technology advances and the sequencing cost continues to drop, it is possible that WGS will become the standard of genetic diagnosis in future (Dunn et al., 2018; Ostrander et al., 2018). For the time being, gene panel is still the cheapest option for genetic diagnosis, but WES is a preferred option if the discovery of novel disease genes is desired. Table 2.4 shows the comparison between gene panel, WES and WGS.

Table 2.4: Comparison between gene panel, WES and WGS.

Aspect	Gene panel	WES	WGS
Cost	Lowest	Moderate	Highest
Coverage (sequencing region)	Targets only specific region in the genome	Targets all exonic regions in the genome	Targets the entire genome
Ability to detect novel disease genes	No, unless the gene is included in the panel	Yes	Yes
Major drawback	Detection of disease-associated variants/genes will only success if the gene is included in the panel	Unable to detect the disease-associated variants in the intronic regions	Lack of bioinformatic tools, high sequencing cost

Abbreviations: WES, Whole exome sequencing; WGS, Whole genome sequencing.

2.4 Bioinformatics and Whole Exome Sequencing (WES) Data Analysis

2.4.1 WES Data Analysis Pipeline

WES, being one of NGS approaches, generates enormous amount of sequencing data. These data need to be analysed with a systematic workflow that involves the use of multiple bioinformatic software (Horner et al., 2010; Magi et al., 2010). Despite lots of software were available for data analysis in the early stage of NGS era, the lack of guidance, best practice and benchmarking of these tools have imposed troubles in data analysis as the validity of the results can be compromised by improper bioinformatic and computational analysis (Aliferis et al., 2011; Magi et al., 2010). In 2012, Altmann et al. (2012) and Dolled-Filhart et al. (2013) published a workflow highlighting the steps in NGS data analysis and suggested some bioinformatic packages for use in each step. In order to provide a clearer guideline to NGS data analyst, Van der Auwera et al. (2013) published a standard workflow known as Genome Analysis Toolkit (GATK) Best Practice which describes in detail about the methods for NGS (including WES) data analysis using Burrows-Wheeler Alignment (BWA) and GATK. Following the announcement of GATK Best Practice, Pirooznia et al. (2014) evaluated the performance

of GATK by validating the variant calls using Sanger sequencing and array genotyping data. They reported that accuracy of more than 99% is achievable if best practice is used in the analysis. Over the years, new variant calling software, such as Platypus, SpeedSeq and DeepVariant, has been developed to aid the data analysis (Chiang et al., 2015; Poplin et al., 2018; Rimmer et al., 2014). Despite this, the results generated from GATK are still very comparable to the new software (Li et al., 2019; Supernat et al., 2018). Even in 2020, many studies are still using GATK Best Practice in their WES data analysis, examples of such studies include Chen et al. (2020b), Isik et al. (2020), Ngoc et al. (2020) and Odumpatta & Mohanapriya (2020). Figure 2.1 shows schematic diagram of GATK Best Practice sourced from Broad Institute (2020). According to Broad Institute, the workflow of the best practice can be divided into three main stages: data pre-processing, variant discovery, and variant annotation. In the first stage, the reads are aligned to the reference sequence using BWA and duplicated reads are removed; in the second stage, variants such as single nucleotide polymorphism are called using variant calling software such as GATK HaplotypeCaller; and in the last stage, the called variants can be annotated using annotation tools like VariantAnnotator (Broad Institute, 2020; DePristo et al., 2011; Van der Auwera et al., 2013). The GATK workflow used in this study is described in Section 3.7.

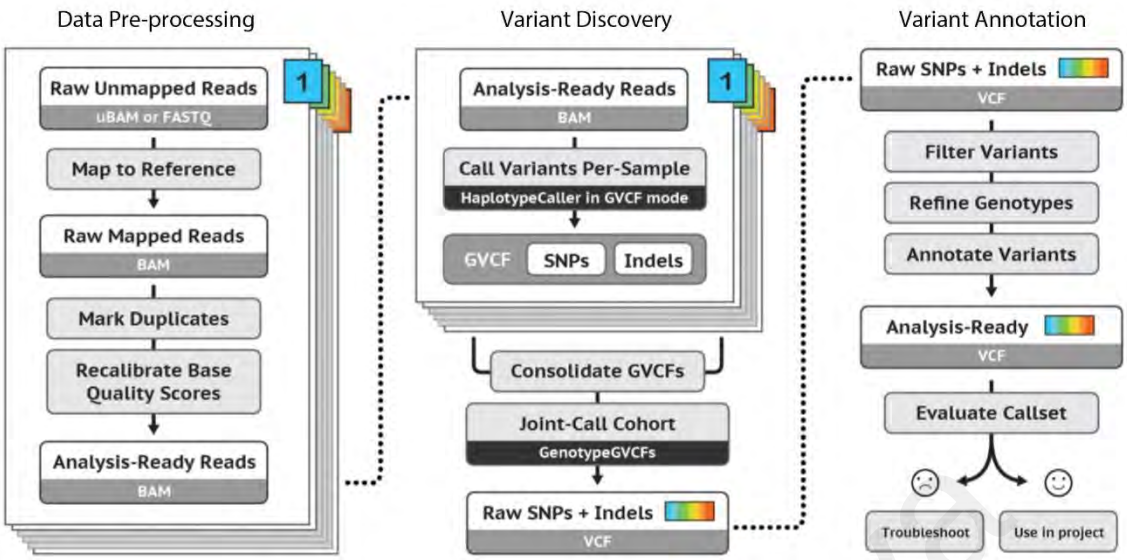


Figure 2.1: Schematic diagram showing the overall workflow of GATK Best Practice for germline short variant discovery (Copyright permission from Broad Institute (2020)).

2.4.2 *In Silico* Prediction for Variant Pathogenicity

Besides identifying genetic variants from WES data, another contribution of bioinformatics in WES is the *in silico* pathogenicity prediction for these variants. Many bioinformatic tools such as SIFT, Polyphen-2 and MutationTaster have been made available to perform this task, these tools utilize specific algorithm, which can be either protein-sequence based, structure based, or both, to calculate the effect score of the variants and predict their pathogenicity (Dong et al., 2015; Mueller et al., 2015; Richards et al., 2015; Tang & Thomas, 2016). Despite many prediction tools are developed, there is no consensus on the best performing prediction tools yet. For instance, SIFT and Provean perform the best in the prediction of missense variants for genes *GJB2*, *GJB6* and *GJB3* (Pshennikova et al., 2019), but Polyphen-2 is the best performer in the predicting pathogenicity of *LDLRAP1* mutation (Shaik et al., 2020). On other hand, both SIFT and Polyphen-2 failed to predict correctly the effect of all the variants in *RYR1* and *CACNA1S* (Schiemann & Stowell, 2016). These studies supported Anderson & Lassmann (2018) that the performance of prediction tools may differ by disease phenotype. Since

no prediction tool is perfect, it is advised to support the findings from prediction tools with additional data such as family segregation analysis, functional studies, literature and minor allele frequency (Coll et al., 2017; Fang et al., 2017; Schiemann & Stowell, 2016). Richards et al. (2015) did provide a very comprehensive guideline on the criteria in determining the pathogenicity of a given variant. While bioinformatic tools cannot absolutely predict the pathogenicity of a variant, they can help to prioritize or shortlist the variants of interest before proceeding with functional studies which are generally very expensive and time consuming (Bonjoch et al., 2019; Borras et al., 2017; Nykamp et al., 2017).

2.5 Gene Prioritization, Functional Characterization and Pathway Analysis

2.5.1 Gene Prioritization

As discussed in Section 2.3.2, one of the added advantages of WES is the ability to detect novel disease genes. However, exome and genome sequencing often produce too many variants to be assessed even after pathogenicity filtering based on prediction software (Dashnow et al., 2019; Lohmann & Klein, 2014). In order to overcome this problem, one of the common strategies is to carry out trio analysis, in which WES is performed on the proband and parents (Dashnow et al., 2019). Trio analysis allows faster identification of *de novo* mutations (Carneiro et al., 2018; Du et al., 2018; Hamanaka et al., 2020; Tong et al., 2018; Zhu et al., 2015). Besides, it also detects compound heterozygous mutations, the key evidence for the identification of autosomal recessive disease genes (Zhang et al., 2020; Zhou et al., 2020). For familial cases, WES can be conducted on multiple affected family members followed by the identification of shared variants between the affect members, these shared variants are potentially associated with the disease (Goh & Choi, 2012), examples of studies using this approach include Basit et al. (2017), Lung et al. (2020) and Wells et al. (2013). Nevertheless, this approach is

generally more expensive compared to singleton sequencing as it involves the library preparation and sequencing of more than one individual (Dashnow et al., 2019). In spite of this, there is strategy to prioritize variants from singleton sequencing. One of the methods is targeted analysis, where variants from known disease genes are prioritized and analysed (Sun et al., 2015). For instance, one can first analyse the variants from epilepsy-associated genes like *SCN1A*, *SCN2A*, *PCDH19* and *DEPDC5* for WES cases involving epilepsy patients (Perucca et al., 2017; Sahli et al., 2019). If no variant of interest is found in the initial analysis, the analysis can be further expanded to discover the novel genes associated with the disease (Sun et al., 2015). In this stage, the novel genes can be ranked using gene prioritization software to assess their likelihood to be associated with the disease, and pathway analysis can then be conducted to discover the possible underlying molecular pathways (Jin et al., 2018; Sevim Bayrak & Itan, 2020; Wang & Xing, 2013).

Gene prioritization software utilizes multiple aspects encompassing functional similarity, molecular pathway, cross-species evidence, expression, disease phenotype and literature to score and rank the novel genes in respect of known disease genes (Bromberg, 2013). Similar to the prediction tools discussed previously, each gene prioritization software has their own advantages and disadvantages (Zolotareva & Kleine, 2019). Despite this, prioritization software can help in shortlisting and proposing the genes which are more worthy for further investigation (Feng et al., 2019; Isakov et al., 2017; Saik et al., 2019). This context is best exemplified by Chen et al. (2017), whom identified the key genes associated with neuropathic pain, the candidate genes were shortlisted from 430 to 206 following gene prioritization. Another study demonstrating the usefulness of gene prioritization is Sevim Bayrak et al. (2020), they ranked the candidate genes identified from pathway and protein-protein interaction analysis through gene prioritization, and they found that 23 out of 95 candidate genes are more likely to be associated with congenital heart disease. Nevertheless, while gene prioritization can help

in shortlisting the candidate genes, it does not provide a clear perspective on how the prioritized genes are related to a disease (Liu et al., 2019). As a result, many studies including Chen et al. (2017) couple gene prioritization with additional analyses such as pathway and protein-protein interaction analysis in order to explore the possible mechanism between the prioritized genes and known disease genes in the pathogenesis of the disease.

2.5.2 Functional Annotation, Protein-Protein Interaction (PPI) Analysis and Pathway Analysis

Functional annotation, protein-protein interaction (PPI) analysis and pathway analysis are three distinct bioinformatic-analysis techniques but are related with each other and can be applied together in human disease research (Jin et al., 2018). In the context of bioinformatics, functional annotation is the prediction of gene function using computational approach, this is best represented by the Gene Ontology (GO) annotation which provides standardized terms describing the function of gene products (Gene Ontology Consortium, 2004; Guzzi et al., 2012; Weichenberger et al., 2017; Zhao et al., 2020b). As demonstrated by the study on the role of *SEMA5A* in infantile spasm (Wang et al., 2019), GO annotation is very useful in predicting the function of newly discovered disease gene. However, it does not reveal a detailed mechanism on how the gene is involved in disease pathogenesis. Under this circumstance, PPI and pathway analysis would be very advantageous as many complex human diseases are triggered by malfunction of biological pathway that involves the interaction between multiple gene products (Furlong, 2013). PPI analysis uses multiple data sources like co-immunoprecipitation experiment, literature, three-dimensional structure, co-expression data and sequence homology to predict the functional relationship, such as dimerization and complex formation, between proteins (Miki et al., 2018; Sonawane et al., 2019;

Stoney et al., 2018). This technique provides researchers a deeper understanding of the molecular mechanism behind a disease (Pattin & Moore, 2009). Despite this, the data obtained from PPI analysis alone might be difficult to interpret due to the lack of functional data which relates the PPI to the underlying biological processes (Liu et al., 2017). Nevertheless, this shortcoming can be overcome by pathway analysis which predicts the overall function or the biological pathway carried out by the group of interacting genes (Reimand et al., 2019).

Functional annotation, PPI and pathway analysis have brought several contributions in disease research. Firstly, these approaches allow the discovery of the underlying molecular pathway behind a disease (Furlong, 2013). For instance, Lin et al. (2017) discovered that the regulatory role of microRNA in cervical squamous cell carcinoma is mainly impacting on the MAPK signalling pathway; and Iglesias et al. (2019) found that the pathogenic role of prior proteins is established through nucleic acid binding and gene regulation. Besides, they facilitate the prediction of disease-associated genes (Jin et al., 2018; Sonawane et al., 2019). As demonstrated in many studies, novel genes have been reported in multiple human diseases such as chronic lymphocytic leukaemia, congenital heart disease, infantile spasms and gastric cancer following the use of functional annotation, PPI and pathway analysis (Kim et al., 2020; Sevim Bayrak et al., 2020; Wang et al., 2019; Yang, 2020). On top of that, the disease genes can be assessed with PPI analysis to study their potential risk, it is hypothesized that key disease genes are having more interactants in the disease pathway (Cai et al., 2010; Feldman et al., 2008; Sonawane et al., 2019). This concept is used by Chen et al. (2017) and Ghatge et al. (2018) in the identification of *CTNNB1* and *EGFR* as the key gene in neuropathic pain and coronary artery disease respectively.

The identification of key genes and biological pathway in human disease is very beneficial in the medical field. On one hand, it explains the relationship between a disease

and its comorbidities (Li & Agarwal, 2009). As shown in the case of polycystic ovarian syndrome, women with the syndrome are prone to develop hypertension, migraine and schizophrenia due to sharing disease genes and pathway (Balqis Ramly et al., 2019). On the other hand, it reveals the potential therapeutic targets for the disease, this will greatly assist in the drug design for disease treatment (Furlong, 2013; Liang et al., 2020; Pattin & Moore, 2009; Stoney et al., 2018). Examples of studies demonstrating the application of functional annotation, PPI and pathway analysis in drug design include Yang et al. (2020) and Zhao et al. (2019). Due to its potential in disease research, many bioinformatic tools have been developed to aid functional annotation, PPI and pathway analysis (Al-Harazi et al., 2019; Jin et al., 2014; Miryala et al., 2018). Despite this, it is still very important to verify the results using experimental data such as those from *in vitro* and *in vivo* studies (Wang et al., 2019; Zhao et al., 2019).

2.5.3 Gene Prioritization and Pathway Analysis in Epilepsy-Genetic Studies

Gene prioritization and pathway analysis have been applied in epilepsy-genetic studies since 2010. Despite a few epilepsy-related pathways have been reported by early studies like Sha et al. (2010), Jia et al. (2011) and Speed et al. (2014), the study subjects in these studies were not grouped based on epilepsy types (generalized epilepsy or focal epilepsy). As a result, the difference in the genetic mechanism between generalized and focal epilepsy may not be distinguishable. In order to tackle this issue, newer studies have given emphasis to epilepsy type and the selection of study subjects is more specific. For instance, Jin et al. (2016) has highlighted that their study focused on tumour-induced epilepsy, and Xiao et al. (2019) focused on the pathway analysis of temporal lobe epilepsy. In the case of genetic generalized epilepsy (GGE), only two pathway analysis studies have been conducted. The first study was conducted by Ozdemir et al. (2019), they identified five Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, namely ‘neurotrophin

signaling pathway', 'pathways in cancer', 'focal adhesion', 'metabolic pathway' and 'MAPK signaling pathway', that are significantly enriched in their study subjects comprised of 15 trios. Meanwhile, the second study, Epi25 Collaborative (2019) divided their analysis into two stages: the gene-set-based analysis and the gene-based analysis, in which the former measured the odd ratio of the gene groups while the latter focused on the odd ratio of individual genes in GGE. From the analysis, Epi25 Collaborative (2019) reported that the GABA-A receptor genes and voltage-gated cation channel genes are closely associated with GGE; and from their gene-based analysis, it is found that the *CACNA1G*, *EEF1A2*, and *GABRG2* are having higher association with GGE compared to other genes in the analysis. Despite the analysis by Epi25 Collaborative (2019) is very comprehensive, the results may not apply to Malaysians. As demonstrated by Wu et al. (2019), Malays, Chinese and Indians have different genetic composition compared to Europeans. Hence, the relative risk of each gene and the associated variants with epilepsy might be different among Malaysians, extra studies are still needed to investigate the genetic basis of GGE in the multi-ethnic Malaysian population.

2.6 Contribution of Genetics in Epilepsy Treatment and Antiepileptic Drug Development

Genetic studies have brought several contributions to the field of epilepsy. Firstly, it may aid the treatment decision (Balestrini & Sisodiya, 2018). Despite antiepileptic drugs are generally reducing seizures by decreasing neuronal excitation signals or by increasing neuronal inhibition signals (Santulli et al., 2016), there are differences in the mechanism of action among the antiepileptic drugs (Rogawski & Löscher, 2004). For instance, lamotrigine, carbamazepine and lacosamide suppress seizures by inhibiting the activity of the voltage-gated sodium channels; bromide and topiramate enhance the effect of inhibitory neurotransmitter gamma-aminobutyric acid; levetiracetam binds to synaptic

vesicle proteins to block neurotransmitter release at the synapse; whereas sodium valproate is expected to target multiple cellular components including the sodium channels, N-methyl-d-aspartate receptors and gamma-aminobutyric acid (Kobayashi et al., 2020; Moavero et al., 2017; Stefanović et al., 2018). Due to different mode of action, the efficiency of each drug in treating genetic epilepsy may vary. As reviewed by Perucca & Perucca (2019), carbamazepine is less effective in treating *SCN1A*-induced Dravet syndrome compared to sodium valproate, the blockage of sodium channels by carbamazepine fails to relieve the symptoms as this disease is caused by the functional loss in sodium channels; however, carbamazepine is an effective treatment for *KCNQ2*-induced benign familial neonatal epilepsy as the mutation in *KCNQ2* is impacting the potassium channels but not the sodium channels, the blockage of sodium channels in this case can help to impede the generation of action potential. The knowledge of genetics allows a better understanding of the cellular defects induced by an epilepsy-causing mutation, this may help the medical practitioners to provide a more effective treatment to patients by administrating the correct medications that tackle the cellular abnormalities caused by the mutation (Naimo et al., 2019; Perucca & Perucca, 2019). As demonstrated in Peng et al. (2019), 52.9% of the patients attained seizure-control after changing antiepileptic drug treatment based on genetic diagnosis.

Besides, many new candidate therapeutic targets have been identified through genetic research (Kambli et al., 2017). For example, Zhang et al. (2019) suggested the potential of *HDAC4* as a therapeutic target following their experiment in mouse models that showed an increased expression of GABA receptors, one of the key components in potentiating inhibitory signal in the brain, after silencing of *HDAC4*; meanwhile, the downregulation of cyclooxygenase-2 in valproate responders promoted the idea to develop new antiepileptic drugs that target cyclooxygenase-2 (Rawat et al., 2020); and the G protein-coupled receptor 40 is proposed as a new therapeutic target attributable to

its regulatory role in the expression of N-methyl-d-aspartate receptors (Yang et al., 2018). The identification of new therapeutic targets may shed some light on the development of new antiepileptic drug in future. As mentioned by Stefanović et al. (2018), the development of new antiepileptic drugs is needed as there are still 20-30% of the patients who suffer from drug-resistant epilepsy.

As a supplement, there are also some recent studies that emphasize the molecular mechanism behind the side effect of antiepileptic drugs. For instance, Zhang et al. (2021) discovered that valproate-induced weight gain is related with the expression of *FTO*, and Chen et al. (2019) proposed that valproate-induced hepatotoxicity is caused by the impairment of farnesoid X receptor signalling pathway. Meanwhile, Walters et al. (2020) identified the effect of vigabatrin on the expression of retinal genes, and Shizu et al. (2020) proposed that phenobarbital is raising the blood triglyceride levels by interfering the interaction between peroxisome proliferator-activated receptor alpha and peroxisome proliferator-activated receptor-gamma coactivator 1-alpha. Despite the findings from these studies are mostly derived from mouse models, they may still provide some clues on the methods to improve the current epilepsy medications in the sense of minimizing side effect.

2.7 Genetic Studies on Epilepsy in Malaysia

Although many genes are found to be associated with epilepsy, there are only a few epilepsy-genetic studies that have been conducted in Malaysia. In 2011, Haerian et al. (2011) showed the lack of association between *SYN2* rs3773364 variant with epilepsy among Malaysians. In 2012, another group of researchers from Universiti Sains Malaysia performed a mutational analysis of *SCN1A* in 36 Malaysian patients diagnosed with generalized epilepsy with febrile seizures plus and reported two *de novo* mutations from their cohort (Emmilia H. Tan et al., 2012; Emmilia Husni Tan et al., 2012). 3 years later,

Haerian et al. (2015a) showed the association of *TIMP4* rs3755724 polymorphism with focal epilepsy among Malaysian Chinese. In the same year, the same research team reported that *RORA* rs12912233 variant can increase the epilepsy-developing risk in Malaysian Chinese by acting synergically with *RORA* rs880626 and *SCN1A* rs3812718 (Haerian et al., 2015b). In 2016, Hidayati Mohd Sha'ari et al. (2016) carried out an association study and found a few *BDNF* variants which may contribute to epilepsy among Malaysian Chinese and Indians. In 2017, Lim et al. (2017) conducted a cross-sectional study to investigate the variation in GGE prevalence among the Malays, Chinese and Indians, and the results showed a significant variation in GGE prevalence among the ethnic groups; besides, the results also demonstrated that about 30% of the GGE patients in Malaysia are having family history, this data has suggested the inheritance of GGE among Malaysians. In the context of familial epilepsy, only one study has been conducted in Malaysia. The study reported a rare *EPM2A* c.758A>T mutation and the *EPM2A* whole gene deletion in a Malay family with Lafora disease (Tee et al., 2019), this finding provided additional evidence that shows the autosomal recessive nature of *EPM2A*. Even though these studies have contributed some knowledge to the genetics behind epilepsy among Malaysians, the findings are still very limited primarily due to the use of targeted sequencing or single variant association studies. As mentioned in Section 2.3.2, targeted analyses may miss out some genetic variants as not every gene is screened during the genetic tests. Furthermore, the lack of pathway analysis has hampered the understanding of molecular pathways underlying epilepsy in Malaysia. Therefore, this project is conducted to study the possible novel genetic factors and the biological pathways associated with GGE among Malaysians.

CHAPTER 3: METHODOLOGY

3.1 Ethical Approval

The study protocol of this project was approved by University Malaya Medical Centre (UMMC) Medical Research Ethics Committee. The details of the ethical approval are as shown below:

Ethics Committee Reference Number: 944.3

Date of approval: 28 April 2016

3.2 Study Population

This study involved 14 unrelated Malaysian probands (4 Malays, 5 Chinese and 5 Indians) fulfilling the inclusion criteria mentioned in Table 3.1. Among the 14 probands, trio whole exome sequencing (WES) was conducted on the 2 Chinese probands and singleton WES was conducted on the remaining 12 probands (Table 3.2). The diagnosis of genetic generalized epilepsy (GGE) was confirmed by epileptologist, and all the probands were recruited from neurology clinic in UMMC. Besides, family members of the probands were also recruited for those agreed to participate in this study, these family members were involved in segregation analysis (Section 3.10). Informed written consent was obtained from all study subjects upon enrolment into this study; for the individuals under 18 years old, the consent was obtained from their parents.

Table 3.1: Inclusion criteria of study subjects involved in this study.

Inclusion criteria of study subjects
1. Diagnosed with genetic generalized epilepsy.
2. Have family history of epilepsy.
3. The epileptic syndromes are caused by unknown/unidentified factor.

Table 3.2: List of samples underwent WES.

Family ID	Sample ID	Ethnicity	Remark
List of samples underwent trio WES			
F-1	GGE-1	Chinese	Proband
	GGE-2	Chinese	Father of GGE-1
	GGE-3	Chinese	Mother of GGE-1
F-4	GGE-4	Chinese	Proband
	GGE-5	Chinese	Father of GGE-4
	GGE-6	Chinese	Mother of GGE-4
List of samples underwent singleton WES			
F-7	GGE-7	Malay	Proband
F-8	GGE-8	Chinese	Proband
F-9	GGE-9	Chinese	Proband
F-10	GGE-10	Indian	Proband
F-11	GGE-11	Malay	Proband
F-12	GGE-12	Indian	Proband
F-13	GGE-13	Indian	Proband
F-14	GGE-14	Chinese	Proband
F-15	GGE-15	Malay	Proband
F-16	GGE-16	Indian	Proband
F-17	GGE-17	Indian	Proband
F-18	GGE-18	Malay	Proband

Abbreviation: WES, Whole exome sequencing.

3.3 Sample Collection

For all the samples underwent WES including the parents of GGE-1 and GGE-4 (Table 3.2), 6 mL of peripheral blood sample was collected from them using dipotassium ethylenediaminetetraacetic acid (K2EDTA) tubes (Becton Dickinson, USA). For other family members who consented to participate in this study, either peripheral blood or buccal cell sample was collected from them.[1] The buccal cell samples were collected using sterile cotton swabs, 3 swabs were collected from each individual.

[1] Peripheral blood sample was collected when sampling was conducted in UMMC. However, for family members under 18 years old or prefer on-site sampling, buccal cell sample was collected.

3.4 Genomic DNA Extraction

3.4.1 Genomic DNA Extraction from Blood Samples

2 mL of peripheral blood from each sample was used for genomic DNA extraction. DNA extraction was performed with QIAamp DNA Blood Midi Kit (Qiagen, Germany) in accordance with the spin protocol provided by the manufacturer (Appendix A). American Chemical Society (ACS) grade ethanol (Emsure ACS, Merck, Germany) was used to precipitate DNA before loading them into QIAamp Midi column (step 5, Appendix A). In the final step of the extraction (step 12, Appendix A), 200 µL of Buffer AE (provided with QIAamp DNA Blood Midi Kit) was used to elute the DNA. For samples undergoing WES (Table 3.2), the DNA quality and quantity were checked using gel electrophoresis and spectrophotometer prior to WES (Section 3.5).

3.4.2 Genomic DNA Extraction from Buccal Cell Samples

For buccal cell samples, DNA extraction was conducted using Presto Buccal Swab gDNA Extraction Kit (Geneaid, Taiwan) under the protocol provided by the manufacturer (Appendix B). ACS grade ethanol (Emsure ACS, Merck, Germany) was used for DNA precipitation before loading into the GD column (step 4, Appendix B). In the elution step (step 6, Appendix B), 50 µL of Elution Buffer (provided with Presto Buccal Swab gDNA Extraction Kit) was used to elute the DNA. In order to increase DNA yield, the column was left standing for 5 minutes after addition of Elution Buffer before proceeding with centrifugation. All centrifugation steps in this protocol were carried out under the centrifugal force of 14,000 x g.

3.5 Genomic DNA Qualification and Quantification

3.5.1 Genomic DNA Qualification Using Gel Electrophoresis

Gel electrophoresis was conducted to check the quality of DNA samples. Before starting gel electrophoresis, 1X tris-borate-EDTA (TBE) buffer was prepared based on the recipe provided by Cold Spring Harbor Protocols ("TBE electrophoresis buffer (10X)," 2010), 12.11 grams of tris (Hoefer Caliber, United States), 6.18 grams of boric acid (Fluka, United States) and 0.74 grams of EDTA, disodium salt, dihydrate (Amresco, United States) were added to 1 L of autoclaved distilled water. The mixture was then stirred using a magnetic stirrer until all the precipitates had dissolved and the solution became clear.

For gel electrophoresis, a 1% agarose gel was first prepared by adding 1.5 grams of agarose (First Base, Malaysia) and 150 mL of 1X TBE buffer into a 250 mL conical flask. The mixture was mixed evenly by shaking it gently and microwaved until all the agarose was dissolved and a clear solution was obtained. 15 µL of FloroSafe DNA Stain (First Base, Malaysia) was added into the agarose solution and mixed gently. The agarose solution was then partially cooled under running tap water and poured into a gel casting tray. Next, an electrophoresis comb was inserted into the gel tray and the gel was left to solidify. Once the gel was solidified, the gel was transferred to an electrophoresis tank and 1X TBE buffer was added into the tank until the gel surface was completely submerged. 5 µL of Lambda DNA/*Hind*III Marker (Thermo Scientific, United States) was loaded into the first well as DNA marker. 4 µL of genomic DNA was then mixed thoroughly with 2 µL of 6X DNA loading dye (AITbiotech, Singapore) on a piece of Parafilm and subsequently loaded into the next well. The gel electrophoresis was conducted using PowerPac Basic (Bio-rad, United States) under a constant voltage of 120 V for 40 minutes.

3.5.2 Genomic DNA Quantification Using Spectrophotometer

The quantity and purity of the DNA samples were measured using NanoDrop 2000c Spectrophotometer (Thermo Scientific, United States). 1 µL of each DNA sample (without dilution) was loaded singly on the pedestal of Nanodrop 2000c for spectrophotometric measurement. The concentration (in ng/µL) and absorbance ratio A260/A280 of each sample were recorded.

3.6 Whole Exome Sequencing (Outsourced)

Whole exome sequencing (WES) for all samples listed in Table 3.2 was outsourced to Macrogen Incorporation (South Korea). In this step, 1 µg of genomic DNA was aliquoted and sent to Macrogen Incorporation for library preparation and WES. According to Macrogen Incorporation, the libraries were prepared with SureSelect Human All Exon V4 (Agilent, United States) and sequencing was carried out using HiSeq 2000 sequencing system (Illumina, United States) with 100 bp paired-end sequencing under a sequencing depth of 30X.

3.7 Whole Exome Sequencing Data Analysis Pipeline

WES data analysis was started with the FASTQ files provided by Macrogen Incorporation following the completion of WES. The analysis pipeline was adapted from Genome Analysis Toolkit (GATK) best practice developed by Broad Institute (Van der Auwera et al., 2013). Poor quality reads were first trimmed using Trimmomatic (Bolger et al., 2014). The trimmed sequence reads were mapped to Genome Reference Consortium Human Build 37 (GRCh37) using Burrows-Wheeler Aligner (Li & Durbin, 2009), duplicated reads were then marked and removed by Picard (http://broadinstitute.github.io/picard) and base quality score recalibration was conducted with the BaseRecalibrator from GATK suite (DePristo et al., 2011). GATK

HaplotypeCaller was used to perform variant calling in each sample, the variants were then filtered using GATK SelectVariants. Under this filter, all variants with QUAL < 30.0 were discarded. The remaining variants were annotated with wANNOVAR (Yang & Wang, 2015).

3.8 *In Silico* Pathogenicity Prediction and Selection of Potential Pathogenic Variants

The selection of potential pathogenic variants was based on the criteria provided by American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015). Table 3.3 shows the ACMG criteria utilized for determining the pathogenicity of the variants in this study. According to ACMG guidelines, variants with minor allele frequency (MAF) of more than 5% or greater than expected for disorder fulfil the classification criteria of benign variants (Table 3.3). The adjusted lifetime prevalence of epilepsy in Malaysia is 0.0078 (Fong et al., 2021). In this study, MAF < 0.01 was set as the cut-off criterion and all variants with MAF ≥ 0.01 were presumed benign. For the variants with MAF < 0.01, the pathogenicity was predicted by Sorting Intolerant from Tolerant (SIFT) (Sim et al., 2012) and Polymorphism Phenotyping v2 (Polyphen-2) (Adzhubei et al., 2010). The usage of *in silico* prediction software allows the identification of variants that fulfil the ACMG criterion "multiple lines of computational evidence support a deleterious effect on the gene or gene product", which is one of the criteria for classifying pathogenic variants (Table 3.3).

Table 3.3: The ACMG criteria for classifying pathogenic and benign variants (Richards et al., 2015).

ACMG criteria for classifying pathogenic variants	ACMG criteria for classifying benign variants
1. Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium. 2. Multiple lines of computational evidence support a deleterious effect on the gene or gene product.	1. Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium. 2. Allele frequency is greater than expected for disorder.

Abbreviation: ACMG, American College of Medical Genetics and Genomics.

In this project, the MAF of the variants was obtained from 1000 Genomes Project (1KGP) (The 1000 Genomes Project Consortium, 2015), Genome Aggregation Database (gnomAD) (Karczewski et al., 2020) and Singapore Sequencing Malay Project (Wong et al., 2013). Due to genetic variation between ethnic groups, different reference populations were used for determination of MAF in each ethnicity (Table 3.4). The variants with MAF < 0.01 in both reference population 1 and 2 (Table 3.4) and predicted damaging by both SIFT (SIFT score ≤ 0.05) and Polyphen-2 ("probably damaging" in HumVar model) were classified as potential pathogenic variants, the genes containing these variants were selected for subsequent functional characterization analysis (Section 3.9).

Table 3.4: Reference populations used for determining MAF of variants among Malays, Chinese and Indians. For 1KGP CDX + KHV, the MAF was derived by dividing the sum of allele count in CDX and KHV by sum of allele number in CDX and KHV.

Ethnicity	Reference population 1	Reference population 2
Malay	1KGP CDX + KHV	Singapore Sequencing Malay Project
Chinese	1KGP EAS	gnomAD EAS
Indian	1KGP SAS	gnomAD SAS

Abbreviations: 1KGP, 1000 Genomes Project; gnomAD: Genome Aggregation Database; CDX, Chinese Dai in Xishuangbanna, China; KHV, Kinh in Ho Chi Minh City, Vietnam; EAS, East Asian; SAS, South Asian.

3.9 Functional Characterization Analysis

A 3-step functional characterization analysis was conducted for the genes with potential pathogenic variants. In the first step, gene prioritization was conducted using ToppGene Suite (Chen et al., 2009). Next, the top 100 prioritized genes and the 34 training genes were subjected to protein-protein interaction (PPI) analysis with Search Tool for Retrieval of Interacting Genes/Proteins (STRING) v11.0 (Szklarczyk et al., 2019). In the final step, functional enrichment was conducted on the interacting genes using Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (Dennis et al., 2003). Sections 3.9.1, 3.9.2 and 3.9.3 describe the procedure involved in each step.

3.9.1 Gene Prioritization

The gene prioritization protocol used in this study was mainly adapted from Carrera et al. (2015), with minor modification in the gene ranking system. In this step, testing genes were ranked using ToppGene based on their functional similarity with training genes. Testing genes were defined as the genes with potential pathogenic variants identified from WES data of all probands listed in Table 3.2, except GGE-9. The reason of removing GGE-9 from this analysis is discussed in Section 5.1. Meanwhile, training genes were defined as the known JME-associated genes retrieved from the list of literature obtained using the search strategy described in the next paragraph. The ToppGene parameters used in determining functional similarity include GO: Molecular Function, GO: Biological Process, GO: Molecular Function, human phenotype, mouse phenotype, domain, pathway, pubmed, interaction, gene family, coexpression and disease.

Before starting gene prioritization, a literature search was conducted with Scopus using the search term ‘"Juvenile" AND "Myoclonic" AND "Epilepsy" AND "Gene"’, a total of 34 JME-associated genes were retrieved from various study types including functional

studies, association studies or case reports,[2] these genes were used as training genes and classified into 15 functional categories based on the protein function (Appendix D). The gene prioritization was conducted in two stages. In stage 1, all the 34 JME-associated genes were used as training genes. For stage 2, only the JME-associated genes from one functional category were used as training genes. Since there were 15 functional categories, gene prioritization was conducted 15 times in stage 2. After that, the testing genes were tiered hierarchically based on their ranking position in stage 1 and stage 2 (Table 3.5). The gene tiers were used to weigh the biological processes enriched in functional enrichment analysis.

Table 3.5: The gene tiering system in gene prioritization.

Gene tier	Description
Tier 1	Ranked top 100 in stage 1 and top 10 in at least 7 categories in stage 2.
Tier 2	Ranked top 100 in stage 1 and top 25 in at least 7 categories in stage 2.
Tier 3	Ranked top 100 in stage 1 and top 50 in at least 7 categories in stage 2.
Tier 4	Ranked top 100 in stage 1 and top 100 in at least 7 categories in stage 2.
Tier 5	Ranked top 100 in stage 1, regardless to their ranking position in stage 2.

3.9.2 Protein-Protein Interaction (PPI) Analysis

In PPI analysis, the testing genes ranked top 100 in stage 1 gene prioritization (Appendix F) and the 34 training genes (Appendix D) were analysed with STRING v11.0 to study the possible protein-protein interaction. The analysis was carried out using default settings. The interacting nodes were then subjected to functional enrichment analysis (Section 3.9.3).

[2] The genes reported in review paper were also considered if the paper is quoting the results from functional studies, association studies or case reports.

3.9.3 Functional Enrichment Analysis

Functional enrichment analysis was conducted primarily on DAVID v6.8, all interacting genes predicted from PPI analysis (Section 3.9.2) were used in the enrichment process. The Gene Ontology (GO) annotation GOTERM_BP_DIRECT was used to annotate the biological processes involved with each gene. EASE score $p \leq 0.05$ was set as the cut-off criterion. The GO terms fulfilling this criterion were mapped to its ancestor term with QuickGO (Binns et al., 2009) and a sunburst chart was generated using R package "plotly" to visualize the relationship between the GO terms. In order to study the potential association of each GO term to GGE, the GO terms were ranked based on the gene tiers using the classification scheme as shown in Table 3.6.

Table 3.6: The classification scheme used for categorizing the GO terms. The GO term categories were arranged in hierarchical order, category A denotes the GO terms having the highest association with GGE, and category E denotes the GO terms having the lowest association with GGE.

GO term category	Description
Category A	GO terms with at least one tier 1 gene.
Category B	GO terms with at least one tier 2 gene.
Category C	GO terms with at least one tier 3 gene.
Category D	GO terms with at least one tier 4 gene.
Category E	GO terms with at least one tier 5 gene.

3.10 Segregation Analysis

Segregation analysis was conducted on families F-1 and F-9 (Table 3.2), 13 family members, 8 from F-1 and 5 from F-9, were involved in this analysis. Genomic DNA of the family members was extracted using protocols as shown in Sections 3.4.1 and 3.4.2. Inheritance of the variants of interest, which are *SCN1A* c.5753C>T in F-1, *GABRA1* c.448G>A and *ERBB4* c.1972A>T in F-9, among family members was then detected using Sanger sequencing.

3.10.1 Polymerase Chain Reaction (PCR)

Before Sanger sequencing, PCR amplifying the variant of interest was first conducted on all the recruited family members including the proband using Q5 High-Fidelity DNA Polymerase (New England Biolabs, United States). PCR primers flanking the variant of interest as shown in Table 3.7 were designed with Primer-BLAST (Ye et al., 2012). The PCR reaction mix was prepared in accordance with Table 3.8, 100-200 ng of genomic DNA was used in each PCR reaction, for no-template control, autoclaved distilled water was added in place of genomic DNA. All PCRs were carried out using Veriti 96-Well Thermal Cycler (Applied Biosystems, United States) under the PCR conditions as shown in Table 3.9. Once the PCR was completed, gel electrophoresis was conducted to verify the PCR results (Section 3.10.2).

Table 3.7: PCR primers for amplifying the *SCN1A*, *GABRA1* and *ERBB4* variants identified in F-1 and F-9.

Family ID	Variant of interest	Primer sequence	PCR product size (bp)
F-1	*SCN1A* c.5753C>T	Forward: 5'-CAGTTTGGCATTGACCTCCT-3' Reverse: 5'-GTGACCGGATCCACTGTCTT-3'	556
F-9	*GABRA1* c.448G>A	Forward: 5'-GCAAAAATTATGCACTGTCTGCG-3' Reverse: 5'-TGCCTACAAGTGGAAGGGAA-3'	371
F-9	*ERBB4* c.1972A>T	Forward: 5'-TGGTAGAGCAAAACCTAATGCAC-3' Reverse: 5'-CTACCCCCTTGAAGCCTTGACT-3'	488

Table 3.8: Reaction mix for PCR. The genomic DNA was first diluted to 50 ng/µL with autoclaved distilled water, and 2 µL of the diluted DNA was then used for PCR. However, for genomic DNA extracted from buccal cell samples in which the starting DNA concentration is less than 50 ng/µL, 4 µL of the DNA was added into the reaction mix and the volume of autoclaved distilled water was reduced to 12.75 µL.

Reagent	Volume (µL)	Final concentration
5X Q5 Reaction Buffer	5	1X
10 mM dNTPs	0.5	200 µM
10 µM Forward Primer	1.25	0.5 µM
10 µM Reverse Primer	1.25	0.5 µM
Template DNA	2 (4)	100-200 ng
Q5 High-Fidelity DNA Polymerase	0.25	0.02 U/µL
Autoclaved distilled water	14.75 (12.75)	-
TOTAL	25	-

Table 3.9: PCR condition for amplifying the *SCN1A*, *GABRA1* and *ERBB4* variants using Q5 High-Fidelity DNA Polymerase.

Variant	Initial denaturation	35 cycles of			Final extension
		Denaturation	Annealing	Extension	
SCN1A c.5753C>T	98°C, 30 s	98°C, 5 s	66°C, 10 s	72°C, 30 s	72°C, 2 min
GABRA1 c.448G>A	98°C, 30 s	98°C, 5 s	62°C, 10 s	72°C, 30 s	72°C, 2 min
ERBB4 c.1972A>T	98°C, 30 s	98°C, 5 s	66°C, 10 s	72°C, 30 s	72°C, 2 min

3.10.2 Gel Electrophoresis

All gel electrophoreses were conducted with 1.5% agarose gel prepared by 100 mL of 1X TBE buffer and 1.5 grams of agarose using the materials and methods mentioned in Section 3.5.1. 10 µL of FloroSafe DNA Stain was added to stain the gel. 4 µL of PCR product was mixed with 1 µL of 6X DNA loading dye before loading into the well. 10 µL of TriDye 100 bp DNA Ladder (New England Biolabs, United States) was used as the DNA marker and the gel electrophoreses were carried out under a constant voltage of 120 V for 60 minutes.

3.10.3 Gel Excision and Purification

Once the PCR results were verified with gel electrophoresis, gel excision was conducted to excise the PCR products from the gel. During the gel excision step, another gel electrophoresis was conducted using the same procedure as mentioned in Section 3.10.2, with the exception that all of the remaining PCR product (approximately 21 µL) was mixed with 4 µL of 6X DNA loading dye before loading into the well. Once the gel electrophoresis was completed, the PCR product bands were excised with a microscope slide and the gel slides were transferred into 1.5 mL microcentrifuge tubes for gel purification.

Gel purification was conducted with QIAquick Gel Extraction Kit (Qiagen, Germany) using the spin protocol provided by the manufacturer (Appendix C). ACS grade isopropanol (Emsure ACS, Merck, Germany) was used to precipitate the DNA during the gel purification (step 4, Appendix C). 30 µL of autoclaved distilled water was used to elute the purified PCR products at the end of gel purification. All centrifugation steps were carried out under the centrifugal force of 14,000 x g.

3.10.4 Sanger Sequencing (Outsourced)

Sanger sequencing was outsourced to Apical Scientific Sdn Bhd (Malaysia). 10 µL of purified PCR products and 5 µL of PCR primers per reaction were aliquoted and sent to Apical Scientific Sdn Bhd for Sanger sequencing. Based on the information provided by Apical Scientific, the Sanger sequencing was conducted with 3730xl DNA Analyzer (Applied Biosystems, United States) using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, United States). The DNA chromatograms provided by Apical Scientific after completion of Sanger sequencing were then analysed to study the inheritance of variants in the families.

3.11 Protein Structure Analysis

Segregation analysis (Section 3.10) allowed the identification of variants which segregate in line with the disease phenotypes, it provided additional evidence on the pathogenicity of the variants. In order to further investigate the mutational impact of the variants on protein structure, three-dimensional structure analysis was conducted on the affected proteins. For the *GABRA1* c.448G>A and *ERBB4* c.1972A>T variants in F-9 (Table 3.7), the three-dimensional structure of gamma-aminobutyric acid type A (GABA-A) receptor (PDB ID: 6HUO) (Masiulis et al., 2019) and transmembrane region of erb-b2 receptor tyrosine kinase 4 (ERBB4) (PDB ID: 2LCX) (Bocharov et al., 2012) was retrieved from RCSB Protein Database Bank (http://www.rcsb.org) (Berman et al., 2000) and used directly in the protein structure analysis. Meanwhile, for the *SCN1A* c.5753C>T in F-1, homology modelling for the C-terminal domain of sodium voltage-gated channel alpha subunit 1 (SCN1A) was first conducted with SWISS-MODEL (Waterhouse et al., 2018) using the three-dimensional structure of sodium voltage-gated channel alpha subunit 2 (SCN2A) (PDB ID: 4JPZ) (Wang et al., 2014) retrieved from RCSB Protein Database as the template. The quality of the predicted SCN1A model was assessed with the QMEAN score from SWISS-MODEL and the Ramachandran plot from PDBsum (Laskowski et al., 2018). Mutant proteins were generated with UCSF Chimera (Pettersen et al., 2004) using the Rotamers tool followed by 500 steps of local steepest descent energy minimization under AMBER ff14SB force field. The hydrogen bonds between amino acid residues and distance between alpha carton atoms were predicted and measured with PyMOL (Schrodinger, 2015). All structural figures in Section 4.6 were prepared using PyMOL.

3.12 Overview of Research Methodology

In general, the methodology of this project can be divided into two main parts: functional characterization analysis and segregation analysis. Figure 3.1 shows the flow chart of the research activities conducted in this study. The functional characterization analysis was set up with a primary objective to identify the potential pathogenic variants (variants of interest) from the probands and a secondary objective to investigate the biological processes which are potentially associated with GGE among Malaysians. Meanwhile, the segregation analysis aimed to study the inheritance pattern of the variants of interest in the proband's family.

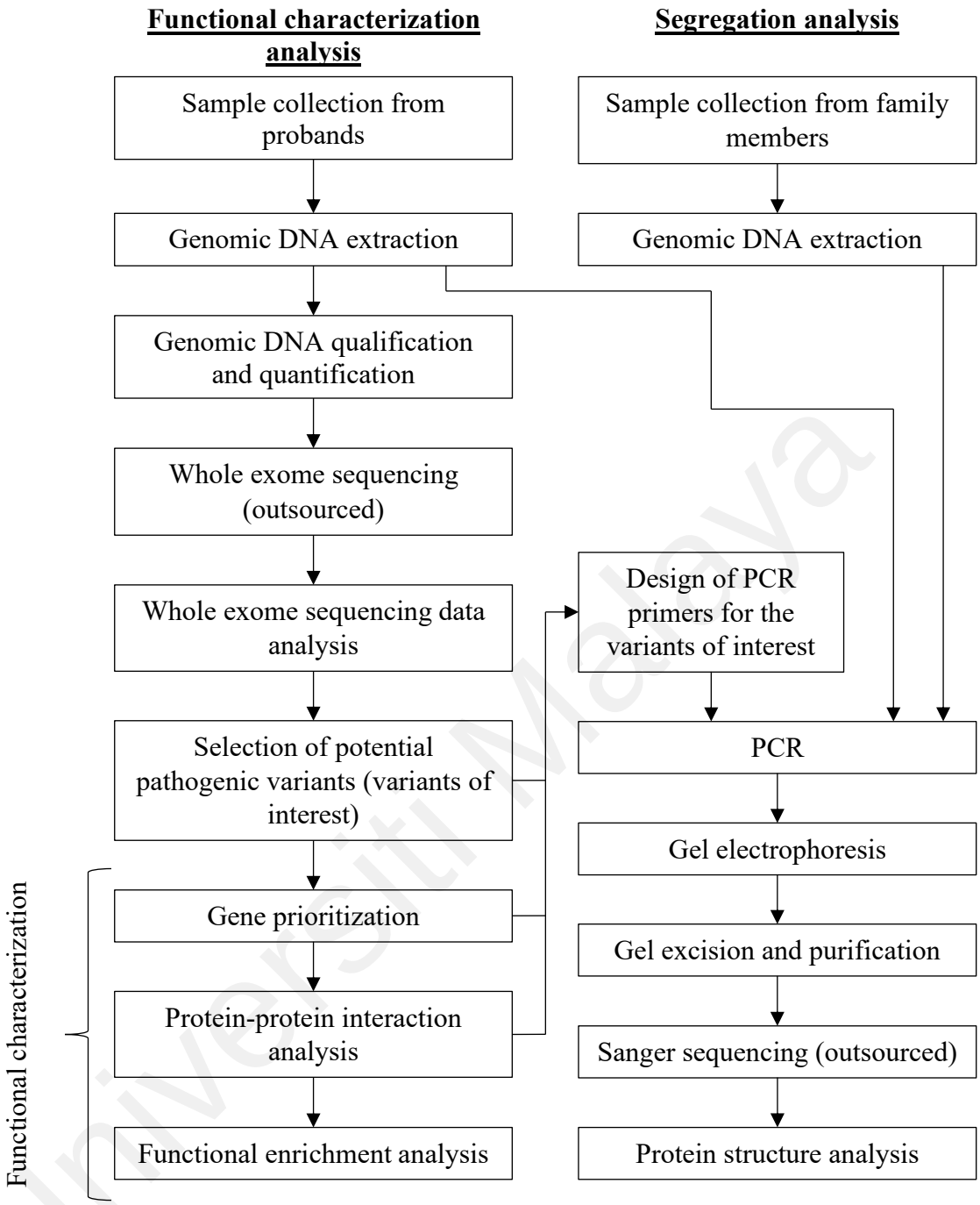


Figure 3.1: Flow chart of research activities conducted in this study.

CHAPTER 4: RESULTS

4.1 Clinical Information and Study Cohort

Table 4.1 shows the clinical information of the 18 individuals that underwent WES in this study, the 18 individuals were composed of 14 unrelated probands (GGE-1, GGE-4, GGE-7, GGE-8, GGE-9, GGE-10, GGE-11, GGE-12, GGE-13, GGE-14, GGE-15, GGE-16, GGE-17 and GGE-18), parents of GGE-1 (GGE-2 and GGE-3) and parents of GGE-4 (GGE-5 and GGE-6). Among the 14 probands, 9 of them were diagnosed with juvenile myoclonic epilepsy (JME), 3 were diagnosed with generalized epilepsy with febrile seizure plus (GEFS+) and 2 probands were having GGE with unclassifiable subtype. Most probands were having seizure onset during adolescence between 10 to 18 years old, except GGE-16 who was having his onset at 8 years old; meanwhile, GGE-17 cannot recall the exact age of onset but she was having her seizure attacks since childhood. Interestingly, GGE-9 exhibited the syndromes of both JME and temporal lobe epilepsy (TLE). Meanwhile, the other probands manifested the syndromes of GGE only. Due to mixed epilepsy type in GGE-9, GGE-9 was removed from functional characterization analysis (Sections 3.9 and 5.1).

For the parents involved in trio WES, GGE-3 (mother of GGE-1) was diagnosed with JME and exhibited similar clinical syndromes as GGE-1. On other hand, GGE-6 (mother of GGE-4) was having febrile seizures but became seizure free before entering primary school. Both GGE-3 and GGE-6 were not under any antiepileptic drug treatment. Meanwhile, there was no history of epilepsy in the paternal family, including the father of GGE-1 and GGE-4. The parents of GGE-1 and GGE-4 were not included in the functional characterization analysis mentioned in Section 3.9.

Table 4.1: Clinical information of the 18 individuals that underwent whole exome sequencing in this study.

Sample ID	Gender	Epilepsy type	Age of onset (years old)	Syndrome	EEG	MRI	Current treatment
GGE-1	Female	JME	12	GTC, myoclonic seizures, febrile seizures	Normal	Normal	VPA, LTG
GGE-4	Female	JME	11	GTC, myoclonic seizures, febrile seizures	PSW, PPR	Not done	VPA
GGE-7	Male	GEFS+	15	GTC, myoclonic seizures, febrile seizures	GSW	Not done	VPA
GGE-8	Female	JME	10	GTC, myoclonic seizures	GSW, PPR	Not done	VPA, LTG
GGE-9	Female	JME and TLE	15	Headache, vertigo, vomiting, facial automatism, head deviation, gelastic episodes, tinnitus, catamenial, aggressive staring, simple partial seizures	Generalized PSW with T4 preponderance	Right hippocampal atrophy and left malrotation	VPA, LTG, LVT
GGE-10	Female	GGE	17	GTC	Normal	Not done	VPA, LTG
GGE-11	Female	JME	12	GTC, myoclonic seizures	Not done	Not done	LTG
GGE-12	Female	GEFS+	13	GTC	PSW, PPR	Not done	VPA
GGE-13	Female	JME	15	Myoclonic seizures	GSW, left temporal sharp	Not done	VPA, LTG
GGE-14	Male	JME	13	GTC, myoclonic seizures	Normal	Not done	VPA
GGE-15	Female	GEFS+	15	GTC, myoclonic seizures, febrile seizures	GSW	Not done	LTG
GGE-16	Male	JME	8	GTC, myoclonic seizures	Polyspikes	Not done	VPA
GGE-17	Female	GGE	Childhood	GTC, febrile seizures	Generalized discharge	Not done	LTG
GGE-18	Female	JME	18	GTC, myoclonic seizures	PSW	Not done	VPA, LTG
GGE-2 (Father of GGE-1)	Male	N/A	N/A	Unaffected (non-epileptic)	N/A	N/A	N/A

Table 4.1, continued.

Sample ID	Gender	Epilepsy type	Age of onset (years old)	Syndrome	EEG	MRI	Current treatment
GGE-3 (Mother of GGE-1)	Female	JME	12	GTC, myoclonic seizures, febrile seizures	Not known	Not known	None
GGE-5 (Father of GGE-4)	Male	N/A	N/A	Unaffected (non-epileptic)	N/A	N/A	N/A
GGE-6 (Mother of GGE-4)	Female	Febrile seizures	N/A	Febrile seizures	Not known	Not known	None

Abbreviations: EEG, Electroencephalography; GEFS+, Generalized epilepsy with febrile seizure plus; GGE, Genetic generalized epilepsy; GSW, Generalized spike and wave; GTC, Generalized tonic-clonic seizures; JME, Juvenile myoclonic epilepsy; LTG, Lamotrigine; LVT, Levetiracetam; MRI, Magnetic resonance imaging; N/A, Not applicable; PPR, Photoparoxysmal response; PSW, Polyspike-waves; TLE, Temporal lobe epilepsy; VPA, Sodium valproate.

4.2 Genomic DNA Qualification and Quantification

4.2.1 Genomic DNA Qualification Using Gel Electrophoresis

Gel electrophoresis was conducted for the 18 DNA samples involved in whole exome sequencing (WES). Figure 4.1 shows the results of gel electrophoresis for the DNA samples. As indicated in the figure, there was slight degradation on the DNA, the slight degradation was probably due to long storage period of the blood samples before genomic DNA extraction. Nevertheless, the overall DNA quality was still acceptable as an intact band was still observable in all samples.

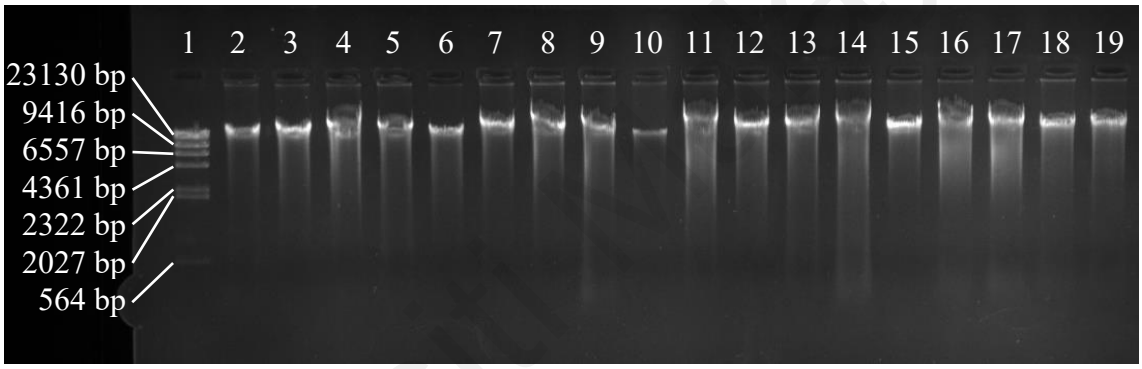


Figure 4.1: Gel electrophoresis of DNA samples involved in whole exome sequencing. Lane 1: Lambda DNA/*Hind*III Marker; Lane 2: GGE-1; Lane 3: GGE-4; Lane 4: GGE-7; Lane 5: GGE-8; Lane 6: GGE-10; Lane 7: GGE-12; Lane 8: GGE-11; Lane 9: GGE-13; Lane 10: GGE-14; Lane 11: GGE-15; Lane 12: GGE-16; Lane 13: GGE-17; Lane 14: GGE-18; Lane 15: GGE-9; Lane 16: GGE-2; Lane 17: GGE-3; Lane 18: GGE-5; Lane 19: GGE-6.

4.2.2 Genomic DNA Quantification Using Spectrophotometer

The spectrophotometric readings of the DNA samples involved in WES are shown in Table 4.2. All samples were having A260/A280 absorbance ratio of between 1.80 and 2.00, indicating the DNA purity was high with minimal or no protein contamination. In terms of DNA yield, the average DNA yield was 61.19 µg, most samples achieved the final yield of at least 30 µg, except GGE-1 and GGE-14 in which the DNA yield was 28.36 µg and 7.34 µg respectively. Nevertheless, the amount of DNA in each sample was still sufficient for WES as only 1 µg of DNA was required for WES.

Table 4.2: Spectrophotometric measurements and overall DNA yield of the DNA samples involved in whole exome sequencing.

Sample ID	Concentration (ng/µL)	A260/A280	Volume (µL)	DNA yield (µg)
GGE-1	141.8	1.92	200	28.36
GGE-4	190.8	1.89	200	38.16
GGE-7	299.7	1.88	200	59.94
GGE-8	297.1	1.90	200	59.42
GGE-9	173.1	1.90	200	34.62
GGE-10	170.3	1.90	200	34.06
GGE-11	289.9	1.87	200	57.98
GGE-12	292.8	1.90	200	58.56
GGE-13	402.9	1.90	200	80.58
GGE-14	36.7	1.80	200	7.34
GGE-15	615.1	1.84	200	123.02
GGE-16	270.5	1.88	200	54.10
GGE-17	446.4	1.90	200	89.28
GGE-18	528.7	1.89	200	105.74
GGE-2	429.2	1.90	200	85.84
GGE-3	465.0	1.90	200	93.00
GGE-5	253.3	1.89	200	50.66
GGE-6	203.4	1.91	200	40.68

4.3 Whole Exome Sequencing (WES) Data Analysis

From the initial analysis of WES data, about 20,000 variants were identified in the exonic regions of each individual, further analysis showed that about 9,500 of the variants were missense variants. Table 4.3 shows the exact number of variants identified in each individual. Despite thousands of missense variants were present in all individuals, subsequent pathogenicity analysis (Section 4.4) predicted that most missense variants were polymorphism. As mentioned in Section 3.8, the polymorphic variants were assumed to be benign and were excluded from analysis in this project.

Table 4.3: Number of variants identified in each individual by whole exome sequencing.

Sample ID	Total number of variants	Number of variants in exonic regions	Number of missense variants
GGE-1	20,085	19,923	9,457
GGE-4	20,149	19,972	9,354
GGE-7	20,138	19,961	9,396
GGE-8	19,723	19,547	9,201
GGE-9	20,010	19,820	9,254
GGE-10	20,537	20,372	9,490
GGE-11	20,028	19,852	9,341
GGE-12	20,637	20,454	9,588
GGE-13	21,165	20,968	9,849
GGE-14	20,221	20,045	9,303
GGE-15	20,120	19,946	9,268
GGE-16	20,598	20,424	9,471
GGE-17	20,079	19,930	9,377
GGE-18	20,614	20,406	9,565
GGE-2	20,087	19,916	9,458
GGE-3	20,513	20,342	9,557
GGE-5	20,053	19,864	9,273
GGE-6	19,953	19,764	9,296

4.4 *In Silico* Pathogenicity Prediction and Selection of Potential Pathogenic Variants

As shown in Table 4.3, thousands of genetic variants can be identified by whole exome sequencing. Analysis of the huge number of genetic variants can be overwhelming as it is not always possible to investigate the pathogenicity of all the variants due to financial and time constraints. Despite this, genetic variants can be shortlisted so that analysis can be focused on those with greater potential to cause a disease. After applying the *in silico* pathogenicity prediction and the filtering strategy mentioned in Section 3.8, a total of 826 genes (Appendix E) from GGE-1, GGE-4, GGE-7, GGE-8, GGE-10, GGE-11, GGE-12, GGE-13, GGE-14, GGE-15, GGE-16, GGE-17 and GGE-18 were found to contain at least one potential pathogenic variant, these genes were used as testing genes in the gene prioritization analysis described in Section 3.9.1. Meanwhile, the 62 potential-pathogenic-variant-containing genes in GGE-9 (Appendix I) were not involved in the

functional characterization analysis. Instead, they were treated as candidate genes for further investigation in segregation analysis (Section 4.6.2). Table 4.4 shows the number of genes containing at least one potential pathogenic variant in each proband, *in silico* pathogenicity prediction was not conducted on GGE-2, GGE-3, GGE-5 and GGE-6 as the WES data from these individuals were only used for aiding the segregation analysis in this project.

Table 4.4: Number of genes containing at least one potential pathogenic variant in each proband.

Sample ID	Number of genes containing at least one potential pathogenic variant
GGE-1	72
GGE-4	64
GGE-7	46
GGE-8	67
GGE-9	62
GGE-10	83
GGE-11	74
GGE-12	59
GGE-13	79
GGE-14	74
GGE-15	55
GGE-16	75
GGE-17	71
GGE-18	71

4.5 Functional Characterization Analysis

4.5.1 Gene Prioritization

Despite *in silico* pathogenicity prediction and minor allele frequency filtering have significantly reduced the number of candidate genes, this method merely helped in predicting the effect of the mutation, it did not intimate the association of the candidate genes with epilepsy. In order to investigate the potential relationship of each candidate gene with epilepsy, gene prioritization was conducted for all the candidate genes listed in Appendix E.

The ToppGene Suite ranked the testing genes (Appendix E) based on their functional similarity with the training genes (Appendix D), with rank 1 indicating the testing gene with the closest similarity with the training genes. Appendix F shows the top 100 prioritized testing genes, with their respective ranking position indicated by Rank A, these genes were used in the protein-protein interaction (PPI) analysis described in Section 3.9.2. Meanwhile, the genes ranked 101 and below in Rank A (not shown) were excluded from PPI analysis.

In order to further categorize the top 100 testing genes, the genes were tiered hierarchically based on their ranking position in each functional category (Rank B-P in Appendix F), the fulfilling criteria of each gene tier are described in Table 3.5. From the analysis, *CACNA1G* and *SCN1A* were having high ranking in many categories including the ion channels, catabolic enzyme, and signal transduction, making them to be classified as tier 1. Meanwhile, 9 genes were categorized as tier 2 and 15 genes were classified as tier 3. Table 4.5 shows the list of genes under each tier, with the presumption that genes under tier 1 were having the highest risk of being associated with juvenile myoclonic epilepsy (JME) and tier 5 indicating the genes with the lowest risk.

Table 4.5: The list of genes under each tier.

Tier	Gene
Tier 1	*CACNA1G, SCN1A*
Tier 2	*ADRA2A, ATP2B4, DMD, GABRA6, LEPR, NTRK1, NTRK2, SLC9A1, TSC1*
Tier 3	*ABCC2, APOB, CACNA1S, CACNB3, CACNG4, CBL, CCND1, ERBB4, JAG1, KCNH1, LRP1, NEDD4, SPTBN2, TRPC1, TRPM2*
Tier 4	*ATP6V0A2, ATP8A2, CDH23, CHRNB1, CHRNE, CSF1R, CYP2D6, DST, FZD4, GCH1, GRIP1, HCRT, HK1, KCNV2, KDR, LRP2, LRP5, MYH6, NOD2, NOTCH3, NXF1, P2RX5, PAM, PLCG2, PRDM16, PTPRD, PTPRS, S100A1, SNTA1, SYNE1, TCOF1, TLR3, TNC, TRPM1, UTRN*
Tier 5	*ABCA13, ABCA5, ABCA8, ABCB4, ABCG8, ACAN, ANO2, ATP4B, ATP6V0A4, CA7, CCKAR, CD4, CDH17, CLN8, CNGB1, CXCR1, DNAJC5, ENTPD2, GJB3, GNA14, HCK, INSRR, IRF4, KCNA10, MUC4, MYH7, RAI1, SLC12A8, SLC26A3, SLC9A3, SPTB, SPTBN4, TAS1R2, TG, TRPM5, TRPV3, TTN, UNC13C, WNK3*

4.5.2 Protein-Protein Interaction (PPI) Analysis

The results from gene prioritization have estimated the potential association of each testing gene with JME. However, it is crucial to study how these genes are interacting with each other during epilepsy pathogenesis. Therefore, PPI analysis was conducted to investigate the possible interaction between the top 100 prioritized testing genes (Appendix F) and the training genes (Appendix D) from the gene prioritization analysis (Sections 3.9.1 and 4.5.1).

The predicted protein interaction network is illustrated in Figure 4.2, the list of interacting nodes in the figure and their corresponding interaction score is shown in Appendix G. Most testing genes were predicted to interact with others, except *ABCA5*, *ABCA8*, *ATP8A2*, *DST*, *INSRR*, *KCNA10*, *KCNV2*, *MUC4*, *NXF1*, *P2RX5*, *PAM*, *RAI1*, *S100A1*, *SLC12A8*, *TSC1* and *UNC13C*. The non-interacting genes consisted mostly of tier 4 and tier 5 genes, except *TSC1* which was classified as tier 2, suggesting that *TSC1* might be associated with JME through a different protein interaction network. For the interacting genes, the tier 1 *SCN1A* was having the highest number of interactants in the network, it was predicted to interact with many training genes including *SCN1B*, *GABRA1*, *EFHC1*, *CACNB4, CLCN2*, *KCNQ3*, *CHRNA4*, *SLC12A5* and *ABCB1*. This result indicated that *SCN1A* might serve as a key gene in JME and should be given priority for further discovery. Meanwhile, *CACNA1G* had only 9 interactants, which were mainly comprised of ion channel genes. For the tier 2 genes, *NTRK2* possessed the highest number of interactants, followed by *GABRA6* and *DMD*. While both *NTRK2* and *GABRA6* were interacting with GABA receptor and ion transporter genes, additional interaction was observed between *GABRA6* with ion channel genes *SCN1A*, *SCN1B* and *CLCN2*. As a result, *GABRA6* might possess a higher risk of being associated to JME compared to *NTRK2*. In contrast, *DMD* was not predicted to interact directly with any training gene except *KCNMA1* and *KCNJ10*. However, its interactants *CACNA1S* and

SNTA1 were found to interact with sodium and calcium ion channel genes, implying that *DMD* might affect the sodium and calcium ion channels through *CACNA1S* and *SNTA1*. Interestingly, the tier 3 gene *CBL* was positioned as second highest in the context of the number of interactants. Nevertheless, its interactants were mainly composed of testing genes. Table 4.6 shows the number of interactants of the testing genes.

Table 4.6: Number of interacting nodes predicted by STRING of the top 100 prioritized testing genes. The testing genes without any interactants are not shown in this table.

Gene	No. of nodes	Gene	No. of nodes	Gene	No. of nodes	Gene	No. of nodes
SCN1A	19	*HCK*	6	*CDH17*	4	*ABCB4*	2
CBL	12	*HCRT*	6	*CHRNB1*	4	*ATP2B4*	2
NTRK2	11	*NOTCH3*	6	*CHRNE*	4	*ATP4B*	2
APOB	10	*SLC9A1*	6	*IRF4*	4	*CDH23*	2
DMD	10	*SLC9A3*	6	*LRP5*	4	*GRIP1*	2
GABRA6	10	*SNTA1*	6	*MYH6*	4	*PRDM16*	2
CACNA1G	9	*SPTBN2*	6	*PTPRD*	4	*SYNE1*	2
CCND1	9	*TTN*	6	*PTPRS*	4	*TNC*	2
ERBB4	9	*ABCC2*	5	*TRPC1*	4	*TRPM1*	2
KDR	9	*ABCG8*	5	*CA7*	3	*TRPV3*	2
UTRN	9	*CACNG4*	5	*CYP2D6*	3	*WNK3*	2
ADRA2A	8	*FZD4*	5	*DNAJC5*	3	*ACAN*	1
ANO2	8	*JAG1*	5	*GJB3*	3	*ATP6V0A2*	1
CACNA1S	8	*NEDD4*	5	*KCNH1*	3	*ATP6V0A4*	1
CD4	7	*NTRK1*	5	*LEPR*	3	*CLN8*	1
LRP2	7	*SLC26A3*	5	*LRP1*	3	*CNGB1*	1
PLCG2	7	*SPTB*	5	*MYH7*	3	*ENTPD2*	1
CACNB3	6	*SPTBN4*	5	*NOD2*	3	*GCH1*	1
CSF1R	6	*TAS1R2*	5	*TG*	3	*HK1*	1
CXCR1	6	*TRPM5*	5	*TLR3*	3	*TCOF1*	1
GNA14	6	*CCKAR*	4	*ABCA13*	2	*TRPM2*	1

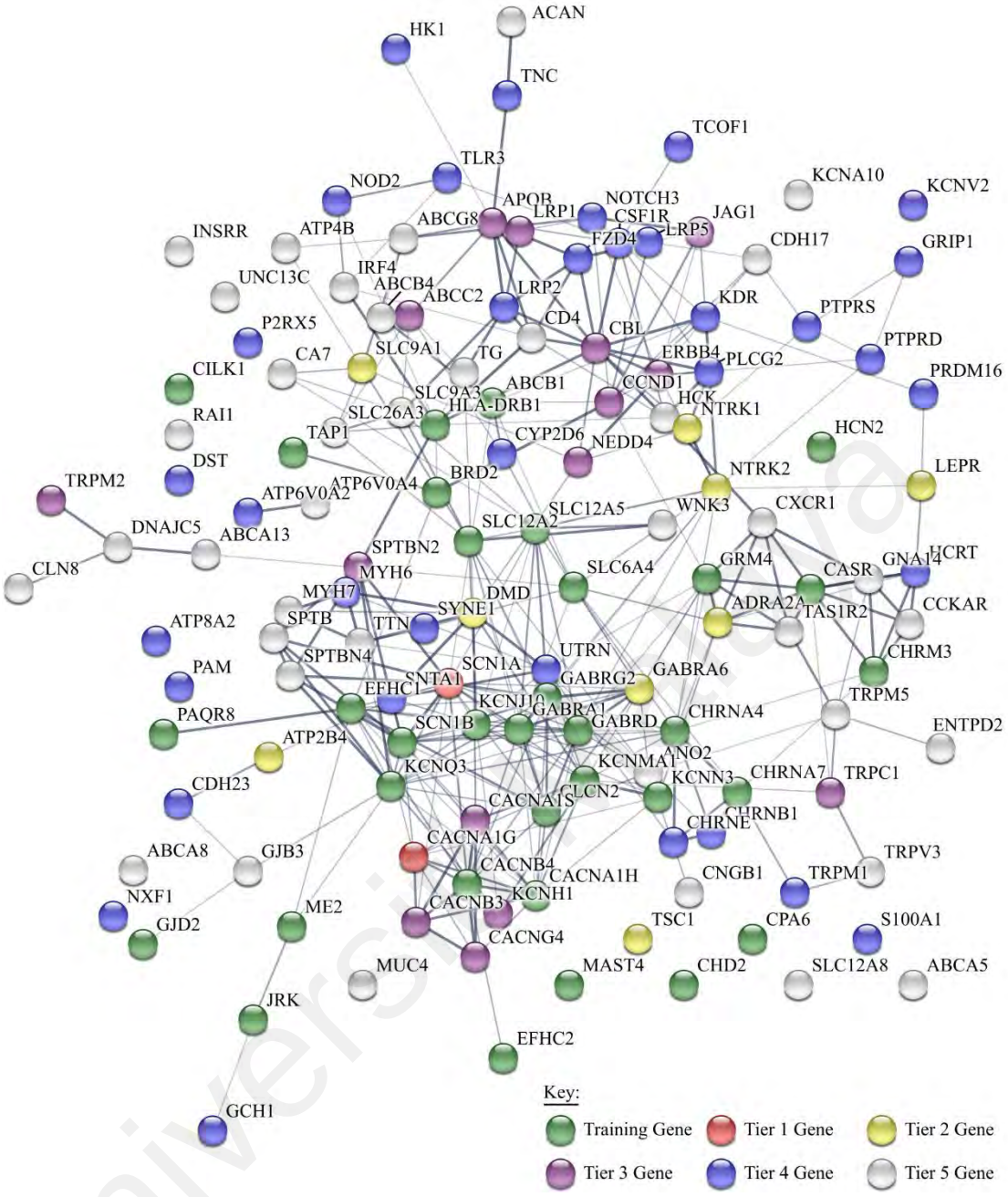


Figure 4.2: The predicted protein interaction network generated by STRING. The thickness of the line between the interacting nodes indicates the interaction score of the prediction, thicker line corresponds to higher interaction score.

4.5.3 Functional Enrichment Analysis

The aim of functional enrichment analysis is to study the possible functions of each gene within an interaction network, this step is very beneficial as it can help in the understanding of the molecular mechanism behind a disease. In this project, functional enrichment focused on the interacting genes predicted from the PPI analysis described in Section 4.5.2. A total of 86 biological processes (expressed as Gene Ontology (GO) terms) were enriched by DAVID v6.8 (Table 4.7), the relationship between the GO terms is illustrated in Figure 4.3. These GO terms were ranked hierarchically from category A to E, with category A refers to the GO terms with the highest association risk with JME and category E refers to the GO terms with the lowest association risk with JME (Table 4.7).

Table 4.7: List of Gene Ontology (GO) terms enriched by DAVID v6.8 and their respective EASE score and category.

Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0006810	transport	19	*SLC12A2, GABRA1, ABCB1, SLC12A5, ABCC2, GABRA6, ABCB4, CLCN2, GABRG2, CACNA1H, TRPM1, CACNB3, CACNB4, TCOF1, CHRNE, GABRD, CDH17, CACNG4, CNGB1*	1.13E-11	Category B
GO:0042391	regulation of membrane potential	11	*CHRNB1, CHRNA4, CHRNA7, CHRNE, NEDD4, KCNMA1, SLC26A3, CACNA1H, CACNA1G, CNGB1, KCNH1*	5.57E-11	Category A
GO:0070588	calcium ion transmembrane transport	11	*TRPM1, TRPM2, CACNB4, TRPC1, ATP2B4, TRPV3, CACNA1S, TRPM5, CACNA1H, CACNG4, CACNA1G*	5.63E-09	Category A
GO:0007268	chemical synaptic transmission	14	*PTPRS, SLC12A5, CHRNA4, SLC6A4, CACNA1G, GJD2, PTPRD, CACNB3, GRM4, CACNB4, KCNQ3, HCRT, GABRD, SCN1B*	5.84E-09	Category A
GO:0034765	regulation of ion transmembrane transport	10	*KCNJ10, NEDD4, KCNMA1, KCNQ3, CACNA1S, TRPM5, CACNA1H, CACNA1G, SCN1A, KCNH1*	4.62E-08	Category A
GO:0006811	ion transport	10	*SLC12A2, CHRNB1, SLC9A3, GABRA1, SLC12A5, CHRNA4, CHRNA7, CHRNE, SLC26A3, SLC9A1*	1.49E-07	Category B
GO:1902476	chloride transmembrane transport	9	*SLC12A2, GABRA1, SLC12A5, GABRA6, GABRD, SLC26A3, CLCN2, GABRG2, ANO2*	1.68E-07	Category B
GO:0006936	muscle contraction	9	*CHRNB1, CHRNE, CACNA1S, UTRN, MYH6, CACNA1H, SNTA1, MYH7, TTN*	4.99E-07	Category C
GO:0007528	neuromuscular junction development	6	*NTRK2, CACNB3, CACNB4, NEDD4, TNC, UTRN*	9.59E-07	Category B
GO:0071805	potassium ion transmembrane transport	9	*SLC9A3, SLC12A5, KCNMA1, KCNQ3, KCNN3, TRPM5, SLC9A1, CNGB1, KCNH1*	1.28E-06	Category B
GO:0034220	ion transmembrane transport	10	*ATP4B, GABRA1, GABRA6, CHRNA7, ATP6V0A2, ATP2B4, ATP6V0A4, CLCN2, GABRG2, ANO2*	9.83E-06	Category B

Table 4.7, continued.

Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0006816	calcium ion transport	7	*TRPM2, CACNB3, CHRNA4, CHRNA7, TRPC1, CDH23, CACNA1S*	1.04E-05	Category C
GO:0061337	cardiac conduction	6	*SPTBN4, CACNB3, CACNB4, CACNA1S, SCN1B, CACNG4*	1.09E-05	Category C
GO:0051899	membrane depolarization	5	*CACNB3, CACNB4, CHRNA4, SCN1B, CACNG4*	1.98E-05	Category C
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	7	*NTRK1, NTRK2, CSF1R, HCK, CD4, ERBB4, KDR*	3.96E-05	Category B
GO:0007271	synaptic transmission, cholinergic	5	*CHRNB1, CHRM3, CHRNA4, CHRNA7, CHRNE*	9.75E-05	Category D
GO:1902083	negative regulation of peptidyl-cysteine S-nitrosylation	3	*ATP2B4, DMD, SNTA1*	1.27E-04	Category B
GO:0006813	potassium ion transport	6	*ATP4B, SLC12A2, KCNJ10, KCNMA1, KCNQ3, KCNH1*	2.01E-04	Category C
GO:0055085	transmembrane transport	9	*GJD2, ABCG8, ABCB1, ABCC2, GJB3, ABCB4, TAP1, ABCA13, SLC6A4*	2.03E-04	Category C
GO:0060048	cardiac muscle contraction	5	*DMD, MYH6, SCN1B, MYH7, TTN*	2.11E-04	Category B
GO:0007214	gamma-aminobutyric acid signaling pathway	4	*SLC12A2, GABRA1, GABRA6, GABRG2*	3.85E-04	Category B
GO:0018108	peptidyl-tyrosine phosphorylation	7	*NTRK1, NTRK2, CSF1R, HCK, ERBB4, KDR, TTN*	5.10E-04	Category B
GO:0051592	response to calcium ion	5	*CCND1, TRPC1, NEDD4, KCNMA1, TTN*	5.63E-04	Category C
GO:0002485	antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-dependent	3	*ABCB1, ABCB4, TAP1*	6.27E-04	Category E

Table 4.7, continued.

Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0002489	antigen processing and presentation of endogenous peptide antigen via MHC class Ib via ER pathway, TAP-dependent	3	*ABCB1, ABCB4, TAP1*	6.27E-04	Category E
GO:0002591	positive regulation of antigen processing and presentation of peptide antigen via MHC class I	3	*ABCB1, ABCB4, TAP1*	6.27E-04	Category E
GO:0086010	membrane depolarization during action potential	4	*CACNA1S, CACNA1H, CACNA1G, SCN1A*	7.95E-04	Category A
GO:0014068	positive regulation of phosphatidylinositol 3-kinase signaling	5	*NTRK2, ERBB4, NEDD4, KDR, CBL*	8.68E-04	Category B
GO:1901385	regulation of voltage-gated calcium channel activity	3	*CACNB3, CACNB4, DMD*	8.74E-04	Category B
GO:0002481	antigen processing and presentation of exogenous protein antigen via MHC class Ib, TAP-dependent	3	*ABCB1, ABCB4, TAP1*	8.74E-04	Category E
GO:0046777	protein autophosphorylation	7	*NTRK1, NTRK2, CSF1R, HCK, ERBB4, WNK3, KDR*	9.41E-04	Category B
GO:0002027	regulation of heart rate	4	*DMD, MYH6, SNTA1, MYH7*	0.001293339	Category B
GO:0001934	positive regulation of protein phosphorylation	6	*NTRK1, NTRK2, CSF1R, CCND1, ERBB4, KDR*	0.001485271	Category B
GO:0035095	behavioral response to nicotine	3	*CHRNB1, CHRNA4, CHRNA7*	0.001485585	Category D
GO:0030049	muscle filament sliding	4	*DMD, MYH6, MYH7, TTN*	0.001952825	Category B
GO:0006821	chloride transport	4	*GABRA1, GABRA6, SLC26A3, GABRG2*	0.002265497	Category B
GO:0006941	striated muscle contraction	3	*MYH6, MYH7, TTN*	0.00316414	Category D

Table 4.7, continued.

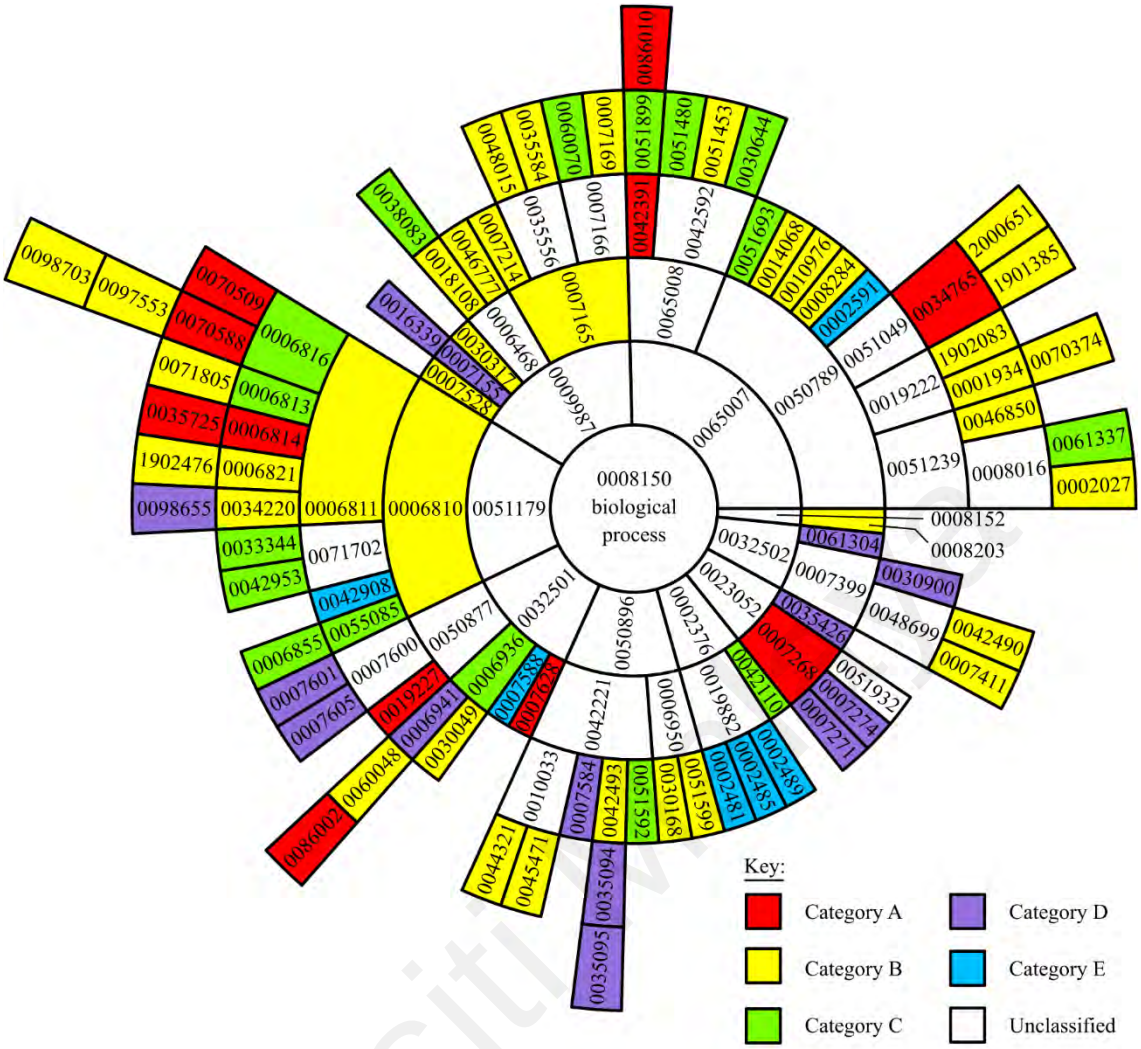
Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0051693	actin filament capping	3	*SPTBN4, SPTB, SPTBN2*	0.00316414	Category C
GO:0098655	cation transmembrane transport	4	*CHRNB1, CHRNA4, CHRNA7, CHRNE*	0.00381852	Category D
GO:0042493	response to drug	8	*NTRK1, ABCG8, ABCB1, SLC12A5, CCND1, ABCB4, SLC6A4, SLC9A1*	0.003867952	Category B
GO:0007411	axon guidance	6	*NTRK1, SPTBN4, CSF1R, SPTB, SCN1B, SPTBN2*	0.003938253	Category B
GO:0042953	lipoprotein transport	3	*LRP1, LRP2, APOB*	0.004223174	Category C
GO:0086002	cardiac muscle cell action potential involved in contraction	3	*SCN1B, CACNA1G, SCN1A*	0.004805921	Category A
GO:0070509	calcium ion import	3	*CASR, CACNA1H, CACNA1G*	0.005423516	Category A
GO:0006855	drug transmembrane transport	3	*ABCB1, ABCC2, ABCB4*	0.006761363	Category C
GO:0035584	calcium-mediated signaling using intracellular calcium source	3	*NTRK2, TRPM2, KDR*	0.007480682	Category B
GO:0008203	cholesterol metabolic process	4	*LEPR, LRP5, APOB, CLN8*	0.010068245	Category B
GO:0007605	sensory perception of sound	5	*SPTBN4, FZD4, GJB3, CDH23, ATP6V0A4*	0.011383171	Category D
GO:0008284	positive regulation of cell population proliferation	9	*NTRK2, CSF1R, HCK, ERBB4, CHRNA7, LRP5, TNC, KDR, ADRA2A*	0.011548433	Category B
GO:0033344	cholesterol efflux	3	*ABCG8, APOB, ABCA13*	0.011562996	Category C
GO:0035725	sodium ion transmembrane transport	4	*TRPM2, TRPM5, SCN1B, SCN1A*	0.012204938	Category A
GO:0007274	neuromuscular synaptic transmission	3	*CHRNB1, CHRNA4, CHRNE*	0.012473479	Category D
GO:0007584	response to nutrient	4	*ABCG8, CCKAR, NOD2, SLC6A4*	0.012661057	Category D
GO:0097553	calcium ion transmembrane import into cytosol	2	*TRPM2, ATP2B4*	0.013058952	Category B

Table 4.7, continued.

Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0016339	calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	3	*CDH23, ME2, CDH17*	0.014384976	Category D
GO:0006814	sodium ion transport	4	*ATP4B, SLC12A2, SCN1B, SCN1A*	0.01612612	Category A
GO:0060070	canonical Wnt signaling pathway	4	*CCND1, FZD4, LRP5, MYH6*	0.017204322	Category C
GO:0007628	adult walking behavior	3	*SPTBN4, KCNJ10, SCN1A*	0.01747218	Category A
GO:0051480	regulation of cytosolic calcium ion concentration	3	*TRPC1, CDH23, CNGB1*	0.018558235	Category C
GO:0010976	positive regulation of neuron projection development	4	*NTRK1, NTRK2, DMD, SCN1B*	0.02067599	Category B
GO:0051453	regulation of intracellular pH	3	*SLC9A3, SLC26A3, SLC9A1*	0.023176531	Category B
GO:0035094	response to nicotine	3	*CHRNA4, CHRNA7, CHRNE*	0.024397635	Category D
GO:0007588	excretion	3	*ABCG8, ATP6V0A4, SLC26A3*	0.024397635	Category E
GO:0061304	retinal blood vessel morphogenesis	2	*FZD4, LRP5*	0.025948899	Category D
GO:0042490	mechanoreceptor differentiation	2	*NTRK1, NTRK2*	0.025948899	Category B
GO:0051599	response to hydrostatic pressure	2	*NTRK1, ATP2B4*	0.025948899	Category B
GO:0035426	extracellular matrix-cell signaling	2	*FZD4, LRP5*	0.025948899	Category D
GO:0070374	positive regulation of ERK1 and ERK2 cascade	5	*NTRK1, CSF1R, ERBB4, KDR, NOD2*	0.028051663	Category B
GO:0038083	peptidyl-tyrosine autophosphorylation	3	*HCK, ERBB4, KDR*	0.028214393	Category C
GO:0007155	cell adhesion	8	*ATP4B, ACAN, HCK, CD4, PTPRS, TNC, CDH17, SCN1B*	0.031283258	Category D
GO:0045471	response to ethanol	4	*NTRK1, CCND1, TNC, CBL*	0.031677531	Category B
GO:0098703	calcium ion import across plasma membrane	2	*TRPM2, ATP2B4*	0.032331173	Category B

Table 4.7, continued.

Accession no.	Name	Gene count	Gene	EASE score	Category
GO:0048015	phosphatidylinositol-mediated signaling	4	*NTRK1, CSF1R, ERBB4, KCNH1*	0.032449019	Category B
GO:0030900	forebrain development	3	*NOTCH3, CCKAR, LRP2*	0.036504608	Category D
GO:0042110	T cell activation	3	*IRF4, CHRNA7, NEDD4*	0.037967175	Category C
GO:2000651	positive regulation of sodium ion transmembrane transporter activity	2	*WNK3, DMD*	0.038672006	Category B
GO:0030644	cellular chloride ion homeostasis	2	*SLC12A5, ABCC2*	0.038672006	Category C
GO:0030168	platelet activation	4	*GNA14, ENTPD2, PLCG2, ADRA2A*	0.039830743	Category B
GO:0007165	signal transduction	14	*CHRM3, CHRNB1, CSF1R, GABRA6, CHRNA4, CHRNA7, ADRA2A, GNA14, CD4, TG, ERBB4, CHRNE, GABRD, TLR3*	0.040121469	Category B
GO:0007601	visual perception	5	*TRPM1, GJD2, KCNJ10, CDH23, CNGB1*	0.043186282	Category D
GO:0019227	neuronal action potential propagation	2	*SCN1B, SCN1A*	0.044971666	Category A
GO:0044321	response to leptin	2	*CCND1, LEPR*	0.044971666	Category B
GO:0046850	regulation of bone remodeling	2	*LEPR, LRP5*	0.044971666	Category B
GO:0042908	xenobiotic transport	2	*ABCB1, ABCB4*	0.044971666	Category E
GO:0051932	synaptic transmission, GABAergic	2	*GABRA1, GABRG2*	0.044971666	No test gene
GO:0030317	flagellated sperm motility	3	*CHRNA7, ATP2B4, APOB*	0.048806337	Category B



Key:

Category A
Category B
Category C
Category D
Category E
Unclassified

Figure 4.3: The sunburst chart illustrating the relationship between the Gene Ontology (GO) terms. The GO terms in the outer rings serve as the child terms of that in the inner rings, the seven-digit numbers in the chart represent the accession number of the GO terms, the name of the GO terms is shown in Appendix H.

As shown in Figure 4.3, all the 86 enriched GO terms can be grouped into nine sectors, which are 'biological regulation' (GO:0065007), 'localization' (GO:0051179), 'cellular process' (GO:0009987), 'multicellular organismal process' (GO:0032501), 'response to stimulus' (GO:0050896), 'immune system process' (GO:0002376), 'signaling' (GO:0023052), 'developmental process' (GO:0032502) and 'metabolic process' (GO:0008152). Among the sectors, both GO:0065007 and GO:0051179 contained 20 enriched GO terms and were the main biological processes with the highest number of GO terms. In GO:0065007, 'regulation of ion transmembrane transport' (GO:0034765),

‘membrane depolarization during action potential’ (GO:0086010) and ‘regulation of membrane potential’ (GO:0042391) were classified as category A, this result was anticipated as epilepsy involves the generation of action potential, which was closely related to GO:0086010 and was regulated by GO:0042391 and GO:0034765. Meanwhile, category A GO terms under GO:0051179 included ‘sodium ion transmembrane transport’ (GO:0035725), ‘sodium ion transport’ (GO:0006814), ‘calcium ion transmembrane transport’ (GO:0070588), and ‘calcium ion import’ (GO:0070509). By comparing the category A GO terms in GO:0065007 and GO:0051179, it can be deduced that sodium and calcium ion transport were playing important roles in membrane depolarization, and the key genes behind this process included *SCN1A*, *CACNA1G*, *CACNA1H* and *CACNA1S*. In line with this deduction, ‘membrane depolarization’ (GO:0051899), the parent term for GO:0086010, was also performed by genes related to sodium ion transport (*SCN1B*) and calcium ion transport (*CACNB3*, *CACNB4*, *CACNG4*, *CHRNA4*). When analysis was expanded beyond GO:0065007 and GO:0051179, another four category A GO terms, namely ‘chemical synaptic transmission’ (GO:0007268), ‘cardiac muscle cell action potential involved in contraction’ (GO:0086002), ‘neuronal action potential propagation’ (GO:0019227), and ‘adult walking behavior’ (GO:0007628), were identified. Analysis of these GO terms found that *CACNA1G* was enriched in GO:0007268, hence suggesting a possible role of *CACNA1G* in synaptic transmission. In total, 11 category A GO terms were observed from this enrichment.

Meanwhile, 38 GO terms were classified as category B. Apart from ion-transport-related processes like ‘potassium ion transmembrane transport’ (GO:0071805) and ‘chloride transport’ (GO:0006821), category B GO terms included a variety of cellular processes such as ‘transmembrane receptor protein tyrosine kinase signaling pathway’ (GO:0007169), ‘gamma-aminobutyric acid signaling pathway’ (GO:0007214), ‘protein autophosphorylation’ (GO:0046777), and developmental processes like ‘axon guidance’

(GO:0007411). The enrichment of non-ion-transport-related processes had highlighted the complexity of molecular mechanism behind epilepsy. Further analysis of the category B GO terms revealed a few genes that were frequently presented in these GO terms, examples of such genes included *NTRK1* (12 GO terms), *NTRK2* (10 GO terms), *CSF1R* (9 GO terms) and *ERBB4* (9 GO terms), these four genes functioned as protein kinase genes and were mainly involved in phosphorylation-related processes such as GO:0046777, GO:0007169 and 'peptidyl-tyrosine phosphorylation' (GO:0018108). However, on some occasions, protein kinase genes can affect other biological processes. For instance, *NTRK1* was involved in 'positive regulation of neuron projection development' (GO:0010976) and GO:0007411, it is possible that *NTRK1* was achieving the neurodevelopmental functions through a complex protein interaction network such as the one as shown in Figure 4.2. This presumption is practicable as *NTRK1* might have an indirect interaction with *SCN1B* through *NEDD4*, *DMD* and *SNTA1*, and *SCN1B* in turn interacts with *SPTBN4*, *SPTB* and *SPTBN2* to perform the biological process GO:0007411. If this presumption is true, then protein kinase genes like *NTRK1*, *NTRK2*, *CSF1R* and *ERBB4* might be associated with epilepsy and should be discovered further.

For the GO terms classified as category C or D, 'calcium ion transport' (GO:0006816), 'potassium ion transport' (GO:0006813), 'cation transmembrane transport' (GO:0098655) and the aforementioned GO:0051899 should be considered in epilepsy-genetic analyses. Despite these GO terms were presumed to have a lower risk of inducing epilepsy, they were closely related to category A and B processes such as GO:0086010 and 'ion transport' (GO:0006811). Hence, there was an increased likelihood that the genes under these GO terms were contributing to epilepsy. Besides, one enriched GO term, 'synaptic transmission, GABAergic' (GO:0051932), was not classified as there was no test gene enriched in this process. However, the presence of training gene had indicated the potential involvement of GO:0051932 in epilepsy, and the genes behind this process

included *GABRA1* and *GABRG2* which were also involved in GO:0007214 and GO:0006821.

In short, functional enrichment has raised four points to consider when analysing the candidate genes in GGE. Firstly, the genes involved in sodium and calcium ion transport should be analysed as they are probably playing important roles in the initialization of membrane depolarization during action potential. Secondly, the genes involved in other ion transport, such as potassium ion transport, should be considered as they may interact with sodium and calcium ion transporters in the regulation of membrane potential. Thirdly, genes involved in synaptic transmission, including those in the gamma-aminobutyric acid signalling pathway and chloride ion transport, should be included in the analysis as synapses are serving as a platform for transmitting electrical impulses throughout the central nervous system. Fourthly and lastly, protein kinase genes may be associated with GGE through multiple mechanisms as they involved in many developmental processes such as GO:0007411, and signalling pathways like GO:0007169. Besides, protein kinases may affect the expression of cellular receptors, Section 4.6.2 describes a case from our cohort in which the GGE phenotype in the family may be induced by the interaction between protein kinase gene *ERBB4* and gamma-aminobutyric acid receptor gene *GABRA1*. The inclusion of protein kinase genes in epilepsy-genetic analyses may shed some light on the novel mechanism behind GGE.

4.6 Segregation Analysis

Segregation analysis was conducted in families F-1 and F-9 (Table 3.2). For F-1, the segregation analysis aimed to investigate the inheritance of the *SCN1A* c.5753C>T variant in the family. Meanwhile, segregation analysis for F-9 focused on the inheritance of *ERBB4* c.1972A>T and *GABRA1* c.448G>A variants in the family.

4.6.1 F-1 Family

4.6.1.1 Clinical and Family Information

GGE-1 was the proband of family F-1, she was a 31-year-old female diagnosed with juvenile myoclonic epilepsy (JME). GGE-1 had her febrile seizure onset at 1 year old and manifested generalized tonic-clonic seizures (GTC) and myoclonic seizures at 12 years old. In medical tests, both electroencephalograms of GGE-1 were reported as normal when electroencephalography (EEG) was conducted during 13 and 30 years old, and magnetic resonance imaging (MRI) on her brain was also reported as normal. In this family, 8 individuals including the proband had manifested the symptoms of epilepsy. Figure 4.4 shows the pedigree of this family and the phenotypes of each member. According to the clinical information of the family members, the mother (II.5, GGE-3) and nephews (IV.9 and IV.10) were diagnosed with JME. Meanwhile, the cousin (III.16) did not present any myoclonic seizures and was diagnosed with genetic generalized epilepsy (GGE). The phenotype of the maternal aunt (II.3) was undetermined due to poor recall; however, her daughter (III.9) was having febrile seizures and her grandsons (IV.9 and IV.10) were diagnosed with JME. Besides, GGE-1's brother (III.15) and maternal uncle (II.7) were also having febrile seizures. Table 4.8 shows the clinical characteristics of all the affected members in this family.

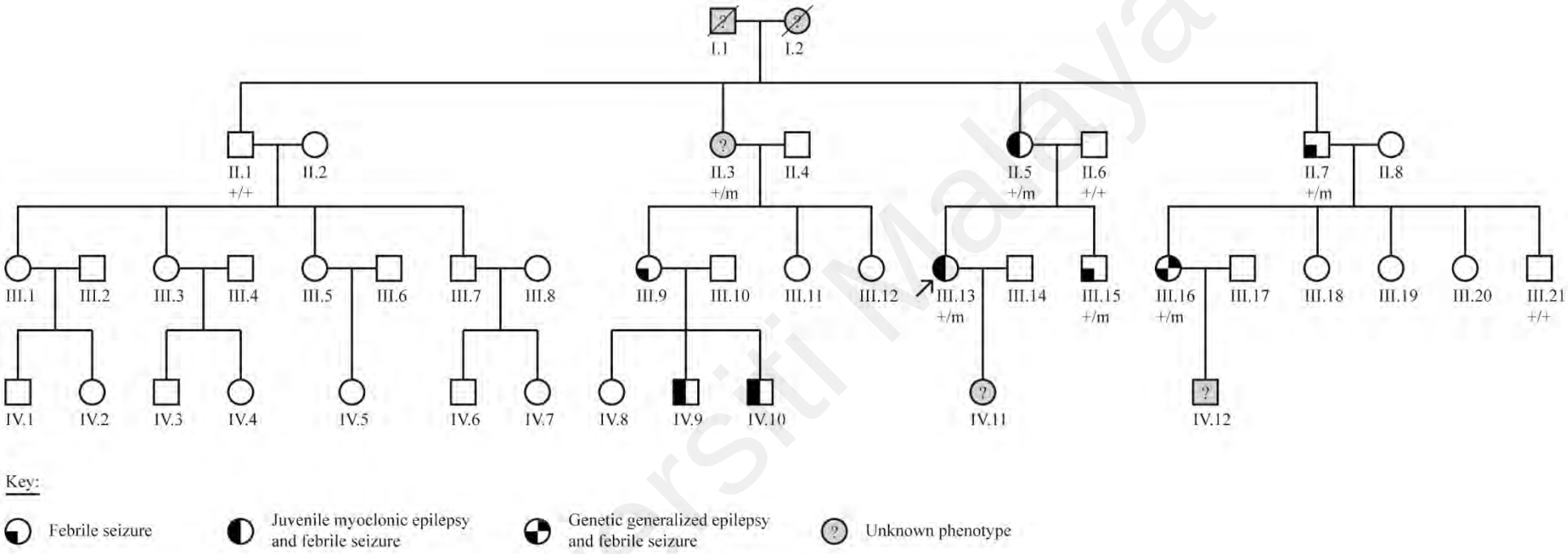


Figure 4.4: The pedigree of F-1 and the phenotypes of each member. GGE-1 (proband) is labelled as "III.13" and pointed by the black arrow. Squares represent male and circles represent female, the slash on the square/circle denotes the deceased member, + denotes the wild type *SCN1A* variant and m denotes the *SCN1A* c.5753C>T variant.

Table 4.8: Clinical characteristics of the affected family members in F-1.

Individual	**II.5**	**II.7**	**III.9**	**III.13 (proband)**	**III.15**	**III.16**	**IV.9**	**IV.10**
Age (years old)	65	62	44	31	30	33	18	14
Age of onset (years old)	12	Not known	Not known	12	Not known	23	14	12
Seizure type	GTC, myoclonic seizures	Not applicable	Not applicable	GTC, myoclonic seizures	Not applicable	GTC	GTC, myoclonic seizures	GTC, myoclonic seizures
Febrile seizure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Diagnosis	JME and FS	FS	FS	JME and FS	FS	GGE and FS	JME and FS	JME and FS
EEG	Not known	Not known	Not known	Normal	Not known	Normal	Not known	Not known
MRI	Not known	Not known	Not known	Normal	Not known	Not done	Not known	Not known
Current treatment	None	Not known	Not known	VPA, LTG	Not known	None	VPA, CBZ	VPA, CBZ
Presence of *SCN1A* c.5753C>T variant	Yes	Yes	Not tested	Yes	Yes	Yes	Not tested	Not tested

Abbreviations: CBZ, Carbamazepine; EEG, Electroencephalography; FS, Febrile seizures; GGE, Genetic generalized epilepsy; GTC, Generalized tonic-clonic seizures; JME, Juvenile myoclonic epilepsy; LTG, Lamotrigine; LVT, Levetiracetam; MRI, Magnetic resonance imaging; VPA, Sodium valproate.

4.6.1.2 PCR and Sanger Sequencing

Polymerase chain reaction (PCR) and Sanger sequencing were conducted on individuals III.13 (proband), II.1, II.3, II.5, II.6, II.7, III.15, III.16 and III.21 (Figure 4.4). The results of the PCR amplification are shown in Figure 4.5. The expected PCR product size was 556 bp (Table 3.7) and PCR amplicons were detected between 500 bp and 600 bp bands of the DNA ladder for all individuals (lane 2-10), these events indicated successful amplification. Furthermore, absence of PCR product band in no-template control (lane 11) indicated that the PCR amplification was valid and there was no contamination in the PCR reaction mix.

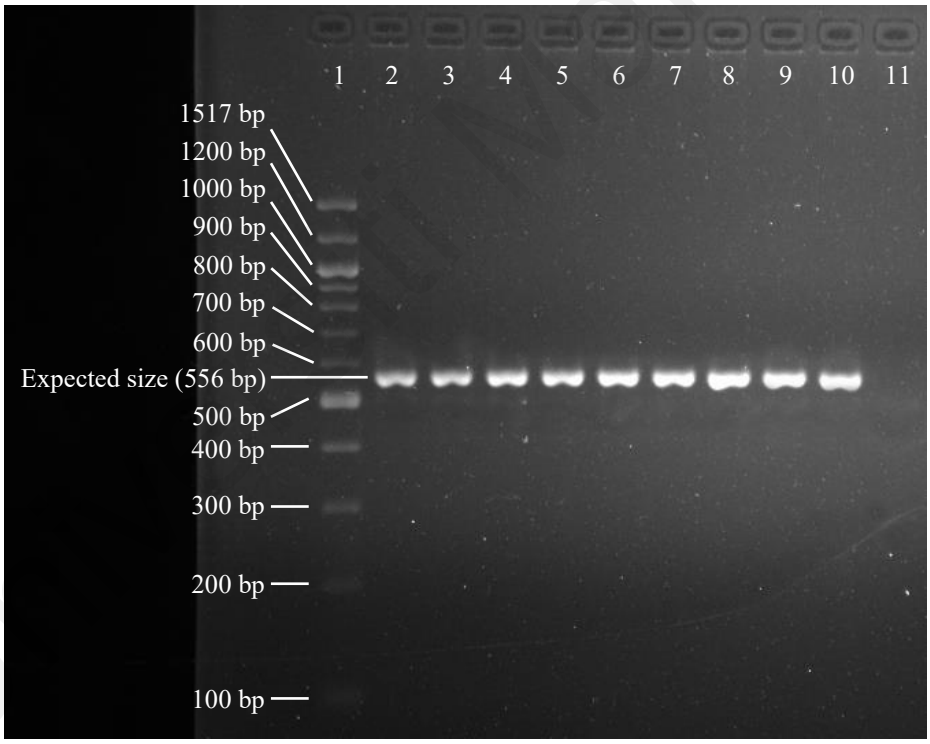


Figure 4.5: The results of PCR amplification for *SCN1A* c.5753C>T variant in F-1 family. Lane 1: 100 bp DNA ladder; Lane 2: III.13 (GGE-1); Lane 3: II.1; Lane 4: II.3; Lane 5: II.5; Lane 6: II.6; Lane 7: II.7; Lane 8: III.15; Lane 9: III.16; Lane 10: III.21; Lane 11: No-template control.

As mentioned in Section 3.10.4, Sanger sequencing was outsourced to Apical Scientific Sdn Bhd, data analysis was based on the DNA chromatograms provided by the company. The proband III.13 (GGE-1), mother (II.5, GGE-3), II.3, II.7, III.15 and III.16

were found to be heterozygous for *SCN1A* c.5753C>T, these members (except II.3 in which her phenotype was undetermined) were having epilepsy in their lifetime. Meanwhile, this variant was absent in the non-epileptic father (II.6, GGE-2), uncle (II.1) and cousin (III.21). The presence and absence of the *SCN1A* variant in the mother (II.5, GGE-3) and father (II.6, GGE-2) were also consistent with the whole exome sequencing results. Figure 4.6 shows the confirmation of the heterozygous *SCN1A* c.5753C>T from the DNA chromatograms of the proband (GGE-1), the presence of the *SCN1A* variant in other family members is indicated in Figure 4.4. From the inheritance pattern of the *SCN1A* variant in this family, it can be deduced that the *SCN1A* c.5753C>T mutation is likely pathogenic and probably autosomal dominant.

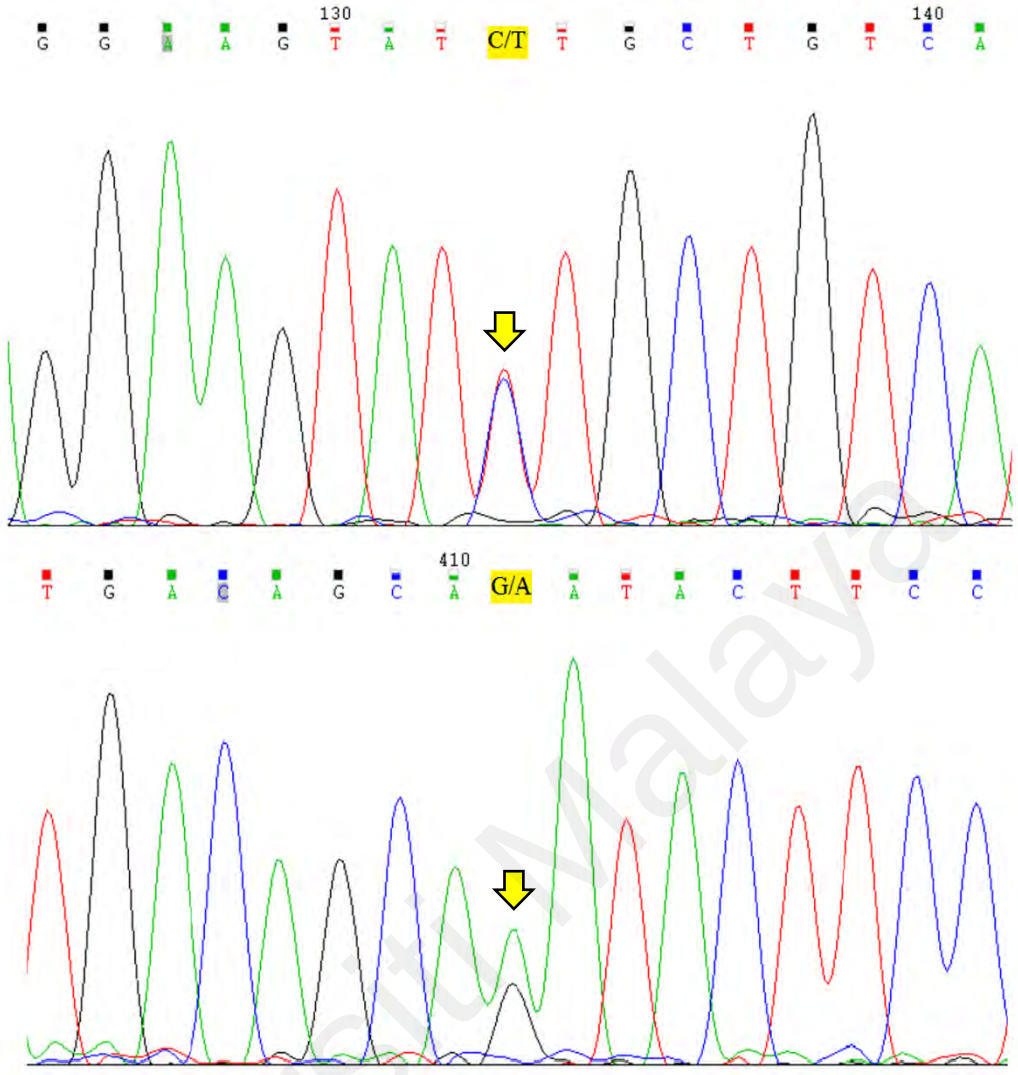


Figure 4.6: DNA chromatograms from Sanger sequencing of the proband (III.13, GGE-1). The heterozygous *SCN1A* c.5753C>T variant is indicated by the yellow arrow. Above: DNA chromatogram of the forward reaction; Below: DNA chromatogram of the reverse reaction.

4.6.1.3 Protein Analysis

The *SCN1A* c.5753C>T variant in this family was predicted to cause damaging effects by both SIFT (SIFT score = 0) and Polyphen-2 (Prediction output: Probably damaging). This variant was expected to induce an amino acid change from serine to phenylalanine at position 1918 (p.S1918F) which corresponds to the C-terminal domain of SCN1A protein. Since there is no known three-dimensional SCN1A protein structure available at present, homology modelling was conducted to model the SCN1A C-terminal domain so that the effect of the p.S1918F mutation can be studied. Figure 4.7 (A) shows the wild

type SCN1A C-terminal domain modelled by SWISS-MODEL. In terms of quality assessment, the modelled SCN1A was having a QMEAN score of -0.62, Ramachandran plot showed 94.5% of the residues fell in the most favoured regions and 5.5% of the residues resided within the additional allowed regions. According to the documentation provided by SWISS-MODEL and PDBsum, the predicted SCN1A model was a good quality model as its QMEAN score had exceeded the minimum threshold of -4.0 and more than 90% of its residues were placed in the most favoured regions of Ramachandran plot (European Bioinformatics Institute, 2013; Swiss Institute of Bioinformatics, 2021). Meanwhile, the mutant SCN1A (Figure 4.7 (B)) was generated with UCSF Chimera by applying the "Rotamers tool" function on the wild type SCN1A (Figure 4.7 (A)). Based on the predicted SCN1A models, there was no observable structural difference between the wild type and mutant SCN1A, both the serine (wild type) and phenylalanine (mutant) residues were not predicted to form any hydrogen bond (Figure 4.7). As a result, it was very unlikely that the p.S1918F mutation would cause any structural change in SCN1A.

Despite p.S1918F mutation was unlikely cause any structural change, this mutation was still classified as likely pathogenic due to high penetrance of epilepsy phenotype among the family members with the mutation. By analysing the template (PDB ID: 4JPZ) used in homology modelling, it was found that the C-terminal of the SCN2A was interacting with calmodulin (results not shown). Hence, it was hypothesized that the SCN1A p.S1918F mutation might have achieved its pathogenic effect through the interaction with calmodulin. In order to test this hypothesis, the binding of calmodulin to SCN1A C-terminal was stimulated with PyMOL. Figure 4.8 shows the results of the stimulation. From the stimulation, steric clash was observed between the calmodulin and the phenylalanine residue in the mutant SCN1A. As a result, the interaction between SCN1A and calmodulin was likely to be interfered by the p.S1918F mutation.

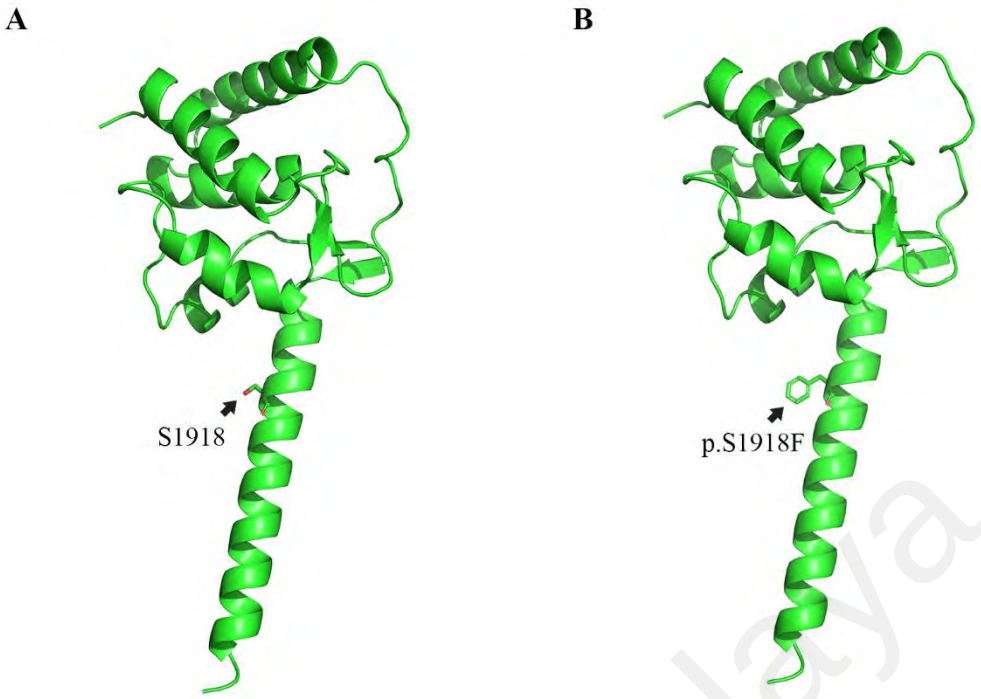


Figure 4.7: The modelled three-dimensional structures of SCN1A C-terminal domain. (A) Wild type SCN1A C-terminal domain, S1918 indicates the serine residue at position 1918. (B) Mutant SCN1A C-terminal domain, p.S1918F indicates the phenylalanine residue at position 1918.

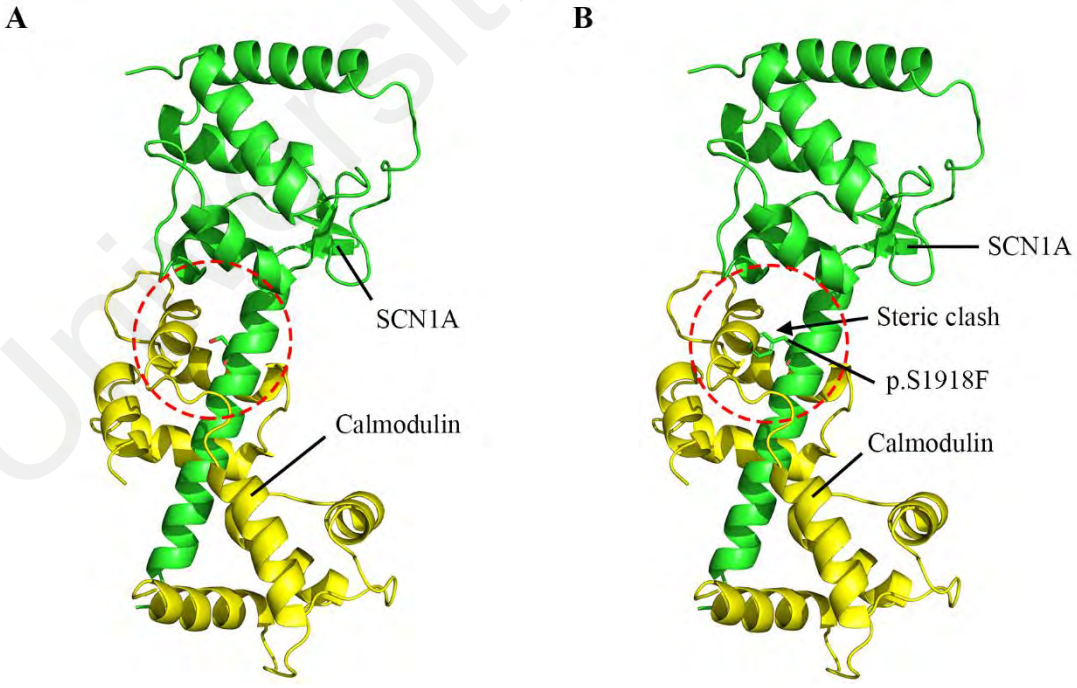


Figure 4.8: Stimulated binding orientation between SCN1A C-terminal (green) and calmodulin (yellow). (A) Wild type SCN1A C-terminal and calmodulin, no steric clash was observed in the structure. (B) Mutant SCN1A C-terminal and calmodulin, steric clash was observed between the phenylalanine residue with calmodulin.

4.6.2 F-9 Family

4.6.2.1 Clinical and Family Information

F-9 was a familial epilepsy case with variable epileptic syndromes, Figure 4.9 shows the pedigree and phenotypes of the family members, GGE-9 (II.3) was the proband of this family. The proband was a 30-year-old female with mixed syndromes of JME and TLE, she had her epilepsy onset at 2 years old. Electroencephalogram of GGE-9 at 22 years old showed generalized polyspike and wave discharges with T4 preponderance, and her MRI results at the same age reported right hippocampal atrophy and left hippocampal malrotation. Meanwhile, her epileptic symptoms included headache, vertigo, vomiting, facial automatism, head deviation, gelastic episodes, tinnitus, catamenial and aggressive staring. The only family members in F-9 with epilepsy were the elder sister (II.1) and the daughter (III.1). The sister (II.1) was diagnosed with genetic generalized epilepsy, presented with generalized tonic-clonic seizures and loss of consciousness. EEG of II.1 reported focal (frontal and occipital) and generalized polyspike and wave discharges. Meanwhile, the daughter (III.1) was diagnosed with focal epilepsy with learning disability. The symptoms of III.1 included vomiting, deviation of eyes and drooling of saliva, and the EEG conducted at 5 years old reported bilateral parietal epileptiform discharges. Interestingly, no family history of epilepsy was reported from both paternal and maternal parents. Table 4.9 shows the clinical characteristics of II.3, II.1 and III.1.

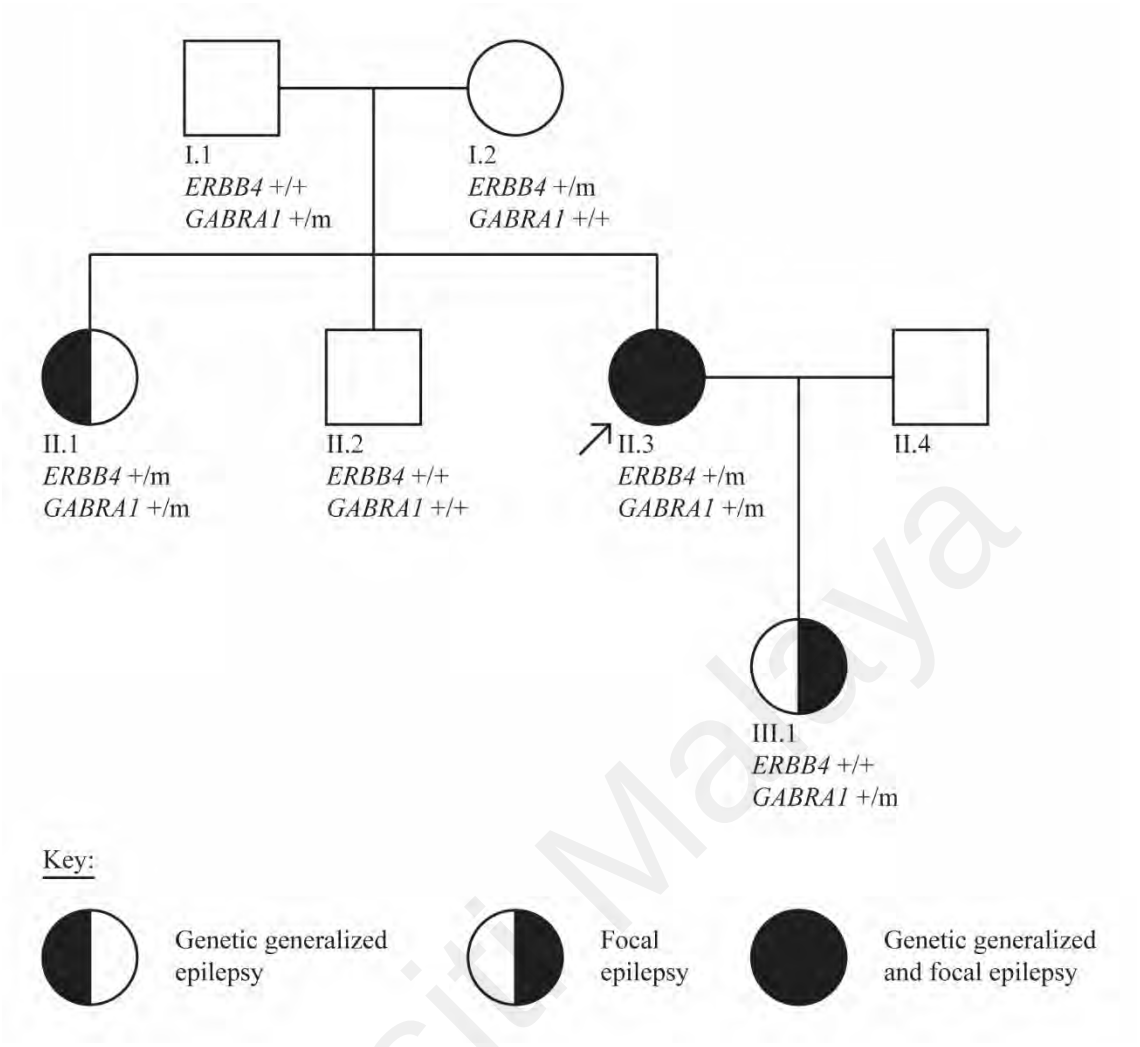


Figure 4.9: Pedigree of F-9 and the phenotypes of each member. GGE-9 (proband) is labelled as "II.3" and pointed by the black arrow. Squares represent male and circles represent female, + denotes the wild type variant and m denotes the mutant variant.

Table 4.9: Clinical characteristics of the affected family members in F-9.

Individual	II.1	II.3 (proband)	III.1
Diagnosis	Genetic generalized epilepsy	Temporal lobe epilepsy and juvenile myoclonic epilepsy	Focal epilepsy with learning disability
Age (years old)	39	30	9
Age of onset (years old)	Between 1-5	2	1
Seizure type	Generalized tonic-clonic seizures	Simple partial seizures	Not classified
Symptom	Generalized tonic-clonic seizures, loss of consciousness	Headache, vertigo, vomiting, facial automatism, head deviation, gelastic episodes, tinnitus, catamenial, aggressive staring, simple partial seizures	Vomiting, eyes deviated to one side and drooling of saliva which lasted for 10-20 minutes
EEG	Focal (frontal and occipital) and generalized polyspike and wave discharges	Generalized polyspike and wave with T4 preponderance	Bilateral parietal epileptiform discharges
MRI	Normal	Right hippocampal atrophy and left hippocampal malrotation	Normal
Current treatment	VPA, TPX	LTG, VPA, LVT	VPA

Abbreviations: EEG, Electroencephalography; MRI, Magnetic resonance imaging; VPA, Sodium valproate; TPX, Topiramate; LTG, Lamotrigine; LVT, Levetiracetam.

4.6.2.2 PCR and Sanger Sequencing

From the results of *in silico* pathogenicity prediction (Section 4.4; Appendix I) and gene prioritization (Section 4.5.1), *ERBB4* was ranked the highest tier (Tier 3) among the genes listed in Appendix I. Hence, *ERBB4* was chosen for further discovery and segregation analysis was conducted for the *ERBB4* c.1972A>T variant identified in this family. Figure 4.10 shows the results of PCR amplification for the *ERBB4* c.1972A>T variant. Due to time difference between sample collection for the family members, PCR was conducted separately for II.2 and III.1; therefore, two gel photos are shown in Figure 4.10. Nevertheless, both PCRs were conducted using the same PCR reagent and under the same PCR condition.

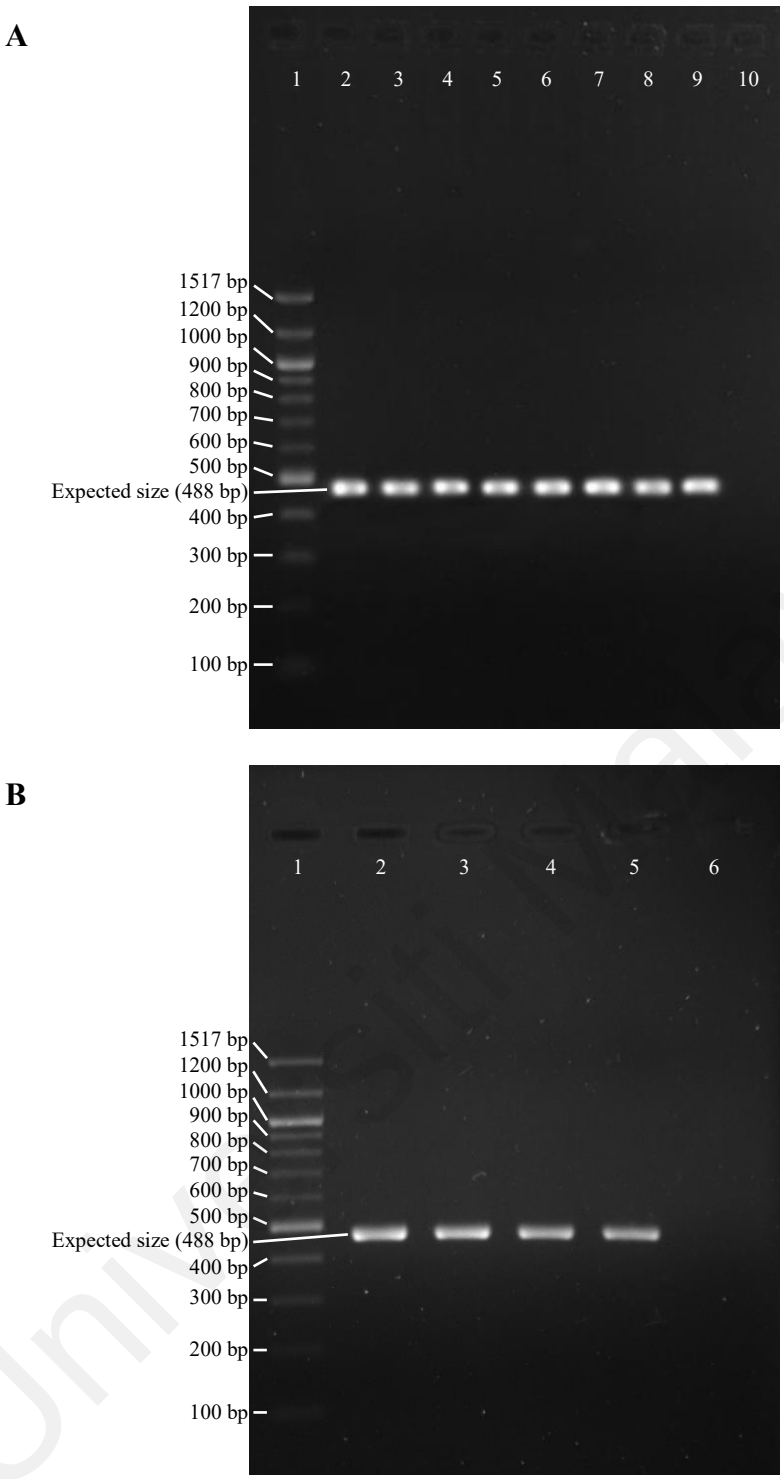


Figure 4.10: The results of PCR amplification for *ERBB4* c.1972A>T variant in F-9. (A) Lane 1: 100 bp DNA ladder; Lane 2 and 3: II.1; Lane 4 and 5: II.3 (GGE-9); Lane 6 and 7: I.1; Lane 8 and 9: I.2; Lane 10: No-template control. (B) Lane 1: 100 bp DNA ladder; Lane 2 and 3: II.2; Lane 4 and 5: III.1; Lane 6: No-template control.

Following the successful PCR amplification of *ERBB4* c.1972A>T variant among the family members (Figure 4.10), Sanger sequencing was conducted to study the inheritance pattern of this variant. The results showed both the affected proband (II.3) and sister (II.1) had inherited the *ERBB4* c.1972A>T variant. While this variant was absent in the unaffected brother (II.2), Sanger sequencing on the parents confirmed the existence of the *ERBB4* variant in the unaffected mother (I.2). The presence of *ERBB4* c.1972A>T in I.2 has questioned the pathogenicity of this variant. Despite *ERBB4* c.1972A>T was predicted as a pathogenic variant in the *in silico* prediction step described in Section 3.8, its pathogenic effect might not strong enough to induce epilepsy. Since PPI analysis showed *ERBB4* was interacting with *GABRA1* (training gene) through *NTRK2* (Section 4.5.2; Appendix G), the exome sequencing data of GGE-9 was screened again for the presence of *GABRA1* and *NTRK2* variants. From the screening, no variant was found in *NTRK2*; however, a novel *GABRA1* variant, c.448G>A, was identified. In order to investigate the potential association of the *GABRA1* c.448G>A variant with epilepsy, segregation analysis was conducted for the *GABRA1* variant. From the results, it was found that the *GABRA1* c.448G>A was present in proband (II.3), sister (II.1), daughter (III.1) and the unaffected father (I.1). Figure 4.9 shows the inheritance of *ERBB4* c.1972A>T and *GABRA1* c.448G>A variants among the family members. From the segregation pattern, it was shown that neither the *ERBB4* c.1972A>T nor the *GABRA1* c.448G>A variant was pathogenic enough to induce epilepsy. Since II.1 and II.3 had inherited both *ERBB4* and *GABRA1* variants, it was hypothesized that the epileptic phenotypes in II.1 and II.3 were caused by interaction between *ERBB4* and *GABRA1*. Meanwhile, the results of PCR amplification for *GABRA1* c.448G>A are shown in Figure 4.11, and the DNA chromatograms confirming the existence of *ERBB4* c.1972A>T and *GABRA1* c.448G>A are shown in Figure 4.12 and Figure 4.13 respectively.

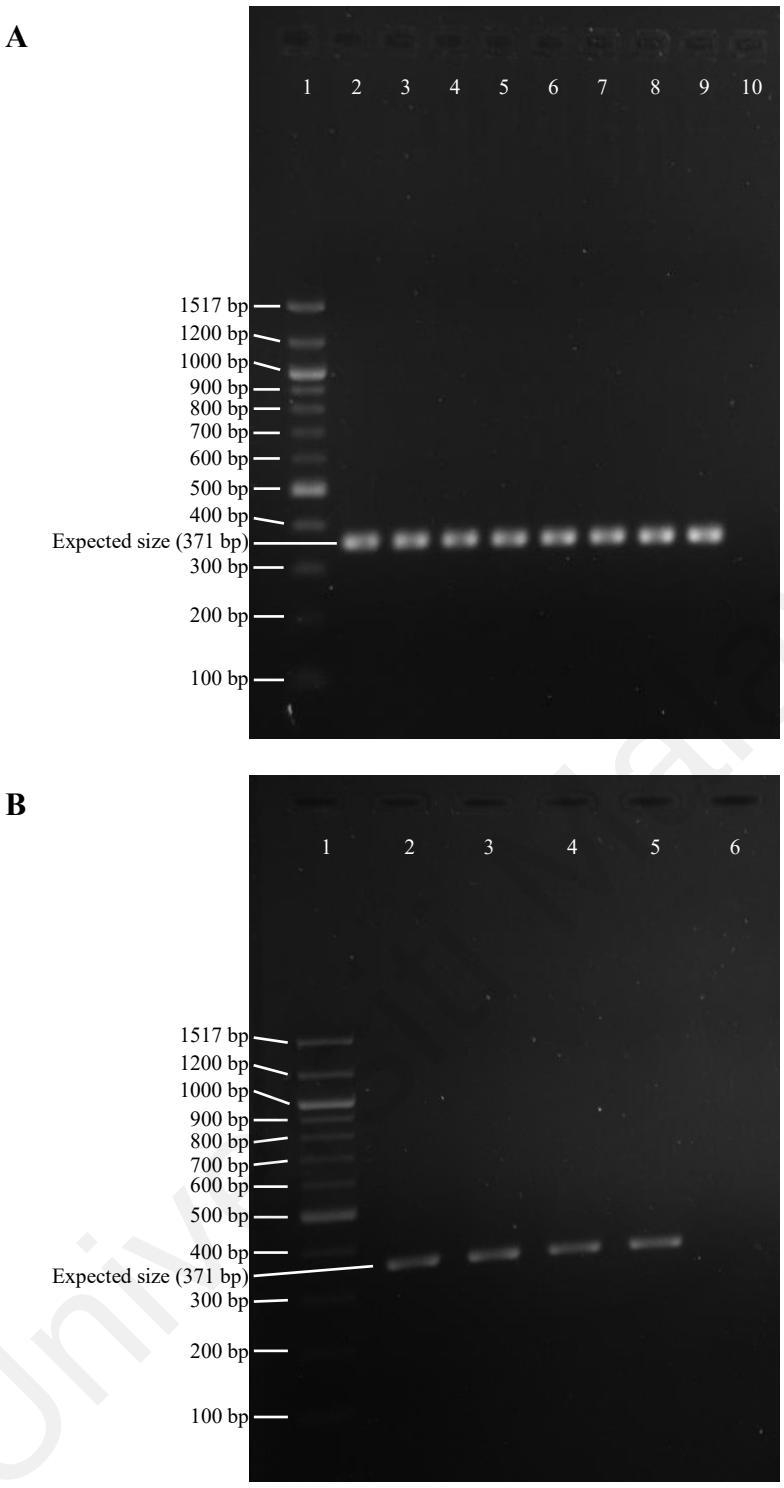


Figure 4.11: The results of PCR amplification for *GABRA1* c.448G>A variant in F-9. (A) Lane 1: 100 bp DNA ladder; Lane 2 and 3: II.1; Lane 4 and 5: II.3 (GGE-9); Lane 6 and 7: I.1; Lane 8 and 9: I.2; Lane 10: No-template control. (B) Lane 1: 100 bp DNA ladder; Lane 2 and 3: II.2; Lane 4 and 5: III.1; Lane 6: No-template control.

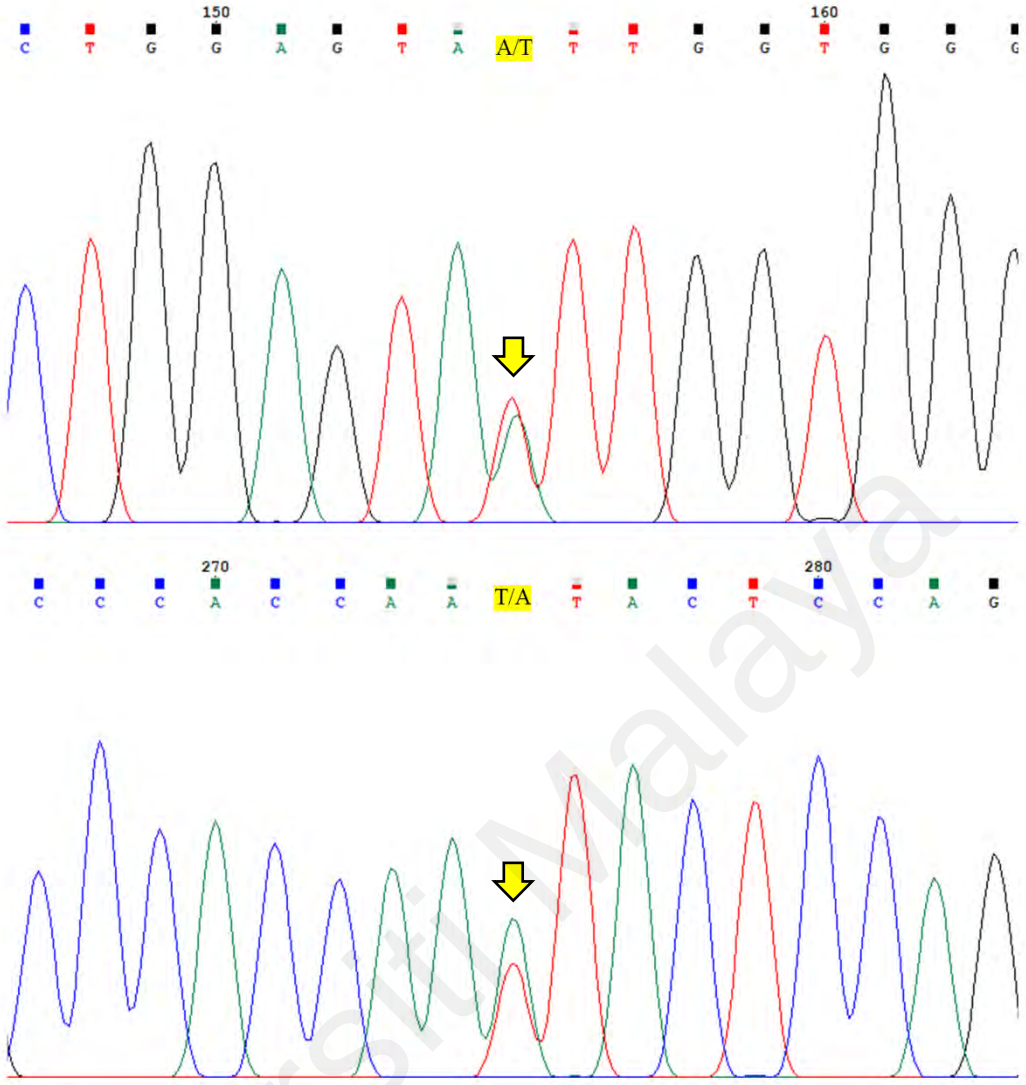


Figure 4.12: DNA chromatograms from Sanger sequencing of the proband (II.3, GGE-9). The heterozygous *ERBB4* c.1972A>T variant is indicated by the yellow arrow. Above: DNA chromatogram of the forward reaction; Below: DNA chromatogram of the reverse reaction.

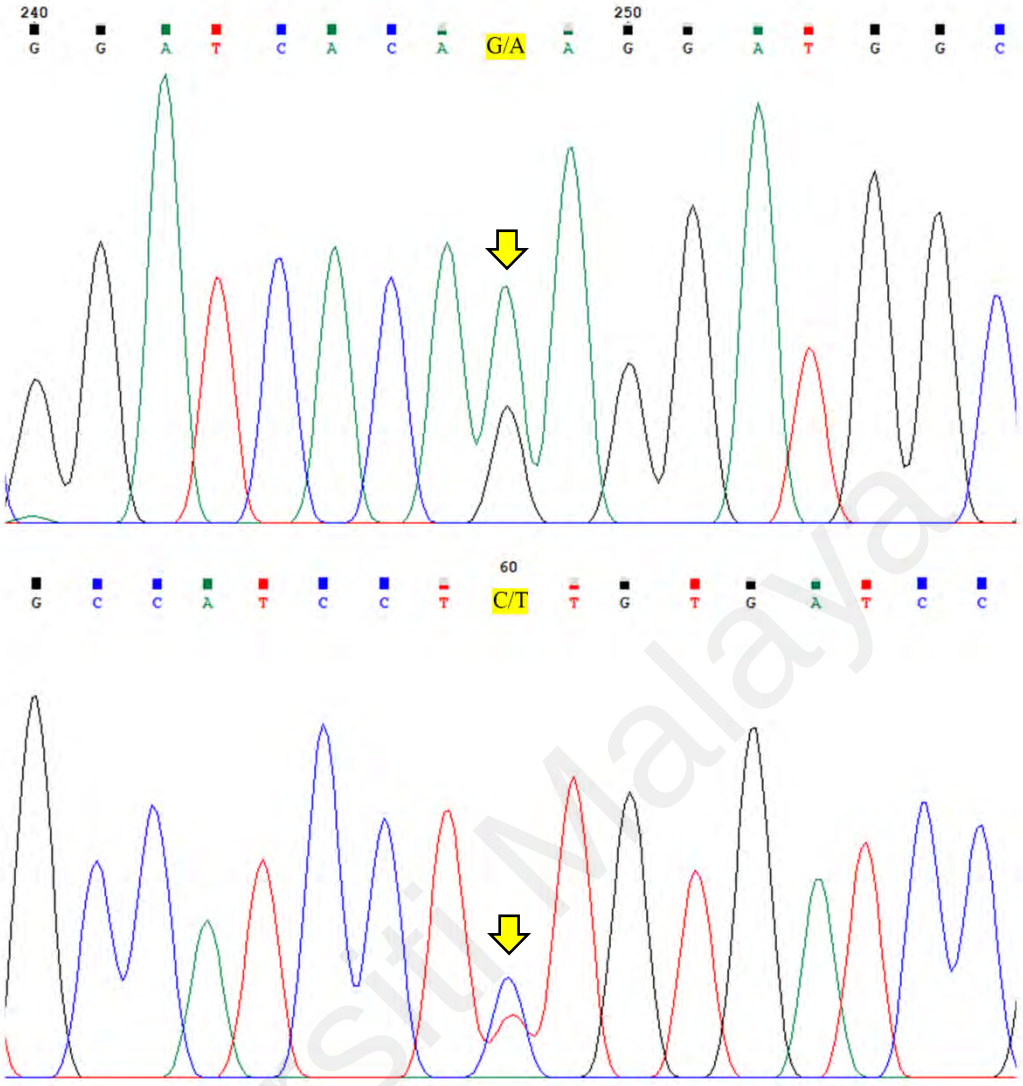


Figure 4.13: DNA chromatograms from Sanger sequencing of the proband (II.3, GGE-9). The heterozygous *GABRA1* c.448G>A variant is indicated by the yellow arrow. Above: DNA chromatogram of the forward reaction; Below: DNA chromatogram of the reverse reaction.

4.6.2.3 Protein Analysis

The novel *GABRA1* c.448G>A variant in this family was predicted as benign variant by SIFT and Polyphen-2; however, it was predicted to be disease causing by MutationTaster (http://www.mutationtaster.org) and Combined Annotation Dependent Depletion (CADD) (Rentzsch et al., 2019). Table 4.10 shows the outcomes of *in silico* pathogenicity prediction for *ERBB4* c.1972A>T and *GABRA1* c.448G>A excerpted from wANNOVAR annotation. In order to investigate the possible effects of the variants on protein structures, structural analysis was conducted for both ERBB4 and gamma-aminobutyric acid type A receptor subunit alpha1 (GABRA1) proteins.

Table 4.10: The details and outcomes of *in silico* pathogenicity prediction for the *ERBB4* and *GABRA1* variants excerpted from wANNOVAR annotation.

Gene symbol	*ERBB4*	*GABRA1*
Genomic position (hg19)	chr2:g.212495294T>A	chr5:g.161300315G>A
cDNA change	c.1972A>T	c.448G>A
Protein change	p.I658F	p.E150K
dbSNP	rs190654033	Not reported
1000 Genomes MAF (EAS)	0.003	Not reported
gnomAD MAF (EAS)	0.0034	Not reported
SIFT score	0.037	0.169
Polyphen-2 prediction	Probably damaging	Benign
MutationTaster prediction	Disease causing	Disease causing
CADD phred score	23.6	22.3

Abbreviations: CADD, Combined Annotation Dependent Depletion; cDNA, Complementary DNA; EAS, East Asian; gnomAD, Genome Aggregation Database; MAF, Minor allele frequency; SIFT, Sorting Intolerant from Tolerant.

As listed in Table 4.10, the *ERBB4* c.1972A>T variant was expected to cause substitution of the isoleucine residue at position 658 with phenylalanine (p.I658F). According to the ERBB4 structure retrieved from RCSB Protein Database Bank (PDB ID: 2LCX) (Figure 4.14), the transmembrane region consisted of a 24-residue fragment with sequence LIAAGVIGGLFILVIVGLTFAVYV that was corresponding to residues 652-675 in ERBB4 protein. This region contained a GG4-like motif AGVIGG (residues

655-660) that served as a polar interaction site for ERBB4 dimerization (Figure 4.14 (A)). When the receptors dimerized, hydrogen bond was formed between alanine residue A655 and glycine residue G659' on the adjacent chain (Figure 4.14 (B)). The p.I658F mutation was expected to cause a minor change in the GG4-like motif of ERBB4 receptors, while the mutation did not disrupt the hydrogen bond between A655 and G659', additional hydrogen bond was formed between A655 and glycine residue G660' (Figure 4.14 (C)).

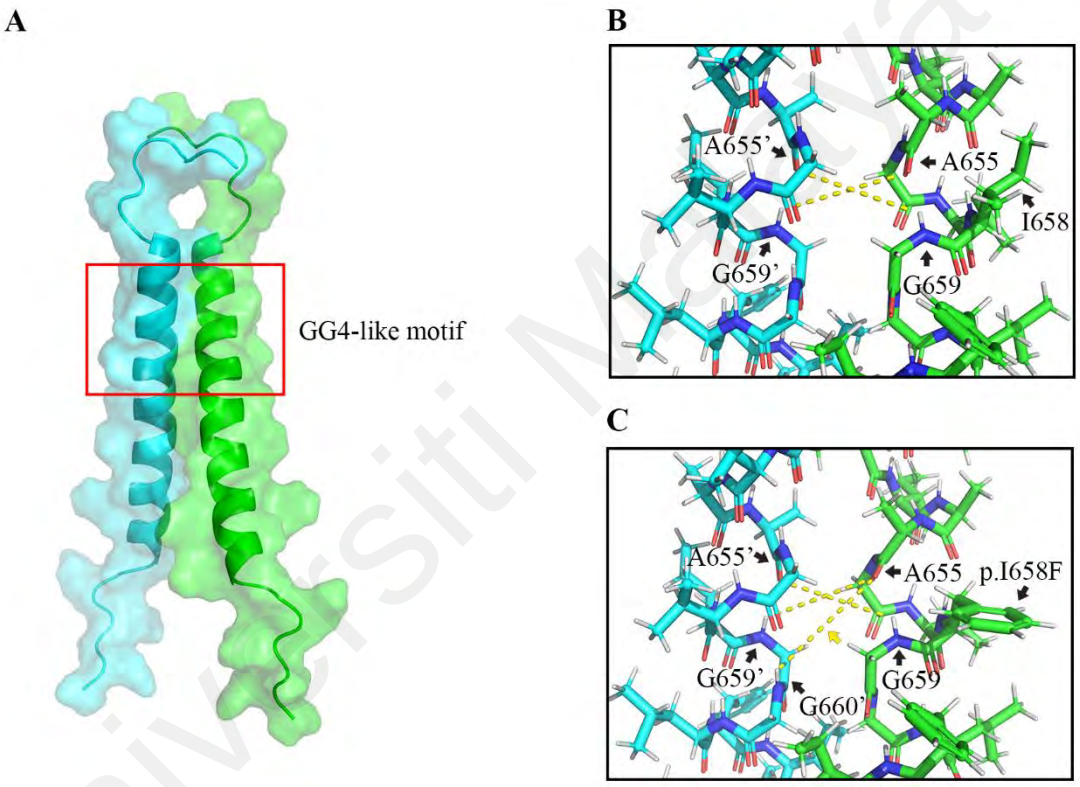


Figure 4.14: Structure of ERBB4 transmembrane domain dimer retrieved from RCSB Protein Database Bank (PDB ID: 2LCX), the individual ERBB4 in the dimer are indicated by green and cyan, yellow dotted lines represent the predicted hydrogen bond. (A) Side view of the ERBB4 transmembrane domain, the red box indicates the position of GG4-like motif. (B) GG4-like motif in wild type ERBB4. (C) GG4-like motif in mutant ERBB4, additional hydrogen bond is pointed by the yellow arrow.

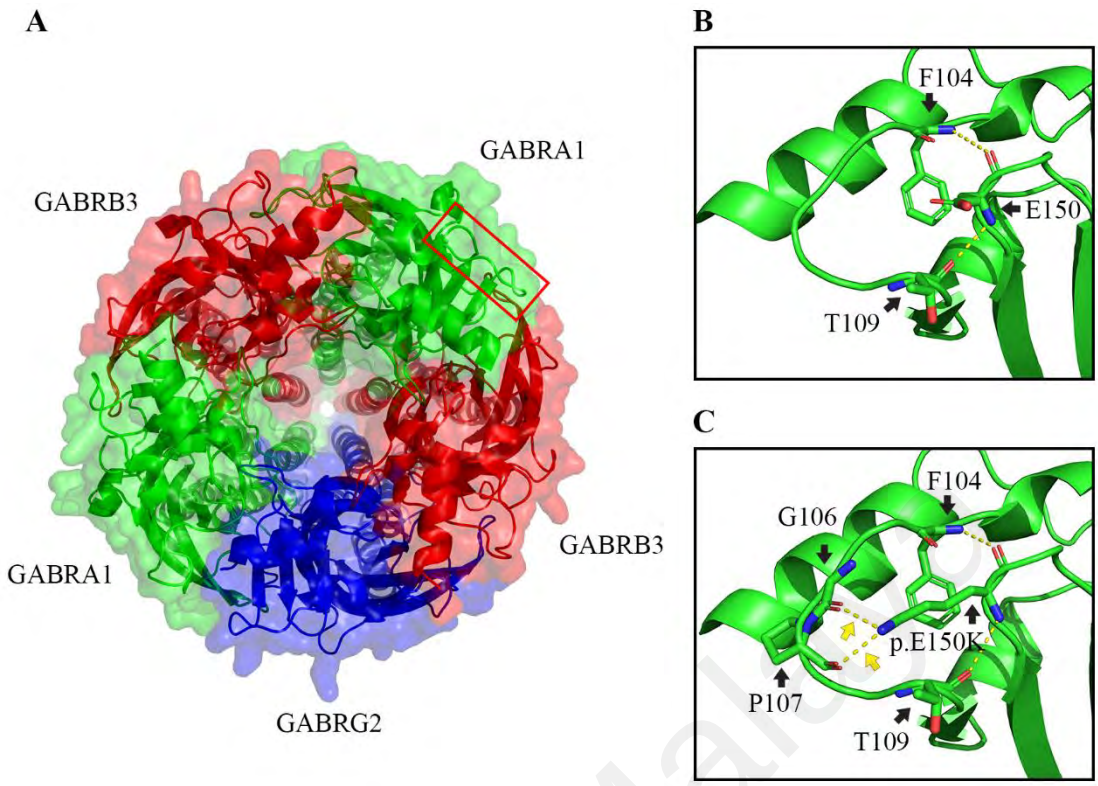


Figure 4.15: Structure of GABA-A receptor retrieved from RCSB Protein Database Bank (PDB ID: 6HUO), yellow dotted lines represent the predicted hydrogen bond. (A) Top view of the GABA-A receptor, it was a pentamer consisting of two GABRA1 subunits (green), two GABRB3 subunits (red) and one GABRG2 subunit (blue). The red box indicates the GABRA1 region shown in (B) and (C). (B) The wild type GABRA1. (C) The mutant GABRA1, additional hydrogen bonds are indicated by the yellow arrows.

Meanwhile, analysis of the published GABA-A structure from RCSB Protein Database Bank (PDB ID: 6HUO) showed the GABA-A receptor was a pentamer consisting of two GABRA1, two gamma-aminobutyric acid type A receptor subunit beta3 (GABRB3) and one gamma-aminobutyric acid type A receptor subunit gamma2 (GABRG2) (Figure 4.15 (A)). The *GABRA1* c.448G>A variant changed the glutamic acid residue at position 150 (E150), which was resided in the N-terminal domain of GABRA1, to lysine residue (p.E150K). The side chain of lysine residue in the mutant GABRA1 was predicted to form hydrogen bonds with the glycine residue G106 and proline residue P107 on the adjacent loop (Figure 4.15 (C)), the formation of the new hydrogen bonds was postulated to induce a minor structural disruption in the loop region of GABRA1. Figure 4.16 shows the distances of the GABRA1 loop from the mutation site and GABRB3, the distances

were measured between the alpha carbon atoms of E150 (or p.E150K in mutant GABRA1), P107, methionine residue M108 in GABRA1, and the glycine residue G58 and proline residue P60 in GABRB3. The decrease in the distances indicated that the loop was drew closer to the mutation site p.E150K in mutant GABRA1.

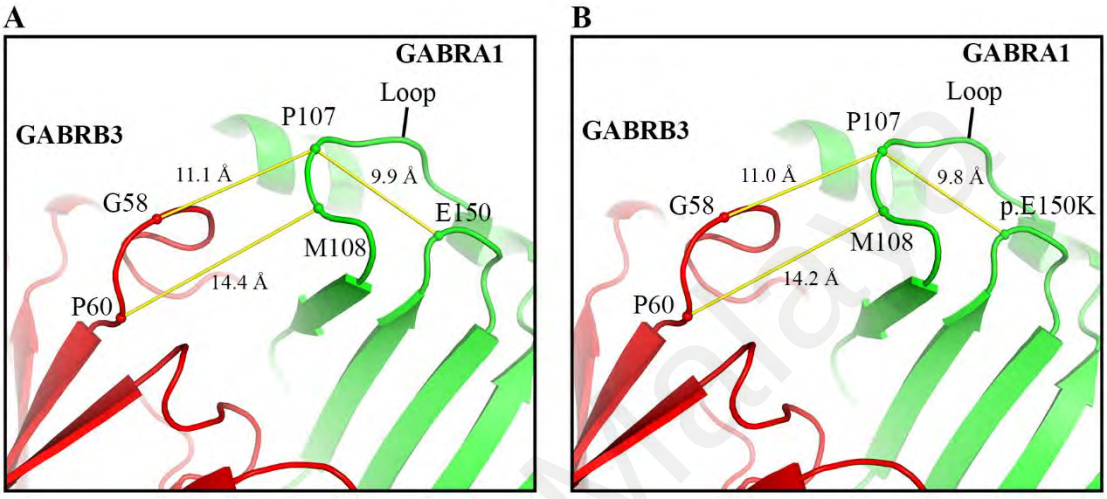


Figure 4.16: The distances of GABRA1 loop from the mutation site and GABRB3 in (A) GABA-A receptor with wild type GABRA1 subunit and (B) GABA-A receptor with mutant GABRA1 subunit.

CHAPTER 5: DISCUSSION

5.1 Clinical Characteristics and Phenotypes of Probands

The phenotypes of all the probands (except GGE-9) were very similar, most of them were diagnosed with JME, followed by GEFS+. In general, the phenotypes of the JME probands in our cohort are consistent with the clinical features mentioned in the ILAE diagnostic manual (Table 2.2). Meanwhile, GEFS+, which refers to the patients who developed febrile seizures after 6 years of age or afebrile seizures (Khair & Elmagrabi, 2015), was presented by GGE-7, GGE-12 and GGE-15. Despite difference in diagnosis, there are overlapping clinical characteristics between JME and GEFS+, these clinical characteristics are best exemplified by the GTC and myoclonic seizures which were observed in both our JME probands and GEFS+ probands (GGE-7 and GGE-15). Meanwhile, GGE-12 was having afebrile seizures that resemble GTC, her phenotype is similar to the GEFS+ case reported by Emmilia H. Tan et al. (2012) but no mutation in *SCN1A* was observed. In terms of EEG, the presence of generalized spike and wave (GSW) and polyspike-waves (PSW) is concurrent with the GGE features described by ILAE (Table 2.2). For GGE-1, GGE-10 and GGE-14, the normal EEG results were probably caused by antiepileptic drugs as normal electroencephalogram can be obtained if EEG is conducted after the drugs are taken. Nevertheless, the presence of GTC and myoclonic seizures has indicated that GGE-1, GGE-10 and GGE-14 were having GGE.

GGE-9 was the most complex case in this cohort. During the recruitment phase of this project, she was diagnosed with JME. However, as the project was ongoing, the diagnosis was changed to JME and TLE due to appearance of focal seizures. Interestingly, her EEG reported generalized PSW, this EEG result fulfilled the diagnostic criteria of GGE. Based on the clinical characteristics, epileptologist confirmed that GGE-9 was having both GGE and focal epilepsy. In contrast, all other probands in this study were manifesting the symptoms of GGE only. As discussed in Qaiser et al. (2020), there are genetic differences

between GGE and focal epilepsy. Since GGE-9 was having both GGE and focal epilepsy, she was removed from functional characterization analysis to prevent the accidental inclusion of focal epilepsy genes which might interfere the relevance of the functional characterization results to GGE.

5.2 Functional Characterization Analysis

5.2.1 Selection of Training Genes

As mentioned in Section 3.9.1, 34 JME-associated genes (Appendix D) were extracted from literature and used as training genes for gene prioritization. The choice of using JME-associated genes instead of other epilepsy-associated genes as training genes was mainly due to the reason that most study subjects involved in the functional characterization analysis were diagnosed with JME.

5.2.2 Selection of Reference Populations for Minor Allele Frequency (MAF) Determination

The high-throughput data from WES has allowed the identification of large number of genetic variants (Section 4.3). However, most of the variants were common variants or polymorphism that may not bring on huge effect in the pathogenesis of epilepsy (Helbig et al., 2016a; Richards et al., 2015). Therefore, the MAF of each variant must be determined so that the rare mutations can be isolated from the common variants for further analysis. In a multi-ethnic population like Malaysian, it is important to note that the MAF of each variant can be different between ethnic groups (Malays, Chinese and Indians) due to genetic variation. The reference populations used in MAF determination must share similar genetic composition with the study population. Since no study has been conducted on the genetic make-up among the ethnic groups in Malaysia yet, the large-scale genome analysis of Singaporean Malays, Chinese and Indians was referred in selecting the

reference population for MAF determination (Wu et al., 2019). According to Wu et al. (2019), the genetic composition of Singaporean Chinese and Indians is similar with East Asian and South Asian respectively; whereas the genetics of Malays is distinct and does not overlap with East Asian or South Asian. The only populations in 1KGP database with close genetic distance ($F_{ST} < 0.01$) to the Malays are KHV and CDX (Wu et al., 2019). Hence, the Singapore Sequencing Malay Project, 1KGP KHV and CDX were chosen as the reference populations for Malays; 1KGP EAS and gnomAD EAS were used as references for MAF in Chinese; 1KGP SAS and gnomAD SAS were selected as the reference populations for Indians (Table 3.4).

5.2.3 Gene Prioritization

Even though hundreds of genes have been reported to be associated with epilepsy (Wang et al., 2017a), the association risk of each gene with epilepsy is different. This phenomenon is best demonstrated by Butler et al. (2017), some genes like *SCN1A*, *KCNQ2*, *CDKL5* and *SCN2A* have higher mutation frequency in epilepsy patients. As shown in Ma et al. (2016), mutation frequency can be a measure to rank the disease-association risk of a gene. However, the calculation of mutation frequency often involves hundreds of samples (Bu et al., 2019; Butler et al., 2017; Gu et al., 2019; Singh et al., 2007). Compared to Butler et al. (2017), the sample size in this study is relatively small (only 13 individuals). Therefore, gene prioritization was used to assess the association risk of each gene with epilepsy instead of mutation frequency.

From the gene prioritization results, *SCN1A* and *CACNA1G* were predicted to have the highest association risk with epilepsy in our cohort (Section 4.5.1). The prediction of *SCN1A* as a high-risk epilepsy gene in our results was consistent with literature. *SCN1A* has been proposed to be the top epilepsy-associated gene with more than 1,700 reported variants (Kluckova et al., 2020), this statement is in line with results of genetic screening from multiple studies that *SCN1A* is always having the highest mutation frequency among

the tested genes in their cohorts (Borlot et al., 2019a; Butler et al., 2017; Miao et al., 2018). While *SCN1A* is deeply explored in epilepsy, the study of *CACNA1G* in epilepsy is less established. Nevertheless, Singh et al. (2007) proposed *CACNA1G* as one of the GGE-associated genes through a mutation analysis in 483 individuals consisted of 360 healthy controls and 123 GGE patients that were mostly made up of Japanese and Hispanics. Besides, another study involving 111-epilepsy-gene-panel screening in 57 Korean GGE patients reported high mutation frequency among the calcium ion channel genes *CACNA1H*, *CACNA1A* and *CACNA1G* (Lee et al., 2018). The high mutation rate in sodium and calcium ion channel genes suggests potential roles of sodium and calcium ion channels in the pathogenesis of epilepsy, this is also consistent with our results in functional enrichment analysis that both sodium and calcium ion channels were involved in the initialization of membrane depolarization, which is a key mechanism in the generation of action potentials and epileptic seizures (Section 4.5.3).

Meanwhile, the genes classified under tier 2 and below (Table 4.5) normally have a low mutation rate in epilepsy patients. Interestingly, two mutations in *GABRA6* were observed in our cohort. *GABRA6* is rarely reported in large-scale epilepsy-genetic screening from other countries like China, Netherlands, and United States (Helbig et al., 2016b; Snoeijen-Schouwenaars et al., 2019; Wang et al., 2017b). As a result, *GABRA6* may be a common genetic factor in epilepsy among Malaysians, further studies with large sample size are needed to confirm this presumption.

5.2.4 Protein-Protein Interaction (PPI) Analysis

In this project, PPI analysis aimed to study the possible interaction between the mutant genes from our patients and the known JME-associated genes. From Figure 4.2, it was shown that the molecular mechanism behind epilepsy involved a complex network of interacting proteins. Among the genes, *SCN1A* was found to have the highest number of

interactants in the network and it was predicted to interact with JME-associated genes from various categories like ion channels (*SCN1B*, *CACNB4*, *CLCN2*, *KCNQ3*, *KCNMA1*, *KCNJ10*), GABA receptor (*GABRA1*, *GABRD*, *GABRG2*), acetylcholine receptor (*CHRNA4*) and ion transporter (*SLC12A5*) (Section 4.5.2). Since ion channels, GABA receptors, acetylcholine receptors, and ion transporters are closely related with epilepsy (Psarropoulou, 2010), the close interaction between SCN1A and these groups of receptors or channels indicates a potential involvement of *SCN1A* in epilepsy. Being a sodium ion channel gene, *SCN1A* was involved in sodium ion transport (Table 4.7). As mentioned in Section 4.5.3, sodium and calcium ion transport were the key processes in the initialization of membrane depolarization. If a pathogenic mutation has occurred in *SCN1A*, there is a high probability that the membrane potential will be dysregulated due to abnormal influx of sodium ions, this will result in random membrane depolarization and subsequently epileptic seizures. Meanwhile, GABA receptors which act as antagonist for sodium ion channels tend to promote resting membrane potential by inducing the influx of chloride ions (Psarropoulou, 2010). The nature of GABA receptors in promoting a resting membrane potential makes it an ideal therapeutic target for epilepsy, many antiepileptic drugs like bromide, felbamate, stiripentol, topiramate and VPA inhibit seizures by enhancing the inhibitory effect of GABA receptors (Kobayashi et al., 2020). Besides, the calcium and potassium ion channels, which aided in the regulation of membrane potential (Table 4.7), may also serve as additional therapeutic targets for antiepileptic drugs. This condition is best demonstrated by LTG which also targets the calcium and potassium ion channels in addition to sodium ion channels (Moavero et al., 2017). The combination of antiepileptic drugs targeting the GABA receptors and ion channels may serve as an optimum treatment option for some epilepsy patients with *SCN1A* mutation, example of such case in our cohort is the *SCN1A*-positive GGE-1 which was treated with VPA and LTG (Table 4.1). The same therapy can also be applied to

patients with mutation in calcium ion channel genes, such as the *CACNA1G*-positive GGE-13 in our cohort (Table 4.1; Appendix E), as calcium ion channels were sharing similar biological processes with sodium ion channels in the sense that they were also involved in the initialization of membrane depolarization during action potential. From this aspect, it is worth noting that dystrophin (*DMD*) was having an indirect interaction with sodium and calcium ion channels, it has been shown to co-express with *SCN1A* and *CACNB4* in the brain (Doorenweerd et al., 2017), the roles of *DMD* in the regulation of voltage-gated calcium channel and sodium ion transmembrane transporter activity render its potential to be used as new therapeutic target for antiepileptic drugs (Table 4.7). Nonetheless, further studies are stilled needed to investigate the practicality of utilizing DMD as the therapeutic target in epilepsy.

Apart from the ion channels or receptors involved in regulation of membrane potential and generation of action potential, *SCN1A* was also predicted to interact with signal transduction gene (*EFHC1*), spectrin genes (*SPTB*, *SPTBN2* and *SPTBN4*), and syntrophin gene (*SNTA1*) (Appendix G). Despite it was not predicted in our functional enrichment analysis (Section 4.5.3), *EFHC1* has been shown to involve in brain development in HEK293 cells and mouse models (de Nijs et al., 2012). Meanwhile, spectrin genes, *CSF1R*, *NTRK1* and *SCN1B* were involved in axon guidance, which is also a part of brain development (Table 4.7). The interaction of *SCN1A* with *EFHC1* and spectrin genes intimates a possible involvement of *SCN1A* in brain development, this may explain the phenomenon reported by Lee et al. (2017) that brain structural alteration was observed in epileptic children with *SCN1A* mutation. Besides, it also provides an explanation on how *SCN1A* is associated with developmental and epileptic encephalopathy (DEE), a condition characterized by early-onset seizures and developmental delay (Scheffer & Nabbout, 2019; Steward et al., 2019). On other hand, *SNTA1* has been shown to be associated with sudden infant death syndrome by increasing

the peak and late sodium current in cardiac muscles (Cheng et al., 2009). Through this mechanism, the interaction between *SNTA1* and *SCN1A* may induce sudden unexpected death in epilepsy (SUDEP), this is also in line with Massey et al. (2014) that *SCN1A* is a risk gene for SUDEP. Nevertheless, these mechanisms do not apply to the *SCN1A* c.5753C>T mutation in GGE-1, as the patient had normal MRI and no symptoms of DEE or SUDEP were observed from her and the family members with the mutation (Section 4.6.1.1).

5.2.5 Functional Enrichment Analysis

As mentioned in Section 2.5.3, there are only two pathway analysis studies that have been conducted on GGE. The results from our functional enrichment analysis are compared with the two studies, Epi25 Collaborative (2019) and Ozdemir et al. (2019), in this section to investigate the similarities and differences in the GGE-associated biological processes between Malaysians and Europeans.

In our results, ion transport like GO:0035725 (sodium ion transmembrane transport), GO:0070588 (calcium ion transmembrane transport), GO:0071805 (potassium ion transmembrane transport), GO:0006821 (chloride transport), and GABA-related processes such as GO:0007214 (gamma-aminobutyric acid signaling pathway) were flagged as important biological processes that are highly associated with GGE (Section 4.5.3). The identification of ion transport and GABA signalling as epilepsy-associated pathways is also reported by Epi25 Collaborative (2019), this result is expected as mutations in the ion channel and GABA receptor genes have been widely reported in epilepsy (Section 2.2.3). Malfunctions in ion channels and GABA receptors can affect the regulation of membrane potential and induce epileptic seizures (Section 5.2.4), the mechanism of ion channels and GABA receptors in epilepsy can be described by the excitation-inhibition (E/I) balance model discussed in Section 5.3.2.

Besides, our functional enrichment analysis has also proposed the potential association of protein-kinase-related biological processes with GGE (Section 4.5.3). While the roles of ion transport and GABA signalling in GGE are well studied, the knowledge on protein kinases in GGE is less established. In comparison with the previous studies, Epi25 Collaborative (2019) did not include kinase genes in their analysis and thus did not reveal any information on the involvement of protein kinases in GGE. However, all the five KEGG pathways enriched in Ozdemir et al. (2019), namely 'neurotrophin signaling pathway', 'pathways in cancer', 'focal adhesion', 'metabolic pathway' and 'MAPK signaling pathway', involve protein kinases as part of the functioning components and can be considered as protein-kinase-related processes. Despite this, the protein-kinase-related processes highlighted by Ozdemir et al. (2019) are mostly different from our study. For instance, the KEGG pathways 'neurotrophin signaling pathway', 'focal adhesion' and 'MAPK signaling pathway', which are equivalent to the GO terms 'neurotrophin signaling pathway' (GO:0038179), 'focal adhesion assembly' (GO:0048041) and 'MAPK cascade' (GO:0000165) respectively, were not enriched in our cohort. Meanwhile, the 'metabolic pathway' in KEGG pathway is equivalent to the GO term 'metabolic process' (GO:0008152), it is too general to compare 'metabolic pathway' (KEGG pathway) in Ozdemir et al. (2019) with GO:0008152 in this study as both terms cover many metabolic activities such as carbohydrate, protein, and lipid metabolism. Nevertheless, the GO term 'cholesterol metabolic process' (GO:0008203), which is a child process of GO:0008152, was enriched in our cohort but its KEGG equivalent 'cholesterol metabolism' was not reported in Ozdemir et al. (2019). On other hand, 'pathways in cancer' in KEGG is an umbrella term that covers numerous cellular processes ranging from cell proliferation regulation to cellular signalling. Since extracellular signal-regulated kinase (ERK) signalling, phosphatidylinositol 3-kinase (PI3K) signalling, and Wnt signalling are components of 'pathways in cancer', there is a

possibility that the biological processes 'positive regulation of ERK1 and ERK2 cascade' (GO:0070374), 'canonical Wnt signaling pathway' (GO:0060070), 'positive regulation of phosphatidylinositol 3-kinase signaling' (GO:0014068) and 'phosphatidylinositol-mediated signaling' (GO:0048015) are serving as GGE-associated pathway in both Malaysians and Europeans. From the comparison of enrichment results between our study and Ozdemir et al. (2019), it can be deduced that the protein-kinase-related processes associated with GGE among Malaysians are different from those of Europeans, except ERK signalling, PI3K signalling, and Wnt signalling pathways. On top of this, the 'transmembrane receptor protein tyrosine kinase signaling pathway' (GO:0007169) and 'calcium-mediated signaling using intracellular calcium source' (GO:0035584) observed from our cohort were not reported in both Epi25 Collaborative (2019) and Ozdemir et al. (2019), this makes GO:0007169 and GO:0035584 worthy to study as they may serve as GGE-associated pathways that are unique to Malaysians or even Asians.

5.3 Segregation Analysis

5.3.1 F-1 (*SCN1A* c.5753C>T Variant)

Since our gene prioritization and PPI analysis demonstrated that *SCN1A* might serve as a key gene in JME (Sections 4.5.1 and 4.5.2), the *SCN1A* c.5753C>T variant in GGE-1 was undoubtably selected as the key variant for further analysis. From the segregation pattern of this variant, it can be deducted that the *SCN1A* c.5753C>T variant is likely pathogenic and autosomal dominant (Section 4.6.1.2). As discussed in Section 5.2.3, *SCN1A* has been proposed to be top epilepsy-associated gene. Apart from DEE and SUDEP (Section 5.2.4), *SCN1A* has been reported in Dravet syndrome and GEFS+ (Binini et al., 2017; Tunçer et al., 2018). However, *SCN1A* is rarely reported in JME. Despite Jingami et al. (2014) has proposed the possible association of *SCN1A* with JME, the family phenotype in their study was GEFS+. According to Richards et al. (2015), the

evidence provided by Jingami et al. (2014) may not strong enough to show the pathogenic role of *SCN1A* in JME due to the lack of segregation and functional analysis data. In contrast, four family members from F-1 in this study were diagnosed with JME (II.5, III.13, IV.9 and IV.10), while Sanger sequencing was not conducted in IV.9 and IV.10, the presence of *SCN1A* c.5753C>T variant in II.5 and III.13 may serve as supportive evidence to show the association of *SCN1A* with JME (Table 4.8). Besides, the *SCN1A* c.5753C>T variant was also inherited in II.7, III.15 and III.16, the symptoms of these individuals were considered milder compared to II.5 and III.13, as III.16 exhibited GTC without myoclonic seizures, while II.7 and III.15 were having febrile seizures only. Pleiotropy among family members with *SCN1A* mutation has been reported in previous studies (Mahoney et al., 2009; Marini et al., 2006). This phenomenon is probably caused by the interaction between *SCN1A* and other genetic as well as environmental factors, the same condition was also observed in epileptic mouse models (Escayg & Goldin, 2010; Salgueiro-Pereira et al., 2019).

SCN1A encodes the alpha subunit of voltage-gated sodium channel Nav1.1, this subunit consists of one N-terminal, one C-terminal, and four homologous transmembrane domains, each transmembrane domain in turn is made up of six transmembrane segments; the alpha subunit binds with beta-1 and beta-2 subunits, which are encoded by *SCN1B* and *SCN2B* respectively, to form a complete sodium channel (Catterall, 2000). In terms of molecular function, the alpha subunit is primarily involved in the regulation of sodium ion intake by changing the permeability to sodium ions through structural reconfiguration, whereas the beta subunits regulate the voltage-gating activity of the alpha subunit (Catterall, 2000; Escayg & Goldin, 2010). Due to direct involvement in sodium ion transport, disruption in the alpha subunit is prone to epilepsy. As shown in Figure 5.1, epilepsy-associated mutations have been reported from almost any domain in the alpha

subunit, suggesting that all the domains in the alpha subunit are important for the normal function of voltage-gated sodium channel Nav1.1.

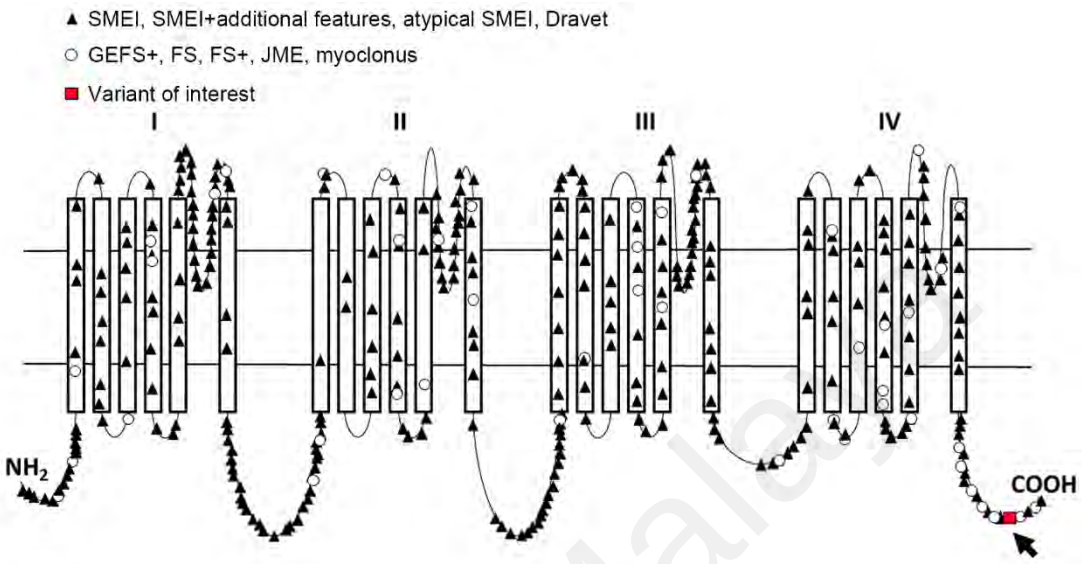


Figure 5.1: Schematic diagram showing the domain organization of the alpha subunit of voltage-gated sodium ion channel Nav1.1 as well as the positions of variants implicated in Dravet syndrome, GEFS+, JME or febrile seizures retrieved from *SCN1A* Variant Database (Claes et al., 2009) (accessed May 2019). The red square marks the position of the *SCN1A* c.5753C>T variant in the structure.

The *SCN1A* c.5753C>T mutation in GGE-1 induced the amino acid change p.S1918F in the C-terminal domain of Nav1.1 alpha subunit (Section 4.6.1.3). According to Parihar & Ganesh (2013), the C-terminal domain serves as an activation-inactivation switch for the sodium channel and about 7% of missense mutations are resided in this region. It is believed that mutations in the C-terminal are causing epilepsy through two mechanisms: first, by weakening the interaction with the beta subunit; and second, by interfering the interaction with calmodulin (Gaudioso et al., 2011; Rusconi et al., 2009; Spampanato et al., 2004). The *SCN1A* c.5753C>T variant identified from this project is a novel variant that has not been reported in 1KGP or gnomAD, and no previous study has been conducted on this variant yet. This mutation was expected to interfere the binding of calmodulin due to steric clash (Figure 4.8). As demonstrated in Gaudioso et al. (2011),

calmodulin tends to boost the inactivation of sodium channels. The interference in calmodulin binding brought on by the p.S1918F mutation is likely to cause delay in the inactivation of sodium ion channels which can subsequently promote the generation of epileptic seizures.

By combining evidence from multiple aspects such as variant rarity, *in silico* prediction, protein analysis, and segregation analysis, we conclude that the *SCN1A* c.5753C>T variant is likely pathogenic and is potentially associated with the epileptic syndromes in family F-1. The findings on the *SCN1A* c.5753C>T variant in this project were published as Chan et al. (2020) in Neurological Sciences (List of Publications and Papers Presented).

5.3.2 F-9 (*ERBB4* c.1972A>T and *GABRA1* c.448G>A Variants)

In comparison with the *SCN1A* c.5753C>T variant in F-1, the pathogenicity of *ERBB4* c.1972A>T and *GABRA1* c.448G>A variants in F-9 was lower. As shown in Figure 4.9, even though the father (I.1) inherited the *GABRA1* c.448G>A variant and the mother (I.2) carried the *ERBB4* c.1972A>T variant, both of them were unaffected. This observation indicated that the *GABRA1* c.448G>A and *ERBB4* c.1972A>T variants were not pathogenic enough to induce epilepsy when present in isolation. On other hand, GGE was observed from the proband (II.3) and sister (II.1) who carried both *GABRA1* c.448G>A and *ERBB4* c.1972A>T variants. This event led to the hypothesis that the GGE phenotype in this family was caused by interaction between the *GABRA1* and *ERBB4* mutations.

GABRA1 encodes the alpha1 subunit of GABA-A receptor, it has been shown to be associated with many GGE syndromes such as JME, GGE with generalized tonic-clonic seizures alone, GEFS+ and CAE (Johannesen et al., 2016). GABA-A receptors are predominant in neurons and serve as major inhibitory neurotransmitter receptors in mammalian brains (Ito et al., 2005). The mechanism of function of the GABA-A

receptors in central nervous system can be described using the "lock-and-pull" model proposed by Masiulis et al. (2019). According to the model, the GABA-A receptor consists of two GABRA1 subunits, two GABRB3 subunits and one GABRG2 subunit (Figure 4.15), the binding of GABA induces a conformational change in the receptor and the GABRB3 subunits are then drawn closer to the GABRA1 subunits. This conformational change leads to activation of GABA-A receptor which will subsequently allow the influx of chloride ions into the cell. Under this mechanism, all GABA-A subunits are important as mutation in any of the subunits, including the GABRA1, GABRB3 and GABRG2, can induce epilepsy (Hernandez & Macdonald, 2019). In this study, the *GABRA1* c.448G>A variant induced the amino acid change p.E150K in the GABRA1 subunit. As shown in Figures 4.15 and 4.16, this mutation allowed the formation of additional hydrogen bonds, the conformation of the GABA-A receptor was slightly disrupted and there was a decrease in the distance between GABRB3 and the loop of GABRA1. The function of this loop is unknown, but previously reported *GABRA1* p.F104C, *GABRB3* p.G32R and *GABRB3* p.V37G mutations affecting the loop region have been shown to be associated with epilepsy (Hernandez & Macdonald, 2019). We hypothesize that the structural disruption in the loop region may reduce the efficiency of GABA-A receptors by interfering the interaction between GABRB3 and GABRA1 subunits in the "lock-and-pull" mechanism.

In contrast to *GABRA1*, studies on *ERBB4* in epilepsy are much more limited. *ERBB4* encodes a receptor tyrosine kinase that serves many biological functions, previous studies on animal models have showed the involvement of *ERBB4* in the development of neuromuscular system as well as the regulation of GABA release and extracellular dopamine level (Deng et al., 2019; Paatero et al., 2019; Skirzewski et al., 2018). In the context of ERBB4 receptor activation, the GG4-like motif in the transmembrane domain functions as the binding site for dimerization of ERBB4 receptors, the dimer is held by

the hydrogen bonds between the GG4-like motifs of the two interacting ERBB4 receptors (Bocharov et al., 2012). The amino acid change p.I658F induced by the *ERBB4* c.1972A>T variant in this family caused the formation of an additional hydrogen bond in the GG4-like motif (Figure 4.14). According to previous studies on p.I658E mutation, additional hydrogen bond in the GG4-like motif can lead to the generation of constitutively active ERBB4 receptors (Bocharov et al., 2012; Vidal et al., 2007), it is possible that the p.I658F mutation in this family is having similar effect as p.I658E and a constitutively active ERBB4 receptor is produced.

The interaction between *GABRA1* and *ERBB4* is very complex, it involves multiple signalling pathways. From the PPI analysis, *GABRA1* and *ERBB4* were predicted to interact indirectly with each other through *NTRK2* (Section 4.5.2, Appendix G). *NTRK2* encodes tropomyosin-related kinase B (TrkB) that functions as a key component in brain derived neurotrophic factor (BDNF) signalling pathway, this pathway alters the expression of GABA-A receptors by modulating the phosphorylation level on the GABA-A receptors (Porcher et al., 2018). Meanwhile, the interaction between *ERBB4* and *NTRK2* is shown to influence the expression of TrkB and BDNF in mouse models (Zhang et al., 2018). If a mutation has occurred in *ERBB4*, it is likely that the BDNF signalling pathway will be affected and this event will eventually impact the phosphorylation process of GABA-A receptors. Besides, other studies in mouse models also revealed the regulatory effect of *ERBB4* on expression of GABA-A receptors through ERBB4-Neuregulin (ERBB4-NRG) signalling pathway (Li et al., 2011; Mitchell et al., 2013; Okada & Corfas, 2004). These lines of evidence sum up the point that *EBRR4* may affect the GABA-A receptors through multiple signalling pathways such as BDNF and ERBB4-NRG signalling.

In mouse models, the downregulation of the expression of GABA-A receptors on neuronal surface can cause a decrease in inhibitory postsynaptic current (iPSC) (Mitchell

et al., 2013). The reduced iPSC is believed to disrupt the E/I balance and give rise to a condition that favours neuronal excitation and epileptic seizures (Stafstrom, 2014). We postulate that the evoked iPSC needs to achieve a certain threshold before an inhibitory signal can be generated, Figure 5.2 illustrates our hypothesized disease model on the pathogenesis of GGE through the interaction between *GABRA1* and *ERBB4*. Under a normal condition where the individuals carry both wild type *GABRA1* and *ERBB4*, the GABA-A and ERBB4 receptors are functioning properly, the E/I balance is well maintained and thus there is no symptom of epilepsy (Figure 5.2 (A)). The structural change in the GABA-A receptor induced by *GABRA1* c.448G>A mutation is expected to decrease the efficiency of the receptor and slightly reduce the iPSC, and the constitutively active ERBB4 receptor produced by *ERBB4* c.1972A>T mutation may reduce the iPSC by downregulating the expression of GABA-A receptors. Nevertheless, each of the mutations will only contribute to a minor disruption in iPSC. When the individuals inherit only one of the mutations (either the *GABRA1* mutation or the *ERBB4* mutation), the evoked iPSC can still achieve the threshold, the E/I balance is still maintained and thus they are asymptomatic (Figure 5.2 (B and C)). However, when the individuals carry both *GABRA1* and *ERBB4* mutations, the pathogenic effects brought by the mutant GABA-A and ERBB4 receptors may act synergically. In this case, the evoked iPSC is greatly reduced and no longer exceeds the threshold. As a result, the inhibitory signal cannot be generated, the E/I balance is tilted to excitation, and these events give rise to GGE syndromes (Figure 5.2 (D)).

Although this hypothesized disease model may explain the pathogenesis of GGE, the individual III.1 diagnosed with focal epilepsy did not inherit the *ERBB4* c.1972A>T variant (Figure 4.9). This observation indeed suggests that our hypothesized disease model is only applicable to the pathogenesis of GGE but not focal epilepsy. It is possible that the focal epilepsy in this family was induced by different set of genes or even

environmental factors. While this disease model may shed some light on the mechanism of *ERBB4* c.1972A>T and *GABRA1* c.448G>A variants in GGE, further investigations such as functional studies are still needed to verify these findings including our hypothesized disease model.

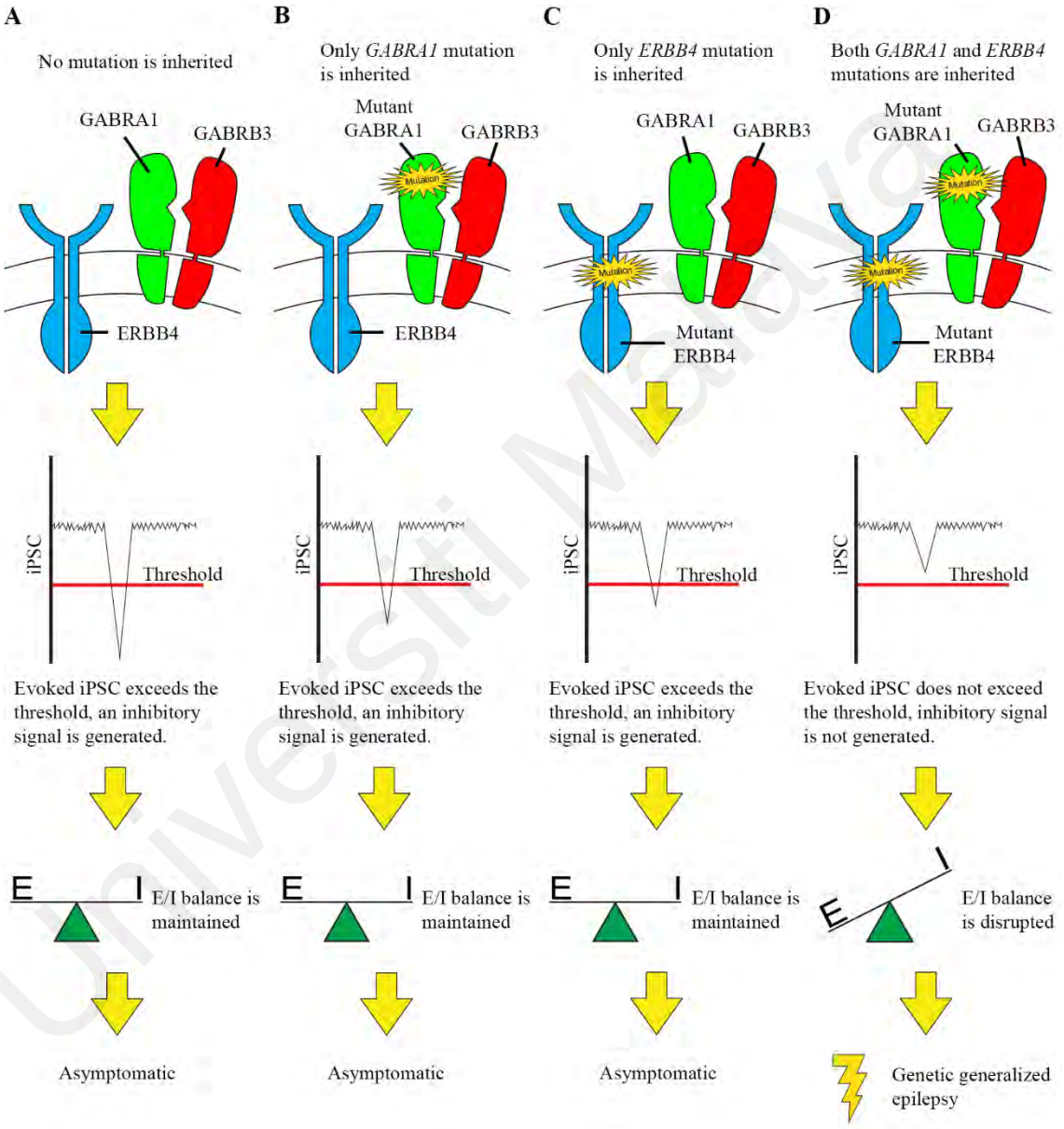


Figure 5.2: Hypothesized disease model on the pathogenesis of genetic generalized epilepsy through the interaction between *GABRA1* and *ERBB4*.

5.4 Usefulness of Functional Characterization Analysis in Whole Exome Sequencing Data Analysis

This study utilized functional characterization analysis to aid the identification of potential pathogenic variants from exome sequencing data. As discussed in Section 2.3.2, WES has an added advantage over gene panels in identification of novel or rare disease genes. For instance, the identification of *ERBB4* as a potential GGE-associated gene in GGE-9. Until 31 December 2020, *ERBB4* has never been tested in commercialized epilepsy panels such as Invitae Epilepsy Panel (Invitae Corporation, 2020), EPSAN (Mayo Foundation for Medical Education and Research, 2020), GeneDx Comprehensive Epilepsy Panel (GeneDx, 2020), Athena Epilepsy Panel (Athena Diagnostics, 2020), and MNG Epilepsy Genetic Testing (Laboratory Corporation of America, 2020). The *ERBB4* c.1972A>T variant would not be discovered if gene panels were opted as the screening method in this project. Despite the added advantage, WES identifies huge number of genetic variants from each test subject (Section 4.3), these variants must be shortlisted so that the variants with greater potential to cause epilepsy can be focused in the analysis. In familial cases, one of the shortlisting strategies is to search for shared variants between affected members, this strategy can be applied if WES was conducted in at least two affected family members (Goh & Choi, 2012), as in the case of trio WES in F-1 and F-4. However, trio WES imposed higher experimental cost as the sequencing was conducted on three individuals from a single family. Due to budget constraint, singleton WES was conducted for probands from families F-7 to F-18 (Table 3.2).

Variant shortlisting via identification of shared variants was not possible for the cases that underwent singleton WES. Under this circumstance, gene prioritization is very helpful as the variant with the highest possibility of being associated with GGE can be deduced from the gene ranks. The effectiveness of gene prioritization in variant shortlisting is best demonstrated in family F-1. The ranking of *SCN1A* as tier 1 gene in

gene prioritization indicated that the *SCN1A* c.5753C>T variant in GGE-1 was highly associated with JME (Section 4.5.1), this prediction was in line with the outcome of the analysis conducted using the shared-variant strategy proposed by Goh & Choi (2012) that the *SCN1A* c.5753C>T variant was called in the WES data of both epileptic GGE-1 and GGE-3 but not the unaffected GGE-2, and segregation analysis involving more family members further supported that *SCN1A* c.5753C>T variant in this family is probably pathogenic (Section 4.6.1.2). The consistency between the results of gene prioritization and segregation analysis has hinted the potential of gene prioritization in the shortlisting of genetic variants identified from WES.

On other hand, PPI analysis may not as useful as gene prioritization in the process of shortlisting genetic variants. However, it is still a useful tool in the analysis of polygenic epilepsy like the one in GGE-9. In family F-9, gene prioritization predicted *ERBB4* c.1972A>T variant as the most probable epilepsy-causing variant in the family, but conflicting results from segregation analysis have questioned the pathogenicity of the *ERBB4* variant (Section 4.6.2.2). This event led to the hypothesis that the epilepsy in this family is polygenic. The interaction between *ERBB4*, *NTRK2* and the GGE-associated *GABRA1* revealed from PPI analysis has facilitated the discovery of the regulatory role of *ERBB4* in the BDNF signalling pathway and its impact on the GABA-A receptors (Section 5.3.2). Our findings in F-9 may provide indirect evidence on the influence of *ERBB4* in human BDNF signalling pathway.

Meanwhile, the usefulness of functional enrichment in WES data analysis may vary depending on application or study design. In the context of individual analysis, such as identifying a pathogenic variant from a proband, functional enrichment might not be useful as the functional effects of each variant can be obtained from literature or by conducting functional studies (Richards et al., 2015). However, in population analysis, functional enrichment on pooled exome sequencing data can help to identify the potential

pathways associated with GGE in the study population (Section 4.5.3). The reason behind functional enrichment of being less useful in individual analysis is that enrichment software like DAVID requires an input gene list that fulfils certain criteria, such as containing appropriate number of key genes, for accurate enrichment (Huang et al., 2009). The gene list generated by WES data from single individual contains too few genes for the enrichment process and this may lead to imprecise results. As demonstrated in this project, even with the pooled WES data from 13 individuals, the GO terms 'brain-derived neurotrophic factor receptor signaling pathway' (GO:0031547) and 'ERBB signaling pathway' (GO:0038127), which are equivalent to the BDNF and ERBB4-NRG signalling pathways discussed in Section 5.3.2, are not enriched in our functional enrichment analysis (Section 4.5.3). Nonetheless, our enrichment results are still acceptably good with the positive enrichment of 'transmembrane receptor protein tyrosine kinase signaling pathway' (GO:0007169) that serves as a parent term covering both GO:0031547 and GO:0038127, but it will be better if GO:0031547 and GO:0038127 are called up in the results.

5.5 Implications/Key Findings of This Study

This project is a preliminary study to investigate the potential genetic factors and biological processes associated with GGE among Malaysians. From the results, it has been shown that ion transport, particularly sodium and calcium ion transport, were playing important role in pathogenesis of GGE. Mutations in the ion channel or GABA receptor genes may disrupt the E/I balance and thus causing epilepsy. Due to active involvement of ion channels and GABA receptors in neuronal-membrane-potential regulation, most antiepileptic drugs are designed to manipulate the function of ion channels or GABA receptors so that epileptic seizures can be suppressed (Kobayashi et al., 2020; Moavero et al., 2017; Stefanović et al., 2018). In this context, our PPI analysis

demonstrated the potential of DMD to be used as therapeutic target for antiepileptic drugs (Section 5.2.4). This discovery may provide some new insights on the design of antiepileptic drugs in future.

Besides, another key finding of this project is the roles of kinase genes in the pathogenesis of GGE. Our functional characterization analysis has demonstrated the potential association of *NTRK1*, *NTRK2*, *CSF1R* and *ERBB4* with GGE among Malaysian patients (Section 4.5). These genes were involved in 'transmembrane receptor protein tyrosine kinase signaling pathway' (GO:0007169) that performs many complex molecular functions. From our analysis in F-9, we proposed a possible mechanism of *ERBB4* in the reduction of iPSC and E/I balance disturbance (Section 5.3.2). On other hand, *NTRK1*, *NTRK2* and *CSF1R* were intimated to play important roles in brain development (Section 4.5.3), these genes were expected to induce a positive feedback on the ERK1/2 signalling cascade (GO:0070374) and PI3K signalling pathway (GO:0014068) during brain development (Table 4.7; Sánchez-Alegría et al., 2018; Sun & Nan, 2017). While there are studies showing the association of brain developmental abnormalities with JME (Gilsoul et al., 2019; Wandschneider et al., 2019), we were unable to confirm whether such association applies to Malaysian patients due to the lack of MRI data. If brain development is associated with JME among Malaysians, then the genes related to 'phosphatidylinositol-mediated signaling' (GO:0048015) such as *CBL* and *NEDD4* should also be considered for further analysis.

In addition, the GO term 'calcium-mediated signaling using intracellular calcium source' (GO:0035584) was enriched in our cohort but not in Europeans, the underlying genes contributing to this GO term included *NTRK2*, *TRPM2* and *KDR*. Similar to tyrosine kinase signalling, calcium-mediated signalling has multiple roles in cellular process. *TRPM2*, one of the key genes in calcium-mediated signalling, has been shown to regulate brain development as well as the activity of voltage-gated calcium ion

channels (Sawamura et al., 2017). Since voltage-gated calcium ion channels are important in the regulation of membrane potential, *TRPM2* maybe a risk factor in GGE among Malaysians. Interestingly, *TRPM1* and *TRPM5* were also predicted to be involved in calcium ion transmembrane transport (Table 4.7). Nevertheless, there is no evidence showing that *TRPM1* and *TRPM5* are highly expressed in human brain. Thus, the association of *TRPM1* and *TRPM5* with GGE remains to be discovered.

The findings from this project have indicated that ion channels and GABA receptors are not the only genetic causes of GGE. While mutations in the ion channels and GABA receptors are still the primary risk factors, genetic analysis should be expanded beyond the ion channels, and kinase genes should be considered as secondary genetic factors in GGE. As discussed earlier, protein kinases may alter the expression of GABA receptors and serve as regulators in brain developmental process. These findings further support the fact that epilepsy is not a pure channelopathy disease and it may involve a complex underlying molecular mechanism, the inclusion of kinase genes in the analysis may help to explore the complex mechanism behind GGE.

5.6 Limitations of Study

Despite this study has highlighted some new ideas on the genetics behind GGE (Section 5.5), there is room for improvement for this project. One of the limitations of this project is the lack of functional study. As mentioned in Section 2.4.2, while *in silico* tools can provide quick prediction on the outcome of a genetic mutation, it is still very important to verify the results using experimental data such as family segregation analysis and functional studies. This condition is best represented by the *ERBB4* c.1972A>T variant in GGE-9, the variant itself was not pathogenic enough to induce epilepsy even though it had been predicted to be damaging by SIFT and Polyphen-2. Although we have proposed a possible disease mechanism for the *ERBB4* variant in the pathogenesis of GGE,

we are unable to confirm whether our disease model is correct until functional study is conducted.

Besides, JME-associated genes were used as training genes during the gene prioritization, this would indicate that the results obtained from functional characterization analysis are more related with JME. Despite JME is a subtype of GGE, the findings from this study would not cover the whole spectrum of genetic factors in GGE. This work can be improved by including additional samples that cover more GGE phenotypes and by incorporating other GGE-associated genes as training genes in the gene prioritization analysis. The inclusion of additional samples with other GGE phenotypes such as CAE and JAE together with other GGE-associated genes in the analysis will provide us a more complete picture of the PPI network and the biological pathways associated with GGE in our country. Nevertheless, current results are still enough for us to draw the conclusion that in addition to ion channel and GABA receptor genes, protein kinase genes are also having the possibility of being associated with GGE, or at least JME in Malaysian population.

CHAPTER 6: CONCLUSION

This study has identified a few genetic factors which are associated with GGE among Malaysian population. From the functional characterization analysis, it is found that ion channels and GABA receptors are crucial in the generation and propagation of action potential. The genes encoding the ion channels, such as *SCN1A*, *CACNA1G*, *CACNA1S*, *CACNB3*, *CACNG4* and *KCNH1*, and GABA receptors, like *GABRA1* and *GABRA6*, should be treated as primary genetic factors as pathogenic mutations in these genes are very likely to induce GGE among Malaysians. Besides, our results have intimated the possible risk of tyrosine kinases in GGE, these kinases are involved in the coordination of numerous cellular processes including brain development and GABA-A-receptor expression. While we are unable to confirm the risk of brain developmental abnormalities in GGE with this study, we have proposed a disease model on the mechanism of tyrosine kinase ERBB4 in pathogenesis of GGE. Due to multiple roles of protein kinases in our body, further studies on the protein-kinase-related pathways will provide noteworthy knowledge on the mechanism behind GGE. Additionally, this study has also identified DMD as a new therapeutic target due to its capability of regulating the sodium transporters and calcium ion channels, this finding may contribute to the design of new antiepileptic drugs in future.

REFERENCES

Aaberg, K. M., Surén, P., Søraas, C. L., Bakken, I. J., Lossius, M. I., Stoltenberg, C., & Chin, R. (2017). Seizures, syndromes, and etiologies in childhood epilepsy: The International League Against Epilepsy 1981, 1989, and 2017 classifications used in a population-based cohort. *Epilepsia, 58*(11), 1880-1891.

Abdel-Mannan, O., & Sutcliffe, A. G. (2020). A national surveillance study of childhood epilepsy mortality in the UK and Ireland. *European Journal of Neurology, 27*(2), 327-333.

Abou-Khalil, B., Ge, Q., Desai, R., Ryther, R., Bazyk, A., Bailey, R., . . . George, A. L. (2001). Partial and generalized epilepsy with febrile seizures plus and a novel SCN1A mutation. *Neurology, 57*(12), 2265-2272.

Acuna-Hidalgo, R., Veltman, J. A., & Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biology, 17*(1), Article#241.

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods, 7*(4), 248-249.

Al-Harazi, O., El Allali, A., & Colak, D. (2019). Biomolecular databases and subnetwork identification approaches of interest to big data community: An expert review. *OMICS: A Journal of Integrative Biology, 23*(3), 138-151.

Aliferis, C. F., Alekseyenko, A. V., Aphinyanaphongs, Y., Brown, S., Fenyo, D., Fu, L., . . . Wang, J. (2011). Trends and developments in bioinformatics in 2010: Prospects and perspectives. *Yearbook of Medical Informatics, 6*, 146-155.

Alsaadi, T., Kassie, S., El Hammasi, K., Shahrour, T. M., Shakra, M., Turkawi, L., . . . Raoof, M. (2017). Potential factors impacting health-related quality of life among patients with epilepsy: Results from the United Arab Emirates. *Seizure - European Journal of Epilepsy, 53*, 13-17.

Altmann, A., Weber, P., Bader, D., Preuß, M., Binder, E. B., & Müller-Myhsok, B. (2012). A beginners guide to SNP calling from high-throughput DNA-sequencing data. *Human Genetics, 131*(10), 1541-1554.

Anderson, D., & Lassmann, T. (2018). A phenotype centric benchmark of variant prioritisation tools. *npj Genomic Medicine, 3*, Article#5.

Arroyo, S., Brodie, M. J., Avanzini, G., Baumgartner, C., Chiron, C., Dulac, O., . . . Serratosa, J. M. (2002). Is refractory epilepsy preventable?. *Epilepsia, 43*(4), 437-444.

Asawavichienjinda, T., Sitthi Amorn, C., & Tanyanont, W. (2002). Prevalence of epilepsy in rural Thailand: A population-based study. *Journal of the Medical Association of Thailand, 85*(10), 1066-1073.

Assadeck, H., Toudou-Daouda, M., Mamadou, Z., Moussa-Konate, M., Hassane-Djibo, F., & Douma-Maiga, D. (2019). Clinical and etiological characteristics of epilepsy in the elderly: A hospital-based study from a tertiary care referral center of Niamey, Niger. *Journal of Neurosciences in Rural Practice, 10*(4), 571-575.

Athena Diagnostics. (2020). Epilepsy. Retrieved on December 31, 2020 from https://www.athenadiagnostics.com/condition-guides/epilepsy

Balestrini, S., & Sisodiya, S. M. (2018). Pharmacogenomics in epilepsy. *Neuroscience Letters, 667*, 27-39.

Balqis Ramly, Nor Afiqah-Aleng, & Zeti-Azura Mohamed-Hussein. (2019). Protein-protein interaction network analysis reveals several diseases highly associated with polycystic ovarian syndrome. *International Journal of Molecular Sciences, 20*(12), Article#2959.

Banerjee, A., Bhatia, V., Didwal, G., Singh, A. K., & Saini, A. G. (2021). ADSL deficiency – The lesser-known metabolic epilepsy in infancy. *The Indian Journal of Pediatrics, 88*(3), 263-265.

Barozzi, C., Galletti, M., Tomasi, L., De Fanti, S., Palazzini, M., Manes, A., . . . Galiè, N. (2019). A combined targeted and whole exome sequencing approach identified novel candidate genes involved in heritable pulmonary arterial hypertension. *Scientific Reports, 9*(1), Article#753.

Basit, S., Albalawi, A. M., Alharby, E., & Khoshhal, K. I. (2017). Exome sequencing identified rare variants in genes HSPG2 and ATP2B4 in a family segregating developmental dysplasia of the hip. *BMC Medical Genetics, 18*(1), Article#34.

Baulac, S., Gourfinkel-An, I., Picard, F., Rosenberg-Bourgin, M., Prud'homme, J.-F., Baulac, M., . . . LeGuern, E. (1999). A second locus for familial generalized epilepsy with febrile seizures plus maps to chromosome 2q21-q33. *The American Journal of Human Genetics, 65*(4), 1078-1085.

Baulac, S., Huberfeld, G., Gourfinkel-An, I., Mitropoulou, G., Beranger, A., Prud'homme, J.-F., . . . LeGuern, E. (2001). First genetic evidence of GABA(A) receptor dysfunction in epilepsy: A mutation in the gamma2-subunit gene. *Nature Genetics, 28*(1), 46-48.

Beck, C., Moulard, B., Steinlein, O., Guipponi, M., Vallee, L., Montpied, P., . . . Malafosse, A. (1994). A nonsense mutation in the alpha4 subunit of the nicotinic acetylcholine receptor (CHRNA4) cosegregates with 20q-linked benign neonatal familial convulsions (EBNI). *Neurobiology of Disease, 1*(1-2), 95-99.

Beghi, E. (2011). New classification proposals for epilepsy: A real advancement in the nosography of the disease?. *Epilepsia, 52*(6), 1197-1198.

Beghi, E. (2019). Social functions and socioeconomic vulnerability in epilepsy. *Epilepsy & Behavior, 100*, Article#106363.

Beran, R. G., Devereux, J., & Buchanan, D. (2020). Some legal aspects of epilepsy. *Epilepsy & Behavior, 111*, Article#107244.

Berg, A. T. (2011). Classification and epilepsy: The future awaits. *Epilepsy Currents, 11*(5), 138-140.

Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., Van Emde Boas, W., . . . Scheffer, I. E. (2010). Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia, 51*(4), 676-685.

Berg, A. T., & Millichap, J. J. (2013). The 2010 revised classification of seizures and epilepsy. *CONTINUUM: Lifelong Learning in Neurology, 19*(3), 571-597.

Berg, A. T., & Scheffer, I. E. (2011). New concepts in classification of the epilepsies: Entering the 21st century. *Epilepsia, 52*(6), 1058-1062.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., . . . Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235-242.

Bezhentsev, V., Ivanov, S., Kumar, S., Goel, R., & Poroikov, V. (2018). Identification of potential drug targets for treatment of refractory epilepsy using network pharmacology. *Journal of Bioinformatics and Computational Biology, 16*(1), Article#1840002.

Binini, N., Sancini, G., Villa, C., Dal Magro, R., Sansoni, V., Rusconi, R., . . . Combi, R. (2017). Identification of two mutations in cis in the SCN1A gene in a family showing genetic epilepsy with febrile seizures plus (GEFS+) and idiopathic generalized epilepsy (IGE). *Brain Research, 1677*, 26-32.

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'Donovan, C., & Apweiler, R. (2009). QuickGO: A web-based tool for Gene Ontology searching. *Bioinformatics, 25*(22), 3045-3046.

Blixen, C., Ogede, D., Briggs, F., Aebi, M. E., Burant, C., Wilson, B., . . . Sajatovic, M. (2020). Correlates of stigma in people with epilepsy. *Journal of Clinical Neurology, 16*(3), 423-432.

Bocharov, E. V., Mineev, K. S., Goncharuk, M. V., & Arseniev, A. S. (2012). Structural and thermodynamic insight into the process of "weak" dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1818*(9), 2158-2170.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120.

Bonjoch, L., Mur, P., Arnau-Collell, C., Vargas-Parra, G., Shamloo, B., Franch-Expósito, S., . . . Castellví-Bel, S. (2019). Approaches to functionally validate candidate genetic variants involved in colorectal cancer predisposition. *Molecular Aspects of Medicine, 69*, 27-40.

Bonzanni, M., DiFrancesco, J. C., Milanesi, R., Campostrini, G., Castellotti, B., Bucchi, A., . . . DiFrancesco, D. (2018). A novel de novo HCN1 loss-of-function mutation in genetic generalized epilepsy causing increased neuronal excitability. *Neurobiology of Disease, 118*, 55-63.

Borlot, F., de Almeida, B. I., Combe, S. L., Andrade, D. M., Filloux, F. M., & Myers, K. A. (2019a). Clinical utility of multigene panel testing in adults with epilepsy and intellectual disability. *Epilepsia, 60*(8), 1661-1669.

Borlot, F., Whitney, R., Cohn, R. D., & Weiss, S. K. (2019b). MEF2C-related epilepsy: Delineating the phenotypic spectrum from a novel mutation and literature review. *Seizure - European Journal of Epilepsy, 67*, 86-90.

Borras, E., Chang, K., Pande, M., Cuddy, A., Bosch, J. L., Bannon, S. A., . . . Vilar, E. (2017). In silico systems biology analysis of variants of uncertain significance in Lynch syndrome supports the prioritization of functional molecular validation. *Cancer Prevention Research, 10*(10), 580-587.

Bosco, D. B., Tian, D. S., & Wu, L. J. (2020). Neuroimmune interaction in seizures and epilepsy: Focusing on monocyte infiltration. *The FEBS Journal, 287*(22), 4822-4837.

Boycott, K. M., Vanstone, M. R., Bulman, D. E., & MacKenzie, A. E. (2013). Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. *Nature Reviews Genetics, 14*(10), 681-691.

Braschi, B., Denny, P., Gray, K., Jones, T., Seal, R., Tweedie, S., . . . Bruford, E. (2019). Genenames.org: The HGNC and VGNC resources in 2019. *Nucleic Acids Research, 47*(D1), D786-D792.

Bray, P. F., Wiser, W. C., Wood, M. C., & Pusey, S. B. (1964). Evidence for a genetic etiology of temporal-central abnormalities in focal epilepsy. *New England Journal of Medicine, 271*(18), 926-933.

Broad Institute. (2020). Germline short variant discovery (SNPs + Indels) – GATK. Retrieved on September 27, 2020 from https://gatk.broadinstitute.org/hc/en-us/articles/360035535932-Germline-short-variant-discovery-SNPs-Indels-

Bromberg, Y. (2013). Chapter 15: Disease gene prioritization. *PLOS Computational Biology, 9*(4), Article#e1002902.

Bu, H., Chen, J., Li, Q., Hou, J., Wei, Y., Yang, X., . . . Kong, B. (2019). BRCA mutation frequency and clinical features of ovarian cancer patients: A report from a Chinese study group. *Journal of Obstetrics and Gynaecology Research, 45*(11), 2267-2274.

Butler, K. M., da Silva, C., Alexander, J. J., Hegde, M., & Escayg, A. (2017). Diagnostic yield from 339 epilepsy patients screened on a clinical gene panel. *Pediatric Neurology, 77*, 61-66.

Cai, J. J., Borenstein, E., & Petrov, D. A. (2010). Broker genes in human disease. *Genome Biology and Evolution, 2*, 815-825.

Camfield, C. S., Striano, P., & Camfield, P. R. (2013). Epidemiology of juvenile myoclonic epilepsy. *Epilepsy & Behavior, 28*(Suppl 1), S15-S17.

Capovilla, G., Kaufman, K. R., Perucca, E., Moshé, S. L., & Arida, R. M. (2016). Epilepsy, seizures, physical exercise, and sports: A report from the ILAE Task Force on Sports and Epilepsy. *Epilepsia, 57*(1), 6-12.

Caprara, A. L. F., Rissardo, J. P., Leite, M. T. B., Silveira, J. O. F., Jauris, P. G. M., Arend, J., . . . Fighera, M. R. (2020). Course and prognosis of adult-onset epilepsy in Brazil: A cohort study. *Epilepsy & Behavior, 105*, Article#106969.

Caputo, D., Iorio, R., Vigevano, F., & Fusco, L. (2018). Febrile infection-related epilepsy syndrome (FIRES) with super-refractory status epilepticus revealing autoimmune encephalitis due to GABAAR antibodies. *European Journal of Paediatric Neurology, 22*(1), 182-185.

Carneiro, T. N. R., Krepischi, A. C. V., Costa, S. S., Tojal da Silva, I., Vianna-Morgante, A. M., Valieris, R., . . . Rosenberg, C. (2018). Utility of trio-based exome sequencing in the elucidation of the genetic basis of isolated syndromic intellectual disability: Illustrative cases. *The Application of Clinical Genetics, 11*, 93-98.

Carrera, P., Di Resta, C., Volonteri, C., Castiglioni, E., Bonfiglio, S., Lazarevic, D., . . . Palatta, S. (2015). Exome sequencing and pathway analysis for identification of genetic variability relevant for bronchopulmonary dysplasia (BPD) in preterm newborns: A pilot study. *Clinica Chimica Acta, 451*(Part A), 39-45.

Carvalho, M. D. C. G., Ximenes, R. A. A., Montarroyos, U. R., da Silva, P. F. S., Andrade-Valença, L. P. A., Eickmann, S. H., . . . Miranda-Filho, D. d. B. (2020). Early epilepsy in children with Zika-related microcephaly in a cohort in Recife, Brazil: Characteristics, electroencephalographic findings, and treatment response. *Epilepsia, 61*(3), 509-518.

Catterall, W. A. (2000). From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. *Neuron, 26*(1), 13-25.

Chahal, C. A. A., Salloum, M. N., Alahdab, F., Gottwald, J. A., Tester, D. J., Anwer, L. A., . . . Somers, V. K. (2020). Systematic review of the genetics of sudden unexpected death in epilepsy: Potential overlap with sudden cardiac death and arrhythmia-related genes. *Journal of the American Heart Association, 9*(1), Article#e012264.

Chatron, N., Møller, R. S., Champaigne, N. L., Schneider, A. L., Kuechler, A., Labalme, A., . . . Lesca, G. (2018). The epilepsy phenotypic spectrum associated with a recurrent CUX2 variant. *Annals of Neurology, 83*(5), 926-934.

Chatterjee, A., Nair, R., Gandeti, R., Puppala, G. K., Chandran, V., Gorthi, S. P., & Radhakrishnan, K. (2020). Socioeconomic consequences of drug-resistant epilepsy in an adult cohort from southern India. *Epilepsy & Behavior, 110*, Article#107173.

Chaves, J., Martins-Ferreira, R., Ferreira, A. M., Brás, S., Carvalho, C., Bettencourt, A., . . . Leal, B. (2020). Immunogenetic protective factors in genetic generalized epilepsy. *Epilepsy Research, 166*, Article#106396.

Chen, C. C., Chen, L. S., Yen, M. F., Chen, H. H., & Liou, H. H. (2012). Geographic variation in the age- and gender-specific prevalence and incidence of epilepsy: Analysis of Taiwanese National Health Insurance-based data. *Epilepsia, 53*(2), 283-290.

Chen, C. J., Liu, D. Z., Yao, W. F., Gu, Y., Huang, F., Hei, Z. Q., & Li, X. (2017). Identification of key genes and pathways associated with neuropathic pain in uninjured dorsal root ganglion by using bioinformatic analysis. *Journal of Pain Research, 10*, 2665-2674.

Chen, J., Bardes, E. E., Aronow, B. J., & Jegga, A. G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Research, 37*(suppl_2), W305-W311.

Chen, J., Zhang, J., Liu, A., Zhang, L., Li, H., Zeng, Q., . . . Zhang, Y. (2020a). CHD2-related epilepsy: Novel mutations and new phenotypes. *Developmental Medicine & Child Neurology, 62*(5), 647-653.

Chen, S., Wang, G., Zheng, X., Ge, S., Dai, Y., Ping, P., . . . Sun, F. (2020b). Whole-exome sequencing of a large Chinese azoospermia and severe oligospermia cohort identifies novel infertility causative variants and genes. *Human Molecular Genetics, 29*(14), 2451-2459.

Chen, Y., Zhou, J., Xu, S., Liu, M., Wang, M., Ma, Y., . . . Zhao, L. (2019). Association between the perturbation of bile acid homeostasis and valproic acid-induced hepatotoxicity. *Biochemical Pharmacology, 170*, Article#113669.

Cheng, J., Van Norstrand David, W., Medeiros-Domingo, A., Valdivia, C., Tan, B. H., Ye, B., . . . Ackerman, M. J. (2009). Alpha1-syntrophin mutations identified in sudden infant death syndrome cause an increase in late cardiac sodium current. *Circulation: Arrhythmia and Electrophysiology, 2*(6), 667-676.

Chiang, C., Layer, R. M., Faust, G. G., Lindberg, M. R., Rose, D. B., Garrison, E. P., . . . Hall, I. M. (2015). SpeedSeq: Ultra-Fast personal genome analysis and interpretation. *Nature Methods, 12*(10), 966-968.

Choudhary, A., Gulati, S., Sagar, R., Sankhyan, N., & Sripada, K. (2018). Childhood epilepsy and ADHD comorbidity in an Indian tertiary medical center outpatient population. *Scientific Reports, 8*(1), Article#2670.

Christiaans, I., Mook, O. R. F., Alders, M., Bikker, H., & Lekanne Dit Deprez, R. H. (2019). Large next-generation sequencing gene panels in genetic heart disease: Challenges in clinical practice. *Netherlands Heart Journal, 27*(6), 299-303.

Claes, L. R., Deprez, L., Suls, A., Baets, J., Smets, K., Van Dyck, T., . . . De Jonghe, P. (2009). The SCN1A variant database: A novel research and diagnostic tool. *Human Mutation, 30*(10), E904-E920.

Coll, M., Striano, P., Ferrer-Costa, C., Campuzano, O., Matés, J., del Olmo, B., . . . Brugada, R. (2017). Targeted next-generation sequencing provides novel clues for associated epilepsy and cardiac conduction disorder/SUDEP. *PLoS One, 12*(12), Article#e0189618.

Commission on Classification and Terminology of the International League Against Epilepsy. (1985). Proposal for classification of epilepsies and epileptic syndromes. *Epilepsia, 26*(3), 268-278.

Cossette, P., Liu, L., Brisebois, K., Dong, H., Lortie, A., Vanasse, M., . . . Rouleau, G. A. (2002). Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. *Nature Genetics, 31*(2), 184-189.

Cvetkovska, E., Kuzmanovski, I., Babunovska, M., Boshkovski, B., Cangovska, T. C., & Trencevska, G. K. (2018). Phenotypic spectrum in families with mesial temporal lobe epilepsy probands. *Seizure - European Journal of Epilepsy, 58*, 13-16.

Dahl-Hansen, E., Koht, J., & Syvertsen, M. (2019). Epilepsy at different ages—Etiologies in a Norwegian population. *Epilepsia Open, 4*(1), 176-181.

Daif, A., Lukas, R. V., Issa, N. P., Javed, A., VanHaerents, S., Reder, A. T., . . . Wu, S. (2018). Antiglutamic acid decarboxylase 65 (GAD65) antibody-associated epilepsy. *Epilepsy & Behavior, 80*, 331-336.

Dashnow, H., Bell, K. M., Stark, Z., Tan, T. Y., White, S. M., & Oshlack, A. (2019). Pooled-parent exome sequencing to prioritise de novo variants in genetic disease. *bioRxiv*, Article#601740.

Davenport, C. B., & Weeks, D. F. (1911). A first study of inheritance of epilepsy. *The Journal of Nervous and Mental Disease, 38*(11), 641-670.

de Nijs, L., Wolkoff, N., Coumans, B., Delgado-Escueta, A. V., Grisar, T., & Lakaye, B. (2012). Mutations of EFHC1, linked to juvenile myoclonic epilepsy, disrupt radial and tangential migrations during brain development. *Human Molecular Genetics, 21*(23), 5106-5117.

Deleo, F., Quintas, R., Pastori, C., Pappalardo, I., Didato, G., Di Giacomo, R., . . . Villani, F. (2020). Quality of life, psychiatric symptoms, and stigma perception in three groups of persons with epilepsy. *Epilepsy & Behavior, 110*, Article#107170.

Deng, W., Luo, F., Li, B. M., & Mei, L. (2019). NRG1-ErbB4 signaling promotes functional recovery in a murine model of traumatic brain injury via regulation of GABA release. *Experimental Brain Research, 237*(12), 3351-3362.

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology, 4*(5), Article#P3.

Department of Statistics Malaysia. (2020). Current Population Estimates, Malaysia, 2020. Retrieved on November 1, 2020 from https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=OVByWjg5YkQ3MWFZRTN5bDJiaEVhZz09

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., . . . Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics, 43*(5), 491-498.

Devinsky, O., Spruill, T., Thurman, D., & Friedman, D. (2016). Recognizing and preventing epilepsy-related mortality: A call for action. *Neurology, 86*(8), 779-786.

Dhiman, V. (2017). Molecular genetics of epilepsy: A clinician's perspective. *Annals of Indian Academy of Neurology, 20*(2), 96-102.

DiFrancesco, J. C., Barbuti, A., Milanesi, R., Coco, S., Bucchi, A., Bottelli, G., . . . DiFrancesco, D. (2011). Recessive loss-of-function mutation in the pacemaker HCN2 channel causing increased neuronal excitability in a patient with idiopathic generalized epilepsy. *The Journal of Neuroscience, 31*(48), 17327-17337.

Dolled-Filhart, M. P., Lee, M., Ou-Yang, C. W., Haraksingh, R. R., & Lin, J. C. H. (2013). Computational and bioinformatics frameworks for next-generation whole exome and genome sequencing. *The Scientific World Journal, 2013*, Article#730210.

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., & Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human Molecular Genetics, 24*(8), 2125-2137.

Doorenweerd, N., Mahfouz, A., van Putten, M., Kaliyaperumal, R., t' Hoen, P. A. C., Hendriksen, J. G. M., . . . Lelieveldt, B. P. F. (2017). Timing and localization of human dystrophin isoform expression provide insights into the cognitive phenotype of Duchenne muscular dystrophy. *Scientific Reports, 7*(1), Article#12575.

Du, X., Gao, X., Liu, X., Shen, L., Wang, K., Fan, Y., . . . Li, F. (2018). Genetic diagnostic evaluation of trio-based whole exome sequencing among children with diagnosed or suspected autism spectrum disorder. *Frontiers in Genetics, 9*, Article#594.

Dubey, D., Alqallaf, A., Hays, R., Freeman, M., Chen, K., Ding, K., . . . Vernino, S. (2017). Neurological autoantibody prevalence in epilepsy of unknown etiology. *JAMA Neurology, 74*(4), 397-402.

Dunn, P., Albury, C. L., Maksemous, N., Benton, M. C., Sutherland, H. G., Smith, R. A., . . . Griffiths, L. R. (2018). Next generation sequencing methods for diagnosis of epilepsy syndromes. *Frontiers in Genetics, 9*, Article#20.

Emmilia H. Tan, Abdul Aziz M. Yusoff, Jafri M. Abdullah, & Salmi A. Razak. (2012). Generalized epilepsy with febrile seizure plus (GEFS+) spectrum: Novel de novo mutation of SCN1A detected in a Malaysian patient. *Journal of Pediatric Neurosciences, 7*(2), 123-125.

Emmilia Husni Tan, Salmi Abdul Razak, Jafri Malin Abdullah, & Abdul Aziz Mohamed Yusoff. (2012). De-novo mutations and genetic variation in the SCN1A gene in Malaysian patients with generalized epilepsy with febrile seizures plus (GEFS+). *Epilepsy Research, 102*(3), 210-215.

Epi25 Collaborative. (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.

Escayg, A., De Waard, M., Lee, D. D., Bichet, D., Wolf, P., Mayer, T., . . . Meisler, M. H. (2000). Coding and noncoding variation of the human calcium-channel beta4-subunit gene CACNB4 in patients with idiopathic generalized epilepsy and episodic ataxia. *The American Journal of Human Genetics, 66*(5), 1531-1539.

Escayg, A., & Goldin, A. L. (2010). Sodium channel SCN1A and epilepsy: Mutations and mechanisms. *Epilepsia, 51*(9), 1650-1658.

Escayg, A., Jones, J. M., Kearney, J. A., Hitchcock, P. F., & Meisler, M. H. (1998). Calcium channel beta 4 (CACNB4): Human ortholog of the mouse epilepsy genelethargic. *Genomics, 50*(1), 14-22.

Espinosa-Jovel, C., Toledano, R., Aledo-Serrano, Á., García-Morales, I., & Gil-Nagel, A. (2018). Epidemiological profile of epilepsy in low income populations. *Seizure - European Journal of Epilepsy, 56*, 67-72.

Espinosa Jovel, C. A., Pardo, C. M., Moreno, C. M., Vergara, J., Hedmont, D., & Sobrino Mejía, F. E. (2016). Demographic and social profile of epilepsy in a vulnerable low-income population in Bogota, Colombia. *Neurologia, 31*(8), 528-534.

European Bioinformatics Institute. (2013). PDBsum documentation PROCHECK - Main Ramachandran plot. Retrieved on July 9, 2021 from http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=n/a&template=doc_procheck01.html

Falardeau, F., Camurri, M. V., & Campeau, P. M. (2017). Genomic approaches to diagnose rare bone disorders. *Bone, 102*, 5-14.

Falco-Walter, J. J., Scheffer, I. E., & Fisher, R. S. (2018). The new definition and classification of seizures and epilepsy. *Epilepsy Research, 139*, 73-79.

Fang, Z., Xu, S., Wang, Y., Sun, L., Feng, Y., Guo, Y., . . . Jiang, W. (2017). Pathogenicity analysis of novel variations in Chinese Han patients with polycystic kidney disease. *Gene, 626*, 433-441.

Fang, Z. X., Hong, S. Q., Li, T. S., Wang, J., Xie, L. L., Han, W., & Jiang, L. (2019). Genetic and phenotypic characteristics of SCN1A-related epilepsy in Chinese children. *Neuroreport, 30*(9), 671-680.

Farghaly, W. M., Abd Elhamed, M. A., Hassan, E. M., Soliman, W. T., Yhia, M. A., & Hamdy, N. A. (2018). Prevalence of childhood and adolescence epilepsy in Upper Egypt (desert areas). *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 54*(1), Article#34.

Feigin, V. L., Nichols, E., Alam, T., Bannick, M. S., Beghi, E., Blake, N., . . . Vos, T. (2019). Global, regional, and national burden of neurological disorders, 1990-2016: A systematic analysis for the global burden of disease study 2016. *The Lancet Neurology, 18*(5), 459-480.

Feldman, I., Rzhetsky, A., & Vitkup, D. (2008). Network properties of genes harboring inherited disease mutations. *Proceedings of the National Academy of Sciences, 105*(11), 4323-4328.

Feng, C., Bai, M., Zhang, H., Zeng, A., & Zhang, W. (2019). Prioritization and comprehensive analysis of genes associated with melanoma. *Oncology Letters, 18*(1), 127-136.

Fiest, K. M., Sauro, K. M., Wiebe, S., Patten, S. B., Kwon, C.-S., Dykeman, J., . . . Jetté, N. (2017). Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies. *Neurology, 88*(3), 296-303.

Fisher, R. S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J. H., Elger, C. E., . . . Wiebe, S. (2014). ILAE official report: A practical clinical definition of epilepsy. *Epilepsia, 55*(4), 475-482.

Fisher, R. S., Cross, J. H., French, J. A., Higurashi, N., Hirsch, E., Jansen, F. E., . . . Zuberi, S. M. (2017). Operational classification of seizure types by the International League Against Epilepsy: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia, 58*(4), 522-530.

Fong, S. L., Lim, K. S., Tan, L., Nabilah Hanis Zainuddin, Ho, J. H., Chia, Z. J., . . . Tan, C. T. (2021). Prevalence study of epilepsy in Malaysia. *Epilepsy Research, 170*, Article#106551.

French, J. A. (2014). ILAE classification redux: Ready for prime time?. *Epilepsy Currents, 14*(2), 84-85.

Furlong, L. I. (2013). Human diseases through the lens of network biology. *Trends in Genetics, 29*(3), 150-159.

Gaudioso, C., Carlier, E., Youssouf, F., Clare, J. J., Debanne, D., & Alcaraz, G. (2011). Calmodulin and calcium differentially regulate the neuronal Nav1.1 voltage-dependent sodium channel. *Biochemical and Biophysical Research Communications, 411*(2), 329-334.

Gene Ontology Consortium. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Research, 32*(suppl_1), D258-D261.

GeneDx. (2020). Neurology Genetics. Retrieved on December 31, 2020 from https://www.genedx.com/test-catalog/medical-specialty/neurology/

Gesche, J., Christensen, J., Hjalgrim, H., Rubboli, G., & Beier, C. P. (2020). Epidemiology and outcome of idiopathic generalized epilepsy in adults. *European Journal of Neurology, 27*(4), 676-684.

Ghatge, M., Nair, J., Sharma, A., & Vangala, R. K. (2018). Integrative gene ontology and network analysis of coronary artery disease associated genes suggests potential role of ErbB pathway gene EGFR. *Molecular Medicine Reports, 17*(3), 4253-4264.

Gilsoul, M., Grisar, T., Delgado-Escueta, A. V., de Nijs, L., & Lakaye, B. (2019). Subtle brain developmental abnormalities in the pathogenesis of juvenile myoclonic epilepsy. *Frontiers in Cellular Neuroscience, 13*, Article#433.

Goh, G., & Choi, M. (2012). Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics & Informatics, 10*(4), 214-219.

Gomez-Ibañez, A., McLachlan, R. S., Mirsattari, S. M., Diosy, D. C., & Burneo, J. G. (2017). Prognostic factors in patients with refractory idiopathic generalized epilepsy. *Epilepsy Research, 130*, 69-73.

Gu, X., Li, C., Chen, Y., Wei, Q., Cao, B., Ou, R., . . . Shang, H. F. (2019). Mutation screening of the KIF5A gene in Chinese patients with amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry, 90*(2), 245-246.

Guekht, A., Zharkinbekova, N., Shpak, A., & Hauser, W. A. (2017). Epilepsy and treatment gap in urban and rural areas of the Southern Kazakhstan in adults. *Epilepsy & Behavior, 67*, 98-104.

Guo, Y., Yu, L., He, B., Li, S., Zhu, Q., & Sun, H. (2018). Aetiological features of elderly patients with newly diagnosed symptomatic epilepsy in Western China. *BioMed Research International, 2018*, Article#4104691.

Guzzi, P. H., Mina, M., Guerra, C., & Cannataro, M. (2012). Semantic similarity analysis of protein data: Assessment with biological features and issues. *Briefings in Bioinformatics, 13*(5), 569-585.

Haerian, B. S., Hidayati Mohd Sha'ari, Fong, C. Y., Tan, H. J., Wong, S. W., Ong, L. C., . . . Zahurin Mohamed. (2015a). Contribution of TIMP4 rs3755724 polymorphism to susceptibility to focal epilepsy in Malaysian Chinese. *Journal of Neuroimmunology, 278*, 137-143.

Haerian, B. S., Hidayati Mohd Sha'ari, Tan, H. J., Fong, C. Y., Wong, S. W., Ong, L. C., . . . Zahurin Mohamed. (2015b). RORA gene rs12912233 and rs880626 polymorphisms and their interaction with SCN1A rs3812718 in the risk of epilepsy: A case-control study in Malaysia. *Genomics, 105*(4), 229-236.

Haerian, B. S., Lim, K. S., Tan, H. J., Wong, C. P., Wong, S. W., Tan, C. T., . . . Zahurin Mohamed. (2011). Lack of association between synapsin II (SYN2) gene polymorphism and susceptibility epilepsy: A case-control study and meta-analysis. *Synapse, 65*(10), 1073-1079.

Hamanaka, K., Imagawa, E., Koshimizu, E., Miyatake, S., Tohyama, J., Yamagata, T., . . . Matsumoto, N. (2020). De novo truncating variants in the last exon of SEMA6B cause progressive myoclonic epilepsy. *The American Journal of Human Genetics, 106*(4), 549-558.

Han, J. Y., & Lee, I. G. (2020). Genetic tests by next-generation sequencing in children with developmental delay and/or intellectual disability. *Clinical and Experimental Pediatrics, 63*(6), 195-202.

Hara, H. S., Gupta, A., Singh, M., Raj, R., Singh, H., Pawar, G., . . . Singh, J. (2015). Epilepsy in Punjab (India): A population-based epidemiologic study. *Neuroepidemiology, 45*(4), 273-281.

Hehir-Kwa, J. Y., Pfundt, R., & Veltman, J. A. (2015). Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Review of Molecular Diagnostics, 15*(8), 1023-1032.

Heijbel, J., Blom, S., & Rasmuson, M. (1975). Benign epilepsy of childhood with centrotemporal EEG foci: A genetic study. *Epilepsia, 16*(2), 285-293.

Heinzen, E. L. (2020). Somatic variants in epilepsy - advancing gene discovery and disease mechanisms. *Current Opinion in Genetics & Development, 65*, 1-7.

Heinzen, Erin L., Depondt, C., Cavalleri, Gianpiero L., Ruzzo, Elizabeth K., Walley, Nicole M., Need, Anna C., . . . Goldstein, David B. (2012). Exome sequencing followed by large-scale genotyping fails to identify single rare variants of large effect in idiopathic generalized epilepsy. *The American Journal of Human Genetics, 91*(2), 293-302.

Helbig, I., Heinzen, E. L., Mefford, H. C., & the ILAE Genetics Commission. (2016a). Primer Part 1-The building blocks of epilepsy genetics. *Epilepsia, 57*(6), 861-868.

Helbig, K. L., Farwell Hagman, K. D., Shinde, D. N., Mroske, C., Powis, Z., Li, S., . . . Helbig, I. (2016b). Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genetics in Medicine, 18*(9), 898-905.

Hernandez, C. C., Gurba, K. N., Hu, N., & Macdonald, R. L. (2011). The GABRA6 mutation, R46W, associated with childhood absence epilepsy, alters 6beta22 and 6beta2 GABA(A) receptor channel gating and expression. *The Journal of Physiology, 589*(23), 5857-5878.

Hernandez, C. C., & Macdonald, R. L. (2019). A structural look at GABAA receptor mutations linked to epilepsy syndromes. *Brain Research, 1714*, 234-247.

Heron, S. E., Khosravani, H., Varela, D., Bladen, C., Williams, T. C., Newman, M. R., . . . Zamponi, G. W. (2007). Extended spectrum of idiopathic generalized epilepsies associated with CACNA1H functional variants. *Annals of Neurology, 62*(6), 560-568.

Hidayati Mohd Sha'ari, Haerian, B. S., Baum, L., Tan, H. J., Mohd Hanip Rafia, Kwan, P., . . . Zahurin Mohamed. (2016). Association of BDNF polymorphisms with the risk of epilepsy: A multicenter study. *Molecular Neurobiology, 53*(5), 2869-2877.

Hirose, S., Okada, M., Kaneko, S., & Mitsudome, A. (2002a). Molecular genetics of human familial epilepsy syndromes. *Epilepsia, 43*(s9), 21-25.

Hirose, S., Okada, M., Yamakawa, K., Sugawara, T., Fukuma, G., Ito, M., . . . Mitsudome, A. (2002b). Genetic abnormalities underlying familial epilepsy syndromes. *Brain and Development, 24*(4), 211-222.

Hoelz, H., Herdl, C., Gerstl, L., Tacke, M., Vill, K., von Stuelpnagel, C., . . . Borggraefe, I. (2020). Impact on clinical decision making of next-generation sequencing in pediatric epilepsy in a tertiary epilepsy referral center. *Clinical EEG and Neuroscience, 51*(1), 61-69.

Horner, D. S., Pavesi, G., Castrignanò, T., De Meo, P. D. O., Liuni, S., Sammeth, M., . . . Pesole, G. (2010). Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. *Briefings in Bioinformatics, 11*(2), 181-197.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols, 4*(1), 44-57.

Husari, K. S., & Dubey, D. (2019). Autoimmune Epilepsy. *Neurotherapeutics, 16*(3), 685-702.

Iffland, P. H., II, & Crino, P. B. (2019). The role of somatic mutational events in the pathogenesis of epilepsy. *Current Opinion in Neurology, 32*(2), 191-197.

Iglesias, V., Paladin, L., Juan-Blanco, T., Pallarès, I., Aloy, P., Tosatto, S. C. E., & Ventura, S. (2019). In silico characterization of human prion-like proteins: Beyond neurological diseases. *Frontiers in Physiology, 10*, Article#314.

Inoue, Y., Ito, M., Kurihara, M., & Morimoto, K. (2004). Epilepsy and driving in Japan. *Epilepsia, 45*(12), 1630-1635.

International League Against Epilepsy. (2020). EpilepsyDiagnosis.org. Retrieved on July 20, 2020 from https://www.epilepsydiagnosis.org/index.html

Invitae Corporation. (2020). Invitae Epilepsy Panel. Retrieved on December 31, 2020 from https://www.invitae.com/en/physician/tests/03401/

Isakov, O., Dotan, I., & Ben-Shachar, S. (2017). Machine learning-based gene prioritization identifies novel candidate risk genes for inflammatory bowel disease. *Inflammatory Bowel Diseases, 23*(9), 1516-1523.

Isik, E., Yilmaz, S., Atik, T., Aktan, G., Onay, H., Gokben, S., & Ozkinay, F. (2020). The utility of whole exome sequencing for identification of the molecular etiology in autosomal recessive developmental and epileptic encephalopathies. *Neurological Sciences, 41*(12), 3729-3739.

Ito, M., Ohmori, I., Nakahori, T., Ouchida, M., & Ohtsuka, Y. (2005). Mutation screen of GABRA1, GABRB2 and GABRG2 genes in Japanese patients with absence seizures. *Neuroscience Letters, 383*(3), 220-224.

Ivanova, N., Peycheva, V., Kamenarova, K., Kancheva, D., Tsekova, I., Aleksandrova, I., . . . Jordanova, A. (2018). Three novel SLC2A1 mutations in Bulgarian patients with different forms of genetic generalized epilepsy reflecting the clinical and genetic diversity of GLUT1-deficiency syndrome. *Seizure - European Journal of Epilepsy, 54*, 41-44.

Jackson, M., Marks, L., May, G. H. W., & Wilson, J. B. (2018). The genetic basis of disease. *Essays in Biochemistry, 62*(5), 643-723.

Jallon, P., & Latour, P. (2005). Epidemiology of idiopathic generalized epilepsies. *Epilepsia, 46*(s9), 10-14.

Jennum, P., Debes, N. M. M., Ibsen, R., & Kjellberg, J. (2021). Long-term employment, education, and healthcare costs of childhood and adolescent onset of epilepsy. *Epilepsy & Behavior, 114*, Article#107256.

Jenssen, S., Sperling, M. R., Tracy, J. I., Nei, M., Joyce, L., David, G., & O'Connor, M. (2006). Corpus callosotomy in refractory idiopathic generalized epilepsy. *Seizure - European Journal of Epilepsy, 15*(8), 621-629.

Jette, N., Reid, A. Y., & Wiebe, S. (2014). Surgical management of epilepsy. *Canadian Medical Association Journal, 186*(13), 997-1004.

Jia, P., Ewers, J. M., & Zhao, Z. (2011). Prioritization of epilepsy associated candidate genes by convergent analysis. *PLoS One, 6*(2), Article#e17162.

Jin, L., Zuo, X. Y., Su, W. Y., Zhao, X. L., Yuan, M. Q., Han, L. Z., . . . Rao, S. Q. (2014). Pathway-based analysis tools for complex diseases: A review. *Genomics, Proteomics & Bioinformatics, 12*(5), 210-220.

Jin, Y., Zhao, C., Chen, L., Liu, X., Pan, S., Ju, D., . . . Wei, B. (2016). Identification of novel gene and pathway targets for human epilepsy treatment. *Biological Research, 49*, Article#3.

Jin, Z. B., Li, Z., Liu, Z., Jiang, Y., Cai, X. B., & Wu, J. (2018). Identification of de novo germline mutations and causal genes for sporadic diseases using trio-based whole-exome/genome sequencing. *Biological Reviews, 93*(2), 1014-1031.

Jingami, N., Matsumoto, R., Ito, H., Ishii, A., Ihara, Y., Hirose, S., . . . Takahashi, R. (2014). A novel SCN1A mutation in a cytoplasmic loop in intractable juvenile myoclonic epilepsy without febrile seizures. *Epileptic Disorders, 16*(2), 227-231.

Johannesen, K., Marini, C., Pfeffer, S., Møller, R. S., Dorn, T., Niturad, C. E., . . . Maljevic, S. (2016). Phenotypic spectrum of GABRA1: From generalized epilepsies to severe epileptic encephalopathies. *Neurology, 87*(11), 1140-1151.

Kambli, L., Bhatt, L. K., Oza, M., & Prabhavalkar, K. (2017). Novel therapeutic targets for epilepsy intervention. *Seizure - European Journal of Epilepsy, 51*, 27-34.

Kananura, C., Haug, K., Sander, T., Runge, U., Gu, W., Hallmann, K., . . . Steinlein, O. K. (2002). A splice-site mutation in GABRG2 associated with childhood absence epilepsy and febrile convulsions. *Archives of Neurology, 59*(7), 1137-1141.

Kang, K. W., Kim, W., Cho, Y. W., Lee, S. K., Jung, K. Y., Shin, W., . . . Kim, M. K. (2019). Genetic characteristics of non-familial epilepsy. *PeerJ, 7*, Article#e8278.

Kapoor, A., & Kinsella, L. (2018). A rare case of isolated myoclonus in an elderly male without a history of epilepsy. *Clinical Neurophysiology Practice, 3*, 96-98.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., . . . MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature, 581*(7809), 434-443.

Kaur, J., Paul, B. S., Goel, P., & Singh, G. (2019). Educational achievement, employment, marriage, and driving in adults with childhood-onset epilepsy. *Epilepsy & Behavior, 97*, 149-153.

Keezer, M. R., Sisodiya, S. M., & Sander, J. W. (2016). Comorbidities of epilepsy: Current concepts and future perspectives. *The Lancet Neurology, 15*(1), 106-115.

Khair, A. M., & Elmagrabi, D. (2015). Febrile seizures and febrile seizure syndromes: An updated overview of old and current knowledge. *Neurology Research International, 2015*, Article#849341.

Kim, M., Lee, E., Zang, D. Y., Kim, H. J., Kim, H. Y., Han, B., . . . Lee, Y. K. (2020). Novel genes exhibiting DNA methylation alterations in Korean patients with chronic lymphocytic leukaemia: A methyl-CpG-binding domain sequencing study. *Scientific Reports, 10*(1), Article#1085.

Kim, S., Cho, C. S., Han, K., & Lee, J. (2016). Structural variation of Alu element and human disease. *Genomics & Informatics, 14*(3), 70-77.

Kimball, O. P., & Hersh, A. H. (1955). The genetics of epilepsy. *Acta Geneticae Medicae Et Gemellologiae, 4*(2), 131-142.

Kind, C. J., Newton, C. R. J. C., & Kariuki, S. M. (2017). Prevalence, risk factors, and neurobehavioral comorbidities of epilepsy in Kenyan children. *Epilepsia Open, 2*(4), 388-399.

Kluckova, D., Kolnikova, M., Lacinova, L., Jurkovicova-Tarabova, B., Foltan, T., Demko, V., . . . Soltysova, A. (2020). A study among the genotype, functional alternations, and phenotype of 9 SCN1A mutations in epilepsy patients. *Scientific Reports, 10*(1), Article#10288.

Kobayashi, K., Endoh, F., Ohmori, I., & Akiyama, T. (2020). Action of antiepileptic drugs on neurons. *Brain and Development, 42*(1), 2-5.

Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., & Mardis, E. R. (2013). The next-generation sequencing revolution and its impact on genomics. *Cell, 155*(1), 27-38.

Koeleman, B. P. C. (2018). What do genetic studies tell us about the heritable basis of common epilepsy? Polygenic or complex epilepsy?. *Neuroscience Letters, 667*, 10-16.

Koh, H. Y., & Lee, J. H. (2018). Brain somatic mutations in epileptic disorders. *Molecules and Cells, 41*(10), 881-888.

Koh, S. (2018). Role of neuroinflammation in evolution of childhood epilepsy. *Journal of Child Neurology, 33*(1), 64-72.

Korotkov, A., Mills, J. D., Gorter, J. A., van Vliet, E. A., & Aronica, E. (2017). Systematic review and meta-analysis of differentially expressed miRNAs in experimental and human temporal lobe epilepsy. *Scientific Reports, 7*(1), Article#11592.

Laboratory Corporation of America. (2020). Epilepsy Genetic Testing. Retrieved on December 31, 2020 from https://mnglabs.com/programs-services/conditions-diagnosed/epilepsy

Lachance-Touchette, P., Brown, P., Meloche, C., Kinirons, P., Lapointe, L., Lacasse, H., . . . Cossette, P. (2011). Novel alpha1 and gamma2 GABAA receptor subunit mutations in families with idiopathic generalized epilepsy. *European Journal of Neuroscience, 34*(2), 237-249.

Lai, S. T., Tan, W. Y., Wo, M. C. M., Lim, K. S., Sherrini Bazir Ahmad, & Tan, C. T. (2019). Burden in caregivers of adults with epilepsy in Asian families. *Seizure - European Journal of Epilepsy, 71*, 132-139.

Landoulsi, Z., Laatar, F., Noé, E., Mrabet, S., Ben Djebara, M., Achaz, G., . . . Leguern, E. (2018). Clinical and genetic study of Tunisian families with genetic generalized epilepsy: Contribution of CACNA1H and MAST4 genes. *Neurogenetics, 19*(3), 165-178.

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein Science, 27*(1), 129-134.

Leal, B., Chaves, J., Carvalho, C., Bettencourt, A., Brito, C., Boleixa, D., . . . da Silva, A. M. (2018). Immunogenetic predisposing factors for mesial temporal lobe epilepsy with hippocampal sclerosis. *International Journal of Neuroscience, 128*(4), 305-310.

Lee, C. G., Lee, J., & Lee, M. (2018). Multi-gene panel testing in Korean patients with common genetic generalized epilepsy syndromes. *PLoS One, 13*(6), Article#e0199321.

Lee, Y. J., Yum, M. S., Kim, M. J., Shim, W. H., Yoon, H. M., Yoo, I. H., . . . Ko, T. S. (2017). Large-scale structural alteration of brain in epileptic children with SCN1A mutation. *NeuroImage: Clinical, 15*, 594-600.

Lenaerts, E., Mandro, M., Mukendi, D., Suykerbuyk, P., Dolo, H., Wonya'Rossi, D., . . . Colebunders, R. (2018). High prevalence of epilepsy in onchocerciasis endemic health areas in Democratic Republic of the Congo. *Infectious Diseases of Poverty, 7*(1), Article#68.

Lennox, W. G. (1947). The genetics of epilepsy. *American Journal of Psychiatry, 103*(4), 457-462.

Li, D., Kim, W., Wang, L., Yoon, K. A., Park, B., Park, C., . . . Won, S. (2019). Comparison of INDEL calling tools with simulation data and real short-read data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16*(5), 1635-1644.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760.

Li, K. X., Lu, Y. M., Xu, Z. H., Zhang, J., Zhu, J. M., Zhang, J. M., . . . Li, X. M. (2011). Neuregulin 1 regulates excitability of fast-spiking neurons through Kv1.1 and acts in epilepsy. *Nature Neuroscience, 15*(2), 267-273.

Li, Y., & Agarwal, P. (2009). A pathway-based view of human diseases and disease relationships. *PLoS One, 4*(2), Article#e4346.

Liang, Z., Zhu, Y., Liu, X., & Hu, G. (2020). Chapter three - Role of protein-protein interactions in allosteric drug design for DNA methyltransferases. In R. Donev (Ed.), *Advances in Protein Chemistry and Structural Biology* (Vol. 121, pp. 49-84): Academic Press.

Lim, K. S., Ng, C. C., Chan, C. K., Foo, W. S., Low, J. S. Y., & Tan, C. T. (2017). Ethnic variation of genetic (idiopathic) generalized epilepsy in Malaysia. *Seizure - European Journal of Epilepsy, 45*, 24-27.

Lim, K. S., Wo, S. W., Wong, M. H., & Tan, C. T. (2013). Impact of epilepsy on employment in Malaysia. *Epilepsy & Behavior, 27*(1), 130-134.

Lin, W., Feng, M., Chen, G., Zhou, Z., Li, J., & Ye, Y. (2017). Characterization of the microRNA profile in early-stage cervical squamous cell carcinoma by next-generation sequencing. *Oncology Reports, 37*(3), 1477-1486.

Lionel, A. C., Costain, G., Monfared, N., Walker, S., Reuter, M. S., Hosseini, S. M., . . . Marshall, C. R. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genetics in Medicine, 20*(4), 435-443.

Liu, L., Wei, J., & Ruan, J. (2017). Pathway enrichment analysis with networks. *Genes, 8*(10), Article#246.

Liu, Y., Fan, P., Zhang, S., Wang, Y., & Liu, D. (2019). Prioritization and comprehensive analysis of genes related to major depressive disorder. *Molecular Genetics & Genomic Medicine, 7*(6), Article#e659.

Liu, Y. T., Lee, Y. C., & Soong, B. W. (2015). What we have learned from the next-generation sequencing: Contributions to the genetic diagnoses and understanding of pathomechanisms of neurodegenerative diseases. *Journal of Neurogenetics, 29*(2-3), 103-112.

Lohmann, K., & Klein, C. (2014). Next generation sequencing and the future of genetic diagnosis. *Neurotherapeutics, 11*(4), 699-707.

Lung, M. S., Mitchell, C. A., Doyle, M. A., Lynch, A. C., Gorringe, K. L., Bowtell, D. D. L., . . . Trainer, A. H. (2020). Germline whole exome sequencing of a family with appendiceal mucinous tumours presenting with pseudomyxoma peritonei. *BMC Cancer, 20*(1), Article#369.

Ma, M., Wang, C., Glicksberg, B. S., Schadt, E. E., Li, S. D., & Chen, R. (2016). Identify cancer driver genes through shared Mendelian disease pathogenic variants and cancer somatic mutations. In R. B. Altman, A. K. Dunker, L. Hunter, M. D. Ritchie, T. A. Murray, & T. E. Klein (Eds.), *Pacific Symposium on Biocomputing 2017* (pp. 473-484): World Scientific.

Magi, A., Benelli, M., Gozzini, A., Girolami, F., Torricelli, F., & Brandi, M. L. (2010). Bioinformatics for next generation sequencing data. *Genes, 1*(2), 294-307.

Mahoney, K., Moore, S. J., Buckley, D., Alam, M., Parfrey, P., Penney, S., . . . Young, T. L. (2009). Variable neurologic phenotype in a GEFS+ family with a novel mutation in SCN1A. *Seizure - European Journal of Epilepsy, 18*(7), 492-497.

Marchuk, D. S., Crooks, K., Strande, N., Kaiser-Rogers, K., Milko, L. V., Brandt, A., . . . Berg, J. S. (2018). Increasing the diagnostic yield of exome sequencing by copy number variant analysis. *PLoS One, 13*(12), Article#e0209185.

Marini, C., Mei, D., Helen Cross, J., & Guerrini, R. (2006). Mosaic SCN1A mutation in familial severe myoclonic epilepsy of infancy. *Epilepsia, 47*(10), 1737-1740.

Marini, C., Porro, A., Rastetter, A., Dalle, C., Rivolta, I., Bauer, D., . . . Depienne, C. (2018). HCN1 mutation spectrum: From neonatal epileptic encephalopathy to benign generalized epilepsy and beyond. *Brain, 141*(11), 3160-3178.

Masiulis, S., Desai, R., Uchański, T., Serna Martin, I., Laverty, D., Karia, D., . . . Aricescu, A. R. (2019). GABAA receptor signalling mechanisms revealed by structural pharmacology. *Nature, 565*(7740), 454-459.

Massey, C. A., Sowers, L. P., Dlouhy, B. J., & Richerson, G. B. (2014). Mechanisms of sudden unexpected death in epilepsy: The pathway to prevention. *Nature Reviews Neurology, 10*(5), 271-282.

Mayo Foundation for Medical Education and Research. (2020). ESPAN - Overview Epilepsy_Seizure Genetic Panels by Next-Generation Sequencing (NGS), Varies. Retrieved on December 31, 2020 from https://www.mayocliniclabs.com/test-catalog/Overview/603346

Mbizvo, G. K., Bennett, K., Simpson, C. R., Duncan, S. E., & Chin, R. F. M. (2019). Epilepsy-related and other causes of mortality in people with epilepsy: A systematic review of systematic reviews. *Epilepsy Research, 157*, Article#106192.

Miao, P., Feng, J., Guo, Y., Wang, J., Xu, X., Wang, Y., . . . Cheng, H. (2018). Genotype and phenotype analysis using an epilepsy-associated gene panel in Chinese pediatric epilepsy patients. *Clinical Genetics, 94*(6), 512-520.

Miki, Y., Iwabuchi, E., Ono, K., Sasano, H., & Ito, K. (2018). Exploring protein-protein interaction in the study of hormone-dependent cancers. *International Journal of Molecular Sciences, 19*(10), Article#3173.

Miryala, S. K., Anbarasu, A., & Ramaiah, S. (2018). Discerning molecular interactions: A comprehensive review on biomolecular interaction databases and network analysis tools. *Gene, 642*, 84-94.

Mital, A. K., Choudhary, P., & Jain, R. B. (2020). Prevalence and risk factors for neurocysticercosis in children with a first-onset seizure in rural North India. *Paediatrics and International Child Health, 40*(3), 158-165.

Mitchell, R. M., Janssen, M. J., Karavanova, I., Vullhorst, D., Furth, K., Makusky, A., . . . Buonanno, A. (2013). ErbB4 reduces synaptic GABAA currents independent of its receptor tyrosine kinase activity. *Proceedings of the National Academy of Sciences, 110*(48), 19603-19608.

Mmbando, B. P., Suykerbuyk, P., Mnacho, M., Kakorozya, A., Matuja, W., Hendy, A., . . . Colebunders, R. (2018). High prevalence of epilepsy in two rural onchocerciasis endemic villages in the Mahenge area, Tanzania, after 20 years of community directed treatment with ivermectin. *Infectious Diseases of Poverty, 7*(1), Article#64.

Moavero, R., Santarone, M. E., Galasso, C., & Curatolo, P. (2017). Cognitive and behavioral effects of new antiepileptic drugs in pediatric epilepsy. *Brain and Development, 39*(6), 464-469.

Mohammadi, M. R., Ghanizadeh, A., Davidian, H., Mohammadi, M., & Norouzian, M. (2006). Prevalence of epilepsy and comorbidity of psychiatric disorders in Iran. *Seizure - European Journal of Epilepsy, 15*(7), 476-482.

Mond, C., Duke, T., & Vince, J. (2019). Epilepsy in Papua New Guinea: A longitudinal cohort study. *Archives of Disease in Childhood, 104*(10), 941-946.

Moshé, S. L. (2011). In support of the ILAE Commission classification proposal. *Epilepsia, 52*(6), 1200-1201.

Mueller, S. C., Backes, C., Haas, J., the INHERITANCE Study Group, Katus, H. A., Meder, B., . . . Keller, A. (2015). Pathogenicity prediction of non-synonymous single nucleotide variants in dilated cardiomyopathy. *Briefings in Bioinformatics, 16*(5), 769-779.

Mula, M., & Sander, J. W. (2016). Psychosocial aspects of epilepsy: A wider approach. *BJPsych Open, 2*(4), 270-274.

Mullen, S. A., Berkovic, S. F., & the ILAE Genetics Commission. (2018). Genetic generalized epilepsies. *Epilepsia, 59*(6), 1148-1153.

Mulley, J. C., Scheffer, I. E., Petrou, S., & Berkovic, S. F. (2003). Channelopathies as a genetic cause of epilepsy. *Current Opinion in Neurology, 16*(2), 171-176.

Naimo, G. D., Guarnaccia, M., Sprovieri, T., Ungaro, C., Conforti, F. L., Ando, S., & Cavallaro, S. (2019). A systems biology approach for personalized medicine in refractory epilepsy. *International Journal of Molecular Sciences, 20*(15), Article#3717.

Ngoc, N. T., Duong, N. T., Quynh, D. H., Ton, N. D., Duc, H. H., Huong, L. T. M., . . . Hai, N. V. (2020). Identification of novel missense mutations associated with non-syndromic syndactyly in two Vietnamese trios by whole exome sequencing. *Clinica Chimica Acta, 506*, 16-21.

Nykamp, K., Anderson, M., Powers, M., Garcia, J., Herrera, B., Ho, Y. Y., . . . Topper, S. (2017). Sherloc: A comprehensive refinement of the ACMG-AMP variant classification criteria. *Genetics in Medicine, 19*(10), 1105-1117.

Odumpatta, R., & Mohanapriya, A. (2020). Next generation sequencing exome data analysis aids in the discovery of SNP and INDEL patterns in Parkinson's disease. *Genomics, 112*(5), 3722-3728.

Okada, M., & Corfas, G. (2004). Neuregulin1 downregulates postsynaptic GABAA receptors at the hippocampal inhibitory synapse. *Hippocampus, 14*(3), 337-344.

Osman, C., Foulds, N., Hunt, D., Jade Edwards, C., & Prevett, M. (2020). Diagnosis of pyridoxine-dependent epilepsy in an adult presenting with recurrent status epilepticus. *Epilepsia, 61*(1), e1-e6.

Ostrander, B. E. P., Butterfield, R. J., Pedersen, B. S., Farrell, A. J., Layer, R. M., Ward, A., . . . Quinlan, A. R. (2018). Whole-genome analysis for effective clinical diagnosis and gene discovery in early infantile epileptic encephalopathy. *npj Genomic Medicine, 3*, Article#22.

Ozdemir, O., Egemen, E., Ugur Iseri, S. A., Sezerman, O. U., Bebek, N., Baykan, B., & Ozbek, U. (2019). Identification of epilepsy related pathways using genome-wide DNA methylation measures: A trio-based approach. *PLoS One, 14*(2), Article#e0211917.

Paatero, I., Veikkolainen, V., Mäenpää, M., Schmelzer, E., Belting, H. G., Pelliniemi, L. J., & Elenius, K. (2019). ErbB4 tyrosine kinase inhibition impairs neuromuscular development in zebrafish embryos. *Molecular Biology of the Cell, 30*(2), 209-218.

Pack, A. M. (2019). Epilepsy overview and revised classification of seizures and epilepsies. *CONTINUUM: Lifelong Learning in Neurology, 25*(2), 306-321.

Pal, D. K., Evgrafov, O. V., Tabares, P., Zhang, F., Durner, M., & Greenberg, D. A. (2003). BRD2 (RING3) is a probable major susceptibility gene for common juvenile myoclonic epilepsy. *The American Journal of Human Genetics, 73*(2), 261-270.

Panayiotopoulos, C. P. (2012). The new ILAE report on terminology and concepts for the organization of epilepsies: Critical review and contribution. *Epilepsia, 53*(3), 399-404.

Parihar, R., & Ganesh, S. (2013). The SCN1A gene variants and epileptic encephalopathies. *Journal of Human Genetics, 58*(9), 573-580.

Pattin, K. A., & Moore, J. H. (2009). Role for protein-protein interaction databases in human genetics. *Expert Review of Proteomics, 6*(6), 647-659.

Pena, S. D. J., & Coimbra, R. L. M. (2015). Ataxia and myoclonic epilepsy due to a heterozygous new mutation in KCNA2: Proposal for a new channelopathy. *Clinical Genetics, 87*(2), e1-e3.

Peng, J., Pang, N., Wang, Y., Wang, X. L., Chen, J., Xiong, J., . . . Yin, F. (2019). Next-generation sequencing improves treatment efficacy and reduces hospitalization in children with drug-resistant epilepsy. *CNS Neuroscience & Therapeutics, 25*(1), 14-20.

Perucca, P., & Perucca, E. (2019). Identifying mutations in epilepsy genes: Impact on treatment selection. *Epilepsy Research, 152*, 18-30.

Perucca, P., Scheffer, I. E., Harvey, A. S., James, P. A., Lunke, S., Thorne, N., . . . Kwan, P. (2017). Real-world utility of whole exome sequencing with targeted gene analysis for focal epilepsy. *Epilepsy Research, 131*, 1-8.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25*(13), 1605-1612.

Pi, X., Cui, L., Liu, A., Zhang, J., Ma, Y., Liu, B., . . . Xiao, B. (2012). Investigation of prevalence, clinical characteristics and management of epilepsy in Yueyang city of China by a door-to-door survey. *Epilepsy Research, 101*(1-2), 129-134.

Pirooznia, M., Kramer, M., Parla, J., Goes, F. S., Potash, J. B., McCombie, W. R., & Zandi, P. P. (2014). Validation and assessment of variant calling pipelines for next-generation sequencing. *Human Genomics, 8*(1), Article#14.

Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., . . . DePristo, M. A. (2018). A universal SNP and small-indel variant caller using deep neural networks. *Nature Biotechnology, 36*(10), 983-987.

Porcher, C., Medina, I., & Gaiarsa, J. L. (2018). Mechanism of BDNF modulation in GABAergic synaptic transmission in healthy and disease brains. *Frontiers in Cellular Neuroscience, 12*, Article#273.

Pranckėnienė, L., Jakaitienė, A., Ambrozaitytė, L., Kavaliauskienė, I., & Kučinskas, V. (2018). Insights into de novo mutation variation in Lithuanian exome. *Frontiers in Genetics, 9*, Article#315.

Psarropoulou, C. (2010). Neurotransmitters, ions, their receptors and channels. In C. P. Panayiotopoulos (Ed.), *Atlas of Epilepsies* (pp. 185-192). London: Springer London.

Pshennikova, V. G., Barashkov, N. A., Romanov, G. P., Teryutin, F. M., Solov'ev, A. V., Gotovtsev, N. N., . . . Fedorova, S. A. (2019). Comparison of predictive in silico tools on missense variants in GJB2, GJB6, and GJB3 genes associated with autosomal recessive deafness 1A (DFNB1A). *The Scientific World Journal, 2019*, Article#5198931.

Ptácek, L. J., & Fu, Y. H. (2003). What's new in epilepsy genetics?. *Molecular Psychiatry, 8*(5), 463-465.

Qaiser, F., Yuen, R. K. C., & Andrade, D. M. (2020). Genetics of epileptic networks: From focal to generalized genetic epilepsies. *Current Neurology and Neuroscience Reports, 20*(10), Article#46.

Rawat, C., Kutum, R., Kukal, S., Srivastava, A., Dahiya, U. R., Kushwaha, S., . . . Kukreti, R. (2020). Downregulation of peripheral PTGS2/COX-2 in response to valproate treatment in patients with epilepsy. *Scientific Reports, 10*(1), Article#2546.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., . . . Bader, G. D. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols, 14*(2), 482-517.

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research, 47*(D1), D886-D894.

Requena, M., Parrón, T., Navarro, A., García, J., Ventura, M. I., Hernández, A. F., & Alarcón, R. (2018). Association between environmental exposure to pesticides and epilepsy. *Neurotoxicology, 68*, 13-18.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine, 17*(5), 405-424.

Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S. R. F., WGS500 Consortium, . . . Lunter, G. (2014). Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature Genetics, 46*(8), 912-918.

Rodan, L. H., Qi, W., Ducker, G. S., Demirbas, D., Laine, R., Yang, E., . . . Berry, G. T. (2018). 5,10-methenyltetrahydrofolate synthetase deficiency causes a neurometabolic disorder associated with microcephaly, epilepsy, and cerebral hypomyelination. *Molecular Genetics and Metabolism, 125*(1-2), 118-126.

Rogawski, M. A., & Löscher, W. (2004). The neurobiology of antiepileptic drugs. *Nature Reviews Neuroscience, 5*(7), 553-564.

Rohena, L., Neidich, J., Truitt Cho, M., Gonzalez, K. D., Tang, S., Devinsky, O., & Chung, W. K. (2013). Mutation in SNAP25 as a novel genetic cause of epilepsy and intellectual disability. *Rare Diseases, 1*(1), Article#e26314.

Rosanoff, A. J., Handy, L. M., & Rosanoff, I. A. (1934). Etiology of epilepsy: With special reference to its occurrence in twins. *Archives of Neurology & Psychiatry, 31*(6), 1165-1193.

Rusconi, R., Combi, R., Cestèle, S., Grioni, D., Franceschetti, S., Dalprà, L., & Mantegazza, M. (2009). A rescuable folding defective Nav1.1 (SCN1A) sodium channel mutant causes GEFS+: Common mechanism in Nav1.1 related epilepsies?. *Human Mutation, 30*(7), E747-E760.

Russell, W. (1889). The influence of hereditary predisposition in epilepsy. *Glasgow Medical Journal, 31*(6), 434-436.

Sahli, M., Zrhidri, A., Elaloui, S. C., Smaili, W., Lyahyai, J., Oudghiri, F. Z., & Sefiani, A. (2019). Clinical exome sequencing identifies two novel mutations of the SCN1A and SCN2A genes in Moroccan patients with epilepsy: A case series. *Journal of Medical Case Reports, 13*(1), Article#266.

Saik, O. V., Nimaev, V. V., Usmonov, D. B., Demenkov, P. S., Ivanisenko, T. V., Lavrik, I. N., & Ivanisenko, V. A. (2019). Prioritization of genes involved in endothelial cell apoptosis by their implication in lymphedema using an analysis of associative gene networks with ANDSystem. *BMC Medical Genomics, 12*(s2), Article#47.

Saint-Martin, C., Gauvain, G., Teodorescu, G., Gourfinkel-An, I., Fedirko, E., Weber, Y. G., . . . Depienne, C. (2009). Two novel CLCN2 mutations accelerating chloride channel deactivation are associated with idiopathic generalized epilepsy. *Human Mutation, 30*(3), 397-405.

Salgueiro-Pereira, A. R., Duprat, F., Pousinha, P. A., Loucif, A., Douchamps, V., Regondi, C., . . . Mantegazza, M. (2019). A two-hit story: Seizures and genetic mutation interaction sets phenotype severity in SCN1A epilepsies. *Neurobiology of Disease, 125*, 31-44.

Sánchez-Alegría, K., Flores-León, M., Avila-Muñoz, E., Rodríguez-Corona, N., & Arias, C. (2018). PI3K signaling in neurons: A central node for the control of multiple functions. *International Journal of Molecular Sciences, 19*(12), Article#3725.

Santulli, L., Coppola, A., Balestrini, S., & Striano, S. (2016). The challenges of treating epilepsy with 25 antiepileptic drugs. *Pharmacological Research, 107*, 211-219.

Sapio, M. R., Vessaz, M., Thomas, P., Genton, P., Fricker, L. D., & Salzmann, A. (2015). Novel carboxypeptidase A6 (CPA6) mutations identified in patients with juvenile myoclonic and generalized epilepsy. *PLoS One, 10*(4), Article#e0123180.

Sarlin, M. B., Kloepfer, H. W., Mickle, W. A., & Heath, R. G. (1960). The detection of carriers in hereditary myoclonic epilepsy. *Acta Geneticae Medicae Et Gemellologiae, 9*(4), 466-471.

Sawamura, S., Shirakawa, H., Nakagawa, T., Mori, Y., & Kaneko, S. (2017). TRP channels in the brain: What are they there for?. In T. L. R. Emir (Ed.), *Neurobiology of TRP Channels*. Boca Raton, Florida: CRC Press/Taylor & Francis.

Scheffer, I. E., Berkovic, S., Capovilla, G., Connolly, M. B., French, J., Guilhoto, L., . . . Zuberi, S. M. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia, 58*(4), 512-521.

Scheffer, I. E., French, J., Hirsch, E., Jain, S., Mathern, G. W., Moshé, S. L., . . . Zuberi, S. M. (2016). Classification of the epilepsies: New concepts for discussion and debate-Special report of the ILAE Classification Task Force of the Commission for Classification and Terminology. *Epilepsia Open, 1*(1-2), 37-44.

Scheffer, I. E., & Nabbout, R. (2019). SCN1A-related phenotypes: Epilepsy and beyond. *Epilepsia, 60*(S3), S17-S24.

Schiemann, A. H., & Stowell, K. M. (2016). Comparison of pathogenicity prediction tools on missense variants in RYR1 and CACNA1S associated with malignant hyperthermia. *British Journal of Anaesthesia, 117*(1), 124-128.

Schrodinger, LLC. (2015). *The PyMOL Molecular Graphics System, Version 2.4.0a0*.

Schuster, E. (1912). A first study of inheritance in epilepsy. *The Eugenics Review, 4*(2), 208-209.

Schwarze, K., Buchanan, J., Taylor, J. C., & Wordsworth, S. (2018). Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genetics in Medicine, 20*(10), 1122-1130.

Sevim Bayrak, C., & Itan, Y. (2020). Identifying disease-causing mutations in genomes of single patients by computational approaches. *Human Genetics, 139*(6-7), 769-776.

Sevim Bayrak, C., Zhang, P., Tristani-Firouzi, M., Gelb, B. D., & Itan, Y. (2020). De novo variants in exomes of congenital heart disease patients identify risk genes and pathways. *Genome Medicine, 12*(1), Article#9.

Sha, Y., Liu, Q., Wang, Y., Dong, C., & Song, L. (2010). Exploring candidate genes for epilepsy by computational disease-gene identification strategy. *Balkan Journal of Medical Genetics, 13*(2), 35-40.

Shaik, N. A., Al-Qahtani, F., Nasser, K., Jamil, K., Alrayes, N. M., Elango, R., . . . Banaganapalli, B. (2020). Molecular insights into the coding region mutations of low-density lipoprotein receptor adaptor protein 1 (LDLRAP1) linked to familial hypercholesterolemia. *The Journal of Gene Medicine, 22*(6), Article#e3176.

Shariff, E. M., & AlKhamis, F. A. (2017). New onset epilepsy in the elderly: Clinical, radiological and electroencephalographic features and treatment responses. *Neurosciences, 22*(2), 102-106.

Sharma, R., Leung, W. L., Zamani, A., O'Brien, T. J., Casillas Espinosa, P. M., & Semple, B. D. (2019). Neuroinflammation in post-traumatic epilepsy: Pathophysiology and tractable therapeutic targets. *Brain Sciences, 9*(11), Article#318.

Sharma, S., & Prasad, A. N. (2017). Inborn errors of metabolism and epilepsy: Current understanding, diagnosis, and treatment approaches. *International Journal of Molecular Sciences, 18*(7), Article#1384.

Shi, X. Y., Tomonoh, Y., Wang, W. Z., Ishii, A., Higurashi, N., Kurahashi, H., . . . The Epilepsy Genetic Study Group, Japan (Chairperson, SK). (2016). Efficacy of antiepileptic drugs for the treatment of Dravet syndrome with different genotypes. *Brain and Development, 38*(1), 40-46.

Shinnar, S. (2010). The new ILAE classification. *Epilepsia, 51*(4), 715-717.

Shizu, R., Otsuka, Y., Ezaki, K., Ishii, C., Arakawa, S., Amaike, Y., . . . Yoshinari, K. (2020). Antiepileptic drug-activated constitutive androstane receptor inhibits peroxisome proliferator-activated receptor alpha and peroxisome proliferator-activated receptor gamma coactivator 1alpha-dependent gene expression to increase blood triglyceride levels. *Molecular Pharmacology, 98*(5), 634-647.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research, 40*(W1), W452-W457.

Singh, B., Monteil, A., Bidaud, I., Sugimoto, Y., Suzuki, T., Hamano, S.-i., . . . Yamakawa, K. (2007). Mutational analysis of CACNA1G in idiopathic generalized epilepsy. *Human Mutation, 28*(5), 524-525.

Skirzewski, M., Karavanova, I., Shamir, A., Erben, L., Garcia-Olivares, J., Shin, J. H., . . . Buonanno, A. (2018). ErbB4 signaling in dopaminergic axonal projections increases extracellular dopamine levels and regulates spatial/working memory behaviors. *Molecular Psychiatry, 23*(11), 2227-2237.

Snoeijen-Schouwenaars, F. M., van Ool, J. S., Verhoeven, J. S., van Mierlo, P., Braakman, H. M. H., Smeets, E. E., . . . Willemsen, M. H. (2019). Diagnostic exome sequencing in 100 consecutive patients with both epilepsy and intellectual disability. *Epilepsia, 60*(1), 155-164.

Sokka, A., Olsen, P., Kirjavainen, J., Harju, M., Keski-Nisula, L., Räisänen, S., . . . Kälviäinen, R. (2017). Etiology, syndrome diagnosis, and cognition in childhood-onset epilepsy: A population-based study. *Epilepsia Open, 2*(1), 76-83.

Sonawane, A. R., Weiss, S. T., Glass, K., & Sharma, A. (2019). Network medicine in the age of biomedical big data. *Frontiers in Genetics, 10*, Article#294.

Song, M. X., & Deng, X. Q. (2018). Recent developments on triazole nucleus in anticonvulsant compounds: A review. *Journal of Enzyme Inhibition and Medicinal Chemistry, 33*(1), 453-478.

Soto-Insuga, V., López, R. G., Losada-Del Pozo, R., Rodrigo-Moreno, M., Cayuelas, E. M., Giráldez, B. G., . . . Serratosa, J. M. (2019). Glut1 deficiency is a rare but treatable cause of childhood absence epilepsy with atypical features. *Epilepsy Research, 154*, 39-41.

Spampanato, J., Kearney, J. A., de Haan, G., McEwen, D. P., Escayg, A., Aradi, I., . . . Meisler, M. H. (2004). A novel epilepsy mutation in the sodium channel SCN1A identifies a cytoplasmic domain for beta subunit interaction. *The Journal of Neuroscience, 24*(44), 10022-10034.

Speed, D., Hoggart, C., Petrovski, S., Tachmazidou, I., Coffey, A., Jorgensen, A., . . . Johnson, M. R. (2014). A genome-wide association study and biological pathway analysis of epilepsy prognosis in a prospective cohort of newly treated epilepsy. *Human Molecular Genetics, 23*(1), 247-258.

Stafstrom, C. E. (2014). Recognizing seizures and epilepsy: Insights from pathophysiology. In J. W. Miller & H. P. Goodkin (Eds.), *Epilepsy* (pp. 1-9). Hoboken, NJ: Wiley-Blackwell.

Stafstrom, C. E., & Tempel, B. L. (2000). Epilepsy genes: The link between molecular dysfunction and pathophysiology. *Mental Retardation and Developmental Disabilities Research Reviews, 6*(4), 281-292.

Stefanović, S., Janković, S. M., Novaković, M., Milosavljević, M., & Folić, M. (2018). Pharmacodynamics and common drug-drug interactions of the third-generation antiepileptic drugs. *Expert Opinion on Drug Metabolism & Toxicology, 14*(2), 153-159.

Steinlein, O. K. (2004). Genetic mechanisms that underlie epilepsy. *Nature Reviews Neuroscience, 5*(5), 400-408.

Steinlein, O. K. (2008). Genetics and epilepsy. *Dialogues in Clinical Neuroscience, 10*(1), 29-38.

Steinman, L. (2018). Blocking immune intrusion into the brain suppresses epilepsy in Rasmussen's encephalitis model. *The Journal of Clinical Investigation, 128*(5), 1724-1726.

Steward, C. A., Roovers, J., Suner, M.-M., Gonzalez, J. M., Uszczynska-Ratajczak, B., Pervouchine, D., . . . Frankish, A. (2019). Re-annotation of 191 developmental and epileptic encephalopathy-associated genes unmasks de novo variants in SCN1A. *npj Genomic Medicine, 4*, Article#31.

Stoney, R., Robertson, D. L., Nenadic, G., & Schwartz, J.-M. (2018). Mapping biological process relationships and disease perturbations within a pathway network. *npj Systems Biology and Applications, 4*, Article#22.

Stosser, M. B., Lindy, A. S., Butler, E., Retterer, K., Piccirillo-Stosser, C. M., Richard, G., & McKnight, D. A. (2018). High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genetics in Medicine, 20*(4), 403-410.

Sugawara, T., Tsurubuchi, Y., Agarwala, K. L., Ito, M., Fukuma, G., Mazaki-Miyazaki, E., . . . Yamakawa, K. (2001). A missense mutation of the Na^+ channel alpha II subunit gene Na(v)1.2 in a patient with febrile and afebrile seizures causes channel dysfunction. *Proceedings of the National Academy of Sciences, 98*(11), 6384-6389.

Sun, J., & Nan, G. (2017). The extracellular signal-regulated kinase 1/2 pathway in neurological diseases: A potential therapeutic target. *International Journal of Molecular Medicine, 39*(6), 1338-1346.

Sun, Y., Ruivenkamp, C. A. L., Hoffer, M. J. V., Vrijenhoek, T., Kriek, M., van Asperen, C. J., . . . Santen, G. W. E. (2015). Next-generation diagnostics: Gene panel, exome, or whole genome?. *Human Mutation, 36*(6), 648-655.

Supernat, A., Vidarsson, O. V., Steen, V. M., & Stokowy, T. (2018). Comparison of three variant callers for human whole genome sequencing. *Scientific Reports, 8*(1), Article#17851.

Suzuki, T., Delgado-Escueta, A. V., Aguan, K., Alonso, M. E., Shi, J., Hara, Y., . . . Yamakawa, K. (2004). Mutations in EFHC1 cause juvenile myoclonic epilepsy. *Nature Genetics, 36*(8), 842-849.

Swiss Institute of Bioinformatics. (2021). Help SWISS-MODEL. Retrieved on July 9, 2021 from https://swissmodel.expasy.org/docs/help

Syrbe, S., Hedrich, U. B. S., Riesch, E., Djémié, T., Müller, S., Møller, R. S., . . . EuroEPINOMICS RES. (2015). De novo loss- or gain-of-function mutations in KCNA2 cause epileptic encephalopathy. *Nature Genetics, 47*(4), 393-399.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., . . . Mering, Christian v. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research, 47*(D1), D607-D613.

Tang, H., & Thomas, P. D. (2016). Tools for predicting the functional impact of nonsynonymous genetic variation. *Genetics, 203*(2), 635-647.

TBE electrophoresis buffer (10X). (2010). *Cold Spring Harbor Protocols, 2010*(6), pdb.rec12231.

Tee, S. K., Ong, T. L., Azman Aris, See, S. M. L., Leong, H. Y., Mohd Khairul Nizam Mohd Khalid, & Viswanathan, S. (2019). Lafora disease in a Malaysian with a rare mutation in the EPM2A gene. *Seizure - European Journal of Epilepsy, 67*, 78-81.

Tetreault, M., Bareke, E., Nadaf, J., Alirezaie, N., & Majewski, J. (2015). Whole-exome sequencing as a diagnostic tool: Current challenges and future opportunities. *Expert Review of Molecular Diagnostics, 15*(6), 749-760.

Thakran, S., Guin, D., Singh, P., Singh, P., Kukal, S., Rawat, C., . . . Kukreti, R. (2020). Genetic landscape of common epilepsies: Advancing towards precision in treatment. *International Journal of Molecular Sciences, 21*(20), Article#7784.

The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74.

Thurman, D. J., Begley, C. E., Carpio, A., Helmers, S., Hesdorffer, D. C., Mu, J., . . . Newton, C. R. (2018). The primary prevention of epilepsy: A report of the Prevention Task Force of the International League Against Epilepsy. *Epilepsia, 59*(5), 905-914.

Timmermans, S., Tietbohl, C., & Skaperdas, E. (2017). Narrating uncertainty: Variants of uncertain significance (VUS) in clinical exome sequencing. *BioSocieties, 12*(3), 439-458.

Tong, W., Wang, Y., Lu, Y., Ye, T., Song, C., Xu, Y., . . . Jin, D. (2018). Whole-exome sequencing helps the diagnosis and treatment in children with neurodevelopmental delay accompanied unexplained dyspnea. *Scientific Reports, 8*(1), Article#5214.

Tran, D. S., Odermatt, P., Le, T. O., Huc, P., Druet-Cabanac, M., Barennes, H., . . . Preux, P. M. (2006). Prevalence of epilepsy in a rural district of central Lao PDR. *Neuroepidemiology, 26*(4), 199-206.

Trinka, E., Kwan, P., Lee, B., & Dash, A. (2019). Epilepsy in Asia: Disease burden, management barriers, and challenges. *Epilepsia, 60*(s1), 7-21.

Tsuboi, T., & Christian, W. (1973). On the genetics of the primary generalized epilepsy with sporadic myoclonias of impulsive petit mal type. A clinical and electroencephalographic study of 399 probands. *Humangenetik, 19*(2), 155-182.

Tsuchida, N., Nakashima, M., Kato, M., Heyman, E., Inui, T., Haginoya, K., . . . Matsumoto, N. (2018). Detection of copy number variations in epilepsy using exome data. *Clinical Genetics, 93*(3), 577-587.

Tumienė, B., Peterlin, B., Maver, A., & Utkus, A. (2018). Contemporary scope of inborn errors of metabolism involving epilepsy or seizures. *Metabolic Brain Disease, 33*(6), 1781-1786.

Tunçer, G. Ö., Teber, S., Albayrak, P., Kutluk, M. G., & Deda, G. (2018). A case of Dravet Syndrome with a newly defined mutation in the SCN1A gene. *Turkish Archives of Pediatrics, 53*(4), 259-262.

van Campen, J. S., Jansen, F. E., Brouwer, O. F., Nicolai, J., & Braun, K. P. J. (2013). Interobserver agreement of the old and the newly proposed ILAE epilepsy classification in children. *Epilepsia, 54*(4), 726-732.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., . . . DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics, 43*(1), 11.10.11-11.10.33.

van Lint, F. H. M., Mook, O. R. F., Alders, M., Bikker, H., Lekanne Dit Deprez, R. H., & Christiaans, I. (2019). Large next-generation sequencing gene panels in genetic heart disease: Yield of pathogenic variants and variants of unknown significance. *Netherlands Heart Journal, 27*(6), 304-309.

Vaz-Drago, R., Custódio, N., & Carmo-Fonseca, M. (2017). Deep intronic mutations and human disease. *Human Genetics, 136*(9), 1093-1111.

Veeramah, K. R., Johnstone, L., Karafet, T. M., Wolf, D., Sprissler, R., Salogiannis, J., . . . Hammer, M. F. (2013). Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia, 54*(7), 1270-1281.

Veeramah, K. R., O'Brien, J. E., Meisler, M. H., Cheng, X., Dib-Hajj, S. D., Waxman, S. G., . . . Hammer, M. F. (2012). De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. *The American Journal of Human Genetics, 90*(3), 502-510.

Veri, K., Talvik, I., Vaher, U., Napa, A., Ilves, P., Uibo, O., . . . Talvik, T. (2018). Incidence of childhood epilepsy in Estonia. *Journal of Child Neurology, 33*(9), 587-592.

Verma, A., & Kumar, A. (2017). Clinical and etiological profile of epilepsy in elderly: A hospital-based study from rural India. *Acta Neurologica Belgica, 117*(1), 139-144.

Vidal, G. A., Clark, D. E., Marrero, L., & Jones, F. E. (2007). A constitutively active ERBB4/HER4 allele with enhanced transcriptional coactivation and cell-killing activities. *Oncogene, 26*(3), 462-466.

Vijai, J., Kapoor, A., Ravishankar, H. M., Cherian, P. J., Kuruttukulam, G., Rajendran, B., . . . Anand, A. (2005). Protective and susceptibility effects of hSKCa3 allelic variants on juvenile myoclonic epilepsy. *Journal of Medical Genetics, 42*(5), 439-442.

Vorderwülbecke, B. J., Kowski, A. B., Kirschbaum, A., Merkle, H., Senf, P., Janz, D., & Holtkamp, M. (2017). Long-term outcome in adolescent-onset generalized genetic epilepsies. *Epilepsia, 58*(7), 1244-1250.

Vučetić Tadić, B., Kravljanac, R., Sretenović, V., Martić, J., & Vukomanović, V. (2020). The features of neonatal seizures as predictors of drug-resistant epilepsy in children. *Epilepsy & Behavior, 106*, Article#107004.

Wallace, R. H., Wang, D. W., Singh, R., Scheffer, I. E., George, A. L., Phillips, H. A., . . . Mulley, J. C. (1998). Febrile seizures and generalized epilepsy associated with a mutation in the Na^{+}-channel ß1 subunit gene SCN1B. *Nature Genetics, 19*(4), 366-370.

Walters, D., Vogel, K. R., Brown, M., Shi, X., Roullet, J. B., & Gibson, K. M. (2020). Transcriptome analysis in mice treated with vigabatrin identifies dysregulation of genes associated with retinal signaling circuitry. *Epilepsy Research, 166*, Article#106395.

Wandschneider, B., Hong, S. J., Bernhardt, B. C., Fadaie, F., Vollmar, C., Koepp, M. J., . . . Bernasconi, A. (2019). Developmental MRI markers cosegregate juvenile patients with myoclonic epilepsy and their healthy siblings. *Neurology, 93*(13), e1272-e1280.

Wang, C., Chung, B. C., Yan, H., Wang, H. G., Lee, S. Y., & Pitt, G. S. (2014). Structural analyses of Ca(2)(+)/CaM interaction with NaV channel C-termini reveal mechanisms of calcium-dependent regulation. *Nature Communications, 5*, Article#4896.

Wang, J., Lin, Z. J., Liu, L., Xu, H. Q., Shi, Y. W., Yi, Y. H., . . . Liao, W. P. (2017a). Epilepsy-associated genes. *Seizure - European Journal of Epilepsy, 44*, 11-20.

Wang, Q., Liu, Z., Lin, Z., Zhang, R., Lu, Y., Su, W., . . . Zheng, X. (2019). De novo germline mutations in SEMA5A associated with infantile spasms. *Frontiers in Genetics, 10*, Article#605.

Wang, S., & Xing, J. (2013). A primer for disease gene prioritization using next-generation sequencing data. *Genomics & Informatics, 11*(4), 191-199.

Wang, Y., Du, X., Bin, R., Yu, S., Xia, Z., Zheng, G., . . . Wang, Y. (2017b). Genetic variants identified from epilepsy of unknown etiology in Chinese children by targeted exome sequencing. *Scientific Reports, 7*, Article#40319.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., . . . Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research, 46*(W1), W296-W303.

Wei, Z., Wang, L., & Deng, Y. (2019). Treatment of myoclonic-atonic epilepsy caused by SLC2A1 de novo mutation with ketogenic diet: A case report. *Medicine, 98*(18), Article#e15428.

Weichenberger, C. X., Palermo, A., Pramstaller, P. P., & Domingues, F. S. (2017). Exploring approaches for detecting protein functional similarity within an orthology-based framework. *Scientific Reports, 7*(1), Article#381.

Wells, Q. S., Becker, J. R., Su, Y. R., Mosley, J. D., Weeke, P., D'Aoust, L., . . . Hong, C. C. (2013). Whole exome sequencing identifies a causal RBM20 mutation in a large pedigree with familial dilated cardiomyopathy. *Circulation: Cardiovascular Genetics, 6*(4), 317-326.

Welton, J. M., Walker, C., Riney, K., Ng, A., Todd, L., & D'Souza, W. J. (2020). Quality of life and its association with comorbidities and adverse events from antiepileptic medications: Online survey of patients with epilepsy in Australia. *Epilepsy & Behavior, 104*, Article#106856.

Winston, G. P., & Jaiser, S. R. (2012). Western driving regulations for unprovoked first seizures and epilepsy. *Seizure - European Journal of Epilepsy, 21*(5), 371-376.

Wo, M. C. M., Lim, K. S., Choo, W. Y., & Tan, C. T. (2015). Employability among people with uncontrolled seizures: An interpretative phenomenological approach. *Epilepsy & Behavior, 45*, 21-30.

Wong, L. P., Ong, Rick T. H., Poh, W. T., Liu, X., Chen, P., Li, R., . . . Teo, Y. Y. (2013). Deep whole-genome sequencing of 100 southeast Asian Malays. *The American Journal of Human Genetics, 92*(1), 52-66.

Wu, C. C., Tsai, M. H., Chu, Y. J., Weng, W. C., Fan, P. C., & Lee, W. T. (2020). The role of targeted gene panel in pediatric drug-resistant epilepsy. *Epilepsy & Behavior, 106*, Article#107003.

Wu, D., Dou, J., Chai, X., Bellis, C., Wilm, A., Shih, C. C., . . . Wang, C. (2019). Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. *Cell, 179*(3), 736-749.

Xiao, W., Wu, Y., Wang, J., Luo, Z., Long, L., Deng, N., . . . Xiao, B. (2019). Network and pathway-based analysis of single-nucleotide polymorphism of miRNA in temporal lobe epilepsy. *Molecular Neurobiology, 56*(10), 7022-7031.

Xie, H., Zhang, Y., Zhang, P., Wang, J., Wu, Y., Wu, X., . . . Jiang, Y. (2014). Functional study of NIPA2 mutations identified from the patients with childhood absence epilepsy. *PLoS One, 9*(10), Article#e109749.

Yang, D., He, Y., Wu, B., Deng, Y., Wang, N., Li, M., & Liu, Y. (2020). Integrated bioinformatics analysis for the screening of hub genes and therapeutic drugs in ovarian cancer. *Journal of Ovarian Research, 13*(1), Article#10.

Yang, H., & Wang, K. (2015). Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nature Protocols, 10*(10), 1556-1566.

Yang, J. (2020). Identification of novel biomarkers, MUC5AC, MUC1, KRT7, GAPDH, CD44 for gastric cancer. *Medical Oncology, 37*(5), Article#34.

Yang, Y., Tian, X., Xu, D., Zheng, F., Lu, X., Zhang, Y., . . . Wang, X. (2018). GPR40 modulates epileptic seizure and NMDA receptor function. *Science Advances, 4*(10), Article#eaau2357.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics, 13*, Article#134.

Yogarajah, M., & Mula, M. (2019). Social cognition, psychiatric comorbidities, and quality of life in adults with epilepsy. *Epilepsy & Behavior, 100*, Article#106321.

Zhang, H., He, X., Mei, Y., & Ling, Q. (2018). Ablation of ErbB4 in parvalbumin-positive interneurons inhibits adult hippocampal neurogenesis through down-regulating BDNF/TrkB expression. *Journal of Comparative Neurology, 526*(15), 2482-2492.

Zhang, H., Lu, P., Tang, H. L., Yan, H. J., Jiang, W., Shi, H., . . . Long, Y. S. (2021). Valproate-induced epigenetic upregulation of hypothalamic Fto expression potentially linked with weight gain. *Cellular and Molecular Neurobiology, 41*(6), 1257-1269.

Zhang, L., Zhang, Z., Bi, X., Mao, Y., Cheng, Y., Zhu, P., . . . Wu, X. (2020). Genetic and preimplantation diagnosis of cystic kidney disease with ventriculomegaly. *Journal of Human Genetics, 65*(5), 455-459.

Zhang, Y., Dong, H. T., Duan, L., Niu, L., Yuan, G. Q., Dai, J. Q., . . . Pan, Y. W. (2019). HDAC4 gene silencing alleviates epilepsy by inhibition of GABA in a rat model. *Neuropsychiatric Disease and Treatment, 15*, 405-416.

Zhang, Y. H., Burgess, R., Malone, J. P., Glubb, G. C., Helbig, K. L., Vadlamudi, L., . . . Scheffer, I. E. (2017). Genetic epilepsy with febrile seizures plus: Refining the spectrum. *Neurology, 89*(12), 1210-1219.

Zhao, J., Lv, C., Wu, Q., Zeng, H., Guo, X., Yang, J., . . . Zhang, W. (2019). Computational systems pharmacology reveals an antiplatelet and neuroprotective mechanism of Deng-Zhan-Xi-Xin injection in the treatment of ischemic stroke. *Pharmacological Research, 147*, Article#104365.

Zhao, L., Liu, H., Yuan, X., Gao, K., & Duan, J. (2020a). Comparative study of whole exome sequencing-based copy number variation detection tools. *BMC Bioinformatics, 21*(1), Article#97.

Zhao, Y., Wang, J., Chen, J., Zhang, X., Guo, M., & Yu, G. (2020b). A literature review of gene function prediction by modeling gene ontology. *Frontiers in Genetics, 11*, Article#400.

Zhou, L., Xu, H., Wang, T., & Wu, Y. (2020). A patient with CAD deficiency responsive to uridine and literature review. *Frontiers in Neurology, 11*, Article#64.

Zhu, L. N., Chen, D., Deng, Y., He, J., He, Y. J., Zhou, D., & Liu, L. (2020). Long-term seizure, comorbidity and socioeconomic outcomes of patients with convulsive epilepsy in rural West China. *Epilepsy Research, 168*, Article#106480.

Zhu, X., Petrovski, S., Xie, P., Ruzzo, E. K., Lu, Y. F., McSweeney, K. M., . . . Goldstein, D. B. (2015). Whole-exome sequencing in undiagnosed genetic diseases: Interpreting 119 trios. *Genetics in Medicine, 17*(10), 774-781.

Zolotareva, O., & Kleine, M. (2019). A survey of gene prioritization tools for Mendelian and complex human diseases. *Journal of Integrative Bioinformatics, 16*(4), Article#20180069.