IDENTIFICATION OF NOVEL GENES INVOLVED IN THE ETIOLOGY OF BENIGN NEONATAL/ INFANTILE EPILEPSY SYNDROMES AND GENETIC EPILEPSY WITH FEBRILE SEIZURES PLUS (GEFS+)

by

Sunay Usluer

B.S., Molecular Biology and Genetics, Bogazici University, 2005

M.S., Molecular Biology and Genetics, Bogazici University, 2008

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Molecular Biology and Genetics
Boğaziçi University
2015

To Oğuz, İlhan and Orhan;
my beloved companions in this way

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor Prof. S. Hande Çağlayan for excellent working environment she provided, and her encouragement in every new step and scientific support during this work.

I express my gratitude to Prof. Nazlı Başak and Prof. Eda Tahir-Turanlı for their precious criticism throughout my thesis work. I am also grateful to Prof. Esra Battaloğlu and Assist. Prof. Özlem Yalçın-Çapan for devoting their time to evaluate this thesis.

I am grateful for the financial support provided by the Boğaziçi University Research Fund through project numbers 10B01D13 and 14B01D4, TÜBİTAK/ESF grant number 110S518 and TUBITAK/IntenC grant number 01DL12011. I was financially supported by TÜBİTAK/ESF under the EuroEpinomics project (grant number 110S518) and the Turkish State Planning Organization (DPT) under the TAM Project, number 2007K120610.

I also appreciate the contributions of. Dilşad Türkdoğan, MD; Uluç Yiş, MD; Sema Saltık, MD; and Bülent Kara, MD, for providing the patient samples and valuable clinical information of BFIS and BFNS patients. I am also thankful to Canan Aykut-Bingöl, MD; Berrin Aktekin, MD; Kadriye Ağan, MD; and Berfu Naz Akbaş, MD, for their efforts in sampling and phenotyping of GEFS+ patients.

I am grateful to Wellcome Trust Sanger Institute and Boğaziçi University Foundation for course and conference supports they provide.

I express my acknowledgements to TUBITAK Advanced Genomics and Bioinformatics Research Center for providing SNP frequency data for identified exome variants.

I would like express my sincere gratitude to M. Murat Kösoğlu, PhD, and Umidahan Djakbarova for their helps in neuronal cell culture and gene expression studies.

I should express my sincere gratitude to all faculty members of Molecular Biology and Genetics Department for their contributions in my scientific and personal improvements. I would like to express my sincere thanks to my friends in Hande Cağlayan`s lab; namely Inanç Değer Fidancı, PhD, Özlem Yalçın-Çapan, PhD, Aslı Gündoğdu Eken, Melek Aslı Kayserili, PhD, Seda Salar, PhD, Yasemin Şen, Gökçe Çakan, Öyküm Kaplan , Hande Özünlü and Derya Dikbıyık and all undergraduate students who somehow contributed to my work.

I thank to my friends in MBG Department for providing warm working environment. I express my since gratitude to Neslihan Zöhrap for teaching me cell culture. I am also grateful to Sibel Aylin Uğur-İşeri, PhD, for her help in linkage analyses.

Lastly I would like to express my appreciation to my husband, Oğuz Usluer and my sons; İlhan and Orhan; for walking this way together with me regardless of how challenging it was and also providing me a real life outside the laboratory. I could not come so far without their support and patience.

ABSTRACT

IDENTIFICATION OF NOVEL GENES INVOLVED IN THE ETIOLOGY OF BENIGN NEONATAL/INFANTILE EPILEPSY SYNDROMES AND GENETIC EPILEPSY WITH FEBRILE SEIZURES PLUS (GEFS+)

Epilepsy is among the most prevalent episodic neurological disorders. Genetic factors play a major role in the etiology of epilepsy. This thesis included analysis of families with distinct epilepsy phenotypes in order to delineate their complex genetic background using advanced and highthroughput current technologies. The first part of the thesis comprised a large family with BFIS phenotype which was analyzed and found to have a synonymous change in the *SCN1B* gene affecting splicing efficiency as shown in neuronal cell culture and by *in silico* tools. It was the first time *SCN1B* gene was shown to be associated with the BFIS phenotype. Several patients in the family also had *KCNQ2* gene copy number gain and a frameshift mutation in the *PRRT2* gene. The lower penetrance of these two BFIS associated gene mutations indicated the oligogenic nature of the disease. Additional families with BFIS phenotype were also analyzed for point mutations in the *SCN1B* and *PRRT2* genes. A frameshifting 2 bp deletion in the *PRRT2* gene was found in one family and rare SNPs in *SCN1B* genes were identified in other families. Five BFNS patients with neonatal disease onset, on the other hand, had inherited *KCNQ2* gene copy number gain mutations suggesting that *KCNQ2* duplications mutations may also be implicated in the etiology of BFIS/BFNS phenotypes. In the second part of the study a large multiplex, multigenerational kindred with epilepsy similar to GEFS+ phenotype and with patients having idiopathic generalized or partial epilepsy subtypes was analyzed by current genomic technologies and found to have a VNTR expansion on the *mir137* gene in significantly higher numbers in individuals with epilepsy phenotypes. The VNTR expansion disrupts the expression of *mir137* that targets all the genes involved in schizophrenia, synapses formation and important ion channel genes involved in epilepsy.

ÖZET

SELİM YENİDOĞAN/BEBEKLİK EPİLEPSİ SENDROMLARI VE ATEŞLİ NÖBETLİ GENETİK EPİLEPSİ ETİYOLOJİSİNDE YER ALAN YENİ GENLERİN BELİRLENMESİ

Epilepsi en sık görülen episodik nörolojik hastalıklar arasındadır. Epilepsi etiyolojisinde genetik faktörlerin rolü büyüktür. Bu çalışma da, epilepsi hastalıklarının etiyolojisinde yer alan genlerin belirlenmesi amacıyla, güncel, gelişmiş ve yüksek verimlilikli teknolojiler kullanılarak, belirli epilepsi sendromları görülen ailelerde genetik analizleri içermektedir. Tezin ilk bölümünde Selim Ailesel Bebeklik Nöbetleri (SABN) görülen büyük bir ailede *SCN1B* geninde gen kırpılma verimini etkileyen sinonim bir değişiklik bulunmuş ve sinir hücresi kültüründe de bu etki gösterilmiştir. *SCN1B* geninin bu fenotiple ilişkisi ilk defa bu çalışma kapsamında gösterilmiş olmaktadır. Bunun yanı sıra bu ailede birçok bireyde *KCNQ2* gen duplikasyonu ve *PRRT2* geninde bir çerçeve kayması mutasyonu bulunmuştur. Bu fenotiple ilişkili bu genlerde görülen mutasyonların düşük penetransı, bu hastalığın çok genli kalıtımını göstermektedir. Benzer fenotip görülen diğer aileler de *SCN1B* ve *PRRT2* genlerinde nokta mutasyonlar açısından analiz edilmiş ve bir ailede de *PRRT2* geninde çerçeve kaymasına yol açan 2 bazlık bir delesyon diğer aillerde ise *SCN1B* geninde patojenik etkisi belirsiz tek nükleotid değişiklikleri bulunmuştur. Ayrıca, Selim Ailesel Neonatal Nöbetler (SANN) Sendromu görülen beş hastada ebeveynlerinden kalıtılan *KCNQ2* kopya sayısı artışı olduğu görülmüştür ki bu da bu gendeki kopya sayısı artışlarının SANN/ SABN hastalıklarının etiyolojisinde yer aldığını göstermektedir. Çalışmanın ikinci bölümünde İdiyopatik Jeneralize ve Parsiyel epileptik nöbetli bireyler olan ve GEFS+ benzeri bir epilepsi sendromu görülen, çok jenerasyonlu ve multipleks bir aile, güncel genomic teknolojilerle incelenmiş ve *miR137* genindeki bir VNTR bölgesinde kopya sayısı artışının epilepsili hastalarda anlamlı bir şekilde yüksek olduğu saptanmıştır. Bu VNTR artışı, şizofreni, sinaps oluşumu ve epilepside yer alan önemli kanal genlerini hedef alan *miR137* geninin anlatımını bozmaktadır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS / ABBREVIATIONS

A	Alanine
A	Adenine
aa	Amino Acid
ABCD3	ATP-Binding Cassette, Sub-Family D (ALD), Member 3
aCGH	Array Comparative Genomic Hybridization
AD	Autosomal Dominant
ADPEAF	Autosomal Dominant Partial Epilepsy with Auditory Features
AGL	Amylo-Alpha-1, 6-Glucosidase, 4-Alpha- Glucanotransferase
AR	Autosomal Recessive
ATP	Adenosine Triphosphate
ATP1A2	Atpase, Na+/K+ Transporting, Alpha 2 Polypeptide
B	Aspartic acid
B	Benign
BDKRB1	Bradykinin Receptor B1
BFIC	Benign Familial Infantile Convulsions
BFIS	Benign Familial Infantile Seizures
BFNIS	Benign Familial Neonatal Infantile Seizures
BFNS	Benign Familial Neonatal Seizures
BGI	Beijing Genomic Institute
C	Cysteine
C	Cytosine
Ca	Calcium
CACNA1A	Calcium Channel, Voltage-Dependent, P/Q Type, Alpha 1A subunit
CACNA1G	Calcium Channel, Voltage-Dependent, T Type, Alpha 1G Subunit
CATSPER2	Cation Channel, Sperm Associated 2
CCG	Cologne Center for Genomics
cDNA	Complementary DNA

CGH	Comparative Genomic Hybridization
Cl	Chloride
CLCN3	Chloride Channel, Voltage-Sensitive 3
CLCN4	chloride Channel, Voltage-Sensitive 4
CNGA3	Cyclic Nucleotide Gated Channel Alpha 3
CNTNAP2	Contactin-Associated Protein-Like 2
CNTNAP3B	Contactin Associated Protein-Like 3B
CNV	Copy Number Variation
COL11A1	Collagen, Type XI, Alpha 1
D	Aspartic acid
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
DPYD	Dihydropyrimidine Dehydrogenase
DRD5	Dopamine Receptor D5
DRLRS	Derivative LogR Spread
dsDNA	Double-Stranded DNA
E	Glutamic acid
EEG	Electroencephalogram
EFHC1	EF-Hand Domain Containing Protein I
EIF2AK3	Eukaryotic Translation İnitiation Factor 2-Alpha Kinase 3
eIF2B5	Eukaryotic Translation Elongation Factor B Subunit Epsilon
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
EtBr	Ethidium Bromide
EtOH	Ethanol
EXTL2	Exostosin-Like Glycosyltransferase 2
F	Phenylalanine
F	Female
F12	Ham’s F12 Nutrient Mixture
FBS	Fetal Bovine Serum
FS	Febrile Seizure

FS+	Febrile Seizure Plus
FsC	Frame-Shift Coding
G	Glycine
G	Guanine
GABA	γ-Aminobutyric Acid
GABRA6	Gamma-Aminobutyric Acid (GABA) A Receptor, Alpha 6
GABRB3	Gamma-Aminobutyric Acid (GABA) A Receptor, Beta 3
GABRD	Gamma-Aminobutyric Acid (GABA) A Receptor, Delta
GABRG2	Gamma-Aminobutyric Acid (GABA) A Receptor, Gamma 2
GEFS+	Genetic Epilepsy with Febrile Seizures Plus
GRAMD1A	GRAM Domain Containing 1A
GRID1	Glutamate Receptor, İonotropic, Delta 1
GRM3	Glutamate Receptor, Metabotropic 3
GRM6	Metabotropic Glutamate Receptor 6
GSEA	Gene Set Enrichment Analysis
GTCS	Generalized Tonic Clonic Seizure
GWAS	Genome-Wide Association Study
H	Histidine
H19	H19, Imprinted Maternally Expressed Transcript
HCl	Hydrochloric Acid
HEK	Human Embryonic Kidney
hERG	Human *Ether-à-go-go*-Related Gene
HPN	Hepsin
HTR1D	5-Hydroxytryptamine (Serotonin) Receptor 1D, G Protein-Coupled
I	Isoleucine
ICEGTC	Intractable Childhood Epilepsy with Generalized Tonic–Clonic Seizures
ID	Identity
Ig	Immunoglobulin
IGE	Idiopathic Generalized Epilepsy
IGV	Integrated Genomic Viewer
ILEA	International League Against Epilepsy

indel	Insertion Deletion Polymorphism
INPP4A	Inositol Polyphosphate-4-Phosphatase, Type I, 107kda
ISCA	International Standards for Cytogenomic Arrays
K	Lysine
K	Potassium
KCNB1	Voltage Gated Potassium Channel, Sharp-Related Subfamily, Member 1
KCNJ10	Inwardly Rectifying Potassium Channel, Subfamily J, Member 10
KCNJ16	Inwardly Rectifying Potassium Channel, Subfamily J, Member 16
KCNK18	Potassium Channel, Two Pore Domain Subfamily K, Member 18
KCNMB3	Potassium Channel Subfamily M Regulatory Beta Subunit 3
KCNQ	Voltage Gated Potassium Channel,QTL-Like Subfamily
KCNQ2	Voltage Gated Potassium Channel,QTL-Like Subfamily, Member 2
KCNQ3	Voltage Gated Potassium Channel,QTL-Like Subfamily, Member 3
KCNQ4	Voltage Gated Potassium Channel,QTL-Like Subfamily, Member 4
KCNV1	Potassium Channel, Voltage Gated Modifier Subfamily V, Member 1
KCNV2	Potassium Channel, Voltage Gated Modifier Subfamily V, Member 2
KCTD15	Potassium Channel Tetramerization Domain Containing 15
KEGG	Kyoto Encyclopedia of Genes and Genomes
$KHCO_3$	Potassium Bicarbonate
Kpn	*Klebsiella pneumoniae*
KV	Voltage Gated Potassium Channel
KV7.2	Potassium Voltage-Gated Channel, KQT-Like Subfamily, Member 2 Protein
Kv7.3	Potassium Voltage-Gated Channel, KQT-Like Subfamily,

	Member 3 Protein
L	Leucine
LB	Luria Broth
LGI4	Leucine-Rich Repeat LGI Family, Member 4
LOD	Logarithm of Odds Ratio
M	Male
M	Months
M	Methionine
MAE	Myoclonic-Astatic Epilepsy
MAF	Minor Allele Frequency
MECP2	Methyl CpG Binding Protein 2
MgCl2	Magnesium Chloride
MINK1	Misshapen-Like Kinase 1
Mir137	MicroRNA 137
Mir137HG	MicroRNA 137 Host Gene
MLPA	Multiple Ligation Point Amplification
MR	Mental Retardation
mRNA	Messenger RNA
N	Asparagine
Na	Sodium
Na_2EDTA	Ethylenediaminetetraacetic Sodium
NaCl	Sodium Chloride
NAHR	Nonallelic Homologous Recombination
Nav1.1	Voltage Gated Sodium Channel Alpha 1 Subunit
Nav1.2	Voltage Gated Sodium Channel Alpha 2 Subunit
Nav1.5	Voltage Gated Sodium Channel Alpha 5 Subunit
Nav1.6	Voltage Gated Sodium Channel Alpha 6 Subunit
Navb1	Voltage Gated Sodium Channel Beta 1 Subunit
NCBI	National Center for Biotechnology Information
NH_4Cl	Ammonium Chloride
NRXN2	Neurexin 2
NSC	Nonsynonymous Coding
OMIM	Online Mendelian Inheritance In Men

P	Proline
PCR	Polymerase Chain Reaction
PE	Partial Epilepsy
PESX	Putative Exon; Splicing Enhancer, Silencers Program
PNH	Peripheral Nerve Hyper Excitability
PrD	Probably Damaging
PRICKLE3	Prickle Homolog 3 (Drosophila)
PRRT2	Proline Rich Transmembrane Protein 2
PsD	Possibly Damaging
pSE	pSpliceExpress
Pwo	*Pyrococcus woesei*
Q	Glutamine
qRT-PCR	Quantitative Real Time PCR
R	Arginine
RNA	Ribonucleic Acid
RNPC3	RNA-Binding Region (RNP1, RRM) Containing 3
RPTPβ	Receptor Tyrosine Phosphatase Beta
Rsa	*Rhodopseudomonas sphaeroides*
RT-PCR	Reverse Transcription PCR
S	Serine
S1PR1	Sphingosine-1-Phosphate Receptor 1
SC	Synonymous Coding
SCN10A	Sodium Channel, Voltage Gated, Type X Alpha Subunit
SCN1A	Sodium Channel, Voltage Gated, Type I Alpha Subunit
SCN1B	Sodium Channel, Voltage Gated, Type I Beta Subunit
SCN2A	Sodium Channel, Voltage Gated, Type II Alpha Subunit
SCN3A	Sodium Channel, Voltage Gated, Type III Alpha Subunit
SCN5B	Sodium Channel, Voltage Gated, Type IV Beta Subunit
SCN7A	Sodium Channel, Voltage Gated, Type VII Alpha Subunit
SCN8A	Sodium Channel, Voltage Gated, Type VIII Alpha Subunit
SDS	Sodiumdodecylsulphate
SEMA4C	Semaphorin 4C
SLC25A25	Solute Carrier Family 25 (Mitochondrial Carrier; Phosphate

	Carrier), Member 25
SLC7A10	Solute Carrier Family 7 (Neutral Amino Acid Transporter Light Chain, Asc System), Member 10
SMEB	SMEI-Borderline
SMEI	Severe Myoclonic Epilepsy of Infancy
SNAP25	Synaptosomal-Associated Protein, 25kda
SNP	Single Nucleotide Polymorphism
SP	Splice Site
T	Threonine
T	Thymine
Taq	*Thermus aquaticus*
TBC1D22A	TBC1 Domain Family, Member 22A
TBE	Tris-Boric Acid-EDTA
TM	Transmembrane
TMEM131	Transmembrane Protein 131
TRNAT	Transfer RNA Threonine (Anticodon AGU)
USA	United States of America
UV	Ultraviolet
V	Valine
VGSC	Voltage Gated Sodium Channel
VNTR	Variable Number of Tandem Repeats
W	Tryptophan
WES	Whole Exome Sequencing
Y	Years
Y	Tyrosine
ZNF792	Zinc Finger Protein 792
ZNF98	Zinc Finger Protein 98

1. INTRODUCTION

1.1. Epilepsy

Epilepsy is a neurological disorder that is characterized by seizures. A seizure is uncontrolled, synchronous and high frequency firing of neurons. These seizures generally start without any sign and end abruptly. The common feature of seizures is the breakdown of mechanisms limiting neuronal activity. Research on etiology of epilepsy continues since ancient times but fundamental features resulting with epilepsy are still not comprehended fully (Eisenstein, 2014). Depending on the origin of seizures in the brain and their types, there are at least a dozen epilepsy types and the differential diagnosis is crucial for the treatment (Savage, 2014).

As most prevalent serious brain disorder epilepsy affects 70 million people in the world, crossing all geographical boundaries. In high-income countries, approximately 6 per 1000 people will develop epilepsy during their lifetime, and 45 people per 100 000 will develop new-onset epilepsy each year. These figures are nearly twice as high in low- and middle income countries. Though it affects people in all ages, the disease onset has two peaks. In developed countries, main peak of onset is in old ages, however, in developing countries infants and children are more affected due to complications before and after birth (Brodie *et al.*, 2012).

Most people with epilepsy have good prognosis, depending on the underlying cause. Childhood epilepsies generally remit, while a small proportion continues throughout the lifetime. Main treatment of epilepsy is achieved through anti-epileptic drugs (AEDs). However, 30-40% of patients do not respond to AED treatment. Standard mortality ratio (ratio of observed deaths to expected deaths) of epilepsy is two or three times above the general population. Seventeen per cent of these deaths are due to sudden unexpected death of epilepsy patients (SUDEP), which is observed more in patients with uncontrolled epilepsy (Brodie *et al*., 2012).

1.1.1. Epilepsy Genetics

Approximately 30% of epilepsy cases have an obvious cause such as trauma or surgery, but remaining 70% lack a defined cause, thus described as idiopathic epilepsy. However, in idiopathic epilepsies genetics is thought to play a significant role though its contribution varies within subtypes. Studies in families suffering from familial types of epilepsies resulted in identification of several genes, most of which are channel genes that are expressed in nervous tissues, or genes associated with channel genes. Thus, idiopathic epilepsies are also termed as channelopathies (Narain, 2014).

Though only 10-40 % of epilepsy cases can be attributed to a gene mutation, researches point out to a genetic contribution ~~to~~ in most of the idiopathic epilepsies. However, epilepsy genetics is complicated with low penetrance and variable expressivity. Two thirds of healthy individuals carry a gene variant associated with epilepsy. Thus, epilepsy is accepted as a complex disorder resulting from the interplay between many genes (Narain, 2014).

Recently developed advanced and highthroughput technologies such as whole genome SNP genotyping, whole genome and whole exome sequencing, together with large projects like 1000 Genomes (Siva, 2008) or Epi4K (McMullin, 2012) resulted in the accumulation of large genomic data yielding the evolution of genetic concepts into genomic concepts. The main outcome of population based studies like 1000 genomes is the detection of unexpected amount of rare variants in normal genomes. Figure 1.1 illustrates the distribution of epilepsy related variants in regard to allele frequency and effect size. Effect size shows the disease contribution of a variant. Variants with low allele frequency and high effect size are located on established Mendelian epilepsy genes. In the other end of the spectrum there are common variants with high allele frequency and low effect size. The most problematic variants in epilepsy genetics are rare variants with low allele frequency and moderate effect size, as established frameworks to assess them are still lacking (Helbig and Lowenstein, 2013).

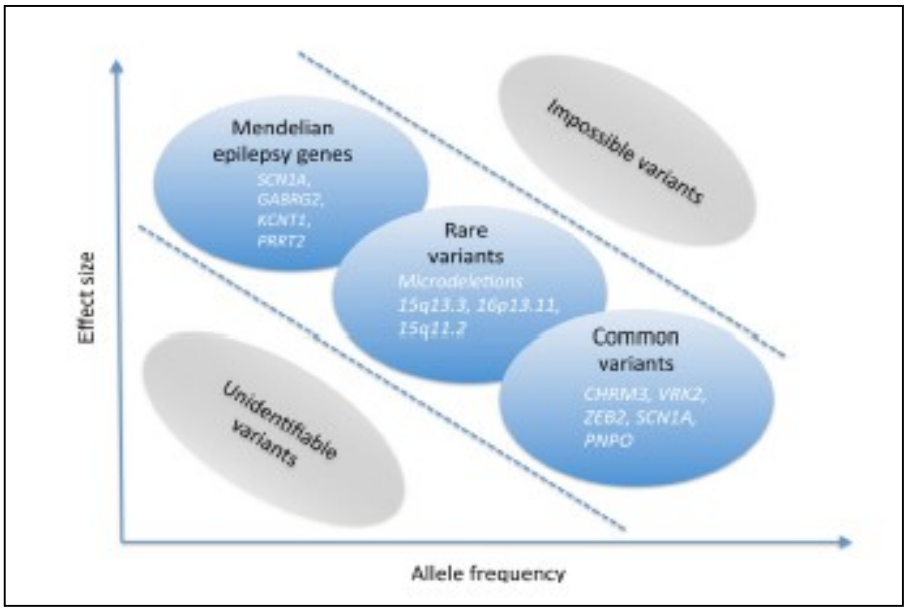


Figure 1.1. Distribution of epilepsy associated variants regarding allele frequency and effect size (Helbig and Lowenstein, 2013).

1.1.2. Epilepsy Syndromes of the First Year

A group of early childhood epilepsy characterized by secondarily generalized focal seizures is defined as benign epilepsy of early childhood. They occur in time dependent manner and disappear after 1 year with benign outcome. There are three different types of this syndrome differentiated according to their age of onset. BFNS starting typically before the fifth day of life, BFNIS (Benign Familial Neonatal Infantile Seizures) occurring between day two and 6 months of age (Kaplan and Lacey, 1983), and benign familial infantile seizures (BFIS) emerging between 3 and 8 months of age (Vigevano *et al*., 1992) Though they have very similar clinical features genetic etiology behind these syndromes is distinct with *KCNQ2* and *KCNQ3* causing BFNS, *SCN2A* driving BFNIS, and *PRRT2* being responsible for BFIS (Maljevic and Lerche, 2014).

Genetic overlap between BFNS, BFNIS and BFIS has become debatable since they have similar disease course but different age of onset. In a large study, Zara *et al*, analyzed 165 patients in 46 families for point mutations in *SCN2A, KCNQ2, PRRT2* and *KCNQ3* genes by sequencing and for deletion/duplication mutations in *KCNQ2* and *KCNQ3* genes by MLPA. Their cohort consisted of BFNS, BFNIS and BFIS patients, the latter making up the majority. They identified mutations in 98% of patients, observed *KCNQ2* mutations in

all three phenotypes and *PRRT2* mutations only in BFIS patients (Zara *et al*., 2013) .

1.1.2.1. Benign Familial Neonatal Seizures (BFNS). BFNS patients generally experience cluster of secondarily generalized focal seizures in first week of life. Generally they do not require treatment and seizures disappear after one month. Seizures often consist of hemi-tonic or –clonic seizures and apnea. Though 15% of patients have recurring seizures later on life the psychomotor development in most cases are benign. Additionally there are increasing number of learning disability cases who suffer of BFNS in early life (Maljevic *et al.,* 2008).

Disease is inherited in autosomal dominant manner and with 85% penetrance and most patients have a mutation in *KCNQ2* gene (87 reported mutation;) generally point mutations but deletions and duplications in some cases. Minority of patients have point mutation in *KCNQ3* gene (6 mutations which forms heterotetrameric channel with *KCNQ2*. These channels particularly expressed in the axon initial segment and nodes of Ranvier of glutamatergic neurons, determines the M-current, and influences subthreshold properties and membrane potential at rest (Maljevic and Lerche, 2014; Maljevic *et al*., 2008).

Though having a mutation on an ion channel implicated in action potential explains the epilepsy phenotype it is still a question why seizures occur in time dependent manner. Possible explanations are (1) increasing number of potassium channel expression on axons during maturation renders neurons less vulnerable to seizures, (2) a link to the developmental switch from GABAergic excitation to inhibition as the M-current might provide a primary inhibitory pathway in the immature brain; and (3) differential expression of splice variants in fetal and adult brain (Weber and Lerche, 2008).

1.1.2.2. Benign Familial Infantile Seizures (BFIS). Benign Infantile Seizures (BIS) is a disorder that is characterized by partial seizures occurring in the first two years of life, which was described by Fukuyama in 1963 (Fukuyama, 1963). In 1992 Vigevano described the familial form of the disease and named the syndrome as Benign Familial Infantile Convulsions/Seizure (BFIC or BFIS) (Vigevano, 2005). Clinical manifestation of the syndrome is such that patients have normal psychomotor development until seizures,

seizures are attributable to no etiology, the outcome is benign and normal psychomotor development continues after remittance of seizures (Specchio and Vigevano, 2006).

The seizures are characterized by clusters of partial seizures with sub-acute onset between 3 to 20 months of age comprising motor arrest, impairment of consciousness, staring, and convulsive movements. Additionally, slow deviation of head and eyes to one side, diffuse hypertonia, cyanosis and unilateral limb jerks are reported (Vigevano, 2005). No status epilepticus is observed. The seizures occur as clusters of 1 to 3 days, lasting 2-5 minutes and the duration becomes shorter with treatment. The disease ceases within two years of life and in treated patients no other seizures are observed. However, in untreated cases isolated brief cluster of seizures are seen within 1 year (Vigevano, 2005).

BFIS is suggested to have autosomal dominant inheritance (Vigevano, 2005). Several linkage analyses in large families identified responsible loci as 19q12-13 (Guipponi *et al*., 1997) 16p12- q12 (Bennett *et al.*, 2000; Caraballo, *et al*, 2002; Tomita *et al*., 1999). Malacarne and colleagues mapped a novel locus to chromosome 2q24 in eight Italian families (Malacarne *et al*., 2001) and Li and colleagues reported linkage to 1p36.12-p35.1 in a large Chinese family, further demonstrating the genetic heterogeneity of the syndrome (Li *et al.*, 2010). In 2009, Ishii *et al* found a linkage between BFIS and *LGI4* gene which is on 19q13.11, a recurrently reported linkage region. However, they could not report the causative mutation in the *LGI4* gene (Ishii *et al.*, 2010). This gene has also been reported to bear claw paw mutation and a role in peripheral nerve development (Bermingham *et al*., 2006). LGI family genes are expressed in the brain (Herranz-Pérez *et al.*, 2010). Another member of the family, *LGI1* is also associated with Autosomal Dominant Partial Epilepsy with Auditory Features (ADPEAF) (Fukata *et al*., 2006).

Malafosse showed that BFIS is not allelic to Benign Familial Neonatal Seizures (BFNS) a similar epilepsy syndrome with seizure onset in the first days of life, excluding the marker on chromosome 20 (Malafosse *et al*., 1994) . *SCN2A* (2q24.3) mutations had been previously identified in BFNIS (Benign Familial Neonatal-Infantile Seizures) patients (Heron *et al*., 2002). In 2006 Striano et al showed that BFNIS and BFIS have overlapping clinical features and identified a novel *SCN2A* mutation in a family with BFIS *(*Striano *et al.*, 2006). Lastly, Zhou *et al.* found a new causative mutation in a Chinese family with

BFIS in the *KCNQ2* (20q13.3) gene which later was shown to be a major gene for BFNS (Heron *et al*., 2007; Zhou *et al*., 2006).

PRRT2 (Proline Rich Transmembrane Protein 2) gene is located on chromosome 16 in the previously identified loci for BFIS and Paroxysmal Dyskinesia and the first report of a *PRRT2* causative mutation for Paroxysmal Dyskinesia was reported in 2011(Chen *et al*., 2011; Wang *et al.*, 2011). Following studies showed that *PRRT2* mutations were found in about 80% of BFIS families in various populations (de Vries *et al*., 2012; Heron *et al*., 2012; Schubert *et al*., 2012). In their 2014 review Nobile and Striano summarized all forty-seven different *PRRT2* mutations that have been reported, including 25 frameshift or nonsense mutations (53%), 16 missense mutations (34%), and six splice site mutations (13%). The result was that 85% of all reported *PRRT* mutations were truncating mutations occurring at c.649 mutation hotspot and 82% of these mutations were the c.649insC, p.R217PfsX8 mutation. In many families unaffected carriers of *PRRT2* mutation were frequently reported. However, this situation was attributed to the difficulty of diagnosis of an infantile seizure (Nobile and Striano, 2014).

Though several mutations of *PRRT2* is reported in a short time after its discovery, not much was known about the function of *PRRT2* gene (112476), except that it is a proline rich protein and contains a putative transmembrane domain. It is located on Chromosome 16p11.2 and contains 3 or 4 exons (Figure 1.2.). It spans approximately 4kb genomic region. It codes for three isoforms at length of 340, 394 and 299 amino acids (Genbank, Swissprot).

First studies demonstrated that it interacts with *SNAP25*, a synaptosomal protein and truncating mutation disrupts this interaction (Lee *et al*., 2012). On the other hand, in a recent study on a PKC mutation, Li *et al* explored the function of *PRRT2* in more detail. They reported that PKC patient with *PRRT2* mutation has elevated glutamate levels in plasma. Furthermore, mice cortex neurons with siRNA silenced Prrt2 gene release glutamate to culture media. So they explored presence of Prrt2 on glutamergic neurons. They conducted double immunostaining on frozen sections of mouse brain and saw that Prrt2 co-localized together with vGlut1, a marker for presynaptic glutamatergic neuronal membrane and also PSD-95, post-synaptic glutamatergic neuronal marker. Furthermore

they demonstrated that truncating mutations disrupt the membrane localization of *PRRT2* protein. They conclude that epilepsy phenotype due to *PRRT2* mutations may be attributed to disrupted glutematergic signaling (Li *et al.,*2015).

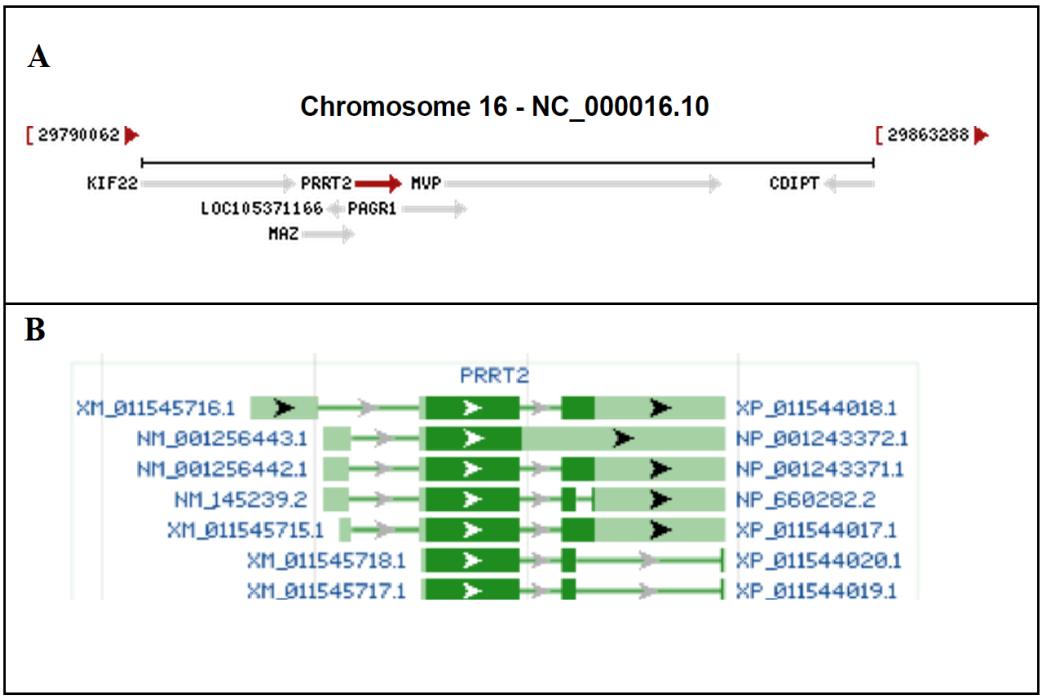


Figure 1.2. Location of *PRRT2* gene and its neighboring genes on chromosome 16 (Panel A) and exon structure of three splice variants (Panel B). Green bars represent exons, dark green represents coding region and light green is non-coding region, green lines represent introns. Reference numbers of cDNA and protein sequences are shown before and after each transcript, respectively, adapted from NCBI, Genbank).

1.1.3. Genetic Epilepsy with Febrile Seizures Plus (GEFS+)

Seizures provoked by fever are common among infants and is the cause of most of the childhood seizures. Normally, febrile seizures ceases by the age of 5, however, some patients have seizures after the critical age of 5 or 6, together with other types of afebrile seizures. This condition is named as Generalized Epilepsy and Febrile Seizures Plus (GEFS +) (Scheffer and Berkovic, 1997).

GEFS+ is an autosomal dominant disorder with 50-60% penetrance. It is characterized by the continuation of generalized febrile seizures beyond the age of 6 and

afebrile generalized convulsions, positive family history of epilepsy with variable phenotypes and benign outcome in most cases (Korff and Nordli, 2006). GEFS+ patients have febrile seizures beyond the age of 6 and in most cases afebrile generalize tonic clonic seizures (GTCS) also accompanies FS or occurs after remission of FS. Even there may be several years of break between febrile seizures and afebrile GTCS. As in the scope of generalized seizures, distinct absence seizures, myoclonic or atonic seizures are also observed. Besides generalized seizures, some focal seizures including temporal lobe epilepsy and frontal lobe epilepsy may be observed in GEFS+ patients or family members. The GEFS+ spectrum also includes more severe phenotypes like epileptic encephalopathies, myoclonic-astatic epilepsy (MAE) and Severe Myoclonic Epilepsy of Infancy (SMEI) *(* Scheffer *et al.,* 2005).

Inheritance of GEFS+ is somehow complicated. The large Australian family reported by Schaeffer and Berkovich demonstrates the complexity associated with genotypic and phenotypic heterogeneity within the syndrome. The family comprised 4 generations, with 25 individuals showing Febrile Seizures plus phenotype. However, some of the family members have afebrile seizures (Scheffer and Berkovic, 1997). They identified a mutation on chromosome 19 in sodium channel β1 subunit gene *SCN1B* in this family and demonstrated that amino acid substitution due to this mutation enhanced sodium current flow through the channel and cause membrane hyperexcitabilty. The family included two individuals showing epilepsy phenotype without the mutation and one asymptomatic individual carrying the mutation. It was remarkable to observe that a single mutation may yield to a variety of phenotypes ranging from benign febrile seizures to more severe myoclonic-astatic epilepsy, implying a major causative mutation and one or more modulatory regions altogether yielding the phenotype (Stafstrom, 2008).

Voltage Gated Sodium Channel Alpha Subunit-1 gene, *SCN1A* (6323), is the most important GEFS+ gene to date and a number of mutations that are scattered all through the gene have been reported on *SCN1A* gene associated with GEFS+ families. However, phenotypic expression of these mutations ranges from benign Febrile Seizures to Severe Myoclonic Epilepsy of Infancy (SMEI) (Schaeffer *et al.*, 2005). Additionally, Sugawara *et al*. (2001) reported a mutation on *SCN2A* gene encoding sodium channel alpha-2 subunit in a GEFS+ patient. Though linkage was not powerful due to small sample size, functional

study highly supported the pathological consequence of *SCN2A* mutation in seizure phenotype (Audenaert *et al.,* 2006).

Table 1.1. Genetic loci and genes identified in GEFS+ families (Adopted from Nakayama, 2009)

Chromosomal Region	**Gene**	**Country**
19q13.1	*SCN1B*	Australia
2q21-q33	*SCN1A*	France
2q24	*SCN2A*	Japan
5q31.1-q33.1	*GABRG2*	France
		Australia
1p36.3	*GABRD*	Australia
2p24	Unknown	Belgium
	GABRA6	(Dibbens *et al*., 2009)
22q13.31	*TBC1D22A*	Tunisian, (Belhedi *et al*., 2013)

$GABA_A$ receptor is a ligand gated chloride channel that has mainly inhibitory function in CNS. There have been mutations associated with GEFS+ both in genes of gamma subunit and delta subunit. Mutations in gamma subunit gene *GABRG2* is also associated with Idiopathic Generalized Epilepsy (IGE) phenotype. On the other hand, one of the *GABRG2* mutations was identified in a family where the proband has SMEI. Recently GEFS+ associated delta subunit gene, *GABRD*, mutations have been reported in two families by Dibbens *et al.* (2004). Electrophysiological studies with mutant delta subunits co-expressed with recombinant $GABA_A$ receptors in Human Embryonic Kidney (HEK) cells show reduced current when exposed to a saturating concentration of GABA for one mutation (E177A) but not for the other (R220C) (Audenaert *et al*., 2006; Scheffer *et al.*, 2005).

In summary, several studies identified mutations on several loci and sodium channel and GABA receptor genes (Listed in Table 1.1).

1.2. Voltage Gated Ion Channels

Intrinsic properties of a neuron determine its current conducting properties. Among these properties the composition of voltage gated ion channel on a neuron is the key feature that modulates membrane potential of any excitable cell. However, intrinsic excitability of a cell is not a static but rather a dynamic process, which gives the plasticity to nervous system as learning, memory, sensory adaptation etc. There are mainly four types of voltage gated ion channels namely Na, K, Cl and Ca channels (Schulz *et al.* 2008).

Voltage gated ion channels are responsible for action potential initiation and propagation. When action potential arrives on the presynaptic membrane it triggers Ca^{2+} influx and cause neurotransmitter release, which binds to ligand gated postsynaptic gates and produces another action potential. Excitatory neurons communicate via glutamate and acetylcholine and inhibitory channels communicate through γ-aminobutyric acid (GABA) and glycine. Furthermore different cells have distinct voltage gated ion channel composition, and this distinct composition determines the biophysical characteristics of a neuron. Not surprisingly, majority of genetic defects identified in idiopathic epilepsy patients rely on channel genes. Any mutations that alter channel characteristics modify signal propagation in central nervous system and could start a synchronous undesired action potential firing (Maljevic and Lerche, 2014).

In normal state, excitability properties of a cell are regulated over regulation of ion channel properties. First of all each ion channel family has a plethora of subtypes which has distinct excitability properties. The expression of these subtypes is temporally and spatially regulated in response to several factors including electrical activity within the cell, injury to the cell, tropic factors and hormonal influences. Furthermore, alternative splicing is another mechanism to produce different types of ion channels. Ion channel properties are also modulated by post-transcriptional modulations, the most profound is phosphorylation and subcellular localization of ion channel within a neuron is also a distinct feature and changes upon channel activity (Schulz *et al*., 2008).

1.2.1. Voltage Gated Sodium Channels

Voltage gated sodium channels are crucial component of neurons since they form the basis of initiation and propagation of action potentials. Due to this fact they are the best candidates in episodic current disorders like epilepsy. There are several sodium channel gene mutations reported as implicated in several types of epilepsy, which are listed in Table 1.2.

Table 1.2. Epilepsy phenotypes caused by voltage-gated sodium channel mutations (Non-epileptic phenotypes are not listed) (Adapted from Steinlein, 2014).

Subunit class	Gene	Channel subunit	Epilepsy phenotypes
a-Subunits	*SCN1A*	Nav1.1	Febrile seizures
			GEFS+
			Dravet syndrome
			SMEB
			West syndrome (infantile spasms)
			Doose syndrome (myoclonic astatic epilepsy)
			Intractable childhood epilepsy with generalized tonic–clonic seizures (ICEGTC)
			Rasmussens's encephalitis
			Lennox–Gastaut syndrome
	SCN2A	Nav1.2	Benign familial neonatal–infantile seizures
			Early infantile epileptic encephalopathy
			Benign familial infantile seizure
	SCN8A	Nav1.6	Infantile epileptic encephalopathy
b-Subunits	*SCN1B*	Navb1	GEFS+

Voltage gated sodium channels are made up of two subunits. Alpha subunits are responsible for channel formation and bear the voltage sensor. Alpha subunit of voltage-gated sodium channel genes comprises nine homologous members (*SCN1A* to *SCN5A* and *SCN8A* to *SCN11A*) that encode the sodium selective ion channel subunits NaV1.1 to NaV1.9. These proteins consist of four domains each made up of 6 transmembrane regions. Alpha subunits are able to form functional channels alone. On the other hand beta subunits are considered as auxiliary and modulate channel trafficking and biophysical characteristics. Beta subunits are coded by 5 genes (*SCN1B* to *SCN5B*) (Steinlein, 2014).

Figure 1.3 represents the structure of voltage gated sodium channels. In this Figure 2 dimentional structure of alpha (shown on the left) and beta (shown on the right) subunits are depicted. Transmembrane domains are shown as colored cylinders. The green rectangles indicate the amino acids important for fast inactivation. The dark blue diamonds represent the mutations that affect slow inactivation (V754I, V787K/C, R1454C and A1529D). The light blue triangles indicate the mutations in Nav1.5 that cause long QT or Brugada syndrome (L567Q, R1232W and Y1795C/H/ insertion). The red circles represent the mutations that cause GEFS+ (C121W in the in the β1 subunit and T875M, W1204R and R1648H in the Nav1.1 a subunit) (Steinlein, 2014).

Sodium channels are closed at resting state and opened upon depolarization of cells. Opening of the cell creates a sodium influx inside the cell rapidly and enforce depolarization. The channels close very rapidly within millisecond but do not close totally. Remaining slow flux continue for tens of seconds. Rapid opening and closing of the channels provide a stringent control over the action potential (Steinlein, 2014).

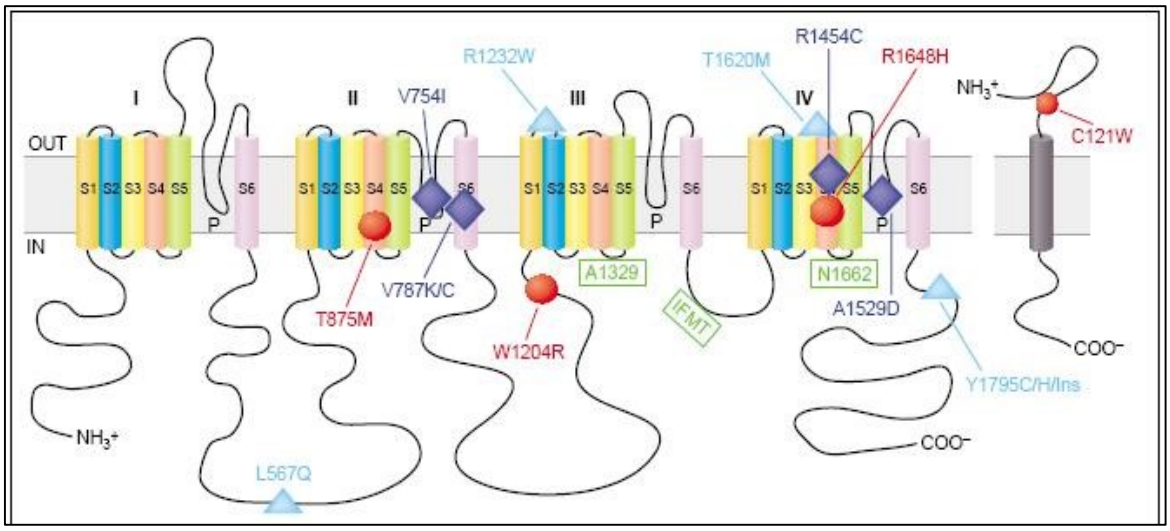


Figure 1.3. Schematic diagram of the sodium channel α and β subunits.

Different subtypes of alpha subunits are expressed at different times and locations. For example while *SCN3A* is predominantly expressed in neonatal brain, *SCN1A*, *SCN2A* and *SCN8A* are expressed in adult brain but in distinct brain and subcellular regions (Steinlein, 2014).

1.2.1.1. Voltage Gated Sodium Channel Beta Subunits. VGSC are heterotrimers that contains one α subunit that covalently bound to β2 or β4 subunits and non-covalently bound to β1 or β3 subunits. β1- β4 subunits are coded by *SCN1B-SCN4B* genes, respectively. They are one span transmembrane proteins with N terminal signal peptide and Ig domain, a transmembrane domain and a C-terminal intracellular domain (Figure 1.3). β1 has a splice variant β1B which is produced by intronic retention after exon 3 and omitting exons 4-6; thus lacking transmembrane and intracellular domains. β1B is thought to be a secreted protein (Patino and Isom, 2010).

VGSC beta subunits are expressed in excitable and non-excitable cells in the nervous system and their expression is temporally regulated. β1B and β3 mRNAs dominate fetal brain which are later replaced by β1 and β2 after birth. Their expression is not dependent on α subunit expression or neuronal signaling. Throughout their Ig domains, beta subunits interact with other cell adhesion molecules and components of extracellular matrix and through their C-terminal they interact with cytoskeletal structures (Pattino and Isom, 2010).

It has been shown that beta subunits are substrates for protein cleaving enzymes. Upon cleavage, TM and extracellular domains are shed into the extracellular space, presumably functioning as soluble ligand of cell adhesion resulting into neurite outgrowth. Inhibition of this cleavage results in reduced cell-cell adhesion and migration. On the other hand, intracellular domains also translocate to nucleus and function as transcriptional regulator of alpha subunits. It is known that β subunits modulate Na^+ currents and also channel cell surface expression. However, this regulation is cell type and subunit type specific (Patino and Isom, 2010).

Though considered as auxiliary subunits of VGSCs, another major function of β subunits are in cellular migration, neurite outgrowth and axonal fasciculation. For example, Scn1b null mice exhibit significant defasciculation of the corticospinal tract, abnormal migration of CGNs, and defasciculation of cerebellar parallel fibers. Interestingly, it is postulated that β subunits modulate migration of cancer cells in similar mechanisms (Patino and Isom, 2010).

1.2.1.2. *SCN1B* Gene and Voltage Gated Sodium Channel Beta 1 Subunit. *SCN1B* (6324) gene is located on chromosome 19q13.q, spans approximately 11kb genomic region. β1 splice variant (NM_001037.4) has 6 exons and codes for 218aa protein. β1B splice variant (NM_199037.3) occurs by retention of intron 3, has 3 exons and codes for 268aa protein. These two protein have identical N-terminal up to 149th amino acid, but C-terminal sequence only has 17% sequence similarity (NCBI, OMIM, (Qin *et al*., 2003).

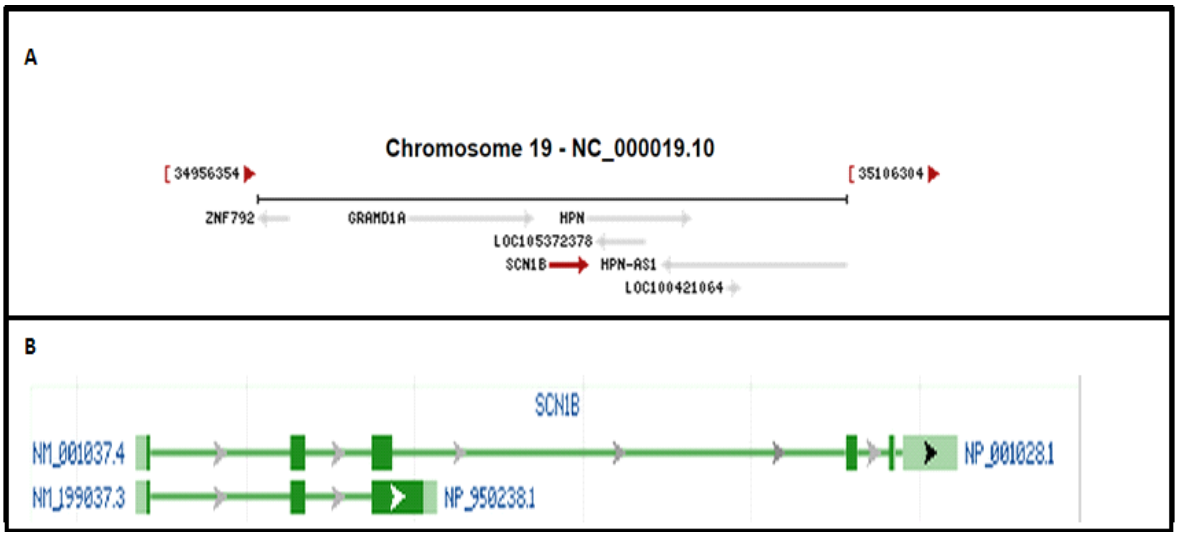


Figure 1.4. Location of *SCN1B* gene and its neighboring genes on chromosome 19 (Panel A) and exon structure of three splice variants, β1 (upper) and β1 (lower) (Panel B) (Adapted from NCBI, Genbank).

Chen *et al.* (2004) produced beta-1-null mice by gene targeting. Knockout mice exhibited ataxic gait, spontaneous seizures, growth retardation, and death around postnatal day 20. Besides these symptoms affected mice had slowing in action potential and abnormalities in alpha subunit gene expression in certain neurons. Furthermore, they showed esophagi, which may be attributed to enteric nervous system involvement (Chen *et al*., 2004).

Pattino *et al.* (2009) reported a homozygous mutation (p.R125C) on *SCN1B* gene in a Dravet syndrome patient, which yielded no cell surface expression though normal cellular expression. The mutation was inherited from healthy parents (Patino *et al*., 2009). Another Dravet patient with homozygous *SCN1B* mutation was reported by Ogiwara *et al.* on 2012. Fendri-Kriaa *et al.* reported another mutation (p.R125L) in heterozygous state in two

patients with Febrile Seizures and one with Febrile Seizures plus and suggested this variation as a susceptibility variant in Tunisian population (Fendri-Kriaa *et al*., 2011). Wallace *et al.* reported homozygous C121W variation in a GEFS+ patient (Wallace *et al*., 1998). There are reports of heterozygous mutations in *SCN1B* gene for temporal lobe epilepsy (TLE) and Genetic epilepsy with febrile seizures plus (GEFS+) (Scheffer *et al*., 2007) (R. Wallace, Scheffer, and Parasivam, 2002). Audenaert *et al.* screened 74 unrelated probands with GEFS+, FS, or FS+ and they identified a 5 nucleotide deletion in the Ig domain (Audenaert *et al*., 2003). Finally, Watanabe *et al.* reported R85H and D153N variations in atrial fibrillation patients (Watanabe *et al*., 2009). All these variants show the diverse nature of *SCN1B* related phenotypes.

All variations listed above are on shared region between β1 and β1B splice variant (Figure 1.5). On the other hand, on 2011 Pattino *et al.* reported variation (p.G257R) unique to β1B variant in two idiopathic generalized epilepsy (IGE) patients, however, they had juvenile onset. The variation resulted into a trafficking deficient variant (Patino *et al*., 2011). β1 contains residues responsible for interaction with α subunit in its intracellular and extracellular domains,which is depicted in Figure 1.5 Mutation sites responsible for causing genetic epilepsy with febrile seizures plus (GEFS + 1), temporal lobe epilepsy (TLE), and Dravet syndrome are located in the extracellular immunoglobulin loop. Alternative splicing site for β1B, putative palmitoylation site, ankyrin interaction, tyrosine phosphorylation site, N-glycosylation sites (ψ), α/β/γ-secretase cleavage sites, receptor protein tyrosine phosphatase β (RPTPβ) interaction, and putative fyn kinase interaction are also marked (Brackenbury and Isom, 2011).

β1 and β1B has complementary expression in both rat (Kazen-Gillespie *et al*., 2000) and human brain (Pattino *et al.,* 2011). In fetal human brain at gestational week 22, β1B dominates over β1. However, as development progresses β1B level stays same but β1 level increases. In postnatal human brain, β1 transcript is 18 fold higher than β1B (Patino *et al*., 2011) (Figure 1.6.).

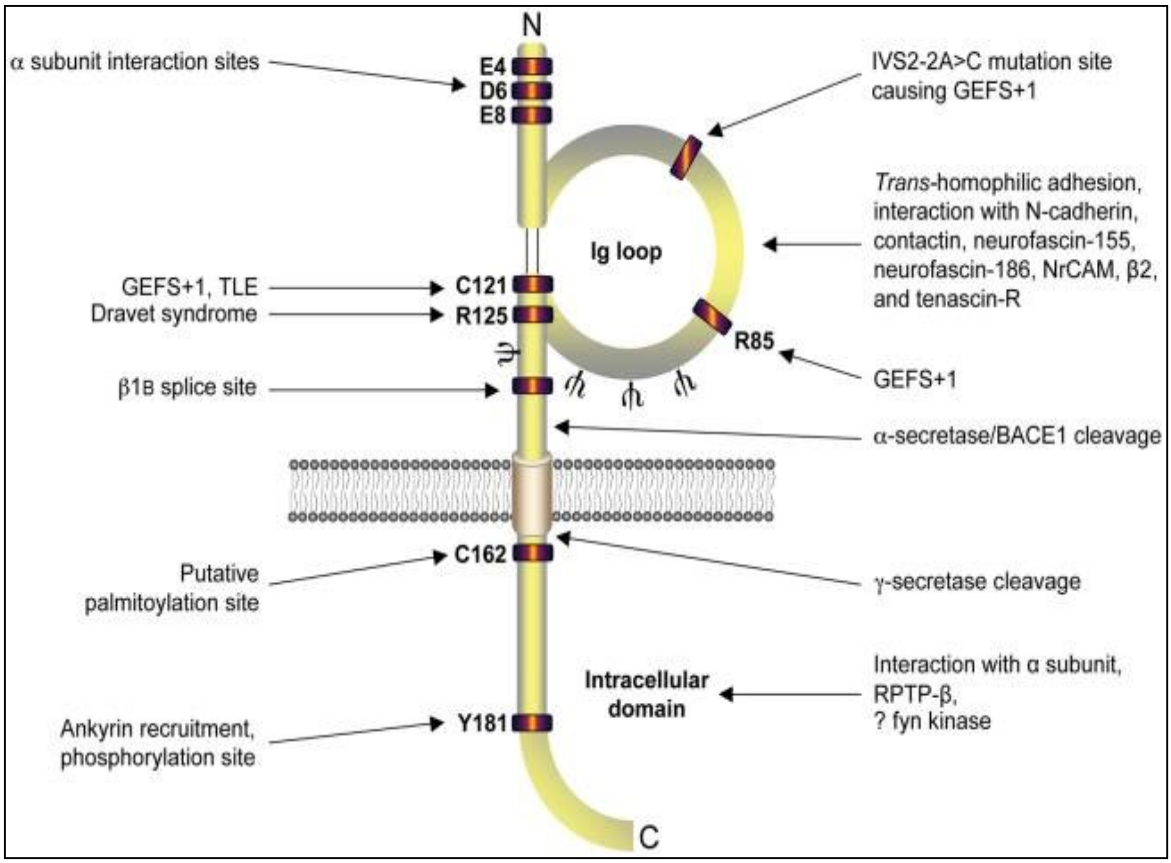


Figure 1.5. Functional architecture of β1/β1B.

1.2.2. Voltage Gated Potassium Channels

There are approximately 80 genes in the human genome that code for a potassium channel and products of these channels are distributed throughout the body. Based on number of transmembrane domains (TM) potassium channels are named as 2TM, 4TM and 6TM or 7TM families, all suspected to evolve from a 2TM ancestral gene through gene duplication. The specificity of these channels for K+ over other cations is defined by a highly conserved amino acid sequence, the so-called GYG signature sequence, which enables selective transmission of K^+ by replacing the six water molecules that surround these ions (Gutman, 2005) (Jegla *et al.*, 2009).

Potassium channels are activated by several different ligands including ATP, Ca^{2+} and also by voltage. The largest family of K+ channels is voltage gated (KV) channels; encoded by 40 genes and they consist four α subunits each containing 6TM regions. α subunits form a single pore and on each subunit 4th TM region contains the voltage sensor

which bears positively charged arginine residues. Upon alteration in membrane potential this region goes into conformational change and alters the state of channel gate (Figure 1.7.). Potassium channel diversity is further increased by forming dimer or tetramer interaction with several auxiliary subunits, besides posttranscriptional modifications and mRNA splicing (Gutman, 2005).

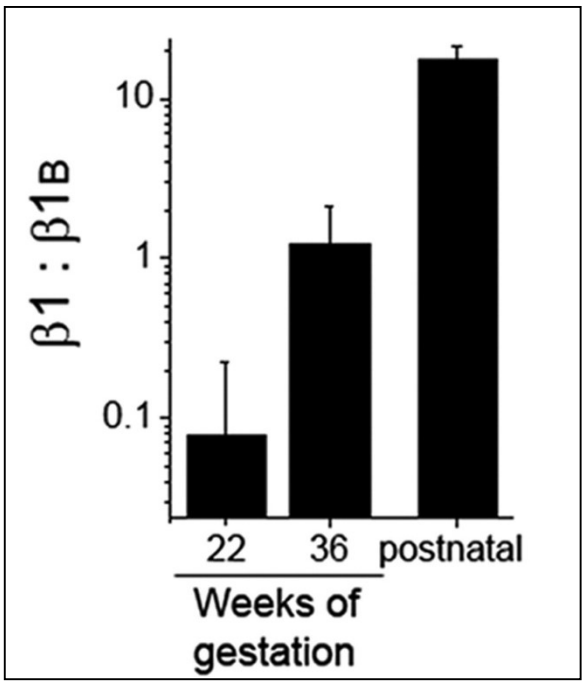


Figure 1.6. β1B is the predominant *SCN1B* splice variant during human fetal brain development (Patino *et al*., 2011).

Kv channels co-localize with Na^+ and Ca^{2+} channels in excitable membranes and are responsible for membrane repolarization or hyperpolarization after depolarization by Na^+ current. This phenomenon is achieved by two currents; faster inactivating A-currents and noninactivating slow M currents, which is activated in subthreshold currents. In a spike train A-currents have higher impact on action potential, whereas M currents determine the response to multiple spikes by increasing the threshold for action potential firing (Bean, 2007). Typical A-type KV channels are found in KV1–KV4 subfamilies, while KV7 (KCNQ) and KV11 (hERG) produce the M-currents (Figure 1.7) (Shieh *et al.*, 2000). Distinct KV subtypes are present in distinct regions of central and peripheral nervous system both in neurons and glial cells and they regulate not only membrane potential also function in Ca^{2+} signaling, secretion, proliferation and migration. Besides, several factors

including interaction with other nerve cells determine the subcellular domain localization of KV channels (Jensen *et al.,* 2011).

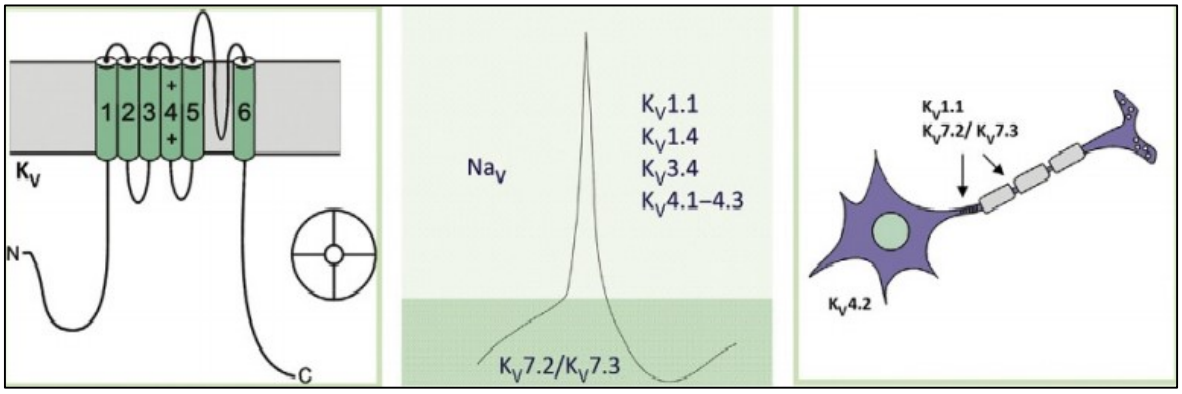


Figure 1.7. Structure and function of Kv channels. Typical structure of Kv channel show on the left, function of different Kv channels on different phases of action potential is shown in the middle and subcellular localization of distinct Kv channels is shown on the right (Maljevic and Lerche, 2014).

<u>1.2.2.1. KCNQ2 (Voltage Dependent K+ Channel, KQT-Like Subtype Member 2) Channels.</u> *KCNQ2* (3785) is one of the five voltage-gated delayed rectifier K+ channels (KV7.1–5), encoded by the KCNQ gene family and often referred to as *KCNQ1–5* channels. As indicated above they are responsible for M-currents. KV7.2 channels are expressed in different parts of the brain including hippocampus, cortex, and thalamus, in both inhibitory and excitatory neurons and can form homo or heterotetrameric channels. Like other voltage gated potassium channels they are 6TM channels with voltage sensor at 4th TM region. Their N and C terminals are positioned intracellularly. C terminus is exceptionally long; it bears assembly region, which provides tetramerization of subunits and many other regulatory regions which involved in the trafficking and gating of these channels (Haitin and Attali, 2008, Maljevic and Lerche, 2014).

KCNQ2 gene is located on 20q13.33 and has at least 18 exons, occupying more than 50 kb of genomic DNA. Splice variants were identified. For example, in fetal brain, exon 8 was absent in all transcripts, while this exon was present in clones derived from adult brain RNA (Biervert and Steinlein, 1999) (Figure 1.8).

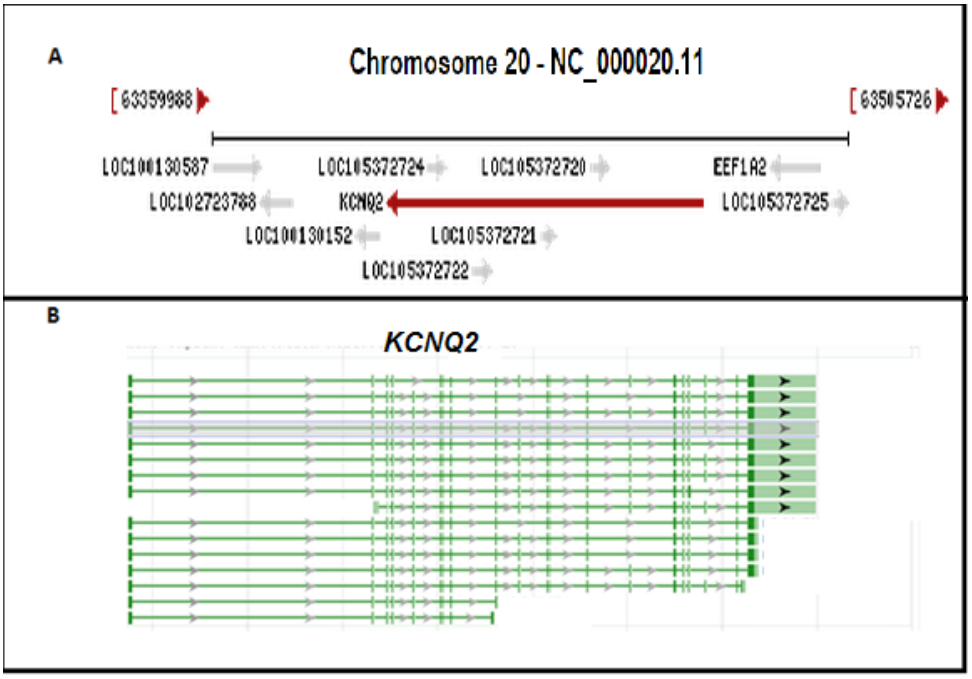


Figure 1.8. Location of *KCNQ2* gene and its neighboring genes on chromosome 20 (Panel A) and exon structure of three splice variants (Panel B), (Adapted from NCBI, Genbank).

Rodent studies showed that kcnq2 and kcnq3 (a common partner) channels are found at low levels at birth but increase within first weeks of development. Homozygous deletion of kcnq2 gene in mouse result in death of pups after birth due to pulmonary atelectasis. Hemizygous mice live after birth and so not have spontaneous seizures. However, they are more sensitive to seizure provoking agents. Interestingly homozygous deletion of *KCNQ3* gene do not have an evident phenotypic effect on mice (Watanabe *et al*., 2000) (Tzingounis and Nicoll, 2008). Both homozygous and heterozygous deletions of kcnq2 gene resulted in increased kcnq3 and kcnq5 mRNA levels, implying a compensation mechanism (Robbins, *et al.*, 2013).

Two knock-in mouse model of Benign Infantile Neonatal Seizures (BFNS) mutations in kcnq2 and kcnq3 genes produced mouse with spontaneous seizures but not confined to a period of time. Additionally these mice were more sensitive to seizure inducing agents (Singh *et al*., 2008). However these phenotypes are sex, strain and seizure test dependent (Otto *et al*., 2009).

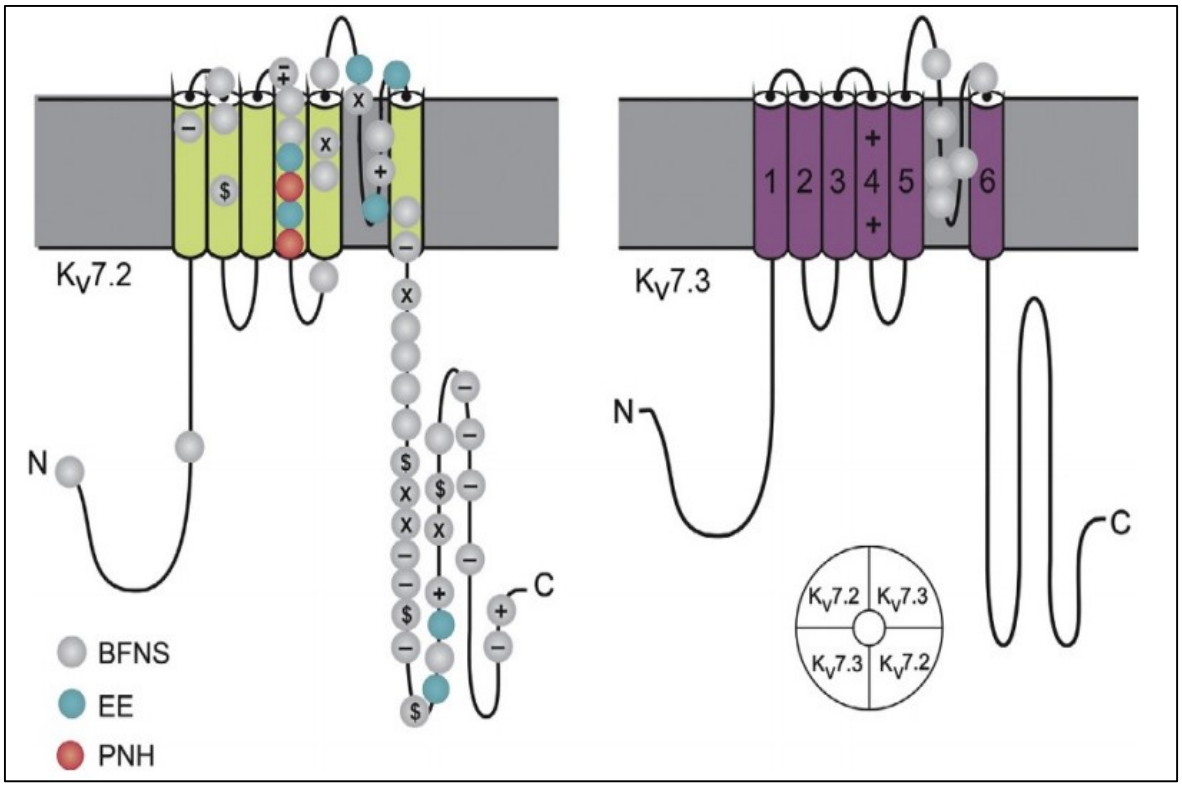


Figure 1.9. Location of disease causing mutations of *KCNQ2* and *KCNQ3* genes represented on Kv7.2 and Kv7.3 channels (Maljevic and Lerche, 2014).

Mutations in *KCNQ2* gene is mainly associated with a benign genetic epilepsy starting in first days of life, BFNS which will be detailed later on. However, there are reports on *KCNQ2* implications in more severe epilepsy phenotypes including peripheral nerve hyper excitability (PNH) and myokymia, Rolandic epilepsy with centrotemporal spikes and Infantile Epileptic Encephalopathy. While BFNS mutations were generally inherited, mutations associated with more severe phenotypes were mostly de novo (Reid *et al.*, 2009; Weckhuysen *et al.*, 2012) (Figure 1.9).

1.3. Current Methods in Epilepsy Genetic Research

Together with the advances in genomic research, concepts in epilepsy research including syndrome definitions and interpretation of pathological mutations have been revolutionized. The most classical method of gene discovery, positional cloning has been replaced by SNP / CNV arrays and linkage/association studies which yielded discovery of many important genes and pathways in epilepsy etiology. However, just after a decade with the advent of massive parallel sequencing technologies, these revolutionary methods

became old technology (Helbig, 2014).

Beside producing large amount of genetic data, new generation sequencing technologies altered perceptions in definitions of disease associated mutations, rare variation and predisposition. Furthermore, large amount of data about our genome added further layer of complexity on its comprehension. At the beginning of 21st century a *de novo* rare variation on a plausible gene was directly accepted as pathological mutation, however, today it is known that every individual have 1-2 % *de novo* variations in their genome and 10% of a healthy genome is copy number variable. This situation necessitates additional methods of mutation confirmation (Helbig, 2014).

1.3.1. Whole Exome Sequencing

Approximately 1 % of the human genome is made of exons and codes for protein. However, 85% of all reported mutations reside in exons. Because of this, current sequencing technologies focused on sequencing only the exons to get maximum information with less~~er~~ effort. This technology is called Whole Exome Sequencing (WES) (Choi *et al*, 2009). However, due to their intrinsic properties the first exons of most genes and some genes could not been targeted. Though conceptually WES targets all exons, in reality 5-10% of human coding sequence cannot be captured by this current technology. However, there is no report of a mutation that is missed by exome sequencing and captured by other methods like whole genome sequencing (Helbig, 2014).

WES method depends on the enrichment of exons using specific probes targeting exons of known genes. After enrichment DNA fragments are sequenced with a massive parallel sequencing platform. Target regions and coverage of all exome varies between different exome enrichment kit providers (Choi *et al*., 2009).

The main challenge of WES is the identification of causative mutation from the background of several benign variants. A WES analysis yields approximately 20000 variant for a Caucasian individual; 95% of which are known variants. But remaining 5%, which are rare or novel variations still makes up a large number (Bamshad *et al*., 2011).

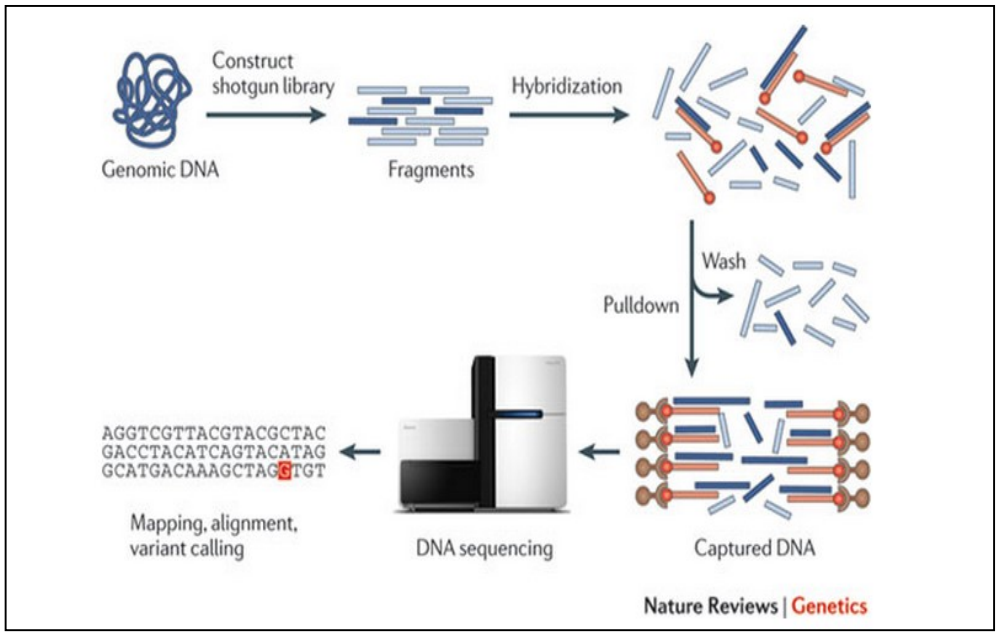


Figure 1.10. Whole Exome Sequencing method (Adapted from Bamshad *et al.,* 2011).

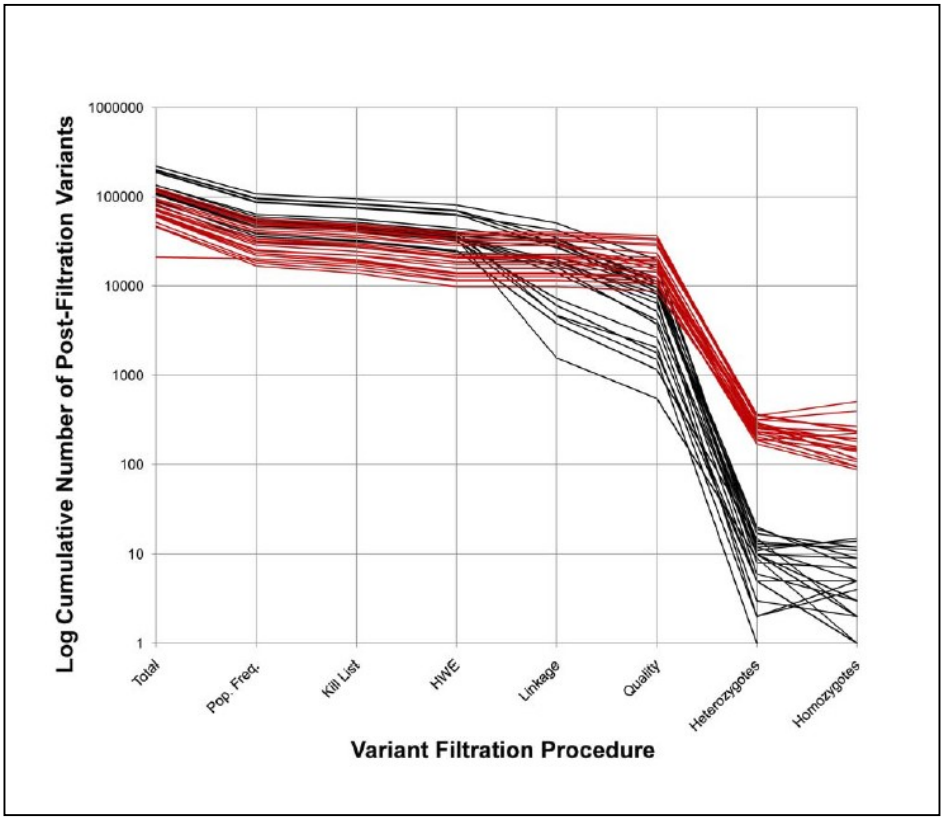


Figure 1.11. Graph of variant reduction by variant filtration procedure. X-axis show different approaches applied in sequential order. Y-axis shows log cumulative number of post-filtration variants. Red lines represent to individual WES data whereas black lines represent family based WES data (Adapted from Adams *et al*. 2012).

There are several filtering strategies proposed for WES data in order to reduce number of variants. Figure 1.11, adopted from Adams *et al* (2012) shows the sequential implementation of several filtering strategies. In this graph individual WES data are depicted as red and family based WES data are in black lines. The striking information in the graph is that at the end of all filters, individual WES data ends up with 100- 800 variants but family based WES data ends up with slightly more than 10 variants. Here the most dramatic reduction in variant number is obtained after linkage analysis (note the logarithmic scale) (Adams *et al*., 2012).

1.3.2. Whole Genome Linkage Analysis

Some diseases are caused by alterations in a single gene and follow the rules depicted by Mendel in the beginning of the twentieth century. However, some other diseases show strong correlation between relatives but their etiology cannot be attributed to single gene alterations, rather multiple genes and non-genetic factors act together giving rise to the disease. As a consequence, while it is easy to identify causative locus for a single-gene disease, some complex experimental approaches and statistical methods are required to identify causative factors underlying the complex diseases. Linkage analysis conducted in large families is the leading tool for the identification of causative genetic patterns inherited together with the disease phenotype (Altshuler *et al.*, 2008).

Linkage analysis was conceived by Sturtevant and applied to fruit flies in 1993 (Altschuler *et al.*, 2008). The basis of the analysis depends on the notion that any loci that are close enough will segregate together in generations. In fruit flies analysis consisted of crossing two flies altering at a single trait and polymorphic markers are analyzed for segregation with the trait. Any marker showing correlated segregation will lie close to the causative locus for the trait. As controlled crosses between humans are not feasible, large families having many affected and healthy individuals are accepted as a large bred analyzed for polymorphic genetic markers segregating together with disease phenotype to identify causative or susceptibility loci (Altshuler *et al*., 2008)

The linkage analysis comprises three steps. The first is the genome-wide analysis by markers to identify causative locus followed by sequencing the identified locus to find causative genetic alterations. Lastly, function of the identified gene at cellular and organismal level is explored to confirm the association of identified genetic alteration with the disease (Altshuler *et al*., 2008).

1.3.3. Array Comparative Genomic Hybridization

Beside changes in DNA sequence, copy number variations (CNV) in genome also cause congenital disorders. Copy number alterations are classified as recurrent and nonrecurrent CNVs. Recurrent CNVs likely arises through nonallelic homologous recombination (NAHR), an unequal crossover between highly similar stretches of DNA referred to as segmental duplications. They are relatively common in the population though speculated to have function in disease predisposition. On the other hand rare nonrecurrent CNVs occur with other mechanisms and they are generally associated with disease (Helbig, 2014).

The most important feature of a CNV detection method is the size of target CNVs. Small CNVs comprising DNA fragments up to a kilo base can be detected by classical sequencing methods. On the other end of the spectrum, karyotype analysis is used to detect deletions or duplications larger than 5-10Mb since 1960s. However detection of micro deletions or micro duplications in lengths between limits of sequencing and karyotype analysis became feasible after the invention of array based comparative genomic hybridization (array-CGH)(Hillman *et al*., 2010).

ArrayCGH method consists of fragmentation of genomic DNA and labeling it with a florescent dye, while reference DNA is also labeled with fluorescent dye in another color. These two labeled DNA samples are mixed in equal ratios and hybridize over an array containing specific probes. The array is scanned by a high resolution camera and resulting image is analyzed for relative intensities of different colors on defined spots containing a known probe. This ratio is then converted into copy number gain or loss data in the analyzed genome (Pinkel *et al*., 1998).

2. PURPOSE

Seventy percent of all epilepsies are classified as idiopathic since they do not appear to have a known cause. Recent advances in human genetics, however, show that the majority of idiopathic epilepsies are genetic. For this reason, genetic research in epilepsy is at upmost importance. On the other hand, due to the intricate structure of the central nervous system and neuronal electrical signaling, besides several complex interactions between genes and proteins, epilepsy is not a straightforward genetic disease. Though some epilepsy syndromes seem to have simple Mendelian inheritance, observed low penetrance and variable expressivity implies more sophisticated biology in the background.

For genetic research in epilepsy, family studies with large pedigrees are the most fruitful, since they produce more robust results with smaller sample numbers. Moreover, family based genetic research has more established tools and concepts. Since epilepsy has more complex inheritance than most genetic diseases, well designed family studies with properly phenotyped individuals is invaluable to reveal disease pathomechanisms. In this thesis we aimed to understand the genetic background in families with three distinct inherited epilepsy syndromes using family based genetic research and current advanced technologies.

In the first part of the thesis we analyzed a multiplex, multigenerational family with BFIS phenotype. BFIS is a benign epilepsy syndrome in which infants between 3 to 7 months of age experience clusters of initially focal, laterally generalized seizures together with eye deviation, apnea and cyanosis. In the second part we analyzed several smaller families with the same phenotype and additionally 5 individuals with BFNS phenotype, which is a similar syndrome with the same clinical manifestations but with a disease onset in the first week of life.

The third part of the thesis comprised analysis of a large kindred with apparent seizure inheritance together with other neuropsychiatric symptoms such as mild intellectual disability, depression and anxiety. The epilepsy syndrome was diagnosed as GEFS+,

which is characterized by febrile seizures beyond the age of 6. The genetic background of several patients in the family, however, manifested other epilepsy subtypes.

In order to elucidate the genetic basis of epilepsy in these families, current methodology such that targeted gene sequencing, family based linkage analysis, whole exome sequencing and copy number analysis with array CGH were applied. Furthermore, minigene analysis to reveal the contribution of a synonymous single nucleotide variation to the splicing efficiency of a putative disease associated gene was conducted.

3. MATERIALS

3.1. Blood Samples and Patient Recruitment

Medical recording and evaluation of BFIS patients and final diagnosis were done by different clinicians. Blood samples of members of 4BF family were provided by Uluç Yiş, MD, from Dokuz Eylül University, Department of Pediatric Neurology. Blood Samples of 5BF family and BFNS samples were provided by Dilşad Türkdoğan, MD, from Neurological Sciences Institute, Marmara University. Blood Samples of other BFIS Families were provided by Sema Saltık, MD, Dilşad Türkdoğan, MD, and Bülent Kara, MD, Department of Pediatric Neurology, Kocaeli University.

Clinical Evaluation of GEFS+ patients were conducted by a team of neurologists including Canan Aykut-Bingöl, MD, Department of Neurology, Berrin Akdoğan, MD, and Berfu Naz Akbaş, MD from Department of Neurology, Yeditepe University Medical School and Kadriye Ağan, MD, from Department of Neurology, Marmara University Medical School.

Blood samples were collected by a nurse under supervision of clinicians. All clinical and genetics studies were conducted according to institutional guidelines of Bogazici University Ethics Committee (granted to project BAP 5722 and TÜBİTAK Project 110S518).

3.2. Oligonucleotide Primers

Primer Sequences used in different PCR reactions all through the study were designed using PrimerBlast tool of NCBI Genbank. After specificity control of each primer pair, they were purchased either from Macrogen Inc., South Korea BIOBASIC, Canada or IDT, USA and listed below. All of the listed primers were purified by desalting only.

Table 3.1. Oligonucleotide primers used in *SCN1B* gene screening.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
SCN1B-Ex1-F	CTCCCGGGGACATTCTAAC	470
SCN1B-Ex1-R	TGTGGGTGTCACATTGCAG	
SCN1B-Ex2-F	CTGACCTGAGCCTGCTGTC	306
SCN1B-Ex2-R	AACAAATAAGCCCCCTCCT	
SCN1B-Ex3-F	CCTTCCCTCCCTGGCTAC	300
SCN1B-Ex3-R	GGCAGGCAGCACCCGACTCA	
SCN1B-Ex3b2-F	GGTAGTGGACAAAGGTGAGTCG	350
SCN1B-Ex3b2-R	GGGCAAGCCCAGGACAACCC	
SCN1B-Ex3b3-F	GGTGGCCTTCTGTCTCTGAGC	295
SCN1B-Ex3b3-R	GGAGTGCAGAGGGGTGTGCG	
SCN1B-Ex4-F	CAGCACACTCAGGCTGTCAT	290
SCN1B-Ex4-R	TTCCTCTCTGGCAAGTGTGA	
SCN1B-Ex5-F	GGTCTGATTGGGGTCACT	190
SCN1B-Ex5-R	GCAAGAGAGGGGGAATTAGG	

Table 3.2. Oligonucleotide primers used in *LGI4* gene screening.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
LGI4-ex01F	AGCGTGCCAGCAGGCGGATG	335
LGI4-ex01R	GGGGTCCCAGGACTTCCGTC	
LGI4-ex02F	GAGTGCTGTCAGGGTCCTTG	349
LGI4-ex02R	ATGCCCACCTGACATCTGTG	
LGI4-ex03F	GGTGGGCATTGATGCACGTG	348
LGI4-ex03R	TCAGGGCCACAGCTCCCATC	
LGI4-ex04F	CTAGAATTGATGGCTGGGAC	334
LGI4-ex04R	GCACCGTCAGCAGCCACAAG	
LGI4-ex05F	ACAATGAGATTGGCTCCATC	339
LGI4-ex05R	CAAATGCATGCACGCCAACC	
LGI4-ex06F	ACTATGGGGCCCCATCTCTG	345
LGI4-ex06R	GAGCCTCTTGGCCCTGGCAG	
LGI4-ex07F	GGGTACCATGAATGGCCTCC	331
LGI4-ex07R	AGGCAGGTCAGGCCAGCCAG	
LGI4-ex08-1F	CTCACATTGTGCTGGCACAG	349
LGI4-ex08-1R	CAGGAGCTCGGCGTCATTGG	
LGI4-ex08-2F	GCGCCTGGCCCCAACGCAGA	342
LGI4-ex08-2R	AAGTGGCGTGTGGCATAGAC	
LGI4-ex08-3F	CGGCCCGTGCTCTTCCACTG	318
LGI4-ex08-3R	GGGGAGTGGTTCGGAGTCAG	
LGI4-ex09-1F	AGATAAGAAAGCTATGGGAG	339
LGI4-ex09-1R	AGTGATGTGGGCAAAGGCAC	
LGI4-ex09-2F	TCATCGCCAGGGACCAGCTG	351
LGI4-ex09-2R	TGAACGTGGCCCACGGTCAG	

Table 3.3. Oligonucleotide primers used in *PRRT2* gene screening.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
PRRT2_Ex1_F	AATCCCGGTGACCGTTGGCG	438
PRRT2_Ex1_R	CACCCGCATTCCCGTGCAGT	
PRRT2_Ex2_F	ATTGGGCCTGCAGTGCTGAGCG	1061
PRRT2_Ex2_R	GCGGGAAGCTGGGAGCCACA	
PRRT2_Ex3_F	CCGCCATCTATGGGGCTGGC	444
PRRT2_Ex3_R	CTTGGGATGCGGGGCAGAGC	

Table 3.4. Oligonucleotide primers used in *SCN1B* cDNA analysis and splicing reporter assay.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
pSE-RNEx3-F	GCCCTGCCCAGGCTTTTGTCA	Depending insert size
pSE-RNEx2-R	GCAGAGGGGTGGACAGGGTAG	
attB1adapterF	GGGGACAAGTTTGTACAAAAAAGCAGGCT	280
attB2adapterR	GGGGACCACTTTGTACAAGAAAGCTGGGT	
attB1-SCN1B_Ex4F	AAAAAGCAGGCTGCAGCCTGGGCTACCCCCTTA	260
attB1-SCN1B_Ex4r	AGAAAGCTGGGTCTGGGTGCCCTCCCACCTCC	
SCN1B mRNA 2F	CTATTAATACCGGCGGCCC	600
SCN1B mRNA 4&5R	ATTACGGCTGGCTCTTCCTT	

Table 3.5. Oligonucleotide primers used to screen variants in 4BF linkage region.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
KCTD15_F	GCGGGGTTGGAAGAGCCGTC	283
KCTD15_R	TCTGTCCCGAGAGCCCCAGC	
SLC7A10_F	GACGCCCCTCGTTCGGAGC	333
SLC7A10_R	CGCTGTCCCGGCTGCGAGG	
ZNF98_2F	AGGCTTTGCCGCATTCTTCACACT	1542
ZNF98_2R	GCTCATCTGGGGCACTGCACA	
ZNF792_F	GCTGCAGTGTGTGGTGCTGACA	305
ZNF792_R	AGGCCTACCCTCATTCGGCA	
GRAMD1A_F	CCAGAGCCTAGGCAGCCGGAA	234
GRAMD1A_R	TTGACTGTTGCAGCCTGGGGC	

Table 3.6. Oligonucleotide primers used to screen variants out of 4BF linkage region.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
EFHC1_F229L-F	GAACTCCGAAAACAGCCTCTTC	165
EFHC1_F229L-R	ACAGCTACCCAAACCTTTTGCC	
EFHC1_R159W-F	TCCCTGTGTGGTTCTGTTTTCA	252
EFHC1_R159W-R	CACAGTCAACAACGCGGAAA	
CACNA1G_Q118H-F	GAGCCAGGAGGTAAACGAGG	371
CACNA1G_Q118H-R	TGGCTGAAACTGAGACGTGG	
GRM6_S829A-F	CCCAGGTTATGGGGGTTGAC	649
GRM6_S829A-R	GTGGTGAGGACTGTGTGGAG	
KCNB1_P775R-F	TGAACTTCGGACTGGTGCTC	639
KCNB1_P775R-R	CTCGCACAGAAGGGGTCATT	
KCNQ4_G192R-F	GCAGACTTCATCGTGTTCGTG	189
KCNQ4_G192R-R	AGGCCTCACCTTGCTATGC	
CACNA1A-EXP-F	AGCCCCGGTAGTAGCCAT	458
CACNA1A-EXP-R	GGGCAGTAGTTCCGTAAGTGG	
CACNA1A-D2136H-F	GTGTCCACATCGGTGTAGCG	236
CACNA1A- D2136H-R	CACTTGGTGCTAGCTGCTGA	
KCNJ10-R18Q-F	ATGCTGGCTGAAACGAATGG	553
KCNJ10-R18Q-R	GATGACGTCAGTTGCCAAGG	
GABRB3-C23X-F	AAACCATGTCGATGCTGGCG	1122
GABRB3-C23X-R	GGGCGGTGGTCGTCGT	
CNTNAP2-S16F-F	CCTGTTCCACCAGAGGAATGAT	1023
CNTNAP2-S16F-R	AGGTACTGTGAAGCAACTACTGG	

Table 3.7. Oligonucleotide primers used to screen GEFS+ WES analysis variants.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
ATP1A2_I260R-F	CCTCACCCAGAACCGCAT	329
ATP1A2_I260R-R	CCCCTACCTTAGACACGGAG	
BDKRB1-G241R-F	GATCCATCCAAGCCGTCCC	362
BDKRB1-G241R-R	CAATTGCAGGCCCAGGTCAA	
KCNJ16-F	CTTGACCGCAAAGCAGTAGC	325
KCNJ16-R	ATGACCTTCGTCTCGCCTTG	
NRXN2-insC-F	TCACATCAGGGCAGCTAATGG	447
NRXN2-insC-R	CTGATCGCCCTCGGTTTGTC	
KCNV1_T388M-F	TCCCTCATTGCTGCCACGTT	346
KCNV1_T388M-R	ACAGTGAGGTCTGCTTGACATTT	
GRM3_D380Y-F	CAGTTCGACCGCTACTTCCA	344
GRM3_D380Y-R	GCTCTTGGCTTACCCGTGAA	

Table 3.8. Oligonucleotide primers used in qRT-PCR of *KCNQ2* and *PRRT2* genes.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
PRRT2-qPCR-F	ACATCGTGGCCTTCGCTTAT	209
PRRT2-qPCR-R	GTGGTGGCTCAGAGGGTTAG	
KCNQ2_Ex4F	CACAGGCAGGGTGGACGATAGTA	552
KCNQ2_Ex4R	AGTCACGAGAAGAAGCACCGAGA	
eIF2B5_Ex15F	CAAGGCCCCTGAGGTCCCCTT	549
eIF2B5_Ex15R	CAGCCCCACTTCCACATCCCA	

Table 3.9. Oligonucleotide primers used to screen variants in GEFS+ linkage region.

Primer Name	**Sequence (5'-3')**	**Product Length (bp)**
	GGGCAATTTCTTTCAGTCATTGAG	261
ABCD3R	TACCTGTTTGTGATGAGCCGA	
DPYD_F	GCAACACCTACCAGACACTCCT	216
DPYD_R	TTTTTCTGGGATGTGAGGGTTTG	
AGL_F	GCAGGAGTTGATGAAGAAACAGG	186
AGL_R	TTCCACAAGTAAGTGCTTTCTTAAA	
COL11A_F	TTTTCTTCGCTACCTTTACCCCT	200
COL11A_R	TGTTTTCCATTCCATCTTTCACCA	
INPP4_F	ATCCTGCAACACGAGCATGG	223
INPP4_R	GTGCCAATAGTTTGACGGGT	
Mir137NF	AGTGCTACCTTGGCAACCAC	310
Mir137NR	TTATGGTCCCGGTCAAGCTC	

3.3. Enzymes and Reagents

3.3.1. Restriction Enzymes

Enzymes used in restriction reactions throughout this study are listed in Table 3.10.

Table 3.10. List of restriction enzymes.

Enzyme	**Vendor**	**Catalog**	**Content**
*Kpn*I	Promega, USA	R634A	• 8-12U/µl KpnI Enzyme in storage buffer • Buffer J 10X Buffer 1 x 1ml • Bovine Serum Albumin, Acetylated 1 x 150µl • MULTI-CORE™ 10X Buffer 1 x 0.25ml
*Rsa*I	Fermentas, USA	ER1121	• RsaI enzyme (10U/µl) • 10X Reaction Buffer Tango

3.3.2 PCR Enzymes and Reagents

Enzymes and reagents used in all PCR Reactions throughout this study are listed in Table 3.11.

Table 3.11. Reagents and enzymes used in PCR Reactions.

Enzyme or Reagent	Vendor	Catalog No	Content
Taq Polymerase	Qiagen, USA	201203	• 250 units *Taq* DNA Polymerase (5 units/µl) • 10x PCR Buffer • 10x CoralLoad PCR Buffer • 5x Q-Solution • 25 mM $MgCl_2$
Taq Polymerase	Roche, Germany	11146165001	• Taq DNA Polymerase, 5 U/µl • PCR Buffer with MgCl2, 10x concentrated
Fast Start High Fidelity *Taq*, dNTP Pack	Roche Germany	04738284001	• FastStart Taq DNA Polymerase, 5 U/µl • PCR Reaction Buffer, 10x concentrated with 20 mM MgCl2 • PCR Reaction Buffer, 10x concentrated without MgCl2 • MgCl2 Stock Solution, 25 mM • GC-RICH Solution, 5x concentrated • PCR Grade Nucleotide Mix
Pwo SuperYield Polymerase	Roche Germany	04340868001	• Pwo SuperYield DNA Polymerase, 5 U/µl • PCR buffer, 10x concentrated, containing 15 mM MgSO4 • GC-RICH Solution,
dNTP Mix	Thermo Scientific, USA	R0193	• dNTP mix (10mM each)
Dimethylsulphoxide (DMSO)	Sigma-Aldrich, Germany	D9170	• Dimethyl sulfoxide PCR Reagent

3.3.3. Eukaryotic Cell Transfection

Eukaryotic cell transfection is conducted using Lipofectamine reagent, properties of which is listed in Table 3.12.

Table 3.12. Reagent used in eukaryotic cell transfection.

Reagent	Vendor	Catalog No	Content
Transfection Reagent	Invitrogen, USA	11668-027	Lipofectamine® 2000

3.4. Kits

3.4.1. Gateway Cloning

Gateway BP Clonase II enzyme mix (Invitrogen, USA) was used for gateway cloning procedure. Contents of the kit are listed in Table 3.13 below.

Table 3.13. Contents of Gateway® BP Clonase™ II Enzyme Mix.

Components Supplied	20 rxns
Gateway® BP Clonase™ II Enzyme Mix	40 µl
Proteinase K Solution (2 µg/µl)	40 µl
30% PEG 8000/30 mM MgCl2 Solution	1 ml
pEXP7-tet Positive Control (50 ng/µl)	20 µl

3.4.2. Nucleic Acid Extraction and Purification

Several kits were used for nucleic acid isolation and purification all through this study. Kits and their contents are listed in Table 3.14.

Table 3.14. List of kits used for nucleic acid extraction of purification.

Kit	Vendor	Content	Quantity
Qiagen Plasmid Mini (100) Kit		QIAGEN-tip 20	100 pieces
		Buffer P1	40ml
		Buffer P2	40ml
		Buffer P3	40ml
		Buffer QBT	2x60
		Buffer QC	2x240
		Buffer QF	140
		RNase A (10mg/ml)	0.4ml

Table 3.14. List of kits used for nucleic acid extraction of purification (cont.).

Kit	Vendor	Content	Quantity
QIAquick PCR Purification Kit	Qiagen, USA	QIAquick Spin Columns	250
		Buffer PB	150 ml
		Buffer PE (concentrate)	55 ml
		Buffer EB	55 ml
		pH Indicator I	800 µl
		Collection Tubes	250
		Loading Dye	550 µl
High Pure Plasmid Isolation Kit	Roche, Germany	Suspension Buffer	80 ml
		RNase A, dry powder	8mg
		Lysis Buffer	80ml
		Binding Buffer	100 ml
		Wash Buffer I	100 ml
		Wash Buffer II	50 ml
		Elution Buffer	40 ml
		High Pure Spin Filter Tubes (containing glass fiber fleece)	250 pieces
		Collection Tubes	250 pieces
Quick-RNA™ MiniPrep	Zymo Research, USA	RNA Lysis Buffer	50 ml
		RNA Prep Buffer	25 ml
		RNA Wash Buffer1 (concentrate)	24 ml
		DNase/RNase-Free Water	6 ml
		DNase I2 (lyophilized)	1 vial
		DNA Digestion Buffer	4 ml
		Spin-Away™ Filters	50 pieces
		Zymo-Spin™ IIICG Columns	50 pieces
		Collection Tubes	100 pieces
QIAamp DNA Blood Maxi Kit	Qiagen, USA	QIAamp Maxi Spin Columns	50
		Collection Tubes (50 ml)	50
		Buffer AL*	2 x 330 ml
		Buffer AW1* (concentrate)	121 ml
		Buffer AW2 (concentrate)	81 ml
		Buffer AE	110 ml
		QIAGEN Protease	5 vials
MagNa Pure Compact Nucleic Acid Isolation Kit I-Large volume	Roche, Germany	Reagent Cartridge	32
		Tip Tray	32
		Sample Tube	35 t
		Elution Tube	35
		Elution Tube Cap	35
MagNA Pure Compact RNA Isolation Kit	Roche, Germany	Reagent Cartridges	32
		Tip Trays	32
		Sample Tubes	2 x 35
		Elution Tubes	35 (2.0 ml) ;
		Tube Caps	35
		DNase Solution	1
		Lysis Buffer (1 bottle)	35 ml

3.4.3. Reverse Transcription PCR

Qiagen Long Range Reverse Transcription kit (Qiagen, USA) was used for cDNA synthesis from total mRNA isolated from patient 4BF14 and his parents and also for RT-PCR part of minigene splicing reporter assay. Kit contents are listed in Table 3.15.

Table 3.15. Contents of the kit used for reverse transcription PCR.

Kit	Vendor	Catalog No	Item	Quantity
Long-Range Two Step RT-PCR Kit	Qiagen, USA	205922	LongRange RT Buffer,	200 µl
			LongRange Reverse Transcriptase (1 reaction/µl)	50 µl
			Oligo-dT, 20 µM	60 µl
			LongRange PCR Enzyme Mix	40 µl
			LongRange RNase Inhibitor	25 µl
			LongRange PCR Buffer, 10x	500 µl
			Q-Solution®, 5x	2 ml
			dNTP Mix, 10 mM each	2x 200 µl
			MgCl2	1.2 ml
			RNase-Free Water	4x 1.9 ml

3.4.4. Quantitative Real-Time PCR

qRT-PCR reactions were conducted on a Light Cycler 480i (Roche, Germany) real time PCR machine using SYBR Green Master Mix kit (Roche, Germany) contents of which is listed in Table 3.16.

Table 3.16. Contents of SYBR Green kit used for quantitative real time PCR.

Kit	Vendor	Catalog No	Content
SYBRGreen Master Mix	Roche, Germany	04707516001	LightCycler® 480 SYBR Green I Master, 2x concentrated Water, PCR Grade

3.4.5. Array-CGH

Kits and reagents used in custom and catalog arrayCGH studies are listed in Table 3.17.

Table 3.17. Materials used for array CGH.

Catalog No	Kit	Contents
AGTG4827A	Agilent SurePrint G3 Human CGH Microarray, 8x60K (Custom)	One glass slide formatted with eight SurePrint G3 60K arrays
G5955A	SurePrint G3 CGH ISCA v2 Microarray Kit, 8x60K	Three glass slides each formatted with eight SurePrint G3 60K arrays
AGT-5190-4240	Agilent SureTag Complete DNA Labeling Kit	Reference DNA, restriction digestion enzymes, labeling reagents, clean- up columns.
AGT-5190-3391	Agilent SureTag Purification Columns	50 spin columns
AGT-5190-3393	Agilent Human Cot-1 DNA	Cot-1 Human DNA
AGT-5188-5220	Agilent aCGH Hybridization Kit	Hybridization solutions Blocking agent
AGT-5188-5226	Agilent aCGH Wash Buffer kit	Wash Buffer I Wash Buffer II
AGT-5185-5979	Agilent Stabilization and Drying Solution	
AGTG2534-60014	Agilent Hybridization Gasket Slide Kit	Package 5 backings 8 array per slide
271004	Acetonitrile (Sigma-Aldrich)	anhydrous, 99.8%

3.4.6. DNA Size Determination and Quantification

DNA size determination was conducted using Agilent DNA 100 Kits (Agilent, USA,) (Table 3.18) in an Agilent Bioanalyzer electronic electrophoresis device. DNA quantification was conducted using Quant-iT™ PicoGreen ® dsDNA (Invitrogen, USA) using QuantiFlour (Promega, USA) fluorimeter.

Table 3.18. Contents of Agilent DNA 1000 Kit.

Item	Content
DNA Chips Kit	25 DNA Chips
	1 Electrode Cleaner
DNA 1000 Reagents	DNA 1000 Ladder DNA 1000 Markers 15/1500 bp (2 vials) DNA Dye Concentrate*(1 vial) DNA Gel Matrix (3 vials) 3 Spin Filters
Syringe Kit	1 Syringe

Table 3.19. Quant-iT™ PicoGreen ® dsDNA Reagent and Kits.

Item	Quantity	Concentration
Quant-iT™ PicoGreen® dsDNA reagent (Component A)	1 mL in 1 vial	Solution in DMSO
20X TE (Component B)	25 mL	200 mM Tris-HCl, 20 mM EDTA, pH 7.5
Lambda DNA standard (Component C)	1 mL	100 µg/mL in TE

3.5. Buffers, Solutions and Media

3.5.1. DNA Extraction

Buffers and solutions used for manual DNA extraction are listed in Table 3.20.

Table 3.20. Buffers and solutions used in manual DNA extraction.

Buffer or Solution	Content	Final Concentration
Lysis Buffer	NH_4Cl $KHCO_3$ Na_2EDTA (pH 7.4)	155 mM 10 mM 0.1 mM
Nuclease Buffer	Tris-HCL (pH 8.0) NaCl Na_2EDTA (pH 8.2)	10 mM 400 mM 2 mM
Sodiumdodecylsulphate (SDS)		10 % SDS (w/v) (pH 7.2)
Proteinase K	Proteinase K in H_2O	20 mg/ml
Sodium Chloride (NaCl	NaCl	2,5 M
Ethanol (EtOH)	Absolute EtOH, Riedel de Haen, Germany	70% and 98%
TE Buffer	Tris-HCl (pH 8.0) Na_2EDTA (pH8.0)	20 mM 1 mM

3.5.2. Cell Cultures

Media and solutions used in bacterial cell culture are listed in Table 3.21 and in SH-SY5Y cell culture are listed Table 3.22.

Table 3.21. Media used for bacterial cell culture.

Media	Content	Vendor	Catalog No	Final Concentration
Luria Broth (LB)	LB Broth Powder	Sigma-Aldrich, Germany	L3022	10g/L Tryptone 5 g/L Yeast Extract 5 g/L NaCl
Luria Agar (LA) Media	LB Broth Powder	Sigma-Aldrich, Germany	L3022	10g/L Tryptone 5 g/L Yeast Extract 5 g/L NaCl
	Agar	Fluka, Germany	05038	1,5 %
Ampicillin	Ampicillin Solution, 100 mg/mL, 0.2 µm filtered	Sigma-Aldrich, Germany	A5354-10ML	100ng/ml

Table 3.22. Media used for SH-SY5Y cell culture.

Media	Content	Vendor	Final Concentration
Transfection media	DMEM (no glucose)	Invitrogen, USA	
	F12	Invitrogen, USA	
	FBS	Invitrogen, USA	10%,
	Penicillin/streptavidin	Invitrogen, USA	1%
	Non-essential amino acids	Invitrogen, USA	1%
Differentiation media	DMEM low glucose	Invitrogen, USA	
	Penicillin/streptavidin	Invitrogen, USA	1%,
	N2 supplement	Invitrogen, USA	1%,
	retinoic acid	Invitrogen, USA	10 µM

3.5.3. Agarose Gel Electrophoresis

Buffers, reagents and dye solutions used in for agarose gel electrophoresis are listed in Table 3.23.

Table 3.23. Buffers and solutions used for agarose gel electrophoresis.

Buffer or solution	Content	Vendor	Final Concentration
10X TBE Buffer	Tris-Base	Fluka, Germany	0.89 M
	Boric acid	Fluka, Germany	0.89 M
	Na_2EDTA (pH 8.3)	Fluka, Germany	20 mM

Table 3.23. Buffers and solutions used for agarose gel electrophoresis (Cont.).

Buffer or solution	**Content**	**Vendor**	**Final Concentration**
Ethidium Bromide (EtBr)		Fluka, Germany	
1 or 2 per cent Agarose Gel	Agarose	Fluka, Germany	1 or 2 %
	TBE Buffer		0.5X
	Ethidium bromide	Fluka, Germany	0.5µg/ml
6X DNA Loading Dye (Ready to use)	10 mM Tris HCl (pH7.6) ; 0.03 % Bromophenol Blue (BPB); 0.03% xylene cyanol FF; 60 % glycerol; 60 mM EDTA	(Fermentas, USA)	1X
GeneRuler 100bp DNA Ladder	DNA Fragment 100bp-1000bp in storage buffer (10 mM Tris-HCl (pH 7.6), 1mM EDTA)	Thermo Scientific, USA	0.1µg/µl
GeneRuler 1kb DNA Ladder	DNA Fragment 250bp-10000bp in storage buffer (10 mM Tris-HCl (pH 7.6), 1mM EDTA)	Thermo Scientific, USA	0.1µg/µl

3.6. Cell Lines and Vectors

Cell lines and vectors used throughout this study are listed in Table 3.24.

Table 3.24. Cell lines and vectors used throughout the study.

Cell Line or Vector	**Vendor**
pSpliceExpress	AddGene plasmid repository
SH-SY5Y	M. Murat Köseoğlu Lab of Fatih University
Z-Competent E.coli DH5α cells	Zymo Research Cooperation, USA

3.7. Equipment and Hardware

Equipment and hardware used throughout this study are listed in Table 3.25.

Table 3.25. List of equipment and hardware.

Equipment or Hardware	Model	Vendor
Autoclave		
Balances	GM 512-OCE	Sartorius, Germany
Centrifuges	5415C	Eppendorf, Germany
	Allegra X-22R	
	Microfuge18	Beckman, Coulter, USA
Deep Freezers (-20 ºC)		Arçelik, Turkey
Ultra-Low Temperature Freezer	Forma 905	Thermo Scientific, USA
Documentation System	GelDoc XR+	BIO-RAD, USA
Electrophoretic Equipment	Mini Sub-Cell GT,	BIO-RAD, USA
	Owl™ D2 Wide Gel System	Thermo Scientific, USA
	Owl™ EasyCast™ B2 Mini Gel Electrophoresis Systems	Thermo Scientific, USA
Heat Block Magnetic Stirrer	Hot-Plate Magnetic Stirrer	Chiltern
Real-Time PCR System	Light Cycler LC 480i 96 well-plate platform	Roche, Germany
Microwave Ovens	Microwave Oven	DBK, Turkey
Power Supplies	EC 250-90	Thermo Scientific, USA
Refrigerator	4ºC Medicool	Sanyo, Japan
	4042T	Arçelik, Turkey
Spectrophotometer	ND 1000 Spectrophotometer,	NanoDrop USA
Thermocyclers	TC-512	Techne, UK
	T100	BioRad, USA
On-Chip Electrophoresis	Bioanalyzer 2100	Agilent Inc. USA
Microarray Scanner	NimbleGen MS200	Roche, Germany
Fluorimeter	QuantiFluor®-ST,	Promega, USA
Array Hybridization Chamber	G2545A	Agilent, USA
Micropipettes	Pipetman (2,20,100,200,1000µl)	Fischer Scientific, USA

3.8. Software

List of software packages and software is given in Table 3.26.

Table 3.26. Software used throughout the study.

Package	Program	Vendor or Reference
Geospiza	Finch TV version 1.4.0	Perkin Elmer, USA
Microsoft Office	Microsoft Excell	Microsoft, USA
CLC	Main Workbench	Qiagen, USA
EasyLinkage Program Package (Hofman and Linder, 2005)	Allegro	Gudbjartsson, 2000
EasyLinkage Program Package (Hofman and Linder, 2005)	SimWalk2	Sobel, Papp and Lange,2002
	Genehunter	Kruglyak *et al*., 1996
LightCycler480i Software version 1.5	SuperLink	Fishelson M, Geiger D. (2002)
	Relative Quantitation	Roche, Germany
Alohomora	Alohomora	Ruschendorf and Nurnberg, 2005
SPSS	SPSS Statistics	IBM, USA
Karyo Studio	Karyo Studio	Illumina Inc. USA
Agilent Microarray Analysis Tools	Feature Extraction	Agilent, USA
Agilent Microarray Analysis Tools	Cytogenomics	Agilent, USA

3.9. Online Tools and Databases

List of online tools and databases is given in the Table 3.27.

Table 3.27. Online tools and databases used throughout the study.

Tool or Database	Link
OMIM	http://www.ncbi.nlm.nih.gov/omim
Pubmed	http://www.ncbi.nlm.nih.gov/pubmed
NCBI: National Center for Biotechnology Information	http://dgv.tcag.ca/dgv/app/home
NCBI_PrimerBlast	http://www.ncbi.nlm.nih.gov/tools/primer-blast/
Ensembl Genome Browser	http://www.ensembl.org/index.html
European Bioinformatics Institute	http://www.ebi.ac.uk/
UCSC Genome Browser	https://genome.ucsc.edu/
ExPasy Bioinformatics Resource Portal	http://www.expasy.org/

Table 3.27. Software used throughout the study (Cont.).

UniProt (SwissProt)	http://www.uniprot.org/uniprot/
Expression Atlas	https://www.ebi.ac.uk/gxa/home
KEGG: Kyoto Encyclopedia of Genes and Genomes	http://www.genome.jp/kegg/
GeneDistiller 2014	http://www.genedistiller.org/
HomozygosityMapper	http://www.homozygositymapper.org/
Agilent SureDesign	https://earray.chem.agilent.com/suredesign/
DECIPHER (DatabasE of genomiC varIation and Phenotype in Humans using Ensembl Resources)	https://decipher.sanger.ac.uk/
DGV: Database of Genomic Variants	http://dgv.tcag.ca/dgv/app/home
Mutation Taster	http://www.mutationtaster.org/
Polyphen 2	http://genetics.bwh.harvard.edu/pph2/
SIFT	http://sift.jcvi.org/
PROVEAN	http://provean.jcvi.org/index.php
ESE-Finder	http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home
ClinVar	http://www.ncbi.nlm.nih.gov/clinvar/
HGMD (Human Gene Mutation Database)	http://www.hgmd.cf.ac.uk/ac/all.php
ISCA (International Standards for Cytogenomic Arrays)	http://www.iscaconsortium.org/

4. METHODS

4.1. Extraction and Quantification of DNA / RNA

Patients DNAs from EDTA-Preserved blood were isolated using three different methods, which are described below. After DNA isolation sample concentration was measured by NanoDrop Spectrophotometer (NanoDrop Inc., USA) according to manufacturer's recommendations. DNA samples were aliquoted in working solutions, which were stored at +4ºC and stock samples were stored at -20 ºC.

For total RNA extraction, blood samples of patients were collected into PAXGene (PreAnalytics, USA) tubes and extracted using MagnaPure RNA Isolation Kit (Roche, Germany) in a MagNA Pure Compact (Roche, Germany) nucleic acid extraction robot as described by the manufacturer. After extraction total RNA was quantified using NanoDrop Spectrophotometer (Thermo Scientific, USA). RNA samples were aliquoted in working solutions and stored at -80ºC.

4.1.1. Manual DNA Extraction

In order to extract adequate amounts of DNA ten ml peripheral blood was collected from patients and their relatives into K'EDTA containing tubes and kept at 4ºC in order to prevent coagulation before extraction. Nine ml of collected blood was used in the procedure and remaining one ml was stored at -70ºC.

In the manual DNA extraction from white blood cells, blood was initially collected into sterile 50 ml Falcon tubes, after addition of 30 ml lysis buffer (stored in 4ºC) the tubes were incubated at 4ºC for 20 minutes for 15 minutes for lysis of leukocyte membranes. Samples were then centrifuged at 5000 rpm for ten minutes at 4ºC to collect the nuclei. The supernatant was discarded and the pellet was washed with lysis buffer. Samples were centrifuged for an additional 10 minutes after addition of 10 ml lysis buffer in order to wash the nuclear pellet.

Three ml nuclei lysis buffer, proteinase K (150 µg/ml) and SDS (0,14 per cent) were added to the nuclear pellet, the samples were incubated at 37ºC overnight for adequate disruption of nuclear envelope and degradation of proteins. Remaining proteins were salted out by addition of ten ml 2,5M NaCl and centrifuged at 5000 rpm at room temperature for 20 minutes. After centrifugation the supernatants were taken into sterile falcon tubes and DNA was precipitated by addition of two volumes of absolute ethanol. DNA forming a white precipitate on the solution surface was fished out by the help of a sterile pipette tip and collected into sterile Eppendorf tubes. After EtOH evaporated, DNA was dissolved in TE buffer and stored at 4ºC for further analysis (Miller, Dykes, and Polesky, 1988).

4.1.2. DNA Extraction with MagNaPure Compact

DNA from some samples were extracted using MagnaPure Large volume DNA Isolation Kit (Roche, Germany) in MagNA Pure Compact Nucleic Acid Isolation Machine (Roche, Germany) according to the manufacturer's recommendations.

Briefly, 1ml of blood samples were put into sample tubes and aligned on the sample rack of the machine. Elution tubes with barcode labels were scanned by barcode reader and also aligned on sample rack of the machine. Patient IDs corresponding to each sample were assigned to each sample tube barcode ID. DNA extraction from 1ml blood protocol was selected. The machine automatically isolates DNA and elutes in 200 µl elution buffer.

4.1.3. DNA Extraction with QIAamp Kit

DNA from some samples were extracted using QIAamp Blood Maxi Kit (Qiagen Inc. USA) according to manufacturer's recommendations. In Summary, 0.5 ml Protease is mixed with 10 ml blood in a 50 ml centrifuge tube and mixed briefly. Then 12 ml of buffer AL was added and the mixture was incubated at 70ºC for 10 minutes. Then, 10 ml absolute Ethanol (Merck, Germany) was added to the mixture and vortexed. The mixture was transferred to Qiagen Maxi Spin columns in 50 ml collection tubes and the tubes were centrifuged at 3000 rpm for 3 minutes in Allegra X22R (Beckman Coulter, Germany). The filtrate was discarded and 5 ml of Buffer AW1 was added on top of the column. The tubes were centrifuged at 4400 rpm for 6 minutes. Without discarding filtrate 5ml of Buffer

AW2 was loaded on top of the tube and centrifuged at 4400 rpm for 20 minutes. After centrifugation the filtrate was discarded, spin columns were transferred into clean 50 ml collection tubes and incubated at 56ºC for 10 minutes. 800 µl Buffer AE was then loaded on the column and centrifuged at 4400 rpm for 5 minutes. This step was applied twice and in total 1600 µl DNA obtained.

4.1.4. Analysis by Agarose Gel Electrophoresis

100 ml of 0.5X TBE buffer was added to adequate amounts of agarose (1 g for 1% gel and 2 g for 2 % gel). One per cent (w/v) agarose gel was used to visualize whole genomic DNA and two per cent (w/v) agarose gel to visualize PCR products. After addition of agarose, the suspension was boiled for preparing a homogenous gel until no filaments of agarose was observed. As soon as the solution cooled down to about 50 ºC, five µl EtBr (10 mg/ml) was added in order to visualize DNA under ultraviolet (UV) light. Hot liquid was poured into the electrophoresis plate and appropriate comb was put into the gel before it polymerize. Polymerization is achieved in 15 minutes at room temperature. 1 µl DNA sample was mixed with 1 µl loading dye and 4 µl H_2O and loaded into the wells on the gel. The gel was submerged into the TBE buffer in the electrophoresis tank and run at 150 volts for 15 minutes.

4.1.5. Visualization of DNA Under Ultraviolet (UV) Light

DNA on the agarose gel was visualized in the GelDoc Documentation System, BioRad, USA and the gel images were stored in the “tiff” format.

4.1.6. Quantification of Genomic DNA and Total RNA by Spectrophotometer

The concentration of genomic DNA and total RNA was measured by NanoDrop spectrophotometer (NanoDrop Inc. USA), calibrated with 2 µl of water after the cleaning of sample pedestal. 1.5 µl of water or elution buffer from MagNA Pure Nucleic Acid isolation Large Volume Kit was used as blank. 1,5 µl of DNA sample was loaded onto the sample pedestal. After measurement of absorption the device the program calculates DNA concentration. If the sample was concentrated beyond the measurement limit of

spectrophotometer it was diluted with sterile distilled H_2O.

4.2. Benign Familial Infantile Seizures (BFIS) and Benign Familial Neonatal Seizures Studies

4.2.1. Target Gene Analysis

4.2.1.1 *SCN1B* Sodium Channel, Voltage-Gated, Type I, Beta Subunit Gene. Coding Region of *SCN1B* gene (Gene ID: 6324) and exon intron boundaries were amplified by PCR in 50µl reaction. PCR reaction conditions for 7 amplicons were shown in Table 4.1 and Thermal cycle of PCR amplification was shown in Table 4.1.

Table 4.1. PCR Reaction mixture to amplify *SCN1B* gene coding regions and exon/intron boundaries.

Reaction Mix	**Stock conc.**	**Exon 1**	**Exon 2**	**Exon 3**	**Exon 3b2**	**Exon 3b3**	**Exon 4**	**Exon 5**
Buffer	10X	1X	1X	1X	1X	1X	1X	1X
Forward Primer	10µM	0.3 µM	0.3 µM	0.3 µM	0.2 µM	0.2 µM	0.2 µM	0.1 µM
Reverse Primer	10µM	0.3 µM	0.3 µM	0.3 µM	0.2 µM	0.2 µM	0.2 µM	0.1 µM
dNTP Mix	10mM each	0.2mM	0.2mM	0.2mM	0.2mM	0.2mM	0.2mM	0.1mM
DMSO		2.5%						
Taq DNA Polymerase (Qiagen Inc.. USA)	5U/µl	0.015 U/µl	0.015 U/µl	0.015 U/µl				
Taq DNA Polymerase (Roche. Germany)	5U/µl				0.025 U/µl	0.025 U/µl	0.025 U/µl	0.025 U/µl
DNA	70ng/ µl	1.4 ng/ µl	1.4 ng/ µl	1.4 ng/ µl	1.4 ng/ µl	1.4 ng/ µl	1.4 ng/ µl	1.4 ng/ µl

Table 4.2. PCR Reaction thermal cycle to amplify *SCN1B* gene coding regions and exon/intron boundaries. (t: Temperature (°C); T: Time; C: Cycles; ID: Initial Denaturation, Den: Denaturation, Ann: Annealing, Elon: Elongation, FE: Final Elongation).

Step	Exon 1		Exon 2		Exon 3		Exon 3b2		Exon 3b3		Exon 4		Exon 5		C
	t	T	t	T	t	T	t	T	t	T	t	T	t	T	
ID	94	5m	94	5m	94	5m	94	5m	94	5m	94	5m	94	5m	1
Den	94	30s	94	30s	94	30s	94	30s	94	30s	94	30s	94	30s	35
Ann	62	30s	62	30s	64	30s	62	45s	62	45s	60	45s	60	25s	
Elon	72	30s	72	30s	72	30s	72	45s	72	45s	72	45s	72	30s	
FE	72	5m	72	5m	72	5m	72	5m	72	5m	72	5m	72	5m	1

4.2.1.2. Leucine-Rich Repeat LGI Family, Member 4 (LGI4)Gene. The coding regions of the *LGI4* gene (gene ID: 163175) were amplified in 12 amplicons. Optimization of the PCR amplification was done within the framework of this thesis. For all amplicons 100 ng genomic DNA was used as template in 50 µl final volume. The reaction conditions were summarized in the Tables 4.3 and 4.4.

Table 4.3. PCR Reaction mixture to amplify *LGI4* gene.

PRRT2 amplification conditions (in total 50 µl)				
Contents	Stock Concentration	Final Concentrations		
		Exons 1. 2. 3. 5. 6. 8-2 and 9-2	Exons 8-1 and 8-3	Exons 4. 7 and 9.1
Buffer (-$MgCl_2$)	10X	1X	1X	1X
dNTP	10mM	0.2 mM	0.2 mM	0.2 mM
MgCl2	25mM	1.5 mM	2 mM	1.5 mM
Forward Primer	10 µM	0.5 µM	0.5 µM	0.25 µM
Reverse Primer	10µM	0.5 µM	0.5 µM	0.25 µM
Roche Taq Polymerase	5U/ µl	0.03 U/ µl	0.03 U/ µl	0.03 U/ µl
DNA	100ng/ µl	2 ng/ µl	2 ng/ µl	2 ng/ µl

Table 4.4. PCR thermal cycle conditions to amplify *LGI4* gene.

<table>
<tr><td colspan="4">Thermal Cycle</td></tr>
<tr><td colspan="2">Temperature ºC</td><td>Time</td><td>cycle</td></tr>
<tr><td colspan="2">94</td><td>5min</td><td>1</td></tr>
<tr><td colspan="2">94</td><td>30s</td><td rowspan="14">35</td></tr>
<tr><td>Exon 1</td><td>65.9</td><td rowspan="12">45s</td></tr>
<tr><td>Exon2</td><td>60.3</td></tr>
<tr><td>Exon3</td><td>60.3</td></tr>
<tr><td>Exon5</td><td>60.3</td></tr>
<tr><td>Exon6</td><td>56</td></tr>
<tr><td>Exon 8-2</td><td>64.6</td></tr>
<tr><td>Exon9-2</td><td>60.3</td></tr>
<tr><td>Exon 8-1</td><td>64.2</td></tr>
<tr><td>Exon 8-3</td><td>66.2</td></tr>
<tr><td>Exon 4</td><td>62.2</td></tr>
<tr><td>Exon 7</td><td>62.2</td></tr>
<tr><td>Exon 9-1</td><td>56</td></tr>
<tr><td colspan="2">72</td><td>45s</td></tr>
<tr><td colspan="2">72</td><td>7min</td><td>1</td></tr>
<tr><td colspan="2">+4</td><td>∞</td><td>1</td></tr>
</table>

The PCR products were visualized on 2% agarose gels, and compared with a 100 bp DNA ladder (Fermentas, Lithuania) for length determination. The gels were checked under UV light and documented.

Adequate amount of PCR products were sent to Macrogen Inc., Korea for PCR purification and automated sequencing. The results were obtained online as ABI document and analyzed in the Finch TV version 1.4.0, (Geospiza Inc. USA) chromatograph analyzing program.

4.2.1.3. Proline-Rich Transmembrane Protein 2 (*PRRT2*) Gene. Coding regions and Exon/Intron boundaries of *PRRT2* (Gene ID: 112476) gene were amplified in 3 amplicons. According to reference cDNA NM_001256442.1 and protein NP_001243371.1 exons were determined and primers were designed using Primer Blast tool in National Center for Biotechnology Information web site (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and purchased from BIOBASIC Inc. Canada. PCR conditions were shown in Table 4.5.

Table 4.5. PCR conditions for *PRRT2* gene amplification.

PRRT2 amplification conditions (in total 50 µl)				
Contents	Stock Concentration	Final Concentrations		
		Exon 1	Exon 2	Exon 3
Buffer (-$MgCl_2$)	10X	1X	1X	1X
dNTP	10mM	0.6mM	0.6mM	0.6mM
MgCl2	25mM	1mM	1mM	1mM
Forward Primer	10 µM	0.4 µM	0.2 µM	0.2 µM
Reverse Primer	10µM	0.4 µM	0.2 µM	0.2 µM
Qiagen Taq Polymerase	5U/ µl	0.03 U/ µl	0.02 U/ µl	0.03 U/ µl
DNA	100ng/ µl	2 ng/ µl	2 ng/ µl	2 ng/ µl
Thermal Cycle				
Temperature ^{o}C	Time		cycle	
94	5min		1	
94	30s		35	
66	30s			
72	60s			
72	7min		1	
12	∞		1	

4.2.2. Evaluation of *SCN1B* c.492 T>C Variant for Splicing Effect

The *SCN1B* c.492 T>C Variant was analyzed by PESX: Putative Exon ;Splicing Enhancer, Silencers Program. Sequence of exon 4 with and without alteration was loaded into program. The program showed putative exon splicing enhancer and silencer. Difference between native sequence and sequence with alteration was inspected visually and found that the variation on exon 4 may abolish binding site of a splicing silencer protein.

4.2.3. Minigene Analysis of the *SCN1B* c.492 T>C Variant

4.2.3.1. Amplification of pSpliceExpress Plasmid DNA. pSpliceExpress Vector was purchased from AddGene plasmid repository. The plasmid arrived as bacterial stab in LB-Agar tube. In order to isolate plasmid required amount of E.coli cells in agar stab were inoculated into liquid LB media containing Ampicillin (100ng/ml) and incubated at 37°C overnight. Afterwards plasmid was isolated from bacteria with Qiagen Plasmid Mini Kit (Qiagen, USA) according to manufacturer's recommendations. Briefly, 50 ml culture was put into a 50 ml Falcon tube and centrifuged at 6000g for 20 minutes. Then the supernatant

was discarded and bacterial pellet was resuspended in 5ml of Buffer P1 and shaken vigorously. Then 5 ml of Buffer P2 was added and shaken again. 5 ml of ice cold Buffer P3 was added and incubated on ice for 30 minutes.

After protein precipitation was observed in the tube, it was centrifuged at 4400 rpm for 30 minutes and supernatant was taken into a fresh tube. Meanwhile, 2 QIAGEN-Tip20 were equilibrated with 1 ml buffer QBT, which allowed resin to enter by gravity flow. Then, supernatant of the previous centrifugate was loaded onto two equilibrated QIAGEN-Tip20 and flowed by gravity.

The resin was washed two times with 2 ml Buffer QC. QIAGEN-Tip20 was taken into a clean 15 ml Falcon tube and DNA was eluted with 0,8 ml Buffer QF. Eluted DNA was precipitated by 0,7 volume of room temperature Isopropanol and centrifugation at 13000 rpm on a benchtop centrifuge. The DNA precipitate was washed with 70% EtOH and left aside to air dry. After drying DNA was dissolved in 50 µl of TE buffer.
Isolated plasmid was sequenced by Macrogen, Korea using M13 Forward primer for confirmation.

4.2.3.2. Amplification of *SCN1B* Exon 4. Exon 4 of *SCN1B* gene was amplified form genomic DNA of sample 4BF14 with sequence specific primers with adaptors according to conditions listed in Table 4.6. Then, required sequences for the recombination reaction were introduced by another PCR using primers with adaptor sequences plus AttB1 sequence. Conditions for the 2nd PCR were listed in Table 4.7 below. Both PCR reactions were done in same thermal cycling conditions which is shown in Table 4.8.

Table 4.6. PCR conditions for amplification of *SCN1B* exon 4 from genomic DNA with sequence specific primers plus adaptor sequences.

SCN1B Exon-4 amplification conditions (in total 50 µl)		
Contents	Stock Concentration	Final Concentrations
Buffer (+$MgCl_2$)	10X	1X
dNTP	10mM	0.2mM
Forward Primer	25 µM	0.25 µM
Reverse Primer	25µM	0.25 µM
Roche Taq Polymerase	5U/ µl	0.02 U/ µl
DNA	100ng/ µl	2 ng/ µl

Table 4.7. Reaction conditions for PCR to introduce AttB1 sequence to the end of amplicon; which was required for recombination.

SCN1B Exon-4 AttB1 PCR conditions (in total 50 µl)		
Contents	Stock Concentration	Final Concentrations
Buffer (+$MgCl_2$)	10X	1X
dNTP	10mM	0,2mM
Forward Primer	10 µM	0,2 µM
Reverse Primer	10µM	0,2 µM
Roche Taq Polymerase	5U/ µl	0,03 U/ µl
Previous PCR product		2 µl

Table 4.8. Thermal cycling conditions for both sequence amplification and adaptor ligation PCR reaction.

Thermal Cycle		
Temperature ^{o}C	Time	cycle
95	5min	1
95	30s	30
64	30s	
72	30s	
72	7min	1
12	∞	1

4.2.3.3 Purification of PCR Products. PCR products were purified in order to get rid of excess primers and most importantly primer dimers which might inhibit recombination reaction. QIAquick PCR Purification Kit (Qiagen, USA) was used for this purpose. In summary, PCR products were mixed with 5 volumes of Buffer PB and applied onto QIAquick spin column, which was placed into collection tube. The tube was centrifuged at 13000rpm on a benchtop centrifuge for 30 seconds. The filtrate was discarded and 0,75ml of Buffer PE was put onto the spin column and centrifuged for 1 minutes. The filtrate was discarded again and the spin column was centrifuged for additional 1 minutes. The spin column was placed into a new, clean micro centrifuge tube and 50 µl of Buffer EB was placed on top of spin column. The tube was centrifuged for 1 minute and eluted DNA was obtained.

4.2.3.4. Gateway Cloning of SCN1B exon 4 into pSpliceExpress (pSE) Vector. Recombination reaction was conducted using Gateway® BP Clonase™ II Enzyme Mix (Invitrogen, USA). Concentration of isolated plasmid and purified PCR product was

measured by NanoDrop spectrophotometer as described above. Reaction mixture was prepared according to Table 4.9 and incubated at room temperature for 1 hour.

Table 4.9. Recombination reaction mix.

Component	**Amount**
SCN1B Exon4 PCR product (20 ng/µl)	7 µl
pSpliceExpress (150 ng/ µl)	1 µl
Gateway® BP Clonase™ II Enzyme Mix	2 µl

After 1 hour, in order to terminate reaction, 1 µl of Proteinase K (supplied with kit) was added to the reaction and incubated at 37ºC for 10 minutes. Afterwards, Z-Competent E.coli DH5α cells (Zymo Research Cooperation, USA) were transformed with recombination reaction product. In brief, 50 µl of Z-Competent E.coli DH5α cells were mixed with total amount of recombination reaction and all of the reaction was spreaded onto LB-Agar plates with Ampicillin (100ng/ml). Plates were incubated at 37ºC overnight and next day bacterial colonies were observed on the plate.

12 liquid 6 ml LB culture were started from 12 different colonies and incubated at 37ºC overnight. The next day plasmid was isolated from bacterial culture using High Pure plasmid Isolation Kit (Roche, Germany). In summary, 6 ml culture was centrifuged at 3000 rpm to sediment bacterial cells. The supernatant was discarded and the pellet was dissolved in 250 µl Suspension Buffer. Then 250 µl Lysis buffer was added and incubated for 5 minutes at room temperature. After that, 350 µl chilled Binding buffer was added and incubated on ice for 5 minutes. Later on, the sample was centrifuged at 13000 rpm for 10 minutes and the supernatant was applied onto the High Pure filter tube, which was placed into a collection tube. After centrifugation at 13000 rpm for 1 minute 500µl of Wash Buffer I was added, which was followed by another 1 minute centrifugation. Then, 700µl Wash Buffer II was added and sample was centrifuged again. Lastly, the filter tube was taken into a clean micro centrifuge tube and 50 µl Elution buffer was applied onto filter tube and centrifuged at 13000 rpm for 1 minute. In result, purified DNA was obtained in filtrate.

4.2.3.5. Confirmation of pSE+*SCN1B* Exon4 Vector Insertion and c.492 T>C Variant. Multiple cloning site of pSE vector was surrounded by KpnI enzyme restriction sequence. Thus digestion with KpnI enzyme and fragment analysis in agarose gel would give an idea whether cloning was successful or not. For KpnI digestion 10µl of Plasmid was incubated with 10 Units of enzyme at 37°C for 1 hour. And later, reaction mix was incubated at 65°C for 10 minutes to stop the reaction. Later restriction digestion was analyzed in 2% agarose gel as described before.

The c.492 T>C Variant created restriction site for RsaI enzyme. To check positive clones with variation KpnI restriction products were also digested with RsaI restriction enzyme. 10 µl of KpnI restriction digestion product was incubated with 10 units of enzyme at 37°C for 3 hours. Later restriction digestion was analyzed in 2% agarose gel.

4.2.3.6. Maintenance of SH-SY5Y Neuronal Cell Line Growth and Differentiation. Growth media of SH-SY5Y cells include 1:1 ratio of DMEM (no glucose, Invitrogen, USA) and F12 (Invitrogen), 10% FBS, 1% penicillin/streptavidin, 1% non-essential amino acids (Invitrogen). Media was refreshed every 2 (or 3) days and cells were splitted every 8 to 10 days. For differentiation, SH-SY5Y cells were plated as 70-80% confluent in growth media. The growth media was changed the day after plating to differentiation media (DMEM low glucose, 1% penicillin/streptavidin, 1% N2 supplement, retinoic acid (final concentration 10 µM)). Media was changed every 3 days and further differentiated for 9 days in culture.

4.2.3.7. Transfection of Differentiated cells. Cells were transfected at the 5th day of differentiation. SH-SY5Y Cells were seeded at 1.2 x 104 cells/cm^2 density in a 12-well dish one day prior to transfection. The transfection procedure was started when the culture had reached 70% confluence. Transfection was carried out according to manufacturer's protocol by using DNA: Lipo ratio of 0,2 µg DNA and 0.5 µl Lipofectamine 2000 (12 well plate). Subsequently 6 hrs. post transfection, the media was changed with the differentiation media. Transfected cells were differentiated for further 72 hrs.

4.2.3.8. Total mRNA Isolation from Transfected SH-SY5Y Cells. 2 days after transfection total mRNA was isolated from differentiated SH-SY5Y cells using Zymo quick RNA

miniprep kit as described by the manufacturer. Briefly, liquid media on the cells were removed by aspiration and 600 µl RNA Lysis Buffer was added on the cell monolayer. Cells were collected by pipetting up and down several times. The lysate was centrifuged at 10000 rpm for 1 minute to clear cell debris. A Spin-Away filter was placed onto the collection tubes, provided in the kit. Supernatant of previous centrifugation was applied on Spin-Away Filter and centrifuged at 13000 rpm for 1 minute. Flow-through was taken into a fresh tube and 1 volume EtOH (98-100%) was added. The mixture was transferred to Zymo-Spin IIICG Column and centrifuged for 30 seconds. Meanwhile DNase reaction mix was prepared as shown in the Table 4.10.

Table 4.10. DNase reaction mix.

Reagent	Amount
DNase I	5µl
10X DNase I reaction Buffer	8 µl
DNase/RNase free water	3 µl
RNA wash buffer	64 µl

400 µl of RNA wash buffer was added onto the column and centrifuged for 30 seconds. Flowingly, 80 µl of the reaction mix described above was put directly on the filter of the column and incubated at room temperature for 15 minutes. Then, centrifuged for 30 seconds.

400 µl of RNA prep buffer was added onto the column and centrifuged and then 700 µl of RNA wash buffer was added onto the column and was centrifuged again for 30 seconds. Finally, 400 µl of RNA wash buffer was added onto the column and centrifuged for 2 minutes. To elute RNA, Filter column was put into a clean tube. 30 µl of DNase/RNase free water was put onto the filter and centrifuged for 30 seconds. Eluted RNA concentration was analyzed by NanoDrop and also on agarose gel. Finally, isolated RNA was aliquoted and stored at -80°C.

4.2.3.9. cDNA Analysis. cDNA analysis was conducted in two steps. By first reaction total mRNA was converted into cDNA. Components in Table 4.11. were mixed and

incubated at 42ºC for 90 minutes and then at 85ºC for 5 minutes in a thermal cycler. After cDNA synthesis target region was amplified by a LongRange PCR step with reaction mix and thermal cycle conditions in shown in Table 4.12.

Table 4.11. cDNA synthesis reaction mix.

Component	**Stock Concentration**	**Volume in 20ul rxn**	**Final conc.**
LongRange RT Buffer	5X	4ul	1X
dNTP (10mM)	10mM	2µl	1mM
oligodT	20 µM	1 µl	1 µM
RNase inhibitor	4U/ µl	0.2 µl	0.04U/ µl
Reverse Transcriptase	1reaction/ µl	1 µl	1 Reaction
dH2O		6.8 µl	
RNA	~5ng/ µl	5 µl	1.25ng/ µl

Table 4.12. LongRange PCR reaction mix for amplification of the minigene product.

Contents	**Stock Concentration**	**Final Concentrations**
Buffer (+$MgCl_2$)	10X	1X
dNTP	10mM	0.5mM
Forward Primer	10 µM	0.4 µM
Reverse Primer	10 µM	0.4 µM
LongRange PCR enzyme mix	Not disclosed	0.4 µl
Reverse Transcription Reaction Product	Not measured	4 µl
Thermal Cycle		
Temperature ºC	Time	cycle
93	3sec	1
93	Hot Start	
93	5min	1
93	15s	35
64	30s	
68	90s	
68	5min	1
12	∞	1

4.2.4. Whole Genome Linkage Analysis for 4BF and 5BF Families

Whole genome linkage analysis was conducted in collaboration with Prof. Thomas Sander's Group at Cologne Center for genomics, University of Cologne, in the scope of the IntenC project. Selected individuals were genotyped using genome-wide SNP genotyping array Illumina HumanLinkage-24 chip and analyzed using EasyLinkage Program (Lindner

and Hoffmann, 2005) with Allegro (Epstein *et al*., 2005) and SimWalk2(Sobel *et al.*, 2002) modules.

4.2.4.1. Analysis of Linkage Region with GeneDistiller Program. Linkage studies often give large intervals with several genes. In order to define the genetic etiology of a disease, the gene in the linkage region that is responsible for the phenotype should be determined. However, it is a cumbersome task to examine all the genes underlying the linkage region manually. On the other hand, automated tools often focus on certain models. Genedistiller program is a useful tool that combines both approaches. It lists all the genes under linkage region, provides limits for directed filtering among these genes and gathers all the information about each gene from well-defined databases such as OMIM and KEGG (Seelow *et al.,* 2008).

4.2.5. Whole Exome Sequencing of 4BF14 and 4BF17 Samples

Patient samples 4BF14 and 4BF17 were sent for Whole Exome Sequencing at Cologne Center for Genomic in collaboration with Prof. Thomas Sander. After quality checking, sample enrichment was conducted with Illumina Exome Enrichment Kit. Next Generation Sequencing was conducted on an Illumina HiSeq2000 machine. Results were analyzed by Costin Leu, PhD, from Cologne Center for Genomics. All variant annotations were conducted according to human genome reference GRCh37 (hg19). Rare variants in the linkage region, shared by two patients were selected using Microsoft Excell Program. Uncovered regions were also checked and seen that there was no common region uncovered in both patients.

4.2.5.1. Sanger Sequencing of WES Variants in Linkage Region. Eight variants in the linkage region were selected for further analysis. Primers for the genomic regions with these variants were designed using Primer Blast tool in National Center for Biotechnology Information web site (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). These 8 variants were amplified in 5 amplicons. Selected variants were listed in Table 4.13. All amplicons were amplified in the same reaction condition using Qiagen Taq polymerase. Reaction mix. and thermal cycling conditions were summarized in Table 4.14. After PCR amplification samples were sent to Macrogen Inc. (South Korea) for Sanger Sequencing.

Table 4.13. Variants in the linkage region selected for Sanger validation.

Chr.19 locus (hg19)	**NCBI gene ID**	**Gene**
22574788	148198	*ZNF98*
22574806		*ZNF98*
33716564	56301	*SLC7A10*
33716570		*SLC7A10*
34304262	79047	*KCTD15*
35448919	126375	*ZNF792*
35448945		*ZNF792*
35500319	57655	*GRAMD1A*

Table 4.14. PCR conditions for *ZNF98*, *SLC7A10*, *KCTD15*, *ZNF 792* and *GRAMD1A* variant amplicons.

Reaction Mix in total 50 µl			
Contents	**Stock Concentration**	**Final Concentrations**	
Buffer (+$MgCl_2$)	10X	1X	
dNTP	10mM	0.2mM	
Forward Primer	10 µM	0.2 µM	
Reverse Primer	10 µM	0.2 µM	
Qiagen Taq Polymerase	5U/µl	0.02 U/µl	
DNA	100ng/µl	2 ng/µl	
Thermal Cycle			
	KCTD15	ZNF 792. SLC7A10. GRAMD1A	
Temperature ºC	Time	Time	cycle
95	5min	5min	1
95	30s	10s	35
66	30s	20s	
72	30s	30s	
72	5min	5min	1
12	∞	∞	1

4.2.5.2. Sanger Sequencing of WES Variants Out of the Linkage Region. Epilepsy or channel gene related variants in exome data were filtered according to the flow listed below.

- Variants that appear in Cologne Center for Genomics database more than 20 times were eliminated
- Mutation position: Non-synonymous coding and splice site variants were selected and and variants in UTRs, intergenic and Intronic regions and synonymous variants were eliminated:
- Benign variants according to Polyphen database were also eliminated.
- Variants with Phred consensus quality smaller than 20 were eliminated

Filtering resulted in 6139 unique variants. Among these we focused on epilepsy or channel related genes. Epilepsy genes were selected according to OMIM or PubMed with keywords of epilepsy or seizure. Channel genes were listed in the report of Klassen *et al.* where they screen all known channel genes in epilepsy patients (Klassen *et al*., 2011).

According to this analysis there were;

- 2 Complex indel
- 43 Splice site
- 3 Aberrant stop gained
- 32 Non-synonymous coding

Among these variants we selected 11 variants according to whether they were expressed in brain, or related to epilepsy syndrome with similar features to BFIS. Selected variants were listed in Table 4.15 below.

These variants were analyzed in 11 amplicons. Primers for the genomic regions with these variants were designed using Primer Blast tool in National Center for Biotechnology Information web site (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and purchased from Macrogen Inc. (South Korea). After PCR optimization patient samples were amplified in conditions summarized in Table 4.16, 4.17, and 4.18. Amplified PCR products were sent to Macrogen Inc. (South Korea) for purification and Sanger sequencing. Results were analyzed with Geospiza FinchTV program (PerkinElmer, USA) and CLC

MainWorkBench Program (Qiagen, USA).

Table 4.15. Selected variants in BFIS samples outside of the linkage region. (EXP: Expansion, AL: Amplicon Length (bp)).

Gene	Variant	Ensembl Gene ID	Ensembl Transcript ID	Ensembl Protein ID	Carrier	AL
EFHC1	F229L	00000096093	00000371068	00000360107	4BF17	165
EFHC1	R159W	00000096093	00000371068	00000360107	4BF14	380
CACNA1G	Q118H	00000006283	00000442258	00000409759	4BF14	371
GRM6	S829A	00000113262	00000319065	00000325675	4BF17	649
KCNB1	P775R	00000158445	00000371741	00000360806	4BF17	639
KCNQ4	G192R	00000117013	00000347132	00000262916	4BF17	189
CACNA1A	EXP	00000141837	00000357018	00000349520	4BF14 /4BF17	458
CACNA1A	D2136H	00000141837	00000357018	00000349520	4BF17	236
KCNJ10-	R18Q	00000177807	00000368089	00000357068	4BF17	553
GABRB3	C23X	00000166206	00000311550	00000308725	4BF17	1122
CNTNAP2	S16F	00000174469	00000455301	00000392208	4BF14	1023

Table 4.16. PCR Conditions for amplification of selected WES variants.

Component	**Stock Concentration**	**Amount in Reaction**	**Final Concentration**
dH2O		Up to 50 µl	
Buffer (+MgCl2)	10 X	5	1X
dNTP	10mM	1	0.2mM
Primer-F	10uM	1	0.2 µM
Primer-R	10uM	1	0.2 µM
Qiagen Taq Polymerase	5U/µl	0.2	0.02 U/µl
DNA	100ng/µl	1	2 ng/µl
Thermal cycle			
95	5min		
95	10	35cycle	
62	20		
72	30		
72	10		

Table 4.17. PCR Conditions for amplification of *CACNA1A*-Exp variant.

Component	**Stock Concentration**	**Amount in Reaction**	**Final Concentration**
dH2O		39.7µl	
Buffer (+MgCl2)	10 X	5 µl	1X
dNTP	10mM	2 µl	0.4mM
Primer-F	10uM	1 µl	0.2 µM
Primer-R	10uM	1 µl	0.2 µM
Qiagen Taq Polymerase	5U/µl	0.3 µl	0.03 U/µl
DNA	100ng/ µl	1 µl	2 ng/µl
Thermal cycle			
95	5min		
95	30	35cycle	
62	60		
72	60		
72	10		

Table 4.18. PCR Conditions for amplification of *GABRB3*-C23X variant.

Component	**Stock Concentration**	**Amount in Reaction**	**Final Concentration**
dH2O		39.7µl	
Buffer (+MgCl2)	10 X	5 µl	1X (1.5mM MgCl2)
dNTP	10mM	2 µl	0.4mM
MgCl2	25mM	2 µl	1mM
QSolution	5X	10 µl	1X
Primer-F	10uM	2 µl	0.4 µM
Primer-R	10uM	2 µl	0.4 µM
Qiagen Taq Polymerase	5U/µl	0.3 µl	0.03 U/µl
DNA	75ng/ µl	1 µl	1.5 ng/µl
Thermal cycle			
95	5min		
95	30	35cycle	
62	60		
72	60		
72	10		

4.2.6. Copy Number Variant Analysis for 4BF Family with Whole Genome Array Comparative Genomic Hybridization (array-CGH)

Targeted Copy Number analysis was conducted with custom Comparative Genomic Hybridization Array, which was designed using Agilent SureDesign tool. Design criteria were as follows:

i. The design included 15MB linkage region on chromosome 19
ii. Also included *PRRT2*, *KCNQ2*, *KCNQ3*, *SCN1A*, *SCN2A* genes.
iii. For regions above each exon had at least two probes. For larger exons there was one probe for every 100 bp.
iv. For introns there was one probe for every 200 bp.
v. ISCA CGH 60K backbone probes.
vi. For randomization, required replication and normalization probes were included into array.

For preparation, hybridization and analysis of arrays, Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis Enzymatic Labeling for Blood, Cells, or Tissues (with a High Throughput option) Protocol Version 7.3 March 2014. In summary, analysis started with 500ng DNA. Concentration of 8 test samples and reference Agilent male and Female DNA samples were set to 50ng/µl and each sample was digested by AluI and RsaI enzymes included in the SureTag Complete DNA Labeling Kit as described in the methods manual. Digestion products were labeled. Random primers were annealed to digested sample then gaps filled with exo(-) klenow enzyme using cyanine 3-dUTP for reference samples and cyanine 5-dUTP for test samples beside standard dNTPs. By this method test and reference samples were labeled in different color.

After labeling reaction, samples were cleaned using Agilent Purification Columns and yield and specific activity of each sample was determined using NanoDrop spectrophotometer. Namely absorbance at 260nm, 550nm (cyanine 3) and 650nm (cyanine 5) was measured. Specific activity and yield were calculated using formulas below:

Test and gender matched reference samples were combined in equal amounts. Then,

hybridization step took place and hybridization master mix was prepared as described in the manual page 67 and mixed with combined gDNA samples. Prepared hybridization mix for each of the 8 samples were applied on 8 different chambers of 8X60K array and put in a rotator in a hybridization oven and left for hybridization for 46 hours.

After hybridization, slide was washed with specific wash buffers and acetonitrile and also with Stabilization and Drying Solution. Afterwards, slide was scanned in NimbleGen MS200 (Roche, Germany) microarray scanner. Obtained .tiff file first analyzed with Agilent Cytogenomics software feature extraction tool. Later on samples were analyzed on CytoGenomics software with default parameters.

4.2.7. Quantitative Real Time PCR Analysis of *PRRT2* and *KCNQ2* Genes

Real Time Quantitative PCR analysis of *KCNQ4* and *PRRT2* genes were conducted on Light Cycler 480i (Roche, Germany) Real Time PCR machine, using SYBRGreen Master Mix (Roche, Germany) kit. For this analysis *eIF2B5* gene was used as reference. Reaction mix consisted of 1X master mix, 3 picomole forward and reverse primers and 20ng DNA. Thermal cycle was as shown in the Table 4.19. For each sample reference and target amplicons were amplified in three replicates and analysis was conducted via advanced relative quantification analysis tool of Roche LightCycler480i Software version 1.5. At the beginning of analysis, after amplification optimization efficiency calculation was conducted with two samples extracted by different extraction methods. Efficiencies of each amplicon were determined and it was seen that efficiency was not changing between samples.

After efficiency calculation, for each sample the target amplicon (*KCNQ2* and *PRRT2*) and the reference amplicon were amplified using the same amounts of DNA for both affected and healthy individuals. Healthy control was used as calibrator and target/reference ratio was normalized to 1 for calibrator taking the efficiency of two reactions into account. After each run specificity of amplification was analyzed by Tm Calling program. The relative quantification was conducted using relative quantification module. During this analysis, previously calculated efficiency of each amplicon was defined and program took this efficiency into account during calculation.

Table 4.19. Reaction mix for quantitative real time PCR (qRT-PCR).

Ingredient	**Concentration**	**Volume (µl)**	**Final Concentration in mix**
SYBR Green Master Mix	2X	10	1X
Forward Primer	10µM	0.3	0.15 µM
Reverse Primer	10µM	0.3	0.15 µM
DNA	5, 10,15,20 ng/ µl	2	0.5, 1, 1.5, 2 ng/ µl

Table 4.20. Thermal cycle conditions for quantitative real time PCR.

Program Name		Pre-Incubation					
Cycles		1		**Analysis Mode**		None	
Target (°C)	**Acquisition Mode**	**Hold (hh:mm:ss)**	**Ramp Rate (°C/s)**	**Acquisitions (per °C)**	**Sec Target (°C)**	**Step size (°C)**	**Step Delay (cycles)**
95	None	00:05:00	4.40	None	0	0	0
Program Name		Amplification					
Cycles		45		**Analysis Mode**		Quantification	
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step size (°C)	Step Delay (cycles)
95	None	00:00:10	4.40	None	0	0	0
66	None	00:00:20	2.20	None	60	0.5	5
72	Single	00:00:30	4.40	None	0	0	0
Program Name		Final Elongation					
Cycles		1		**Analysis Mode**		Quantification	
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step size (°C)	Step Delay (cycles)
72	Single	00:05:00	4.40	None	0	0	0
Program Name		Melting Curve					
Cycles		1		**Analysis Mode**		Melting Curves	
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step size (°C)	Step Delay (cycles)
95	None	00:00:05	4.40	None	0	0	0
65	None	00:01:00	2.20	None	0	0	0
97	Continuous	none	0.11	5	0	0	0
Program Name		Cooling					
Cycles		1		**Analysis Mode**		None	
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step size (°C)	Step Delay (cycles)
40	None	00:00:30	2.20	None	0	0	0

4.2.7.1. Size Determination Using Agilent Bioanalyzer 2100. After amplicons were optimized, in order to check specificity efficiently, samples were analyzed using Agilent Bioanalyzer 2100 machine with DNA 1000 chips and DNA 1000 kits as described by the manufacturer. In summary, the gel was prepared by mixing with the dye solution and centrifuging through a filter supplied in the kit. After the gel was ready, chip was first loaded with gel then 5µl marker solution was applied on each well of the chip. 1µl ladder solution was applied on the ladder well and 1 µl of test samples were applied on remaining wells. The chip was mixed using specific vortexed for 4 minutes and loaded on the machine. After DNA1000 assay program was chosen, the instrument started analysis.

4.3. Genetic Epilepsy with Febrile Seizures Plus (GEFS+) Study

4.3.1. Pedigree Drawing

Index patient for GEFS+ phenotype belonged to a large kindred of a village in North West Anatolia. In order to establish family relationships individuals in the village were interviewed and each individual was given an ID number. Smaller pedigrees showing the relationship between individuals were drawn by the clinicians. Pedigree information was converted into a list of individuals with ID numbers and parent information. In the context of this thesis, using this data a large pedigree was drawn showing the relationship of all affected individuals with each other. For this purpose, Progeny (Progeny Inc. USA) program was used.

During the first visit blood samples were also collected but samples were tagged with patient names. This caused ambiguity in samples. So for a reliable resampling, small sub-pedigrees showing IGE and PE phenotypes were selected. Individuals in these pedigrees were identified during another visit to the village. Resampling included selectively targeted individuals also.

4.3.2. Blood Samples and DNA Extraction

In order to follow inheritance of certain phenotypes the large pedigree divided into

three branches. One large branch shows Idiopathic Generalized Epilepsy (IGE) and two smaller pedigrees show Partial Epilepsy (PE). Due to the unreliability of samples obtained during first sampling, another sampling scheme was conducted.

Before visiting the village, individuals who will be included in the study were identified. They were assigned unique ID numbers and an EDTA blood tube was prepared for every selected individual to prevent sample loss. The village was visited again and new samples from selected individuals were collected as well as better phenotype data.

In the result of second visit 73 blood samples were collected. Genomic DNA was isolated from EDTA preserved bloods using QIAmp DNA isolation kit (Qiagen) according to manufacturer's description. Sample concentrations were determined by NanoDrop spectrophotometer.

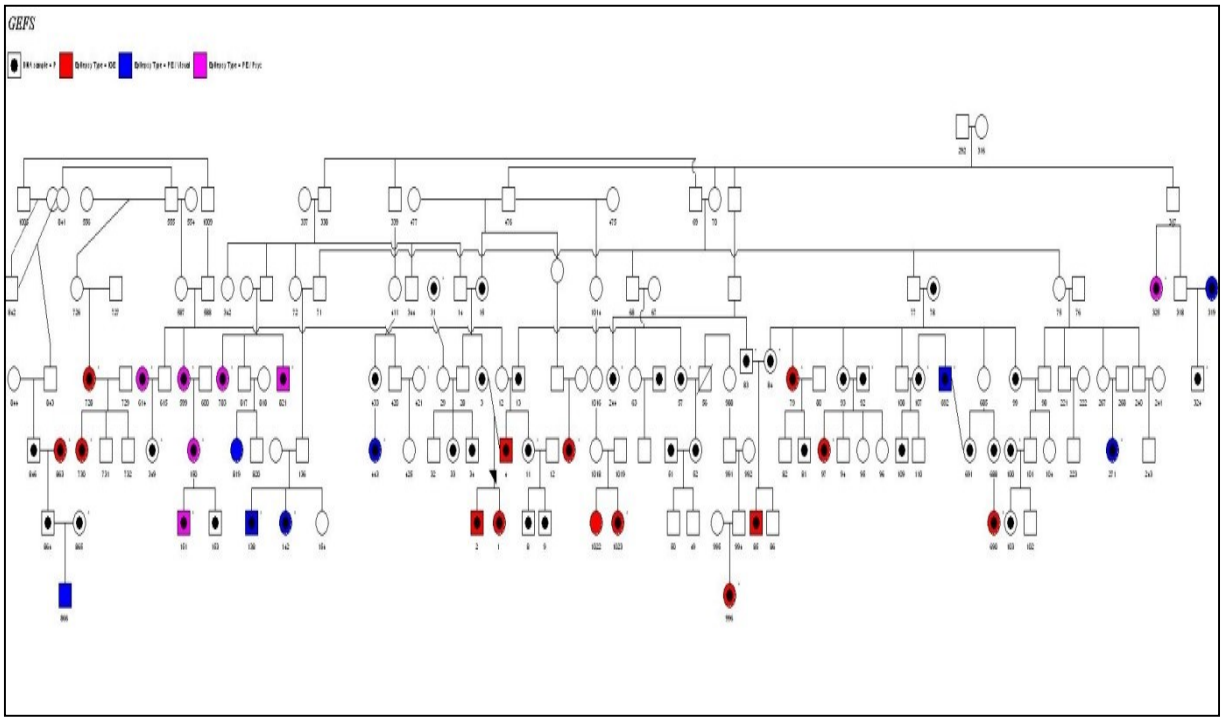


Figure 4.1. Representative family pedigree for GEFS+ patients. Different epilepsy types are shown in different color codes. Red represents IGE, blue represent visual partial epilepsy, and purple represent psychiatric partial epilepsy. Black spot in individual icons represents sampled individuals.

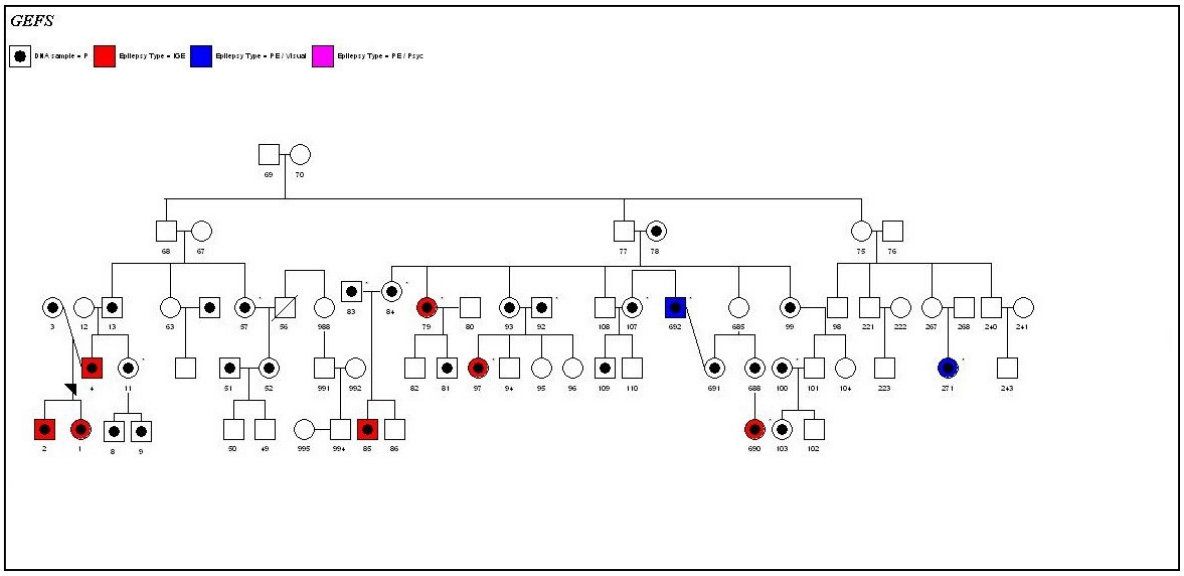


Figure 4.2. Sub-pedigree showing the relation of individuals with IGE. Different epilepsy types are shown in different color codes. Red represents IGE, blue represent visual partial epilepsy. Black spot in individual icons represents sampled individuals.

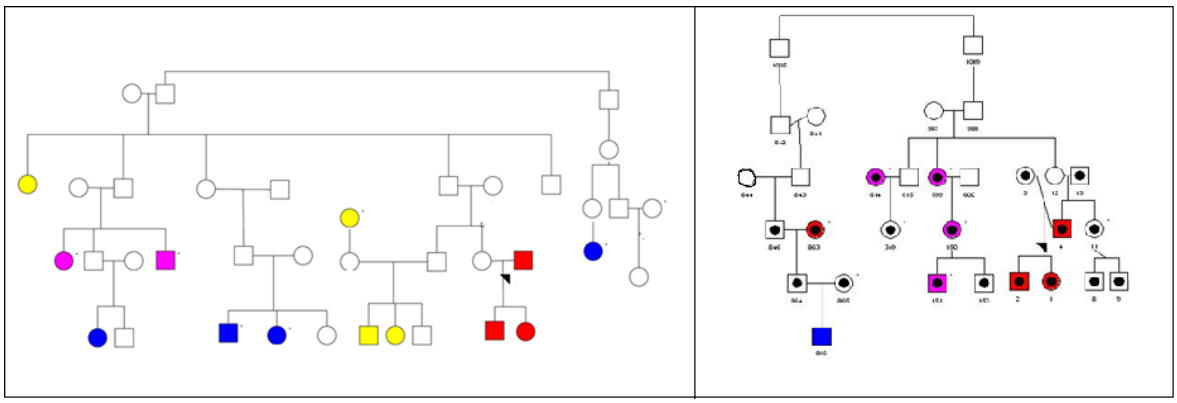


Figure 4.3. Two sub-pedigrees showing the relation of individuals with partial epilepsy. Different epilepsy types are shown in different color codes. Red represent IGE, blue represent visual partial epilepsy, purple represents psychiatric partial epilepsy and yellow represent motor partial epilepsy. Black spot in individual icons represents sampled individuals.

4.3.3. Whole Exome Analysis

Whole exome sequencing service was obtained from BGI-Hong Kong, initially for the samples GEFS1 and GEFS4 which were the index case and his affected daughter.

Since the exome data of two individuals were uninformative samples GEFS2 and GEFS3 were also sequenced. Whole Exome Enrichment was conducted with NimbleGen EZ Whole Exome Enrichment Kit and Sequencing was done in Illumina HiSeq2000 in 100X coverage. Figure 4.4 shows the basic bioinformatics analysis done by BGI-Hong Kong.

4.3.3.1 Filtering of Whole Exome Variants. Variants obtained from WES were analyzed in order to find plausible candidates to further investigate by Sanger sequencing. As first step 1807 high confidence variants were selected among all variants (2201). Next filtering step was selection of rare variants. Standard bioinformatics analysis of BGI includes a verification proposal column, where variants are classified variants into three groups according to their presence in BGI database. Using this data 1444 novel or rare variants were selected. After elimination of putative false positives there remained 1260 rare and high quality variant. Among these variants there were 3 variants on epilepsy related genes and 14 variants on channel genes and 4 variants on channel related genes. Six variants were selected for validation by Sanger sequencing, 3 of them were in epilepsy genes and chosen regardless of their share by patients. And 3 variants in channel genes were selected because they were expressed in the brain and were shared by two patients. All annotations were conducted according to human genome reference GRCh37 (hg19).

4.3.3.2. Sanger Validation of Selected Variants. PCR primers targeting genomic regions harboring these variants were designed using NCBI Primer Design Tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and purchased from Macrogen Inc. (Korea). PCR amplifications were conducted using Roche *Pwo* SuperYield Polymerase (Roche, Germany) in conditions described below.

Table 4.21. WES variants that were selected for validation.

Gene	**Variant**	**Product Length**
ATP1A2	I260R	329
BDKRB1	G241R	362
KCNJ16	R309M	325
NRXN2	insC;Het,frameshift	447
KCNV1	T388M	346
GRM3	D380Y	344

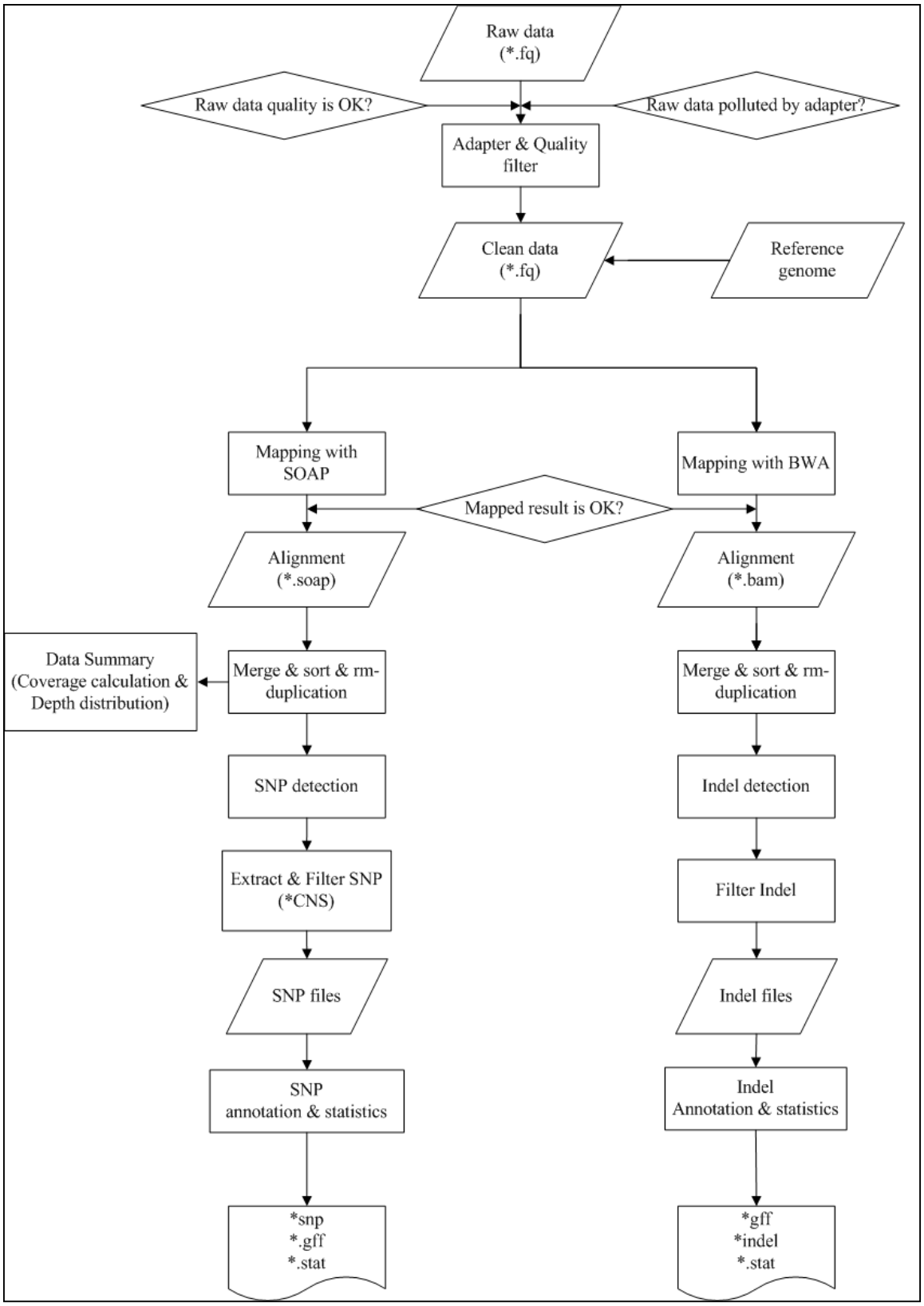


Figure 4.4. Whole Exome Sequencing standard bioinformatics analysis pipeline of BGI.

Amplicon for *ATP1A2* variant did not amplify by the conditions above. It was amplified using Roche Fast Start High Fidelity dNTP Pack Kit (Roche Germany) and in conditions in Table 4.22.

The PCR products were visualized on 2% agarose gels, and compared with a 100 bp and 1000bp DNA ladder (Fermentas, Lithuania) for length determination. The gels were checked under UV light and documented.

Adequate amount of PCR products were sent to Macrogen Inc., South Korea for PCR purification and automated sequencing. The results were obtained online as ABI document and analyzed in the Finch TV version 1.4.0, (Geospiza Inc., USA) chromatograph analyzing program.

Table 4.22. PCR amplification conditions for amplicons *BDKRB1*, *KCNJ16*, *NRXN2*, *KCNV1*, *and GRM3* variant amplicons.

Contents	**Stock Concentration**	**Amount (µl)**	**Final Concentrations**
Buffer (1.5mM $MgCl_2$)	10X	5	1X
dNTP	10mM	1	0.2mM
Forward Primer	10 µM	2	0.4 µM
Reverse Primer	10µM	2	0.4 µM
PWO SuperYield Taq Polymerase	5U/ µl	0.3	0.03 U/ µl
DNA	100ng/ µl	1	2 ng/ µl
Thermal Cycle			
Temperature ^{o}C	Time		cycle
94	5min		1
94	15s		35
66	15s		
72	30s		
72	10min		1
12	∞		1

Table 4.23. Reagents and conditions for PCR amplification of *ATP1A2* variant.

Contents	**Stock Concentration**	**Amount (µl)**	**Final Concentrations**
Buffer	10X	5	1X
dNTP	10mM	1	0.2mM
$MgCl_2$	25mM	5	2.5mM
Forward Primer	100 µM	0.3	0.6 µM
Reverse Primer	100µM	0.3	0.6 µM
Fast Start High Fidelity Taq Polymerase	5U/ µl	0.5	0.05 U/ µl
DNA	100ng/ µl	1	2 ng/ µl
Thermal Cycle			
Temperature ºC	Time		cycle
94	5min		1
94	30s		35
66	30s		
72	30s		
72	5min		1
12	∞		1

4.3.4. Whole Genome Linkage Analysis

For whole genome linkage analysis, SNP genotyping was done in two parts. First 15 individuals were selected and genotyped with Illumina Human CytoSNP 12 DNA Analysis Beadchips. The genotyping service was taken from Done Genetik (İstanbul, Turkey). After initial linkage analysis it was seen that these samples were not adequate for a significant LOD score, so another group 14 individuals were also selected and genotyped using the same microarray.

Before starting linkage analysis, in order to check sample relations Alohomora program (Ruschendorf and Nurnberg, 2005) was used. The program checked allele sharing frequencies between individuals. So in normal relations parent-offspring should have maximum allele sharing frequencies, then sib-pairs cluster together and un-relatives show much lesser allele sharing frequencies.

Whole Genome Linkage Analysis of 29 individuals is very difficult for any linkage program or computer. So 29 individuals were divided into 4 small pedigrees and was analyzed by Genehunter (Kruglyak, Daly, Reeve-Daly, and Lander, 1996) module of

EasyLinkage Program (Lindner and Hoffmann, 2005) with different inheritance models, namely autosomal recessive and dominant, also with different penetrance ratios. Whole genome linkage analysis was done using 1 marker in every 1 cM and analyzing these markers in sets of 100 and 75. For chromosomes with LOD score above 2, analysis was repeated for selected chromosomes in 1 marker in 0,1 cM resolution. And, lastly, for analysis of the linkage region on a chromosome every marker in the range was used. Other parameters used in the analysis were shown in the Figure 4.5.

4.3.4.1. Whole Genome Homozygosity Mapping. Besides linkage analysis homozygosity mapping was also conducted using the tool in homozygositymapper.org web site. This tool shows large homozygous blocks in whole genome SNP data and gives a score and is also a visual tool to scroll between different homozygosity blocks.

4.3.4.2. Analysis of Variants in the Linkage Region. After linkage analysis, linkage data was compared with the WES data and variants in the linkage region were selected for verification. Furthermore, miR137 microRNA region, which was also in the linkage region but not covered by WES, also analyzed for any sequence alterations by Sanger sequencing. PCR primers targeting genomic regions harboring these variants were designed using NCBI Primer Design Tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and purchased from Macrogen Inc., South Korea. PCR amplifications were conducted using Roche *Pwo* SuperYield Polymerase (Roche, Germany) in conditions described in Table 4.24.

4.3.4.3. miR137 Screen in Patients. The analysis of copy number of rs58335419, (-/CGCTGCCGCTGCTAC) was conducted on Agilent Bioanalyzer 2100 (Agilent Inc. USA) machine with DNA 1000 chips and DNA 1000 kits as described above. The length of the amplicon with 3 copies of VNTR is 310bp and each copy of VNTR adds 15 bp to the total length.

miR137 target region was amplified by PCR for every reliable DNA sample (109 samples in total) and analyzed first on 2% agarose gel and exact fragment length was determined using Agilent Bioanalyzer. Copy number of VNTR was deduced from the fragment size of amplicon which is automatically determined by the instrument program in

comparison to the intrinsic DNA ladder.

GeneHunter Options - easyLINKAGE Plus v5.08

Please adjust the following fields according to your needs

Inheritance model (parametric part of GeneHunter)

Dominant | Recessive | >1 liability class | LC file: NONE

GeneHunter programs:	GeneHunter v2.1r5
MOD options (only in effect for GeneHunter MODscore):	Global
Allele frequency algorithm:	Codominant
Used sib pair combination:	All pairs of affected / phenotyped sibs
Scoring function:	All
Compute sharing statistics:	On
Dump IBD:	Off
Recombination counting:	On
Eliminating less informative subjects:	Off
Frequency of the disease allele	0.0010
Max BITs (= max. number of meioses in the largest pedigree)	19
Analysis steps between markers	5

Autosomal genotypes	wt/wt	mt/wt (paternal) (only imprinting)	wt/mt (maternal)	mt/mt
Penetrance vectors	0.0000	0.7000	0.7000	0.7000

X chromosomal genotypes	wt/wt	mt/wt (paternal) (only imprinting)	wt/mt (maternal)	mt/mt
Penetrance vector female	0.0000	0.7000	0.7000	0.7000
Penetrance vector male	0.0000	0.7000	0.7000	

Advanced GeneHunter options for output handling

Haplotyping: Off | Analyze sets of markers | Markers per set: 100

Display only TOTALS | Display all family plots (very time consuming)

Note: All fields are required for the analysis.
Missing fields result in program termination.

Reset to predefined options | Ok | Exit

Figure 4.5. Parametric multipoint whole genome linkage analysis options for GeneHunter program.

Table 4.24. Reaction mixture and PCR conditions for the variants in GEFS+ family whole genome linkage region (A: Amount (µl); FC: Final Concentration).

Contents	Stock Conc.	*DPYD. AGL. COL11A1. INPP4A*		*AGL. MIR137*		*SEMA4C.*	
		A	FC	A	FC	A	FC
Buffer (1.5mM $MgCl_2$)	10X	2.5	1X	2.5	1X	5uL	1X
dNTP	10mM	1	0.4mM	0.5	0.2mM	1uL	0.2 mM
Forward Primer	10 µM	1	0.4 µM	0.5	0.2 µM	1uL	0.2 µM
Reverse Primer	10µM	1	0.4 µM	0.5	0.2 µM	1uL	0.2 µM
PWO SuperYield Taq Polymerase	5U/ µl	0.3	0.06 U/ µl	0.3	0.06 U/ µl	0.3 uL	0.03 U/ µl
DNA	60 ng/ µl	0.5	1.2 ng/ µl	0.5	1.2 ng/ µl	1uL	0.6 ng/ µl
Total Volume		25µl		25µl		50µl	
	Thermal Cycle						
Temperature ºC		Time	cycle				
94		5min	1				
94		15s	35				
AT		30s					
72		30s					
72		10min	1				
12		∞	1				

Table 4.25. Selected variants in GEFS+ family whole genome linkage region and product length of corresponding amplicons and annealing temperature for PCR reaction.

Gene	**Variant**	**Product Length (bp)**	**Annealing Temperature (oC)**
ABCD3	NM_002858.3:c.685-9T>C	261	58oC
DPYD	NM_000110: c.2194 A:p.V732I	216	58oC
AGL	NM_000645:c.3707 G>A:p.R1236H	186	60oC
COL11A1	NM_080629: c. 828A>C:p.K276N	200	60oC
INPP4A	NM_001134224.1:c.2801+22C>T	223	60oC
SEMA4C	NM_017789: c.1792G>A:p.E598K	344	56oC
miR137		310	64oC

4.3.4.4. Statistical Analysis of miR137 VNTR Pathogenicity Using SPSS. SPSS program was used for statistical analysis of *miR137* VNTR pathogenicity. Eighty-six samples with precise medical record were selected for analysis. VNTR copy number data were converted into total expansion score by summing up the VNTR copy number of each allele for each patient. Affection score of patients were evaluated in two different settings. In the first setting only epilepsy is considered as the disease phenotype and individuals with epilepsy were considered as affected. In second setting the phenotype was expanded with depression, anxiety or headache besides epilepsy, and individuals with either of the phenotypes are considered as affected. Total affection score was obtained by counting the number of observed phenotypes namely epilepsy, depression, anxiety or migraine. On the other hand, a two modal affection status was obtained by assuming individuals who carried one of the phenotype as affected and assuming individuals who had neither of these phenotypes as unaffected.

Before starting the analysis descriptive statistics of the cohort was produced by using descriptive statistics tool of the program. Then, the significance of association between miR137 VNTR expansion and affection status was analyzed by cross tabulation (chi-square tests), independent samples t-test and one-way ANOVA option of SPSS software in different settings.

First, a chi-square test between epilepsy status (w/o epilepsy) and expansion status (have *miR137* VNTR expansion or not) was conducted. Later on in order to check the difference of expansion length between individuals with or without epilepsy, independent samples t-test was conducted with 99% confidence interval. Lastly, one way ANOVA was used to determine if there was a difference in mir137 VNTR length between different epilepsy types.

For the second type of disease setting, chi-square test was conducted to test difference of *miR137* VNTR expansion status (expanded or not) between affected (individuals having at least one symptom) and unaffected (individuals without any symptom). Then, in order to check the difference of expansion length between affected and unaffected individuals independent samples t-test was conducted with 99% confidence interval. Lastly, one way ANOVA was used to determine if there was a significant

difference of *miR137* VNTR length between individual groups with different total affection score.

4.3.5. Whole genome CNV Analysis

Whole genome copy number analysis was conducted with two methods; with Illumina Human CytoSNP 12 DNA Analysis Beadchip and with Agilent SurePrint ISCA 8X60K Human Comparative Genomic Hybridization (CGH) array.

Illumina Human CytoSNP 12 DNA Analysis Beadchips has about 300,000 SNP probes distributed around known disease associated genes and CNVs. Beside SNP genotype data it produced copy number variance data including duplications, deletions, amplifications, copy-neutral LOH, and mosaicism. Experimental part of SNP and CNV genotyping was conducted by DoneGenetik (İstanbul, Turkey) and results were provided as SNP annotation file and CNV calling file. SNP annotation file was used for linkage analysis. Copy number analysis was conducted using Illumina Karyo Studio (Illumina Inc. USA) with CNV calling file. Results were exported as tab delimited text file, which was also analyzed with Microsoft Excel program (Microsoft Inc. USA). This analysis was conducted for samples selected for linkage analysis.

Another copy number analysis was conducted with Agilent SurePrint ISCA 8X60K human Comparative Genomic Hybridization (CGH) Array for samples GEFS1 and GEFS4. The experimental procedure and analysis was described in Section 4.1.10.

5. RESULTS

5.1. Genetic Analysis of 4BF Family

The index patient (4BF14) was diagnosed with BFIS at age of 5 months, had typical symptoms of BFIS namely seizures with deviation of head and eyes to one side, cyanosis, apnea and general motor arrest. Several members of the family had similar symptoms during infancy and remission about a year later. 14BF17 was deceased during the course of the study. For other patients in generations I and II clinical information were obtained during interviews with other family members especially with the member 4BF24 (Figure 5.1). Blood samples from 13 patients and 15 unaffected individuals were collected with informed consent of the family members. Blood samples of patients 4BF14, 4BF15 and 4BF16 were also collected in PAXgene (PaxAnalytics, USA) tubes for RNA isolation. The pedigree was analyzed by PedCheck (O'Connell and Weeks, 1998) to confirm family relationships.

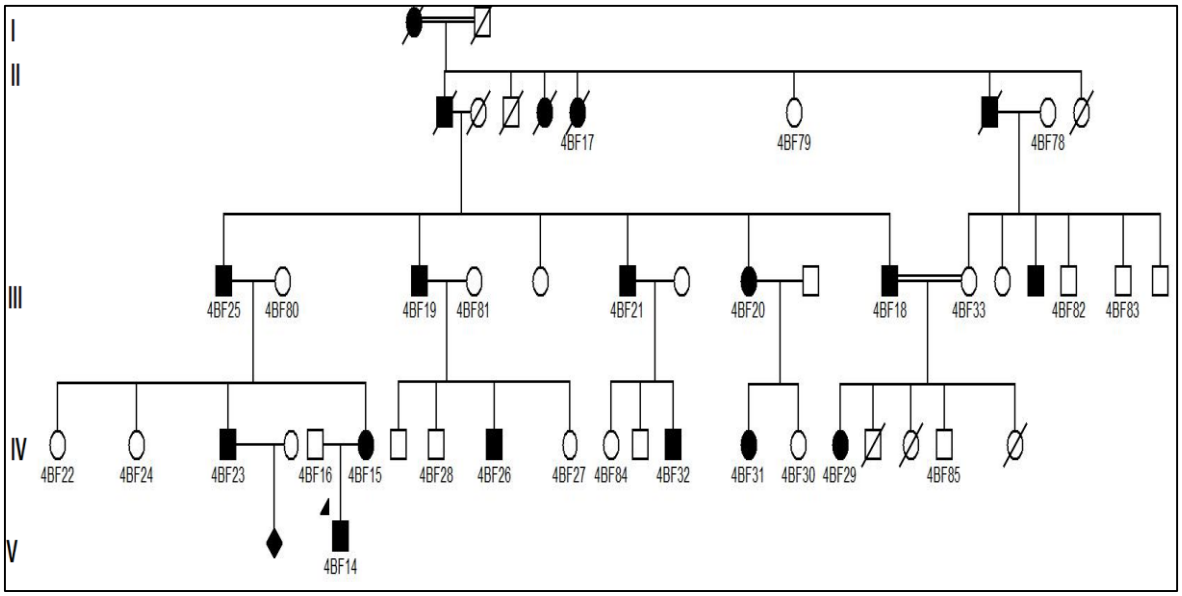


Figure 5.1. The 4BF family pedigree. Affected individuals are in black. Arrow denotes the index patient.

5.1.1 Target Gene Sequencing

The index patient of the 4BF family was screened for mutations in established candidate genes namely *SCN2A* and *KCNQ2* and part of putative *SCN1B* gene by Melek Aslı Kayserili in the course of her M.S. thesis work. One exon of *SCN1B* gene and *LGI4* gene were analyzed in the framework of this study. All genes above were negative for mutations in this family except *SCN1B*.

5.1.1.1. LGI4 Gene Analysis. *LGI4* was another gene found to be linked to BFIS in a previous study (Ishii *et al*, 2010) and was in approximately 90 kilobases away from *SCN1B* gene. Thus *LGI4* gene coding sequence and exon/intron boundaries were screened for any alteration in the the index patient by PCR amplification and direct sequencing of these amplicons. All of the amplicons were sequenced yielding no alteration (Appendix I).

5.1.1.2. *SCN1B* Gene Analysis. In the result of the analysis one novel synonymous change was found in the *SCN1B* gene (NM_001037.3 c.492T>C; NP_001028.1 p.T164T). Analysis of the family members for the novel variant by RsaI restriction digestion showed its co-segregation with the BFIS phenotype with a two-point LOD-score of 2.81 (AD inheritance, 70% penetrance). All of the 13 affected family members carried the mutation along with 3/15 healthy members of the family.

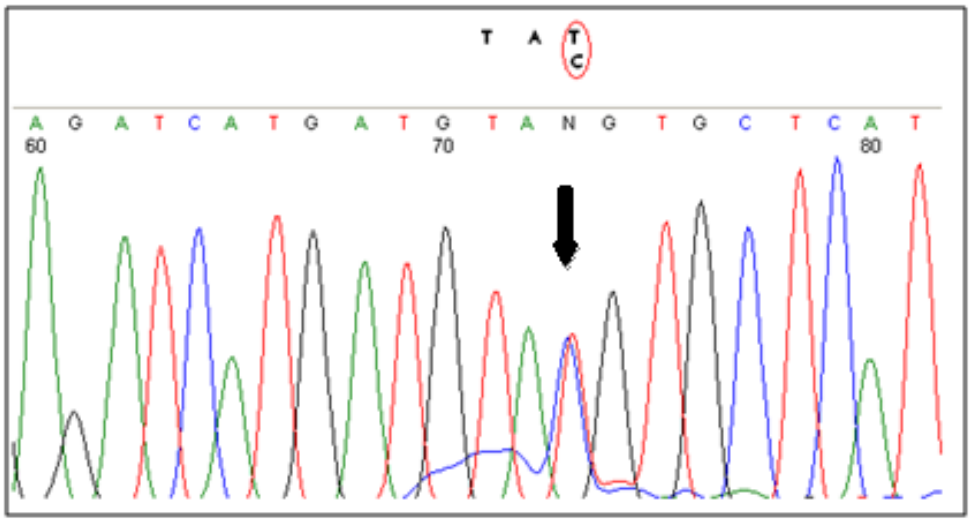


Figure 5.2. NM_001037.3, c492T>C transition in the BFIS index patient (4BF14). Arrow indicates the presence of the variant in heterozygous state (T/C).

The variant conserved in mammals was not observed in 461 unrelated and unaffected individuals in the Turkish population. ESE Finder analysis showed that the mutation created an additional binding site for splicing regulator SRp55 (data not shown) and PESX analysis showed that it abolished a silencer sequence (Figure 5.3) (Cartegni,*et al.*, 2003; Smith *et al.,* 2006).

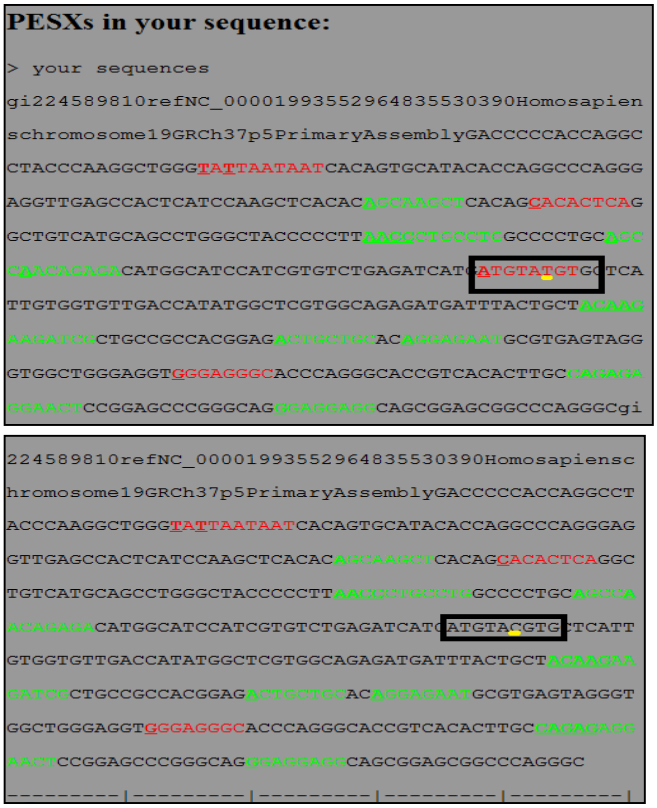


Figure 5.3. PESX analysis result. Red sequences show splicing silencers and green sequences show splicing enhancers. Upper panel is the normal sequence. Lower panel is the sequence with alteration. Altered nucleotide is underlined with yellow line.

5.1.2. Genome-Wide Linkage Analysis

The analysis was conducted by genome-wide SNP genotyping array, Illumina HumanLinkage-24 chip of the 28 family members. In the result of whole genome multipoint linkage analysis conducted by Allegro module with autosomal dominant inheritance and 70% penetrance, LOD score of 3.48 was found on 19p12-q13. (Figure5.4A). Fine mapping was carried out again by SimWalk module and resulted again with a 3.48 LOD score for the region between markers rs959419 (GRCh 37:20826692) and

rs870379 (GRCh 37: 35543756) (Figure 5.4B). The haplotypes in the linkage region is given in Figure 5.5. The linkage region included the *SCN1B* (19q13) gene together with 74 other protein coding genes. *LGI4* gene was not in the linkage region. Mutational analysis of the *LGI4* gene in the index patient did not reveal any alteration.

5.1.2.1. Analysis of the Linkage Region with GeneDistiller Program. The significant linkage region was delimited by: rs10410777 (GRCh37 Pos: 20403541) and rs870379 (GRCh37 Pos: 35543756) = 15.1Mbp on 19p12-q13. The region was very large because it contained the centromere.

- 146 genes (59 of which were pseudogenes).
- 78 protein coding genes.
  i. 14 genes for different zinc finger proteins.
  ii. Only *SCN1B* gene had epilepsy remark in its OMIM Report.
  iii. Three of them; *SLC7A10* (solute carrier family 7, (neutral amino acid transporter, y+ system) member 10), *SCN1B* (sodium channel, voltage-gated, type I, beta)and *HPN* (Hepsin) had nervous system phenotype.
- 1 t-RNA gene (*TRNAT-AGU* tRNA transfer RNA threonine (anticodon AGU)
- 7 miscRNA gene, 6 of them were hypothetical and one was microRNA 1270

5.1.3. Whole Exome Sequencing and Analysis of Variants in the Linkage Region

Although *SCN1B* mutation segregated together with the disease it was suspected that it could be due to the basic haplotype in the region and could mask another mutation in the linkage region. Since, there was not any other strong candidate in the region, patient samples 4BF14 and 4BF17 were sent for Whole Exome Sequencing at Cologne Center for Genomic in collaboration with Prof. Thomas Sander.

After selection of unique variants (explained in the Section 4.2.5.1) in the linkage region shared by two patients 6 variants on four genes, *ZNF98, ZNF792, SCN1B* and *GRAMD1A.* Three other variants in *SLC7A10* and *KCTD15* genes were also selected due to their functional similarity to epilepsy related genes. Table 5.1 shows all variants in 9 genes.

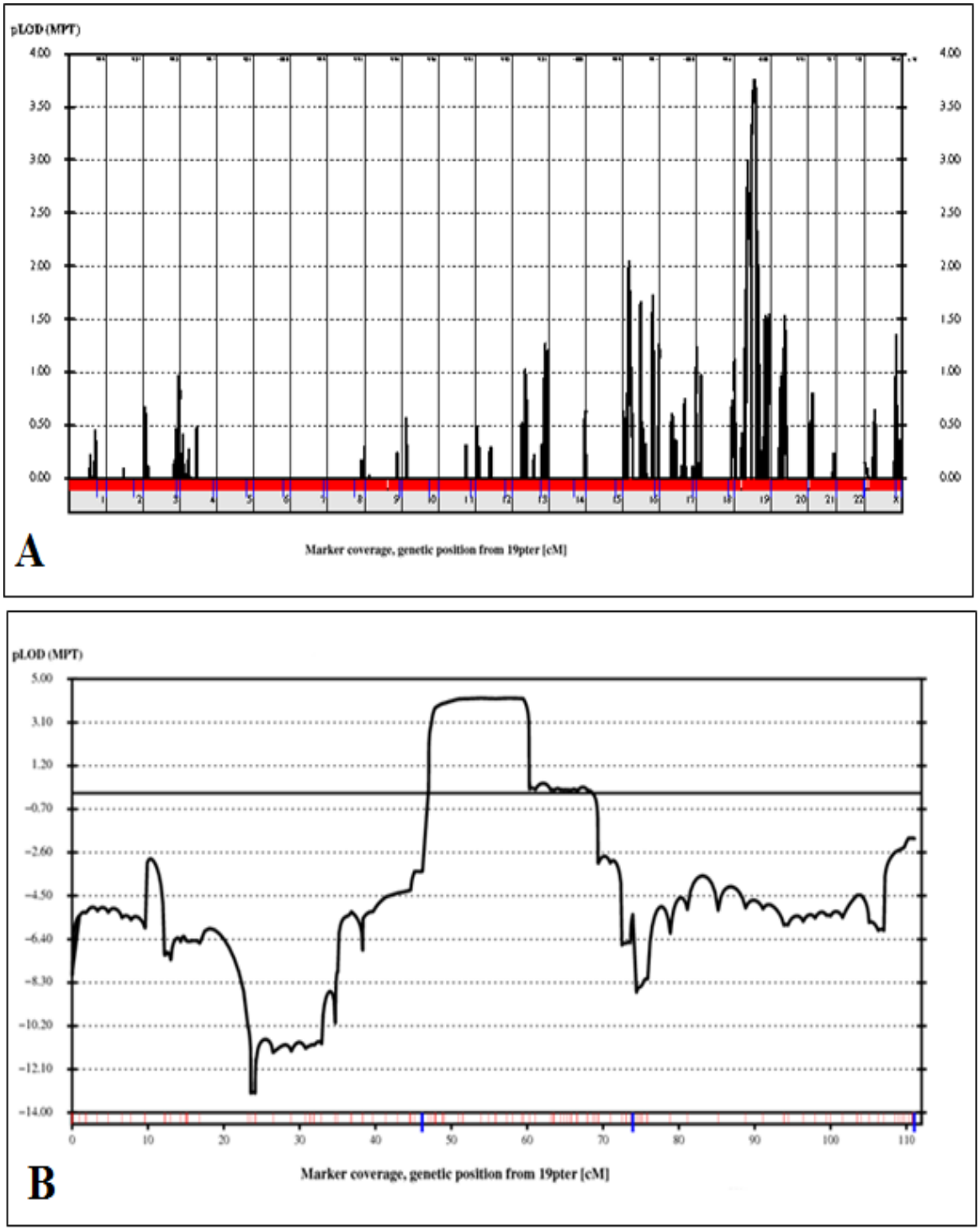


Figure 5.4. Genome-wide Linkage analysis result of the BFIS family. A. Whole genome linkage results obtained from Allegro module with AD inheritance and 70% penetrance model. B. Linkage analysis result of SimWalk module with AD inheritance and 70% inheritance model focusing on chromosome 19 linkage region.

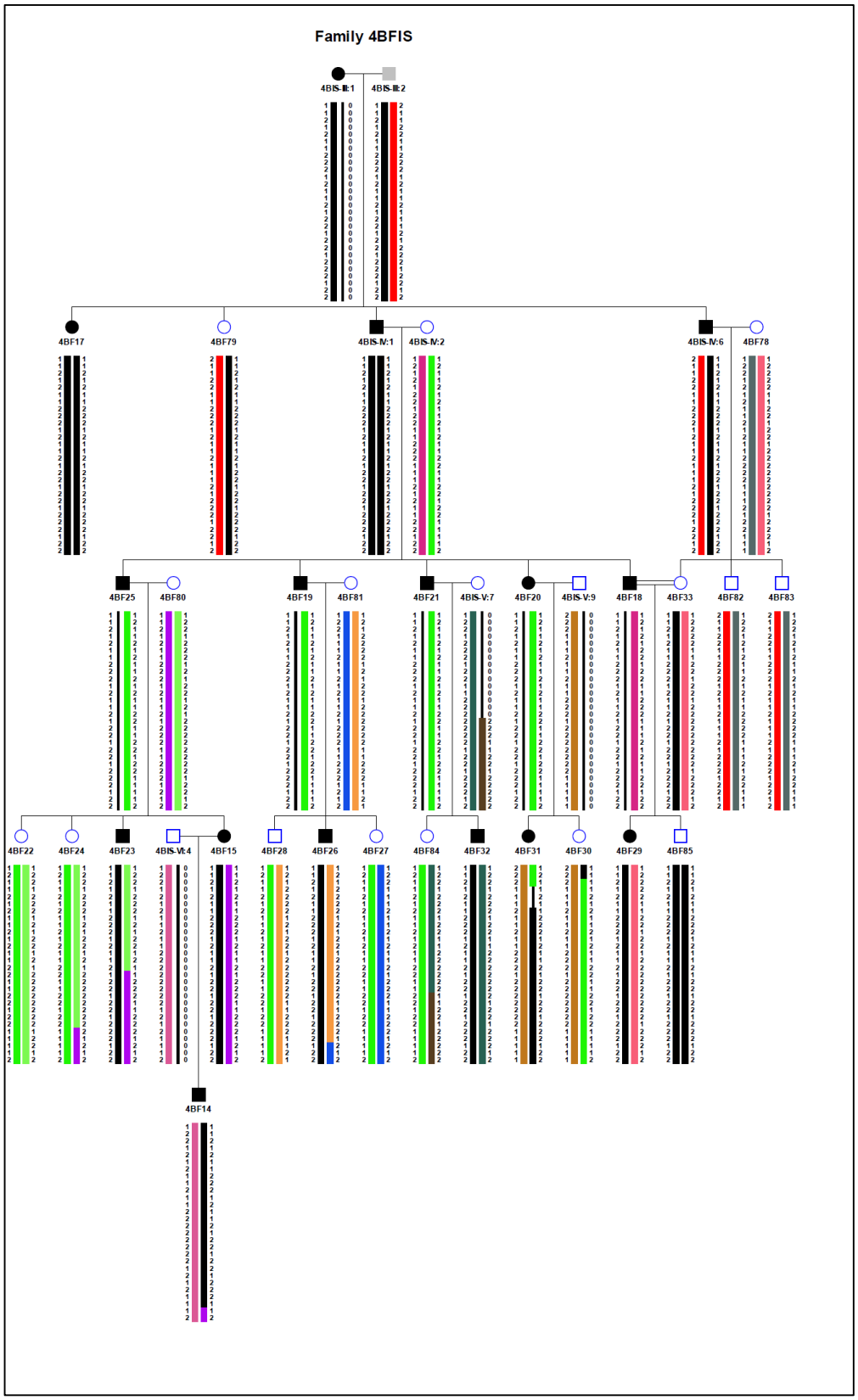


Figure 5.5. Haplotypes of 4BF family at linkage region, 19p12-q13. Co-segregating risk haplotype is shown as black bar. Individuals with a pedigree number are actually genotyped.

In the result of Sanger validation variants on *ZNF98* and *SLC7A10* were found to be false positive. Remaining variants were also analyzed in other family members. Indel variant on *KCTD15* gene showed low LOD score, on the other hand variants on *ZNF792* and *GRAMD1A and SCN1B* segregated together. However, variants on *ZNF792* were reported as benign by polyphen. Additionally, *GRAMD1A* splicing variant was not further evaluated since there was low expression of Gramd1a protein in the brain.

Table 5.1. Nine variants on six genes detected by WES analysis and confirmed by sequencing. (NSC: Nonsynonymous Coding, FsC: Frame-Shift Coding, SP: Splice Site, SC: Synonymous Coding, PrD: Probably Damaging; PsD: Possibly Damaging; B: Benign).

Chr.19 locus	**NCBI gene ID**	**Gene**	**Amino acid Change**	**Change Biotype**	**Polyphen Result**
22574788	148198	*ZNF98*	Arg417Trp	NSC	PrD
22574806		*ZNF98*	Cys411Ser	NSC	PrD
33716564	56301	*SLC7A10*	Pro16Thr	NSC	PrD
33716570		*SLC7A10*	Pro14Thr	NSC	PsD
34304262	79047	*KCTD15*	Gln297ProfsX74	FsC	
35448919	126375	*ZNF792*	Ala614Thr	NSC	B
35448945		*ZNF792*	Ser605Thr	NSC	B
35500319	57655	*GRAMD1A*	c.220-8C>T	SP	
35530064	6324	*SCN1B*	p.Thr164Thr	SC	

5.1.4. Analysis of Variants Outside the Linkage Region

In order to find a modifier variant, variants outside the linkage region were also analyzed. Filtering scheme was summarized in the Section 4.2.5.2. In the result of filtering 11 variants were selected according to whether they are expressed in brain, or associated with epilepsy syndromes with similar features to BFIS. Selected variants are listed in Table 5.2.

Specific Primers to amplify variants listed above were designed and co-segregation of these variants in the pedigree was analyzed. It was observed only the variants in *EFHC1* and *KCNJ10* genes were shared by other members of the family but two point LOD score for these variants were 0. Furthermore they did not showed any evident modifying effect. *PRRT2* mutation was, however, observed in other affected individuals also and analyzed in detail.

Table 5.2. Selected epilepsy or channel genes outside the linkage region validated by Sanger sequencing. (Chr.: Chromosome, Ref. All: Reference Allele Alt. All: Alteration Allele, HGMD: Human Gene Mutation Database, PrD: Probably Damaging, PsD: Possibly Damaging).

Variant	**Carrier**	**Chr.**	**Ref. All**	**Alt. All**	**SNP/HGMD ID**	**Polyphen**
*EFHC1*_F229L	4BF17	6	T	C	CM042023	PsD
*CACNA1G*_Q118H	4BF14	17	G	T		PsD
*GRM6*_S829A	4BF17	5	A	C		PsD
*KCNB1*_P775R	4BF17	20	G	C		PsD
*KCNQ4*_G192R	4BF17	1	G	C		PrD
CACNA1A-EXP	4BF14	19	CTG	-	rs16054	
	4BF17	19	CTG	-	rs16054	
CACNA1A-D2136H	4BF17	19	C	G		
KCNJ10-R18Q	4BF17	1	C	T	rs115466046	PsD
GABRB3-C23X	4BF17	15	G	T		
CNTNAP2-S16F	4BF14	7	C	T	rs74898808	
PRRT2 c.649dupC	4BF14	16	-	C		

5.1.4.1. *PRRT2* Gene Mutational Analysis. *PRRT2* (Proline Rich Transmembrane Protein 2) gene is located on chromosome 16, in a previously identified locus for BFIS and Paroxysmal Dyskinesia. However, the gene was associated with BFIS only very recently. First report about *PRRT2* as causative mutation for Paroxysmal Dyskinesia was in Journal of Medical Genetics at early 2011 by Liu *et al*. Later on 2012, reports of *PRRT2* mutations in BFIS came one after another (Heron *et al.*, 2012).

Although there were some other reported mutations, the most common mutation was a C insertion after a long 9 nucleotides stretch of cytosines. This was the main reason why

it was so lately discovered, since in all NGS data the variant was accepted as false positive, (which was frequently observed after long homopolymers).

Table 5.3. Explicit gene names and OMIM Records for selected variants outside the linkage region.

Gene	**Name**	**OMIM Record**
EFHC1	EF-hand domain containing protein I	Myoclonic juvenile epilepsy
CACNA1G	Alpha 1G subunit of T-type Voltage Dependent Calcium Channel	
GRM6	metabotropic glutamate receptor	Congenital night blindness
KCNB1	Voltage Gated Potassium Channel, Sharp-Related subfamily, Member 1	
KCNQ4	Voltage Gated Potassium Channel, QTL-Like subfamily, Member 4	Autosomal Dominant Deafness
CACNA1A	Alpha 1A subunit of P/Q type Voltage Dependent Calcium Channel	Spinocerebellar ataxia Episodic Ataxia Familial Hemiplegic Migraine
KCNJ10	Inwardly Rectifying Potassium Channel, Subfamily J, Member 10	SESAMES syndrome
GABRB3	GABA Receptor, Beta-3	Susceptibility to Childhood Absence Epilepsy; Insomnia
CNTNAP2	Contactin-Associated Protein-Like 2	Cortical dysplasia-focal epilepsy syndrome; Pitt-Hopkins like syndrome 1; Autism susceptibility 15
PRRT2	Proline Rich Transmembrane Protein 2	Convulsions, familial infantile, with paroxysmal choreoathetosis; Episodic kinesigenic dyskinesia; Benign familial infantile seizures, ,

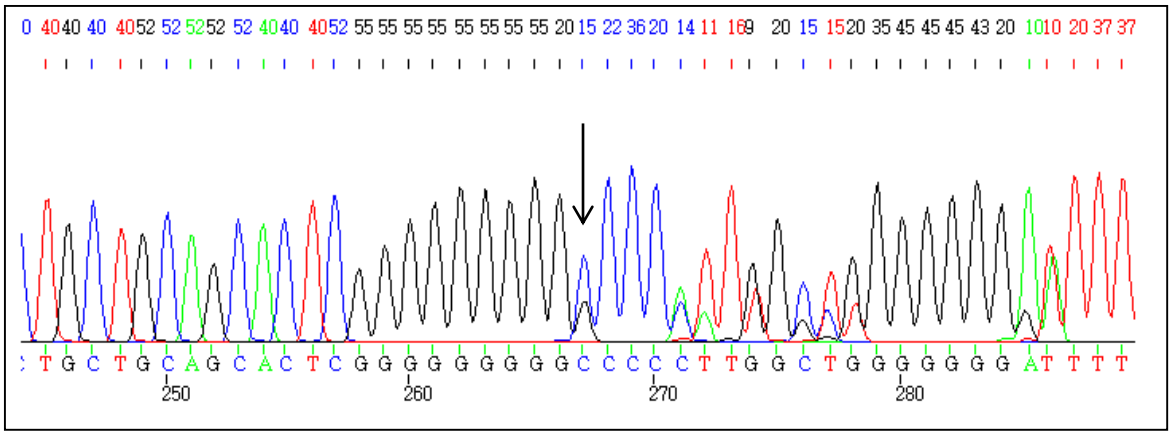


Figure 5.6. Sequencing result for common BFIS mutation; *PRRT2* c.649dupC, p.R217Pfs*8. Variant is shown by an arrow.

We identified the same mutation in the result of our WES data in index patient of 4BF family (4BF14). Then mutation was validated by Sanger sequencing (Figure 5.6) and co-segregation analysis was conducted in all family members. LOD score analysis for all variants done using SuperLink program under assumption of AD inheritance and 80% penetrance. Results are shown in the Table 5.4.

Taking linkage and WES analysis into account synonymous variant in *SCN1B* gene is the most plausible variant due to its high linkage score and association with channel gene. This variant was selected for functional analysis.

Table 5.4. LOD score analysis results for variants in linkage region and *PRRT2* variant on chromosome 16. SuperLink, AD inheritance, 80% penetrance.

Gene	**Variant**	**Two point LOD score (SuperLink)**
PRRT2	p.R217Pfs*8	1.0824
KCTD15	Gln297ProfsX74	2.0232
ZNF792	Ala614Thr	2.4955
ZNF792	Ser605Thr	3.0067
GRAMD1A	-8_8 splice site	3.0067
SCN1B	c.492 T>C	3.006

5.1.5. *SCN1B* cDNA Analysis

To see the effect of the silent mutation on splicing, total mRNA was isolated from the blood sample of the index patient (4BF14) and his parents (4BF15 and 4BF16). By two-step RT-PCR, cDNA was produced and using nested mRNA specific primers targeting half of exon 2 and half of exon 6 (570 bp), *SCN1B* coding sequence was amplified and sequenced. When RT-PCR products were analyzed on 2% agarose gel, all samples had the same expected size. These samples were also analyzed by Sanger sequencing that showed silent *SCN1B* c.492 T>C variation in heterozygous state. Overall evaluation of RT-PCR results of mRNA isolated from blood show that silent mutation did not have any effect on splicing and expression of the gene (Figure 5.7.).

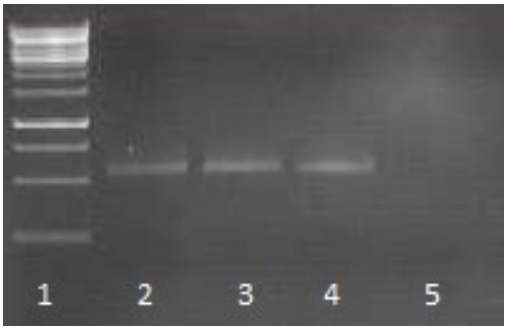


Figure5.7. mRNA analysis of *SCN1B* gene. Left: *SCN1B* partial mRNA amplified from index patient (lane 2), his mother (Lane 3) and his father (Lane 4). Lane 1 is 1 kb DNA Ladder and Lane 5 is negative control.

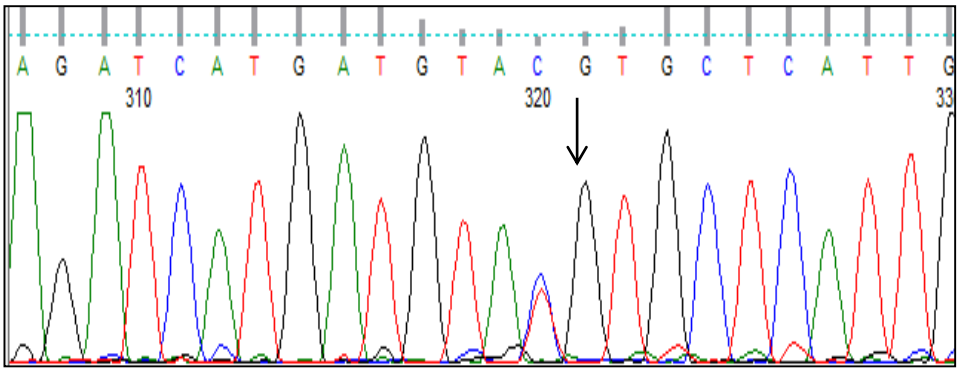


Figure 5.8. mRNA analysis of index patients showing *SCN1B* c.492 T>C variation in heterozygous state (shown by arrow).

5.1.6. Minigene Splicing Assay Analysis of *SCN1B* Variant

The search for any other mutation in the linkage region revealed that, *SCN1B* was the most plausible candidate in the region. Though synonymous the mutation, it may still have an effect on mRNA splicing as revealed by *in slico* analysis. Analysis of mRNA obtained from patient was not conclusive since the protein is expressed in the brain and mRNA is not present in adequate amounts in plasma. In order to conduct the analysis in more controlled environment an *in vitro* splicing assay by using *SCN1B* minigene was developed.

Exon 4 of *SCN1B* gene bearing the mutation was amplified with PCR and cloned into pSpliceExpress vector by a homologous recombination (Kishore, Khanna, and Stamm, 2008).

Plasmid was obtained as LB-Agar stab of *E. coli* containing the plasmid. Bacteria from the stab were grown in LB-Ampicillin medium and plasmid was isolated from the bacteria by using QIAGEN Plasmid Mini Kit as described by the manufacturer. Meanwhile, *SCN1B* gene exons 3 and 4 were amplified from genomic DNA of index patient and a healthy control. With another PCR reaction with adaptor primers AttB sequence was introduced into amplicons.

After confirming amplicons and plasmid sequences they were quantified with Picogreen dsDNA quantification Kit (Invitrogen) on Quantiflour fluorimeter (Promega). The recombination experiment was set using Invitrogen BP clonase enzyme mix along with its control PCR product as described by the manufacturer. After transformation of Z-Competent™ E.coli-DH5a cells (Zymo Research) with constructed plasmid and growing cell O/N in media with Ampicillin, constructed plasmid was isolated using Qiagen plasmid isolation kit and Roche miniprep kits.

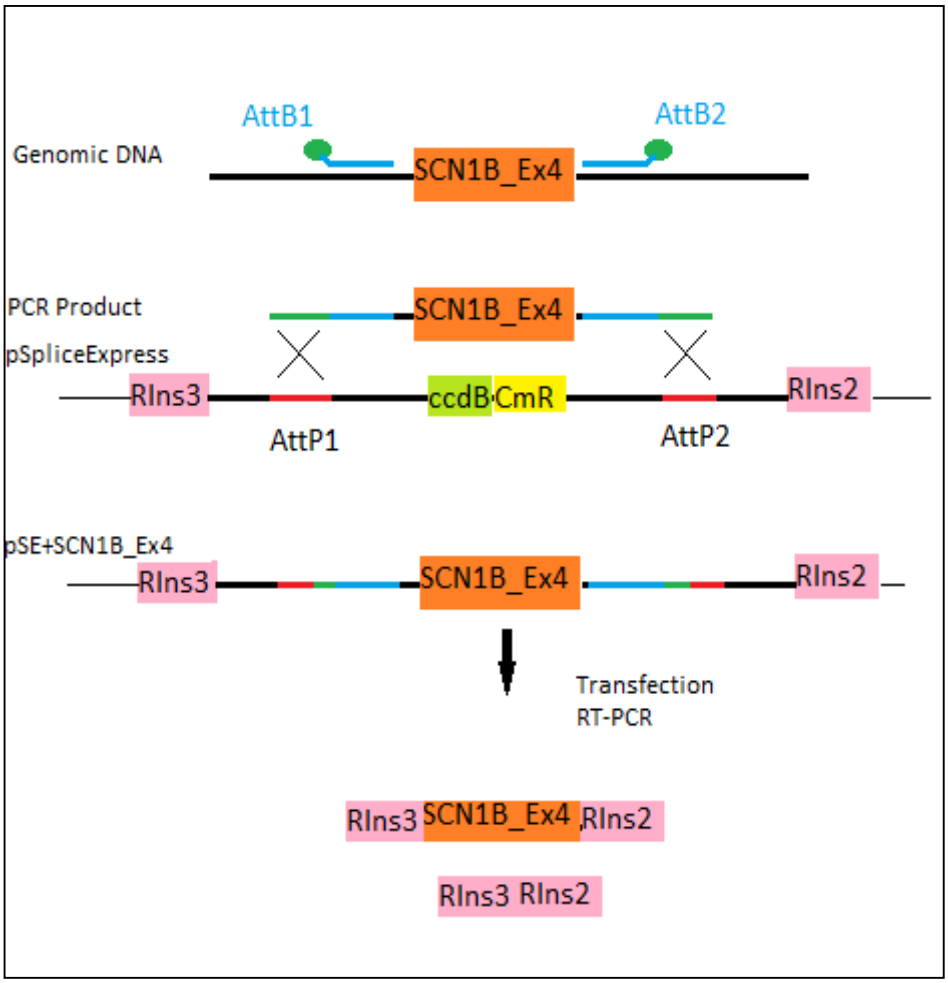


Figure 5.9. Workflow of splicing analysis with pSpliceExpress (AttB1&AttB2 translocation sites on primers; blue bars: sequence specific primers; Orange box: SCN1B gene exon 4; Pink boxes: Rat insulin gene exons 3 and 2 in order; AttP1 & AttP2 translocation sequences on vector; Green box: ccdB gene; Yellow box: CmR gene).

Recombination site was surrounded by KpnI restriction site and the *SCN1B* variation created restriction site for RsaI enzyme. Isolated plasmids were first cut by KpnI and later by RsaI enzymes and analyzed on agarose gel. After Sanger sequencing two clones were confirmed to bear desired mutation on *SCN1B* exon 4 inserted into pSE vector.

Table 5.5. Expected result of restriction cut of vector a plasmid with KpnI.

Expected fragments	Fragment 1	Fragment 2
Native plasmid:	2843bp	3916bp
Plasmid with SCN1B_Ex4	896bp	3916bp

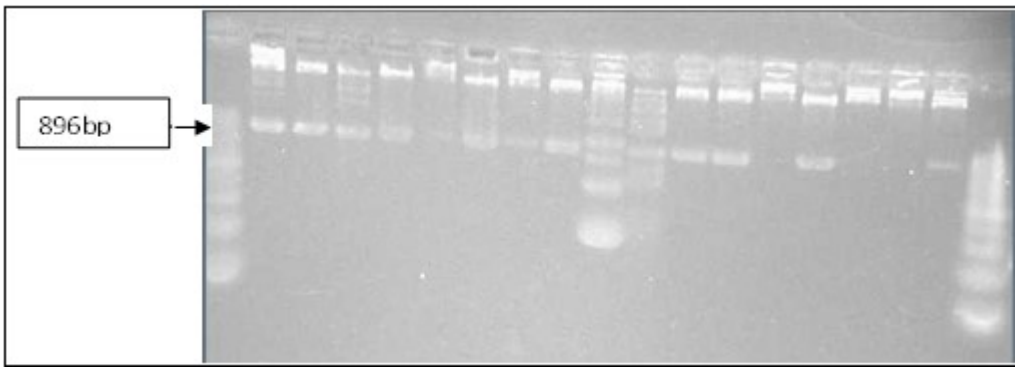


Figure 5.10. KpnI restriction digestion of recombinant plasmids with *SCN1B* exon 4 inserts. Arrow show expected band at 896bp.

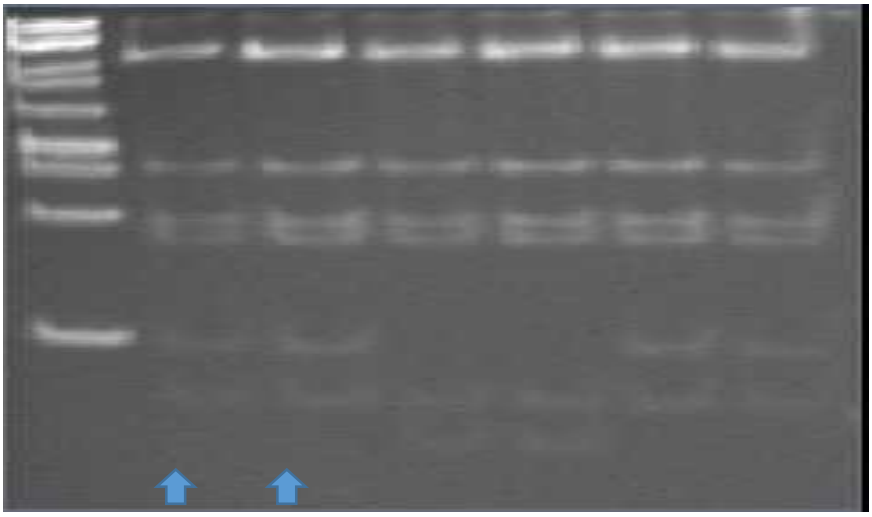


Figure 5.11. KpnI and RsaI restriction digestion of recombinant plasmids. Blue arrows shows the clones that have *SCN1B* exon 4 insert bearing c.492T>C mutation.

Three plasmids, one with *SCN1B* exon 4 with c.492T>C mutation, one with wild type *SCN1B* exon 4 insert and empty vector were used to transfect differentiating SH-SY5Y Neuroblastoma cells. Cells were incubated in differentiation media for further 2 days. After mRNA isolation, total mRNA was used as template for further reverse transcription and cDNA amplification. RT-PCR products were analyzed in 2% agarose gel and two bands for mutant and wild type *SCN1B* vectors and only one band for empty vector were observed. The upper band (450 bp) represents *SCN1B* exon4 surrounded by rat insulin exons and lower band (300 bp) represents rat insulin exons only (Figure 5.13).

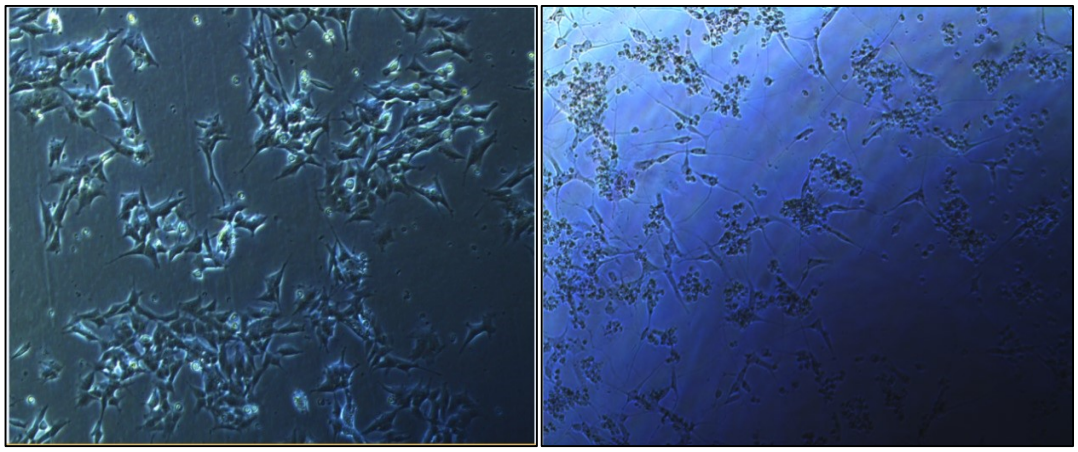


Figure 5.12. SH-SY5Y cells before (left) and after (right) differentiation.

In Figure 5.13, left panel shows the agarose gel picture. cDNA amplification of mRNA isolated from SH-SY5Y cells transfected by pSpliceExpress plasmid with *SCN1B* exon 4 with p.T164T mutation (lane 2), wild type *SCN1B* exon 4 (lane 3) and 300 bp empty vector with the two rat exons (lane 4). Lane 1 shows 100 bp ladder (Fermentas) and lane 5 shows no plasmid control. Right panel shows the schematic representation of the plasmid with the insert and the empty vector alone. Yellow and green arrows represent Rat Insulin 2 gene exon 3 and exon 2, respectively. Blue arrow represents human *SCN1B* gene exon 4. Pink and light blue arrows represent primers used for PCR amplification.

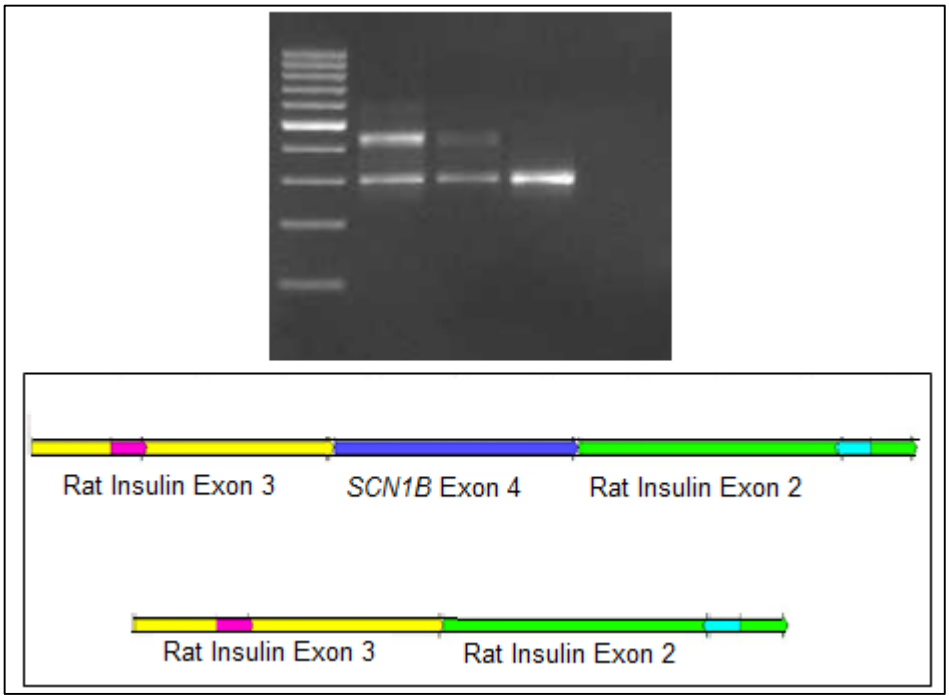


Figure 5.13. cDNA analysis of total mRNA isolated from transfected SH-SY5Y cells.

Despite equal amount of mRNA templates were used for all three reactions, we observed much fainter band at 450 bp in wild type sample compared to the mutant. We did not observe any band at 450 bp in empty vector sample. Overall the Minigene Assay indicates that the synonymous mutation p.T164T in *SCN1B* promoted exon 4 splicing in differentiating neuroblastoma cells, confirming *in silico* findings.

5.1.7. Copy Number Variation (CNV) Analysis of 4BF family

5.1.7.1. Custom array Comparative Genomic Hybridization (Array-CGH) Analysis. Although the *PRRT2* gene mutation was identified in 11/13 affected individuals and 6/12 healthy relative individuals in the pedigree, low penetrance of variant implied possible modulatory effect of another variant on chromosome 19 linkage region. Due to this fact, a custom Comparative Genomic Hybridization Array was designed using Agilent SureDesign tool for structural analysis of the genome. Design criteria were listed in the Section 4.2.6 and number of probes in each design group is listed on Table 5.6.

Table 5.6. Number of different probe groups on custom design CGH array.

Probe Group	**Number of probes**
Chr19 Exons	2759
Chr19 ıntrons	3958
PRRT2, KCNQ2, KCNQ3, SCN1A, SCN2A, F8	1536
ISCA CGH 60K backbone	40208
Human_CGH_1k_Agilent Replicate Probe Group	1000
Human_CGH_1k_Agilent Normalization Probe Group	1262

Seven samples from 4BF family and one positive control with known CNV were selected for CGH analysis with custom Agilent 8X60K CGH array. Experiment was carried out as described by manufacturer's manual. After conduction of assay, only results for sample 4BF18 gave LogR derivative spread result in a reliable interval (Figure 5.15). It became more obvious at chromosome 19 linkage region where probes were densely distributed.

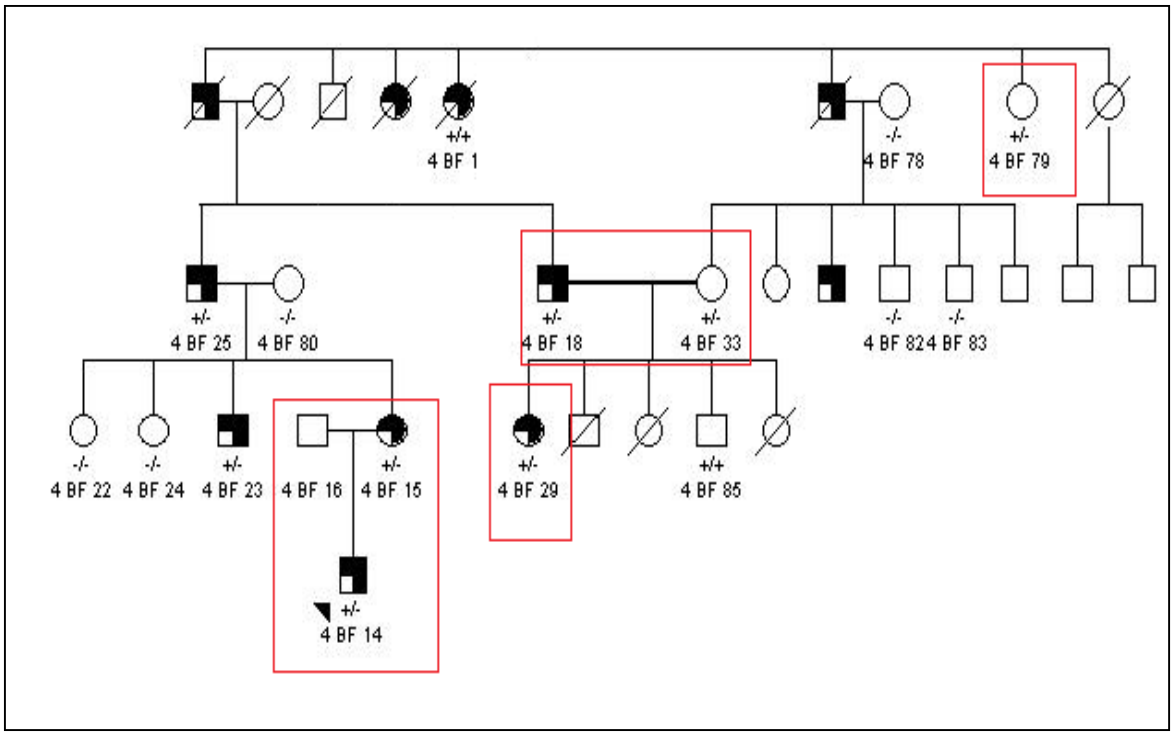


Figure 5.14. Samples selected for CGH analysis (individuals in red boxes).

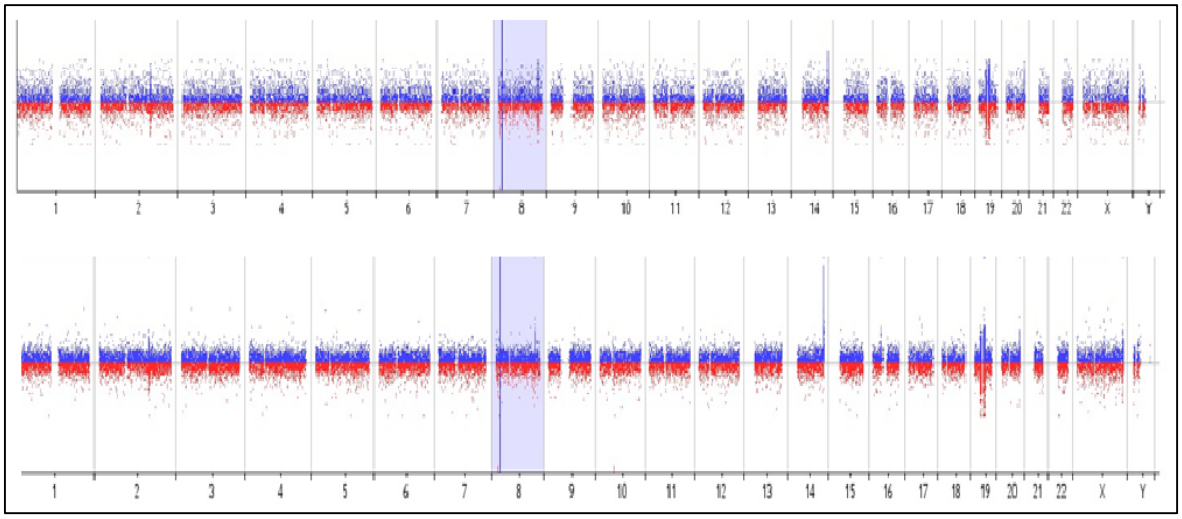


Figure 5.15. Whole genome LogR value distributions of probes. (Upper panel: sample 4BF14 and below panel: sample 4BF18).

In the result of CGH analysis of sample 4BF18 there were copy number alterations in several regions (Table 5.7). Some of them were due to copy number differences in the reference DNA, showing the sensitivity of analysis. Although there were some variants with low LogR value that were most probably false positive. However there were 2 regions

deserve further investigation since they harbored genes implicated in BFIS namely *PRRT2* and *KCNQ2*. Figure 5.16 shows the LogR value distribution of probes on *KCNQ2* and *PRRT2* genes. Since sample 4BF18 was affected but do not carry frameshift mutation on *PRRT2* gene, dosage alteration could be underlying the etiology of the phenotype in this patient.

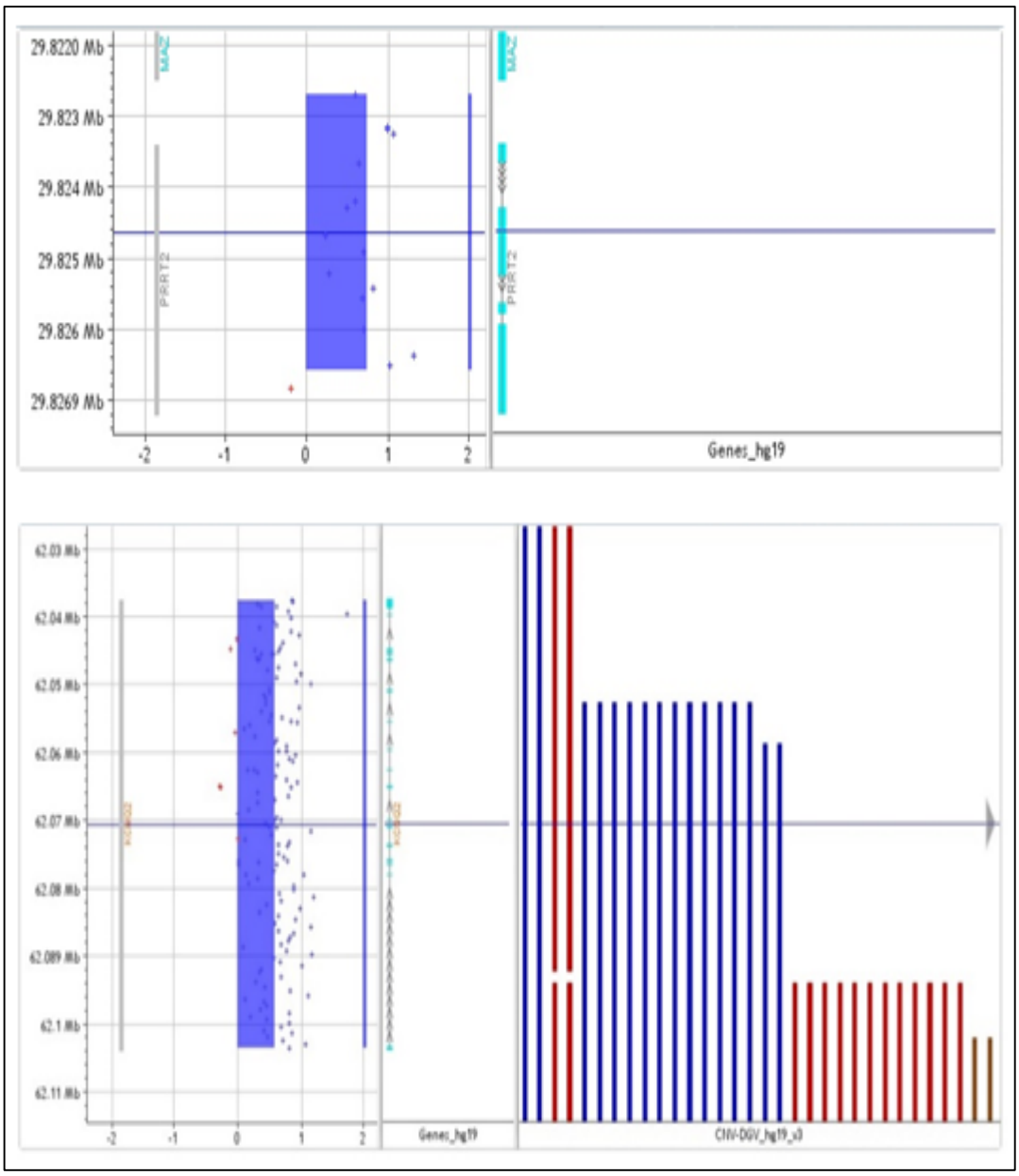


Figure 5.16. LogR value distribution of probes on genes *PRRT2* (upper panel) and *KCNQ2* (lower panel) for sample 4BF18.

Table 5.7. Result of CGH analysis for sample 4BF18 (Amp/Del: Amplification or Deletion, positive values denote amplification and negative values denote deletion).

Chr	**Start-Stop (bp)**	**Size(bp)**	**Amp/Del**	**Annotations**	**Comment**
Chr2	166128213-166165930	37,718	0.268926	*SCN2A*	Low logR value
Chr8	166128213-166165930	961,200	-0.692122	*FAM66B, DEFB109P1B, ZNF705G...*	Amplification in reference
Chr8	166128213-166165930	915,134	0.287457	*KCNQ3, HPYR1, LRRC6...*	Low logR value
Chr10	46949255-47655146	705,892	-0.671925	*SYT15, GPRIN2, PPYR1*	Amplification in reference
Chr14	106405703-107124579	718,877	1.379820	*ADAM6, NCRNA00226,*	Deletion in reference
Chr14	106405703-106716404	310,702	3.554276	*ADAM6, CpG:18*	Deletion in reference
Chr14	106957891-107070613	112,723	0.684158		Deletion in reference
Chr16	29822705-29826578	3,874	0.735788	*PRRT2, CpG:57*	Should be evaluated
Chr19	29999922-30022116	22,195	0.808485	*VSTM2B, CpG:447*	Population CNV
Chr19	33473865-33504623	30,759	0.540914	*RHPN2*	Population CNV
Chr19	33663661-34309015	645,355	0.409189	*WDR88, LRP3, SLC7A10...*	High probe density on linkage region
Chr19	34838775-34895921	57,147	0.419698	*KIAA0355, GPI, PDCD2L*	High probe density on linkage region
Chr19	34971899-35034473	62,575	0.489833	*WTIP, CpG:180, CpG: 14...*	High probe density on linkage region
Chr19	35487216-35537559	50,344	0.557573	*GRAMD1A, SCN1B, HPN*	High probe density on linkage region
Chr20	62037542-62103647	66,106	0.578032	*KCNQ2, CpG:20, CpG: 68...*	Should be evaluated
ChrX	154113924-154117574	3,651	0.836290	*F8, F8A1, F8A2...*	May be population CNV

5.1.7.2 Quantitative Real Time PCR (qRT-PCR) Analysis of *PRRT2* and *KCNQ2* Genes. CGH results for BF18 sample were confirmed and other samples were also analyzed by Quantitative Real Time PCR. For this analysis, relative quantification with calibrator

method was used. Eukaryotic translation elongation factor B subunit epsilon (*eIF2B5*) gene was used as reference since its amplification or deletion has major effect on phenotype, if not lethal.

At the beginning of the analysis, PCR amplification of the amplicons was optimized. Specific amplification was checked by Tm Calling program and also in Agilent Bioanalyzer electronic electrophoresis instrument. Presence of a single peak in melting curve and a single band in Bioanalyzer traces indicates the specific amplification (Figures 5.17-5.21). Besides, to check whether correct amplicon was amplified, qPCR products were Sanger Sequenced.

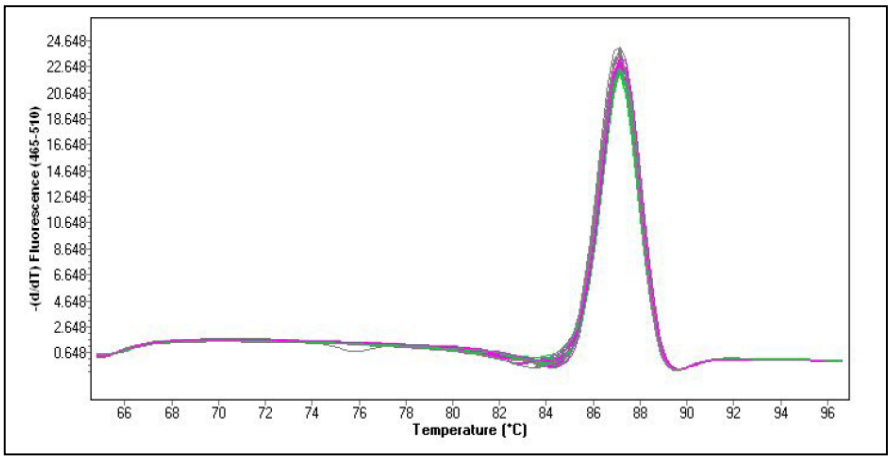


Figure 5.17. Melting peak for *eIF2B5* Exon 15.

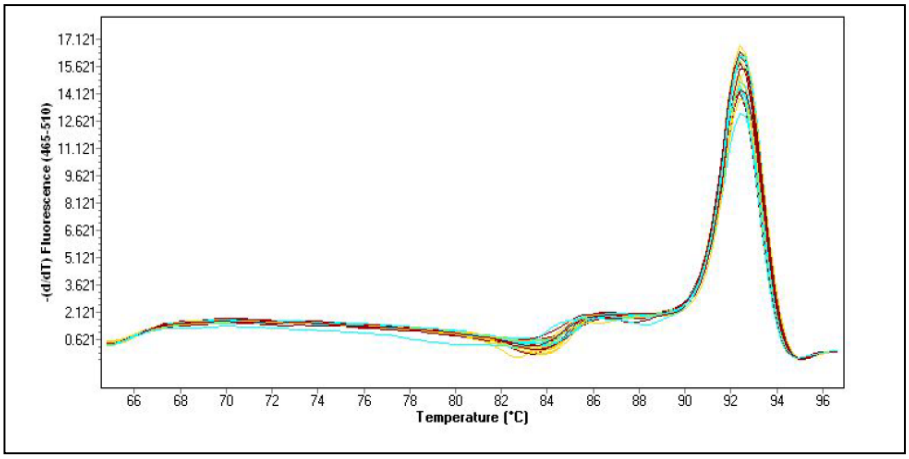


Figure5.18. Melting peak for *KCNQ2* Exon 4.

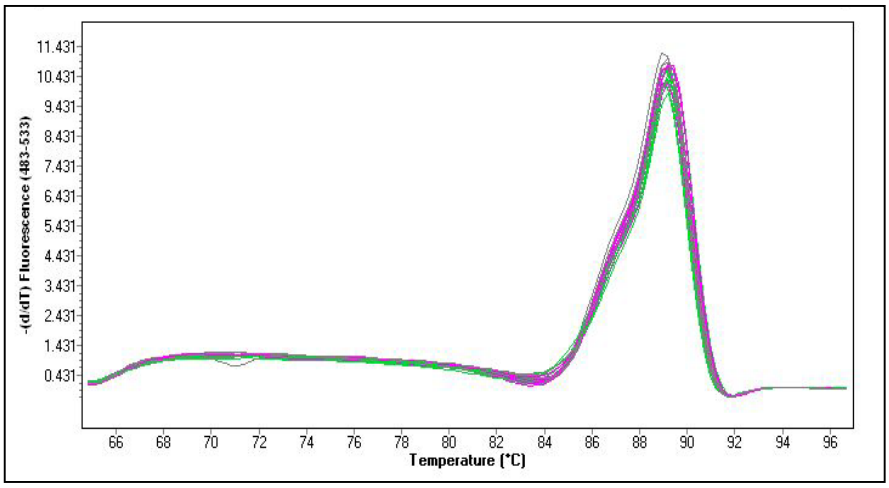


Figure5.19. Melting peak for *PRRT2*.

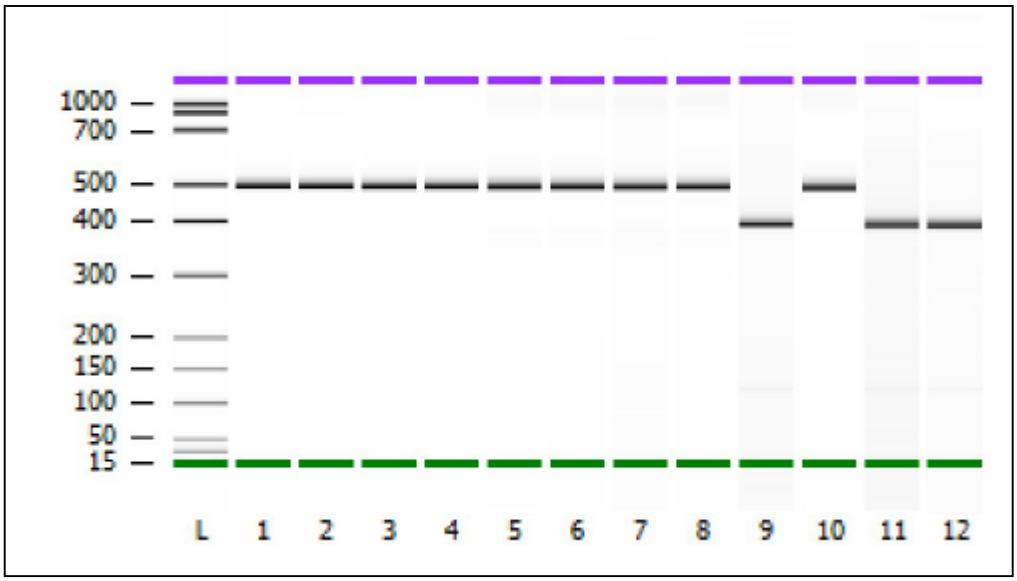


Figure 5.20. Representative gel image produced by Agilent Bioanalyzer. Each line corresponds to a qRT-PCR product. First lane is the ladder and band sizes are indicated on left. Green and purple lines on every lane show the marker band.

After amplification optimization efficiency calculation was conducted with two samples extracted by different extraction methods. Efficiency of each amplicon was determined and it was seen that the efficiency was not changing between samples (Figures 5.22-5.24).

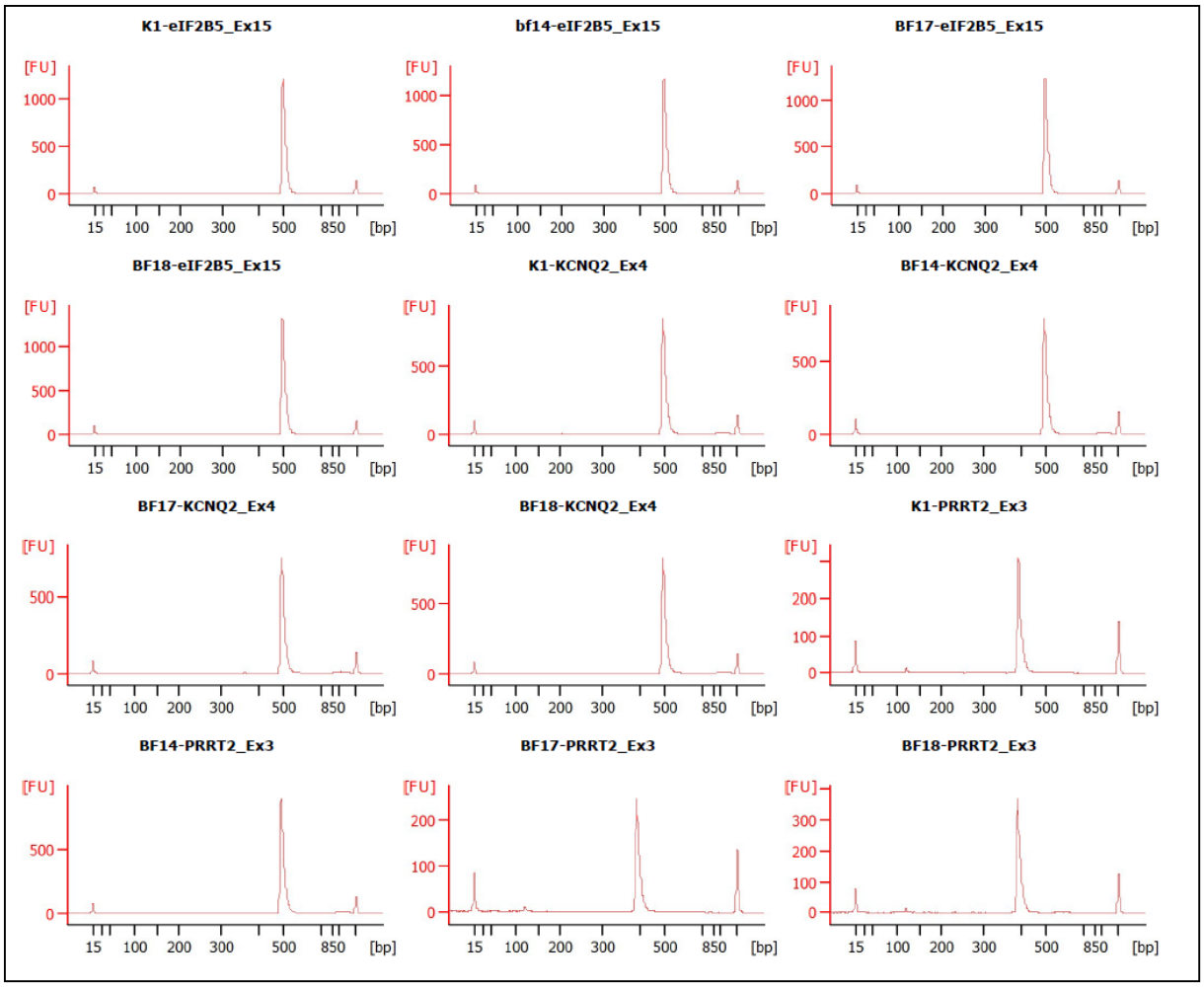


Figure 5.21. Representative Electrophoregraph produced by Agilent Bioanalyzer. Each graph represents a qRT-PCR product.

After efficiency calculation, for each sample, target amplicon (*KCNQ2* and *PRRT2*) and reference amplicon were amplified using same amount of DNA for both affected individuals and a healthy individual. Healthy control was used as calibrator and target/reference ratio was normalized to 1 for calibrator taking the efficiency of two reactions into account. Each sample was analyzed in triplicate in a run and each reaction repeated at least three times for every sample.

In the result of the analysis for *KCNQ2*, copy number gain in 4BF18 sample was confirmed and 5 other affected individuals found to have copy number gains together with two unaffected individuals. Father of the index patient (4BF16) was found to have only two copies of *KCNQ2* gene. qRT-PCR result is depicted in the Figure 5.25. which also

shows the relative ratio for each individual in the table below the graph.

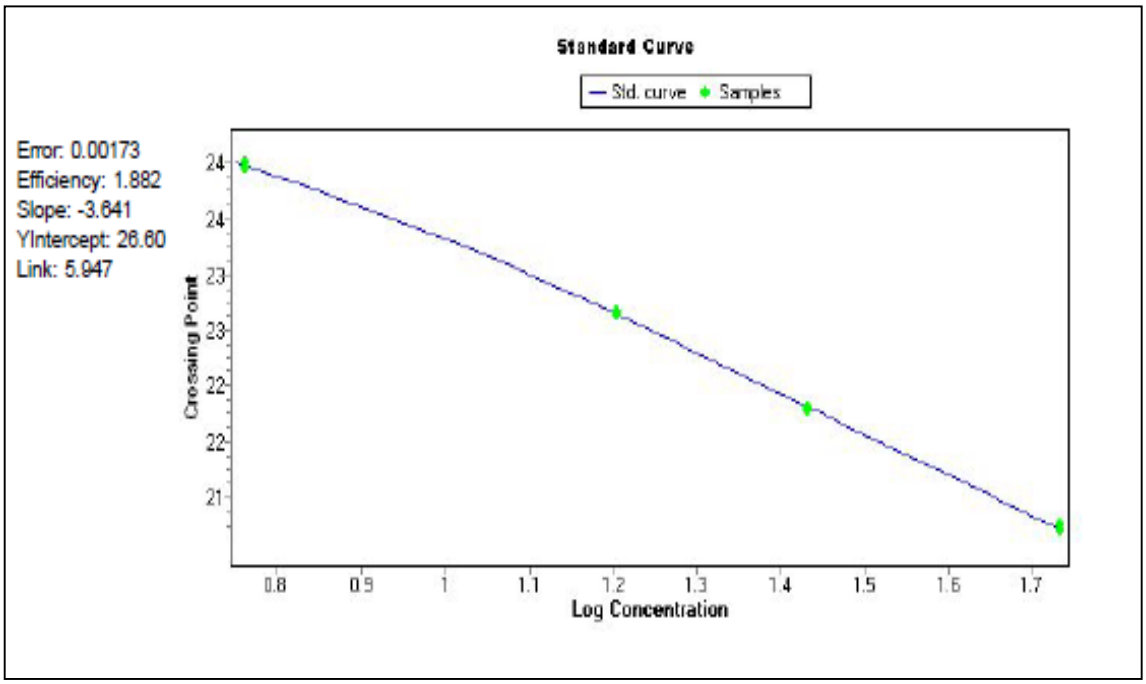


Figure 5.22. Efficiency calculation for *eIF2B5*_Ex15.

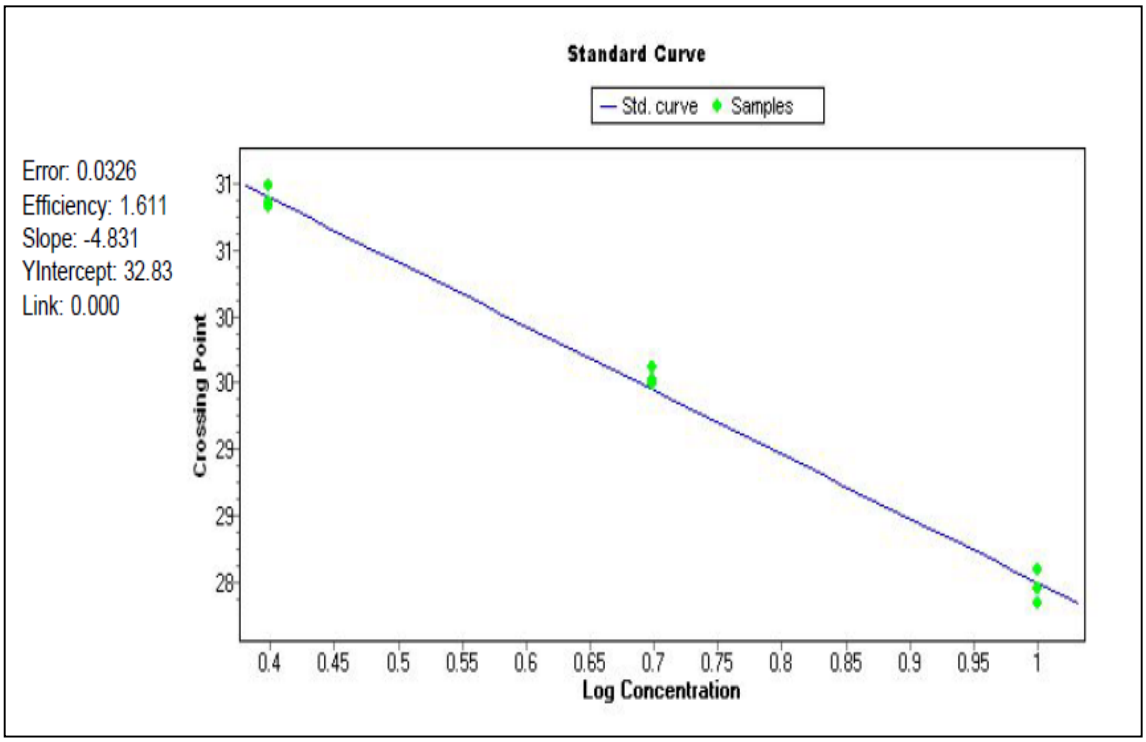


Figure 5.23. Efficiency calculation for *KCNQ2*_Exon4.

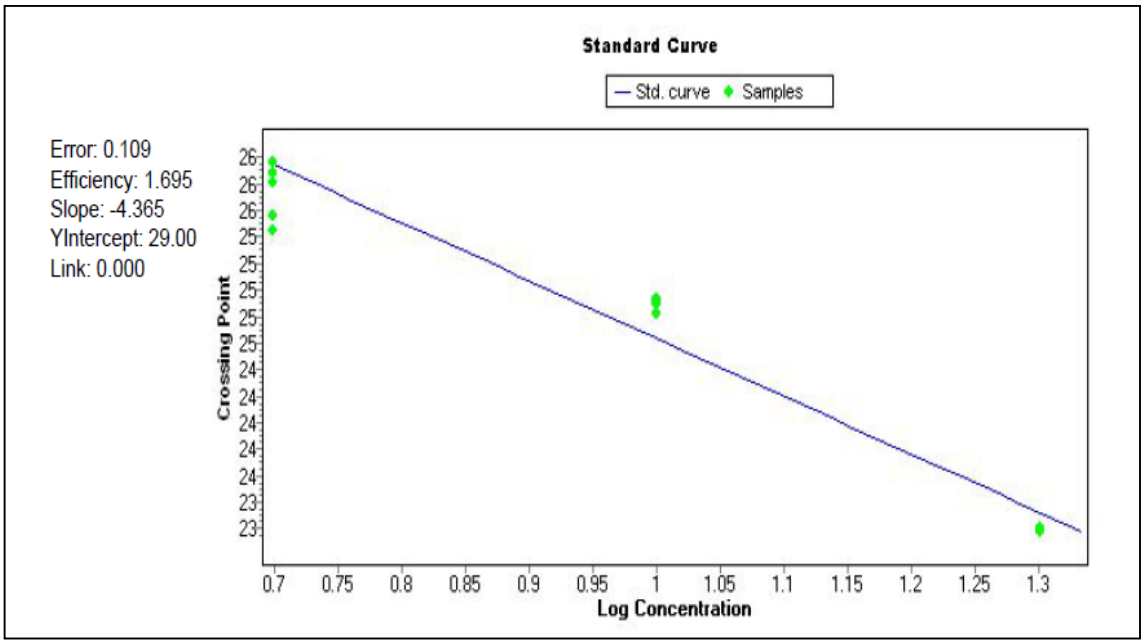


Figure 5.24. Efficiency calculation for *PRRT2*.

On the other hand, array CGH result showing copy number gain for *PRRT2* gene was found to be false positive since all samples were normalized to 1 in the result of relative quantification. Figure 5.26. is a result pane produced by Light Cycler software and represents three separate reactions which produced same results.

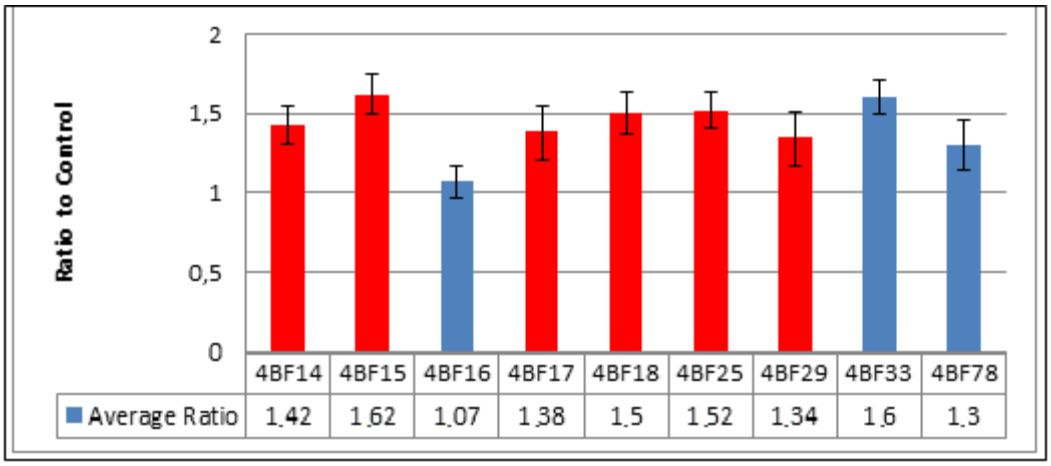


Figure 5.25. *KCNQ2* gene relative quantification results for 4BF family. Each bar represents an individual. Affected members are depicted in red and unaffected members are shown in blue. Average relative ratio is shown in below of each individual ID number.

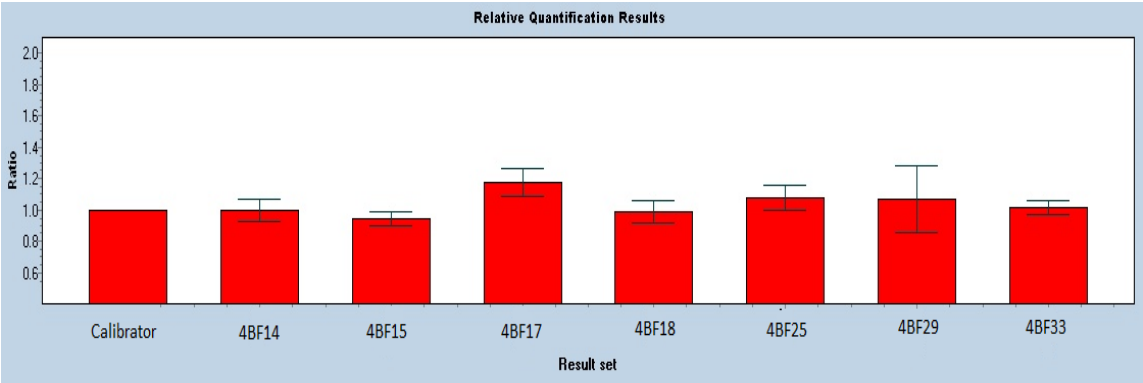


Figure 5.26. *PRRT2* gene relative quantification results for 4BF family. Each bar represents an individual identified by the ID numbers below the bar.

All possible pathologic candidates namely; *PRRT2* on chromosome 16 and *SCN1B* variation on chromosome 19 and *KCNQ2* duplication on chromosome 20 and their segregating genotypes the family are shown in Appendix B..

5.2. Analysis of Other BFIS and BFNS Families

Besides 4BF family, 16 other BFIS families and 5 BFNS families were analyzed. Seven index patients with BFIS phenotype were screened for *SCN1B* gene and 24 affected individuals from 16 families were screened for *PRRT2* gene alterations. Additionally 5 BFNS families, who were screened for sequence variations on *SCN2A* and *KCNQ2* genes and found to be negative, were screened for *PRRT2* gene alteration. Also these patients were analyzed by array CGH and Quantitative Real Time PCR. Phenotypes of all BFIS patients are summarized on Table 5.8. and BFNS patients' on Table 5.10.

Table 5.8. Phenotypes of BFIS families other than 4BF (Dev. Symp.: Developmental symptoms, MR: Mental Retardation).

Patient	Onset	Seizure type	End	Dev. Symp.	Family History
5BF34					
6BF45	45 days	Left tonic unilateral	5Y	Normal	FS in mother and cousin, Epilepsy in Sister
7BF49	3M	Myoclonic, tonic eye deviation loss of consciousness	4M	Normal	Positive

Table 5.8. Phenotypes of BFIS families other than 4BF (Cont.).

Patient	Onset	Seizure type	End	Dev. Symp.	Family History
8BF55	9M	Grand mal tonic, staring, stiffening, eye deviation, loss of consciousness	11M	Normal	Positive
9BF59	6M	Febrile Grand Mal, loss of consciousness	3.5Y	Moderate MR, profound language delay	Consanguineous parents, brother epilepsy
10BF64	2M			Normal	Epilepsy in mother& grandmother
11BF71	7M	Hypo motor seizure, staring, loss of consciousness		Normal	Epilepsy in mother
12BF74	3Y	Head drop, staring, seizure cluster, feeling dizzy, faint		Normal	Same history at Father and father's cousins
14BF88	6M	Petit Mal, Grand Mal, loss of Cons.	11Y	Normal	same phenotype in sister and brother
15BF93	9M	Tonic, temporal lobe, loss of Con.		Normal	mother, maternal aunt, and brother same epilepsy history
16BF104	7M	Tonic, clonic, drop attack, mostly focal seizures			sister same

5.2.1. Linkage and Mutational Analysis of 5BF Family

5BF is another family with BFIS phenotype. It consists of 11 individuals in three generations, 7 healthy and 4 affected individuals. SNP genotyping of the family members were done with Illumina 6K array at Cologne Center of Genomics. Simulation analysis with Autosomal Dominant disease phenotype with 99% penetrance produced 2.09 maximum LOD score for the family. Whole genome linkage analysis revealed 4 loci with the same disease model reaching the defined LOD score. They were;

- Chr 6, LOD=2.09 (rs7617113-rs6929015, 11 SNPs, GRCh37-Pos: 100609338-104108684=3.5MBp, 7Genes in the region; 2 with Neuronal Phenotype)
- Chr 10, LOD=2.09 (rs111875-rs1891760, 29 SNPs, GRCh37-Pos:94462882-115259106=20.8MBp, 242 Genes in the region; 26 with Neuronal Phenotype)
- Chr 13, LOD=2.09 (rs9523070-rs1626391, 16 SNPs, GRCh37-Pos: 91419809-99331568=7.9MBp, 45 Genes in the region; 3 with Neuronal Phenotype)

- Chr 16, LOD=2.09 (rs1424110-rs424074, 2 SNPs, GRCh37-Pos: 79167945-79234587=66.6KBp, 0Genes in the region)

Same analyses were also done with Autosomal Recessive disease model with 99% penetrance and simulation analysis produced 1.79 maximum LOD score. There were two loci in the result of linkage analysis reaching to this LOD score. They were;

- Chr 1, LOD=1.79 (rs3102460-rs2878079, 13 SNPs, GRCh37-Pos: 244576717-246390090=1.8MBp, 45Genes in the region; 1 with Neuronal Phenotype)
- Chr 22, LOD=1.79 (rs374225-rs5762174, 22 SNPs, GRCh37-Pos:20029878-27865344=7.8MBp, 291Genes in the region; 11 with Neuronal Phenotype)

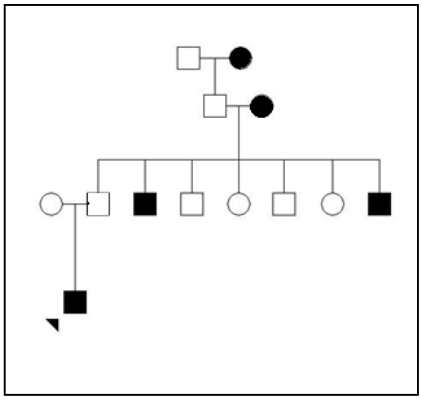


Figure 5.27. Pedigree for 5BF family. Affected individuals are shown in black and index patient is indicated with an arrow.

The family was also analyzed with homozygosity mapper, a tool to determine shared homozygous genotypes of affected individuals. It resulted two large homozygous region on

- Chr. 2,(rs6430398-rs20349, GRCh37-Pos: 133921456-134576249=654KBp)
- Chr. 3,(rs1996562-rs477078,, GRCh37-Pos:40135746-42370594 =2.2MBp). Linkage results were not further analyzed in this family.
-

PRRT2 gene although is on chromosome 16 was outside the linkage region found to chromosome 16 in the family. The alteration was a 2 nucleotide deletion (c.215-216 delCA, T72FsX133) and was inherited together with disease phenotype with 2 exceptions (Figures 5.30 and 5.31). The *PRRT2* mutation was concluded to be the pathological variation in the family since there were no linkage score above 2.09.

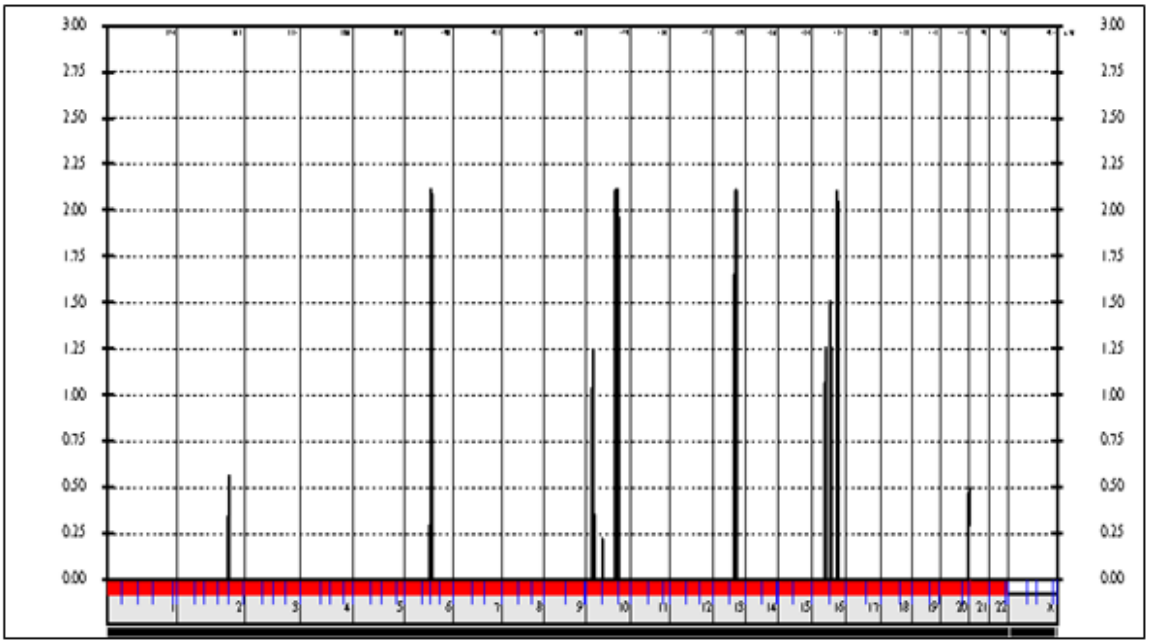


Figure 5.28. Parametric multipoint whole genome linkage results for 5 BF family. Red line shows marker density and blue lines show marker analysis groups. Vertical black lines show LOD score which is scaled in y- axis.

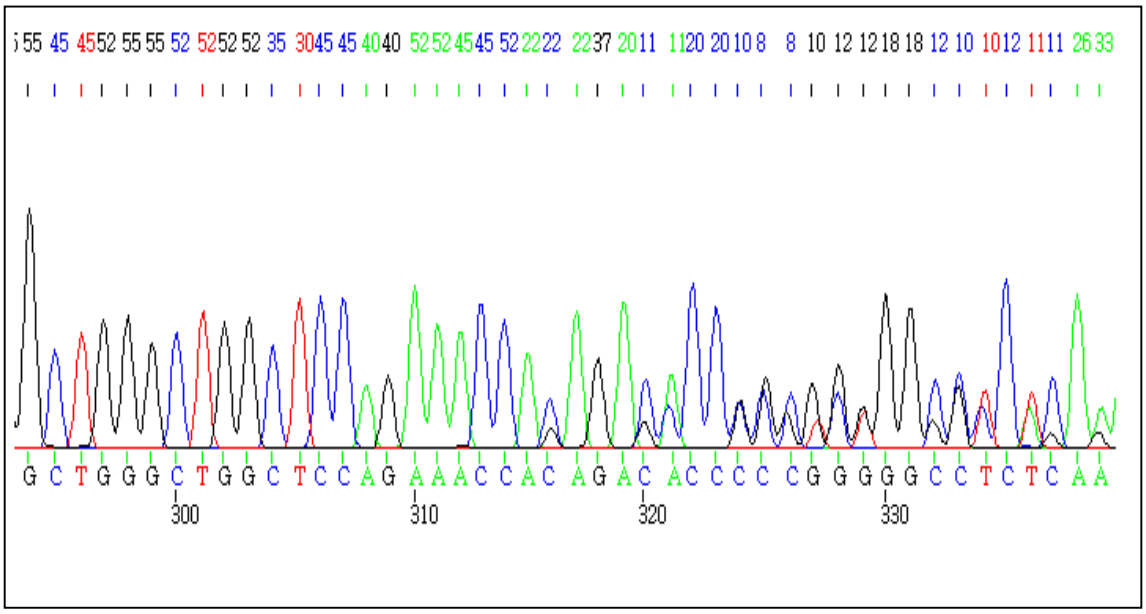


Figure 5.29. Sequencing result for *PRRT2* c.215-216 delCA, T72FsX133 alteration which is observed in 5BF family.

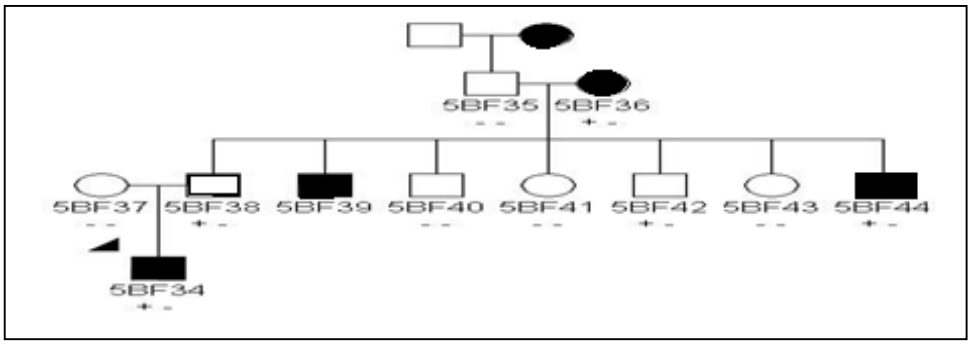


Figure 5.30. Inheritance of for *PRRT2* c.215-216 delCA alteration in 5BF family. (+ represents the altered allele, - represents the normal allele).

5.2.2. *SCN1B* Screen in 6 BFIS Index Patients

As described above coding region and exon/intron boundary of the *SCN1B* gene was amplified by PCR and analyzed by Sanger sequencing for 6 BFIS index patients (5BF 34, 6BF45, 7BF 49, 8BF 55, 9BF 59, 10 BF 64, 11 BF 71). No novel alteration was identified in these patients' genotypes. However 6 known SNPs with low minor allele frequencies were detected. Two of these variations were on splice site, one variant were synonymous and 3 varianuts were nonsynonymous. SIFT score for nonsynonymous variants were damaging (Table 5.9).

Table 5.9. *SCN1B* gene SNP genotyping results for selected BFIS index patients. Minor allele genotypes of each individual is shown in bald.

Ex	SNP ID	Variant	MAF	SIFT	5BF34	6BF45	7BF49	8BF55	9BF59	11BF71
1	rs72556351	c.40+15G>T	T= 0.1482 /742	NA	**G/T**	**G/T**	G/G	G/G	G/G	G/G
3B	rs55742440	*c.629T>C p.Leu210Pro	C= 0.3776 /1891	D	**C/C**	**T/C**	T/T	T/T	**T/C**	**T/C**
3B	rs67701503	*c.744C>A p.Ser248Arg	A= 0.1274 /638	D	**C/A**	C/C	C/C	C/C	**C/A**	C/C
3B	rs67486287	*c.749G>C p.Arg250Thr	C= 0.1192 /597	D	**G/C**	G/G	G/G	G/G	**G/C**	G/G
4	rs16969930	c.501T>C p.Ile167Ile	C= 0.0156 /78	N	T/T	T/T	T/T	**T/C**	T/T	T/T
5	rs28365109	c.591-14C>A	A= 0.0080 /40	N	C/C	C/C	C/C	**C/A**	C/C	C/C

5.2.3. Analysis of BFNS Patients (BF19-BF23)

A common phenotype was observed in 5 BFNS patients with seizure onset between 0-70 days and occurring many times/day. Breast feeding initiates/activates seizures and the baby seems to be suffocating and the chin gets stiff; Ictal and interictal EEG show temporal lobe origin; development normal; drug response and prognosis is very good. Table 5.10 shows additional features and exact age of onset and age at diagnosis.

Table 5.10. Phenotypes observed in BFNS families.

BFNS patient ID	**Gender**	**Relatives**	**Additional features**	**Age of onset**	**Age at Diagnosis**
19 BF 135	M	Parents unaffected		6 days	31 days
20BF 139	M			70 days	80 days
21BF 142	F	Parents unaffected		10 days	2 mo
22BF 151	M	Parents unaffected	Apneic seizures	7 days	25 days
23 BF 155	M	FS in sister		35 days	5 weeks

These 5 patients were screened for *SCN2A*, *KCNQ2* and *PRRT2* gene alterations and found to be negative. In order to evaluate gene copy number in these patients, index samples were also analyzed with array CGH. In the result of analysis, amplification was observed in the *KCNQ2* gene in 4 samples. However, DRLRS (Derivative LogR Spread) value was above 0.3 for three samples, which made the array result unreliable (Figure 5.31). However these samples were also selected for qPCR validation. qRT-PCR was conducted as explained above. In the result of analysis, there were evident gene duplications in *KCNQ2* gene which was inherited from both healthy parents in three patients. Interestingly, these three BFNS patients had the same age of disease onset (Table 5.11 and Figure 5.32). Array CGH result showed duplication of *PRRT2* gene also but in the result of qRT-PCR analysis it was seen as false positive since all samples normalized approximately to 1 in the end of the analysis (Figure 5.33)

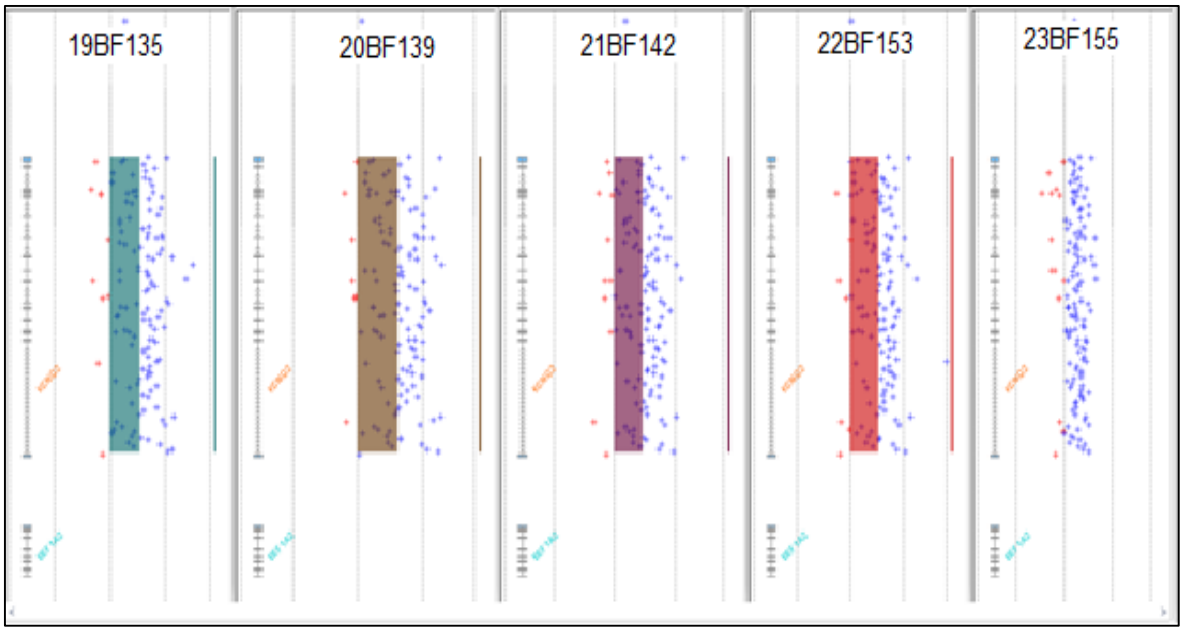


Figure 5.31. array CGH results for BFNS patients for *KCNQ2* gene. Each box represents a patient and patient ID is written on top of each box.

Table 5.11. Relative quantification results for *KCNQ2* gene in BFNS patients.

Individual ID	**Average fold increase**
19BF135(index)	1.73
19BF136(father)	2.3
19BF137(mother)	1.8
20BF139(index)	1.15
21BF142(index)	1.9
21BF143(mother)	1.9
21BF144(father)	2.1
22BF151(index)	1.8
22BF153(mother)	1.8
22BF154(father)	1.8
23BF155(index)	1.18
23BF156(sister)	1.1

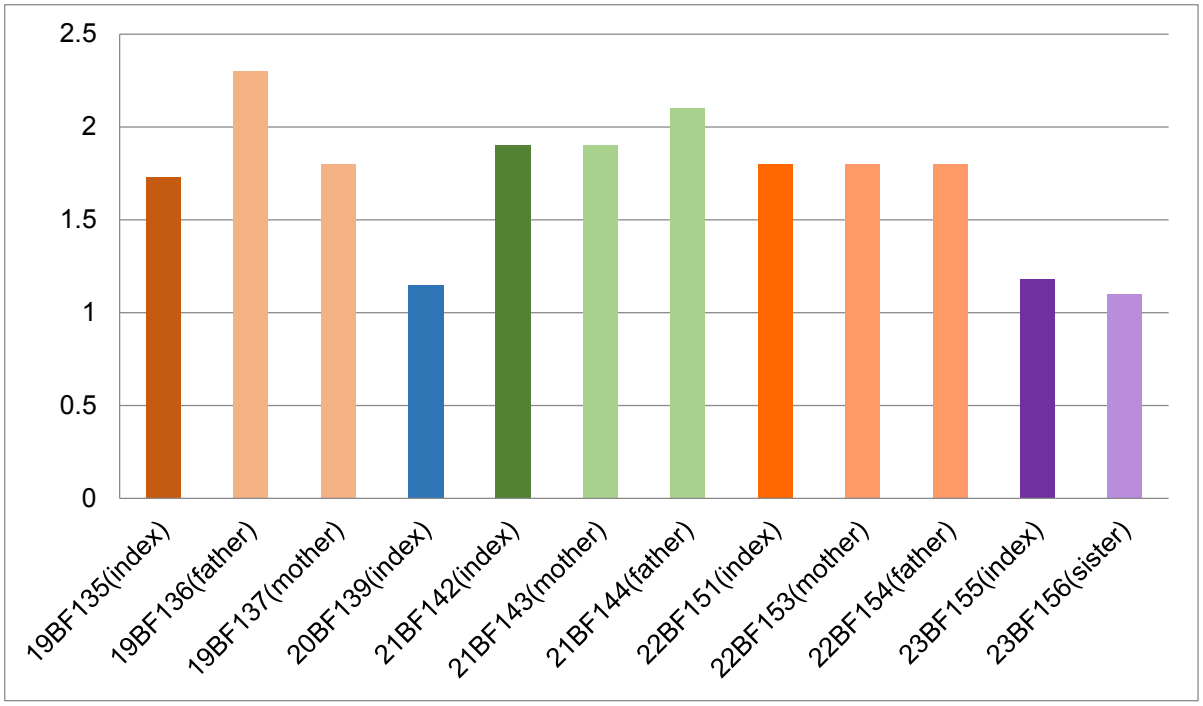


Figure 5.32. Relative quantification of *KCNQ2* gene for BFNS patients and family members. Each bar represents an individual, each family is depicted in different color where index patients are darker than other family members.

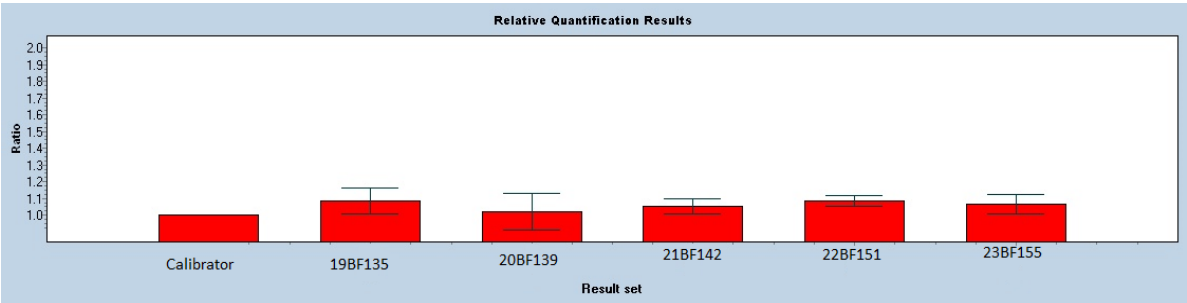


Figure 5.33. *PRRT2* gene relative quantification results for BFNS patients. Each bar represents an individual identified by the ID numbers below the bar.

5.3. Genetic Epilepsy with Febrile Seizures plus (GEFS+) Study

GEFS+ index patient belongs to very large kindred, about 1700 individuals from a village with 3000 inhabitants. The family relationships are very complex thus only nuclear family was investigated initially. The index patient belongs to a large multiplex and multigenerational family with many consanguineous marriages. Additionally sampled individuals live in an isolated populations due to cultural constrains.

5.3.1. Target Gene Analysis

Mutational analysis was carried on GEFS+ index patient for *GABRG2* (Özlem Yalçın Çapan), *SCN1A* (Seda Salar), *SCN1B* and *SCN2A* (M.Aslı Kayserili) genes, by PCR amplification of coding sequence and exon-intron boundaries followed by direct DNA sequencing. These genes were all negative for point mutations.

5.3.2. Whole Exome Sequencing

Whole genome linkage analysis was proposed for this family in the beginning of the study. However, when the whole pedigree was examined it was seen that there were various phenotypes among affected individuals. The family was too large for any program to analyze. On the other hand affected individual number in sub-pedigrees was not enough to capture different phenotypes by whole genome linkage analysis.

When these facts are taken into account, a relatively more direct approach was thought to be more feasible. Index patient and one of his children are sent for whole exome sequencing. Numbers of total variants and variants shared by two patients were shown in Table 5.12.

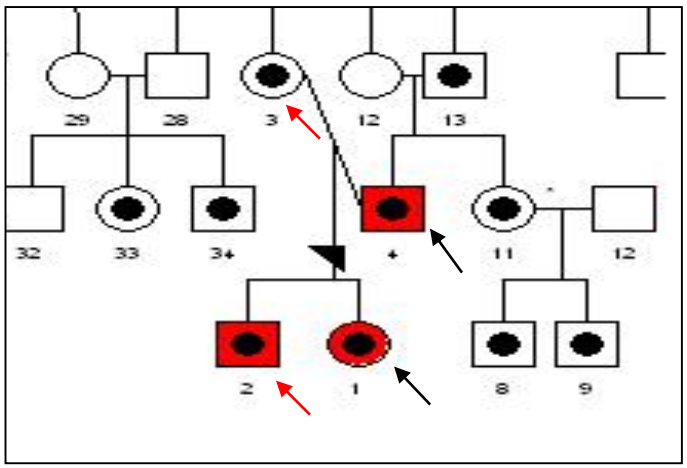


Figure 5.34. Index patients selected for Initial Whole Exome Sequencing (shown by black arrow) and affected and unaffected family members selected for second Whole Exome Sequencing (shown by red arrow).

Table 5.12. Summary of WES variant analysis of individuals GEFS1 and GEFS4.

SNP		
Total	2253	1911 with high confidence
		341 with low confidence
Common to both patients	795 with high confidence (among these 112 epilepsy associated variants)	642 missense
		16 nonsense
		2 read-through
		135 splice site
INDEL		
Total	508	
Common to both patients	167 (among these 21 epilepsy associated variants epilepsy)	29 3'UTR
		67 5'UTR
		24 cds-indel
		20 frameshift
		36 splice site

Additionally two other individuals from core family (individuals 2 and 3, shown by red arrow in Figure 5.34) were selected for another WES analysis in order to reduce the amount of insignificant variants. Variants obtained from WES were analyzed in order to find plausible candidates to further investigate by Sanger sequencing. Selected variants were shown in Table 5.15.

5.3.2.1. Analysis of Whole Exome Variants. WES analysis statistics for the 4 patients are given in Table 5.13.

PCR primers for a set of variants selected according to their expression profile, location, function and verification proposal by the service supplier (BGI) were designed. These exome variants were analyzed in a selected portion of the pedigree. In the result of this analysis none of the variants showed significant co-segregation in the family and two point LOD score. *ATP1A2* and *BDKRB1* variants were observed only in two samples and *GRM3* and *KCNV1* variants were observed only in a small family. *KCNJ16* and *NRXN2* variants were observed in other healthy family members (Table 5.16).

In the result of WES variant analysis, selected variants did not segregate together with disease phenotype. Other variants had lower probability as causative variation since they were not shared between patients or they were not expressed in nervous system. so in orser to pinpoint the locaton of disease causing variant research continued with linkage analysis.

Table 5.13. WES SNP data analysis statistics.

Sample	**GEFS1**	**GEFS2**	**GEFS3**	**GEFS4**
Total	91015	106346	111014	93493
1000genome and dbsnp135		102930	107464	
1000genome specific		70	74	
dbSNP135 specific		1674	1670	
dbSNP rate		98.36%	98.31%	
Novel		1672	1806	
Homozygous	32644	47573	47855	36240
Heterozygous	58371	58773	63159	57253
Synonymous	10246	10159	10311	10095
Missense	12495	8766	8809	12443
Stop gain	159	64	60	163
Stop loss	92	29	28	90
Exonic	22992	18754	18960	22791
Exonic and splicing	25743	264	248	25474
Splicing	2751	87	90	2683
NcRNA		3025	3220	
UTR5	6433	2101	2205	6604
UTR5 and UTR3		8	6	
UTR3	6983	16433	16787	7316
Intronic	49373	55712	59032	51559
Upstream		1903	2043	
Upstream and downstream		130	143	
Downstream		2746	2918	
Intergenic	2483	5183	5362	2540

Table 5.14. WES indel data analysis statistics.

Sample	GEFS1	GEFS2	GEFS3	GEFS4
Total	6966	17192	18202	7278
1000genome and dbsnp135		6287	6598	
1000genome specific		1167	1229	
dbSNP135 specific		5491	5672	
dbSNP rate		68.51%	67.41%	
Novel		4247	4703	
Homozygous	2778	7823	7903	3062
Heterozygous	4188	9369	10299	4216
Frameshift	240	228	227	239
Non-frameshift Indel	206	268	271	201
Frameshift block substitution		0	0	
Non-frameshift block substitution		0	0	
Stop gain		3	5	

Table 5.14. WES indel data analysis statistics (Cont.).

Sample	GEFS1	GEFS2	GEFS3	GEFS4
Stop loss		1	1	
Exonic		493	497	
Exonic and splicing	406	7	7	411
Splicing		69	67	
NcRNA		556	575	
UTR5	462	303	324	502
UTR5 and UTR3		2	2	
UTR3	541	3995	4201	581
Intronic	4961	9729	10411	5188
Upstream	37	326	352	35
Upstream and downstream		24	27	
Downstream		643	676	
Intergenic	113	1045	1063	121

Table 5.15. WES variants selected for validation by Sanger sequencing (Variants in bald have primers designed.).

Chromosome	**Gene Name**	**Substitution**	**Share Number**
chr1	***ATP1A2***	**I260R**	**1**
chr14	***BDKRB1***	**G241R**	**2**
chr11	***NRXN2***	**insC;Het,frameshift**	**1**
chr4	*CLCN3*	A707V	1
chrX	*CLCN4*	A556T	1
chr10	*GRID1*	G695D	1
chr10	*KCNK18*	P175R	1
chr9	*KCNV2*	G117S	1
chr2	*SCN7A*	Splice Site	1
chr4	*DRD5*	E327K	1
chr1	*HTR1D*	Splice Site	1
chr3	*KCNMB3*	Q87R	1
chr17	*MINK1*	E837D	1
chr3	*SCN10A*	S1337T	1
chr7	***GRM3***	**D380Y**	**2**
chr17	***KCNJ16***	**R309M**	**2**
chr8	***KCNV1***	**T388M**	**2**
chrX	*PRICKLE3*	S550C	1
chrX	*PRICKLE3*	P504R	1
chr9	*SLC25A25*	R147Q	1
chr9	*CNTNAP3B*	A435T	2

Table 5.16. Results for variant analysis in selected members of GEFS+ family. Affected individuals are shown in bald ("+" represents mutated, "-" represents healthy allele).

Patient ID	***ATP1A2***	***BDKRB1***	***GRM3***	***KCNJ16***	***KCNV1***	***NRXN2***
1	-/-	+/-	+/-	+/+	+/-	+/-
2	-/-	-/-	+/-	-/-	+/-	-/-
3	-/-	-/-		+/-	-/-	+/-
4	+/-	+/-	+/-	+/-	+/+	-/-
12	+/-	-/-	+/-	-/-	+/-	-/-
13	-/-	-/-	+/-	+/-	+/-	+/-
57	-/-	-/-	+/-	+/-	-/-	+/-
63	-/-	-/-	+/-	+/-		+/-
77	-/-	-/-	-/-	-/-	-/-	+/+
78	-/-	-/-	-/-	-/-		-/-
79	-/-	-/-	-/-	-/-	-/-	+/-
92	-/-	-/-		-/-	-/-	-/-
93	-/-	-/-	-/-	-/-	-/-	+/-
97	-/-	-/-	-/-	-/-	-/-	+/-
690	-/-	-/-	-/-	+/-	-/-	-/-
31						-/-
32						+/-
33						+/-
107						-/-
138						+/-
151						-/-
325						+/-
373						-/-

5.3.3. SNP Genotyping and Whole Genome Linkage Analysis

In order to pinpoint genomic region segregating together with the disease phenotype whole genome linkage analysis was conducted in a selected sub-pedigree showing IGE phenotype. For this purpose selected members are genotyped with a 300K array. The family members selected for linkage analysis are shown in red rectangles in the Figure 5.35.

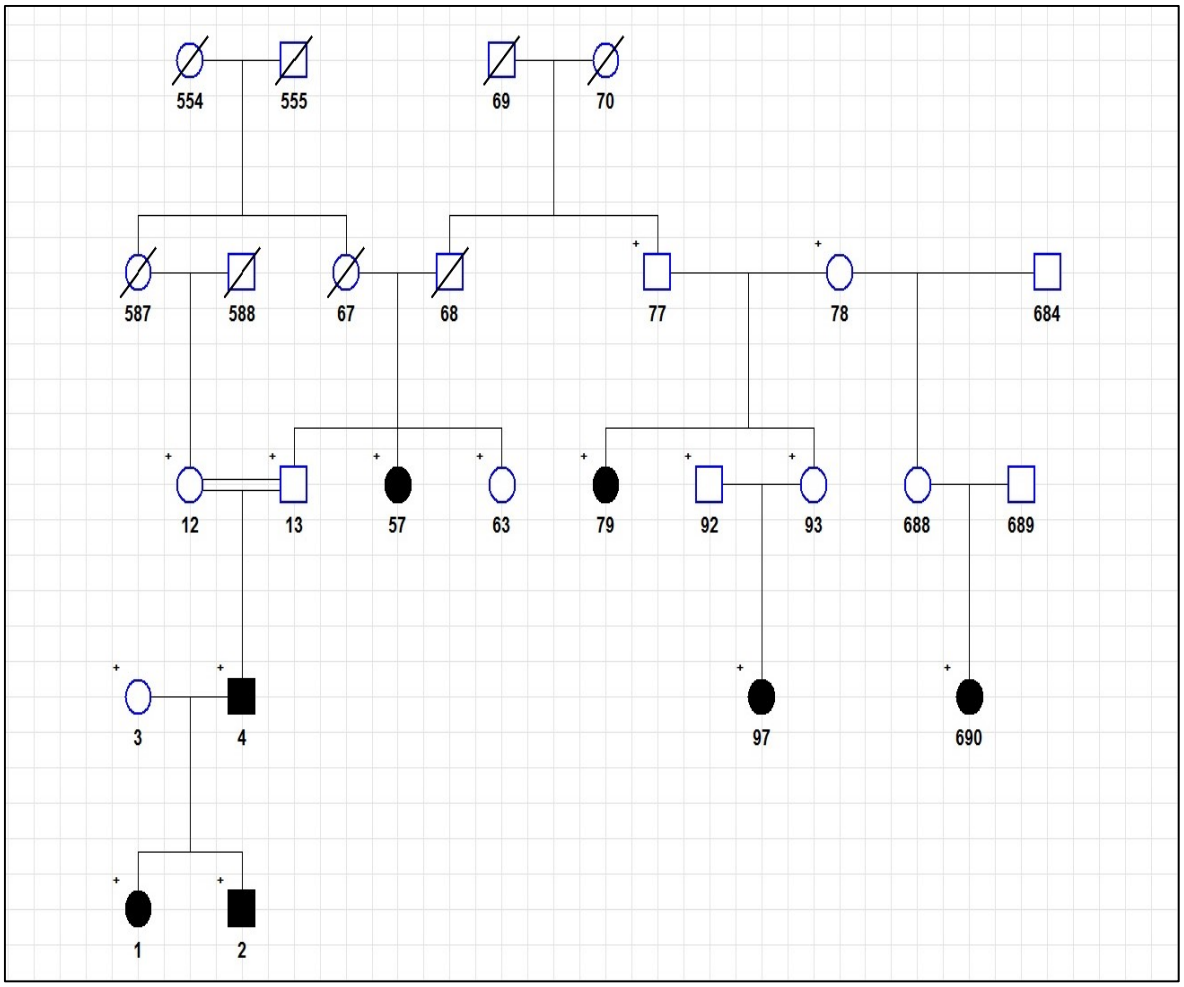


Figure 5.35. Selected sub pedigree for initial linkage analysis. Affected members are shown in black and have IGE phenotype. Individuals with asterix are genotyped for linkage analysis.

Whole genome SNP genotyping with Illumina CytoSNP 300K was followed by multipoint linkage analysis with different models. Namely, Autosomal Dominant with 70% penetrance and Autosomal Recessive with 100% and 80% penetrance.

In the result of analyses, maximum LOD score of 1.89 was obtained by AR inheritance with 100% penetrance for a region on Chromosome 7 (Figure 5.36). Decreasing penetrance did not increase the LOD score (Figure 5.37). Autosomal Dominant inheritance model did not produce significant linkage on any genomic region, neither (Figure 5.38).

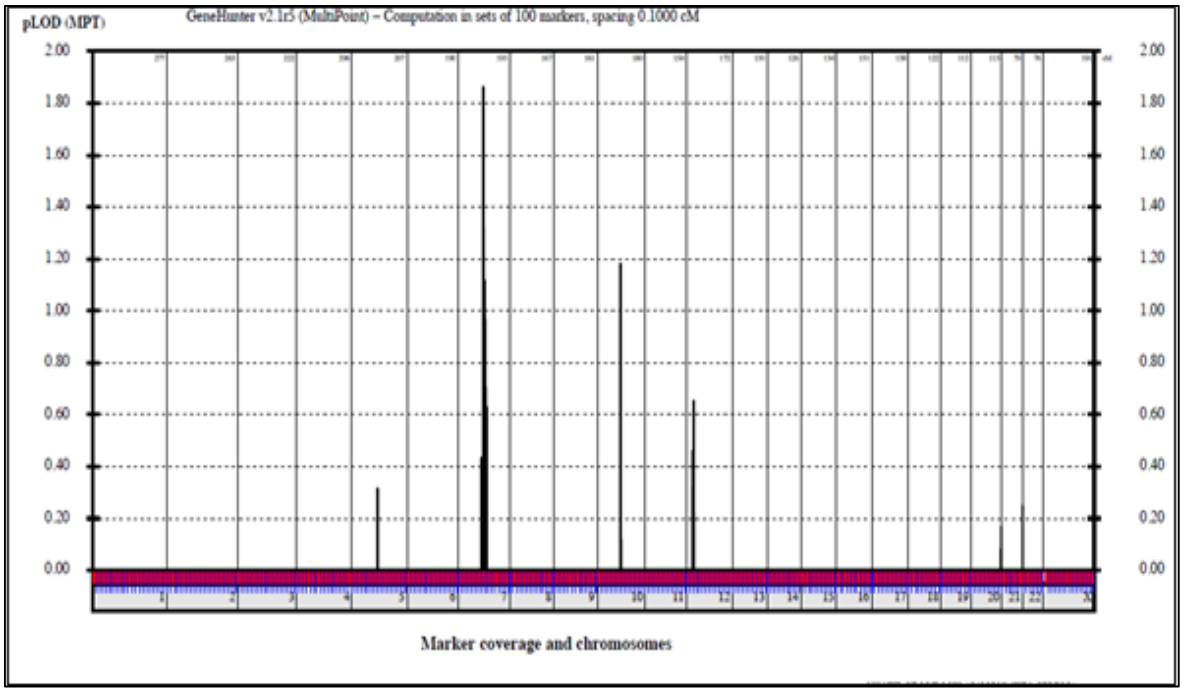


Figure 5.36. Result whole genome linkage analysis with AR inheritance and 100% penetrance.

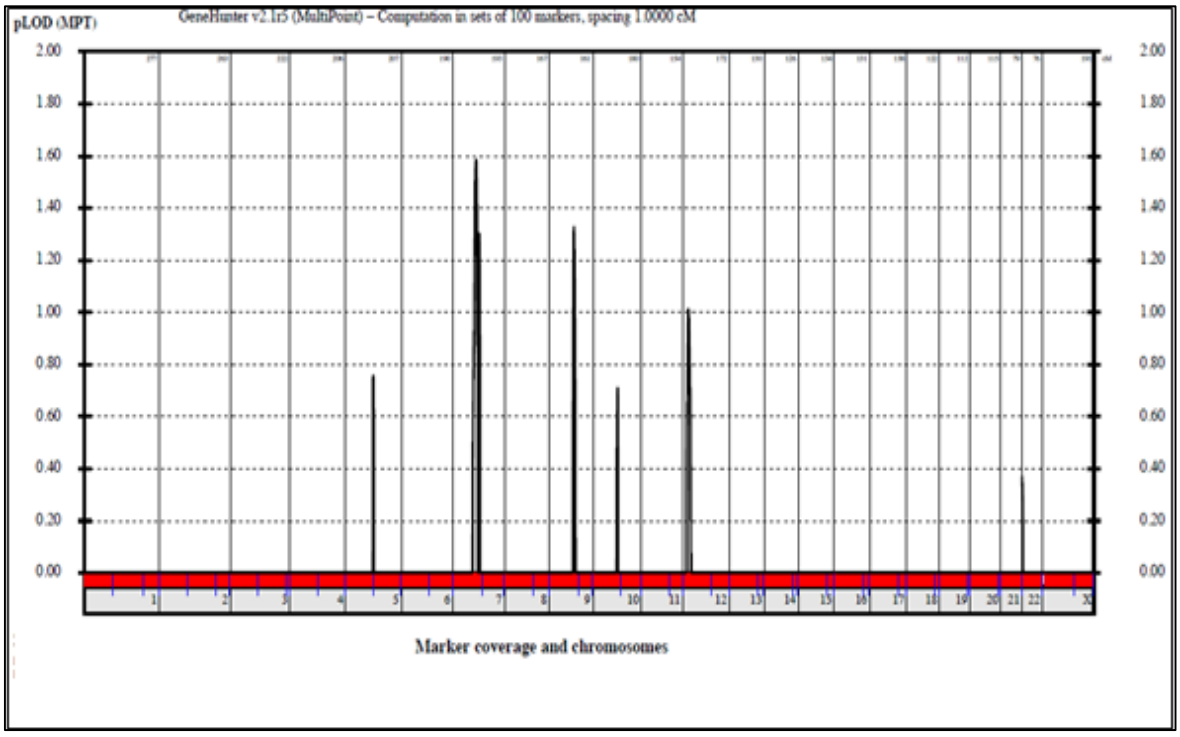


Figure 5.37. Result of whole genome linkage analysis with AR inheritance and 70% penetrance.

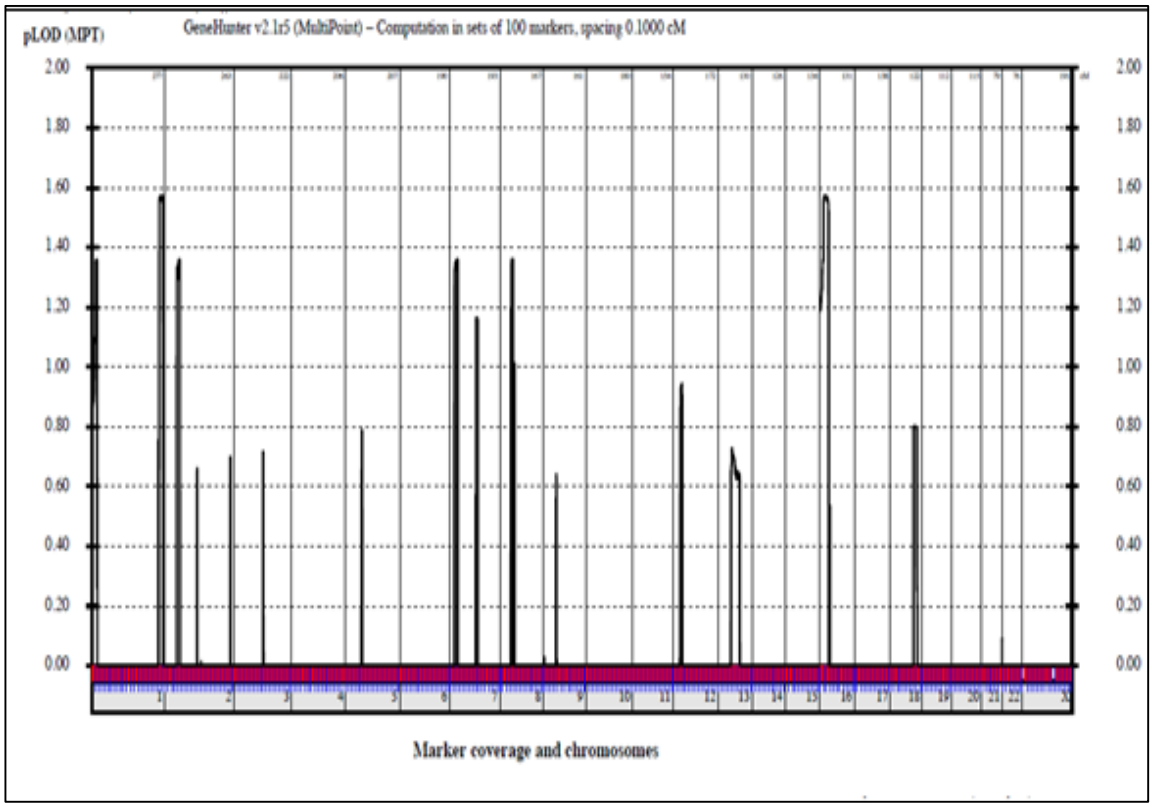


Figure 5.38. Result of whole genome linkage analysis with AD inheritance and 70% penetrance.

In summary highest LOD score was obtained with AR disease model on chromosome 7. The linkage region was analyzed by Genedistiller program (www.genedistiller.org). Out of 22 genes in the region only the *GRM3* gene had a variation as was detected in WES analysis. Due to the lack of a significant Lod score 14 other samples were analyzed with Illumina CytoSNP 300K array. Total number of genotyped individual increased to 29 (Figure 5.39), which was a large number for any linkage program to analyze with bit size of 40. For this reason, genotyped individuals were grouped into 4 smaller pedigrees and total LOD score was investigated (Figure 5.40).

Before Linkage analysis, in order to check sample relations Alohomora program was used. In the result of Alohomora analyses all family relationships were confirmed, since parent-offspring and sib-sib pairs are clustered together (Rüschendorf and Nürnberg, 2005) (Figure 5.41).

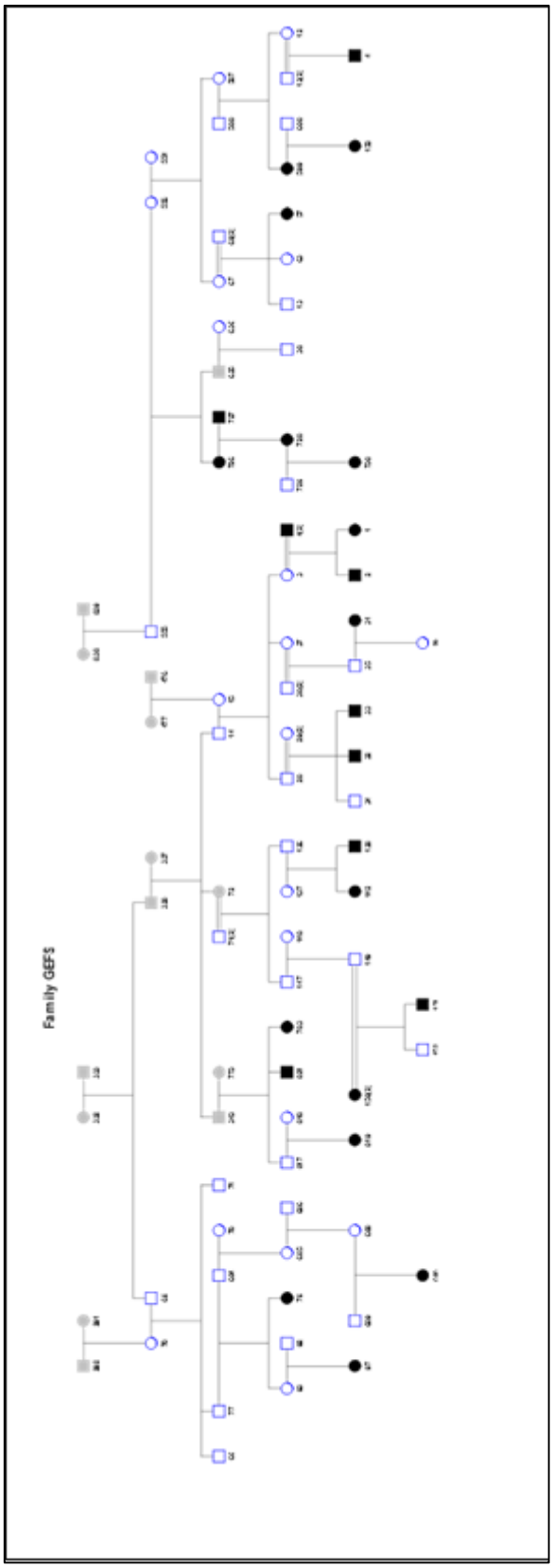


Figure 5.39. Enlarged sub-pedigree for second round of linkage analysis.

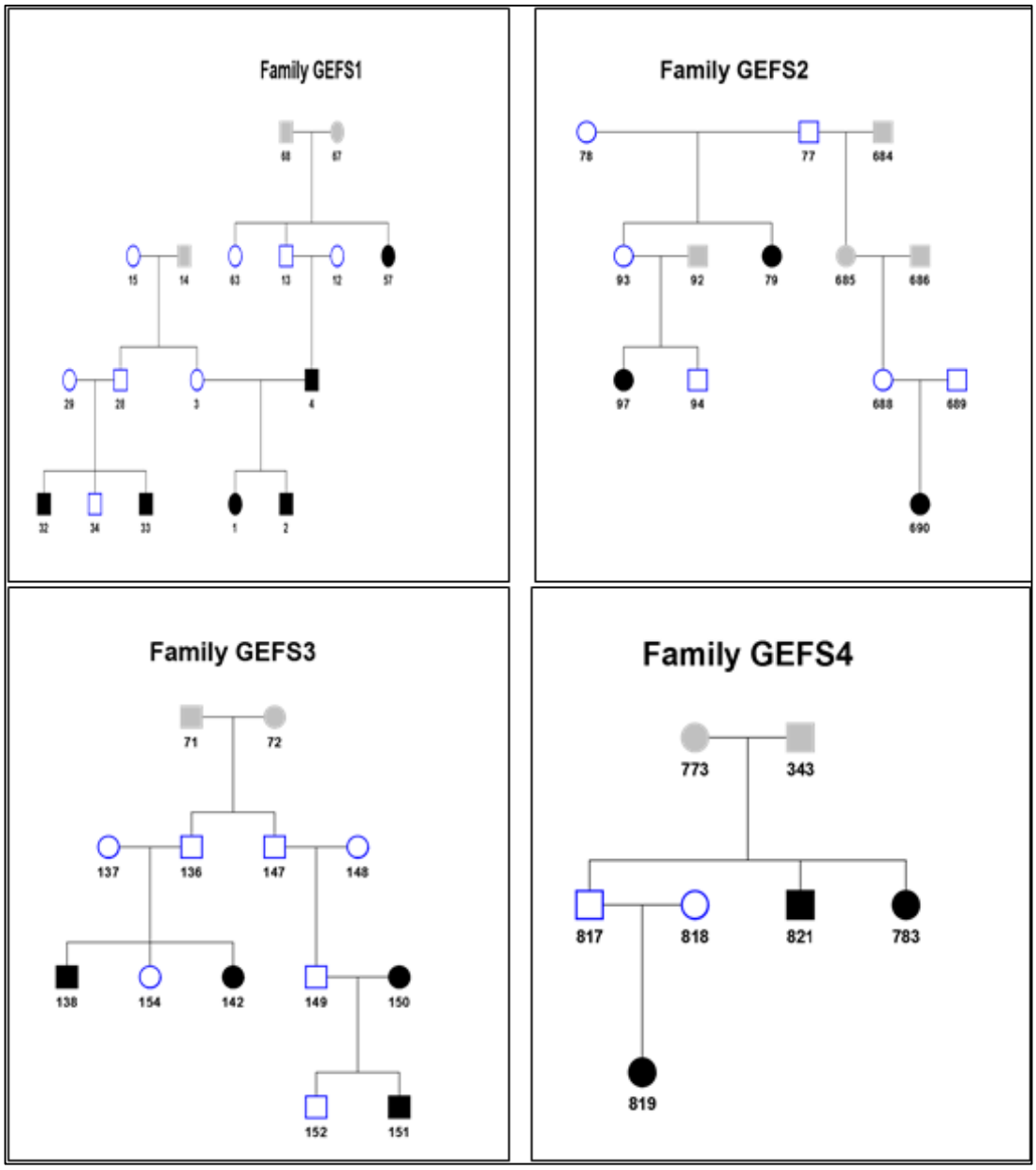


Figure 5.40. Small pedigrees of mixed phenotype in which GEFS+ sub-pedigree divided.

Whole Genome Multipoint Parametric Linkage Analysis was conducted with different inheritance models both with autosomal recessive and autosomal dominant and also with different penetrance values from 100% to 70 %. Whole genome was analyzed in 1 marker in 1cM resolution in sets of 100 markers. Then on chromosomes where there were LOD score larger than 2, linkage analysis was repeated in 1 marker in 0.1 cM resolution in sets of 100 and 75 markers to define borders of linkage region. Flowingly, every marker on linkage region was analyzed in sets of 100 and 75 markers.

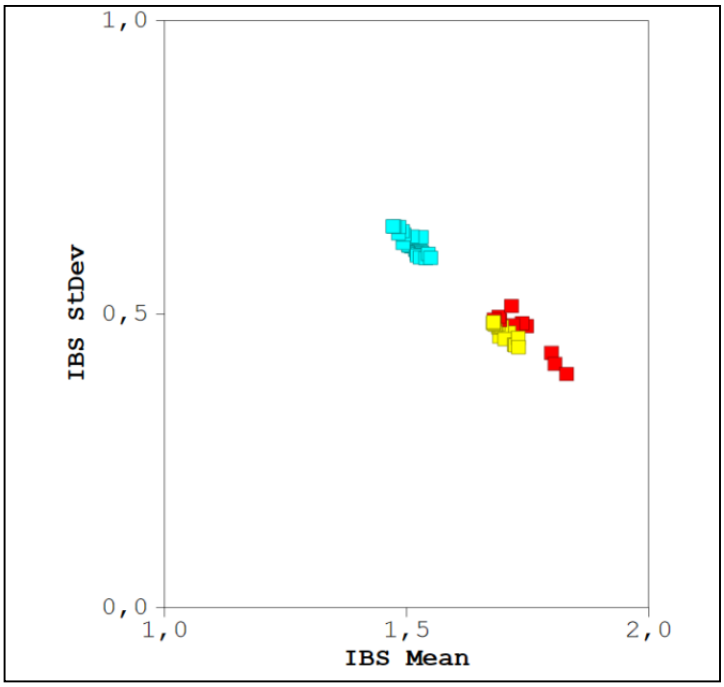


Figure 5.41. Allele sharing analysis results. (Red denotes parent-offspring, yellow denotes sib-sib pairs and blue denotes individuals who are not relatives.

5.3.3.1. Whole Genome Linkage Analysis Assessed by Autosomal Dominant Inheritance Model. In the result of linkage analysis with Autosomal Dominant (AD) inheritance model, in 1 marker in 1 cM resolution and in sets of 100 markers, a significant LOD score above 2 was found on a region on chromosome 1, 2, 3 and 9. Figure 5.42 shows the analysis result. The analysis was repeated for chromosomes, 1, 2, 3, and 9 with 1cM resolution and sets of 75 markers. A significant linkage score approaching 3.5 was observed on chromosome 1 (Figure 5.43). This region was analyzed again including every genotyped marker in sets of 75 markers and 90% penetrance and the LOD score approximated 4. (Figure 5.44)

When markers with LOD score above 2 on chromosome 1 were included; linkage region was between markers rs236285 and rs338466 ranging from 94cM to 109 cM. When analyzed with Genedistiller program (Seelow *et al*., 2008), it was seen that there were 116 genes and 47 of them were protein coding. When markers with LOD score above 3 were chosen, the linkage region was between markers rs7543296 and rs12137571 ranging from 101 cM to 107 cM. When analyzed with Genedistiller program, it was seen that there were 24 genes and 9 of them were protein coding. Genes in the linkage region was analyzed in different aspects which were summarized in Table 5.17.

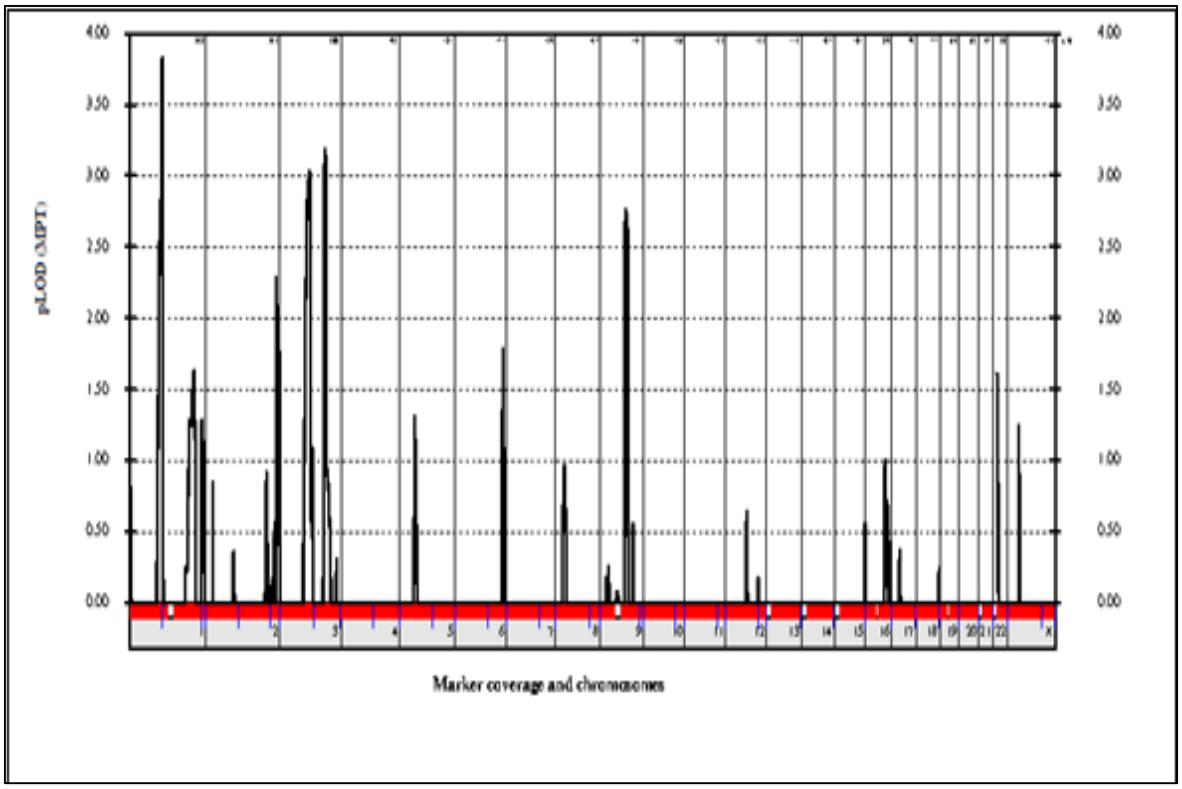


Figure 5.42. Result for parametric multipoint whole genome linkage analysis with AD and 90% penetrance model and in 1 marker/cM resolution and 100 markers per group.

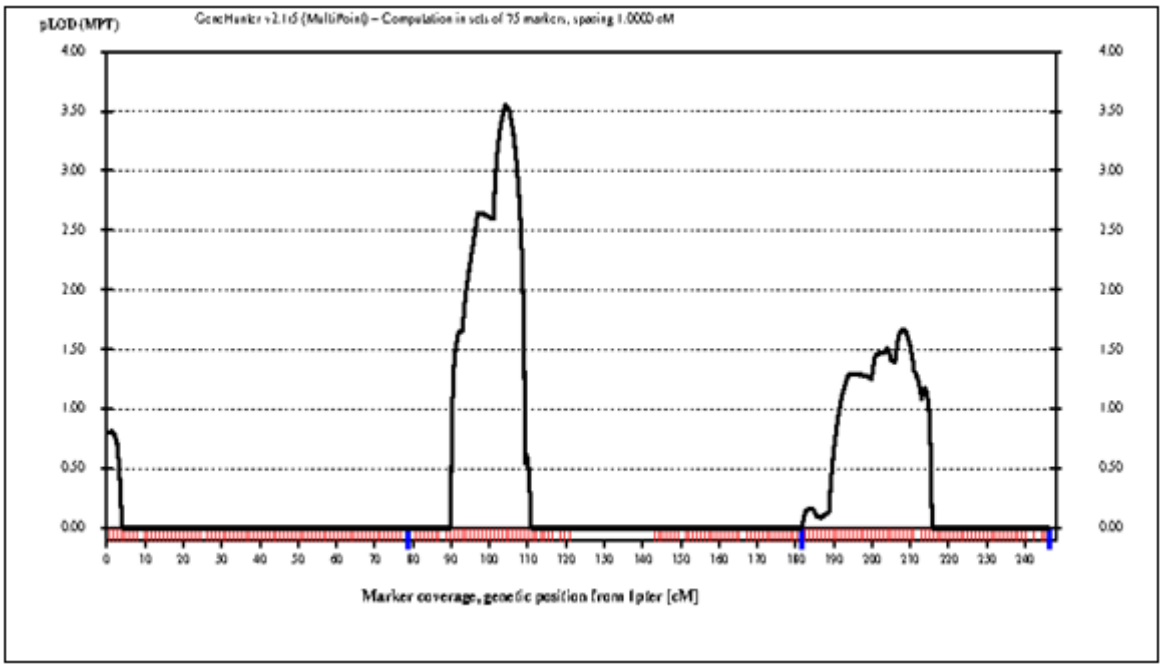


Figure 5.43. Result for parametric multipoint linkage analysis on chromosome 1 with AD and 90% penetrance model and in 1 marker/cM resolution and 75 markers per group.

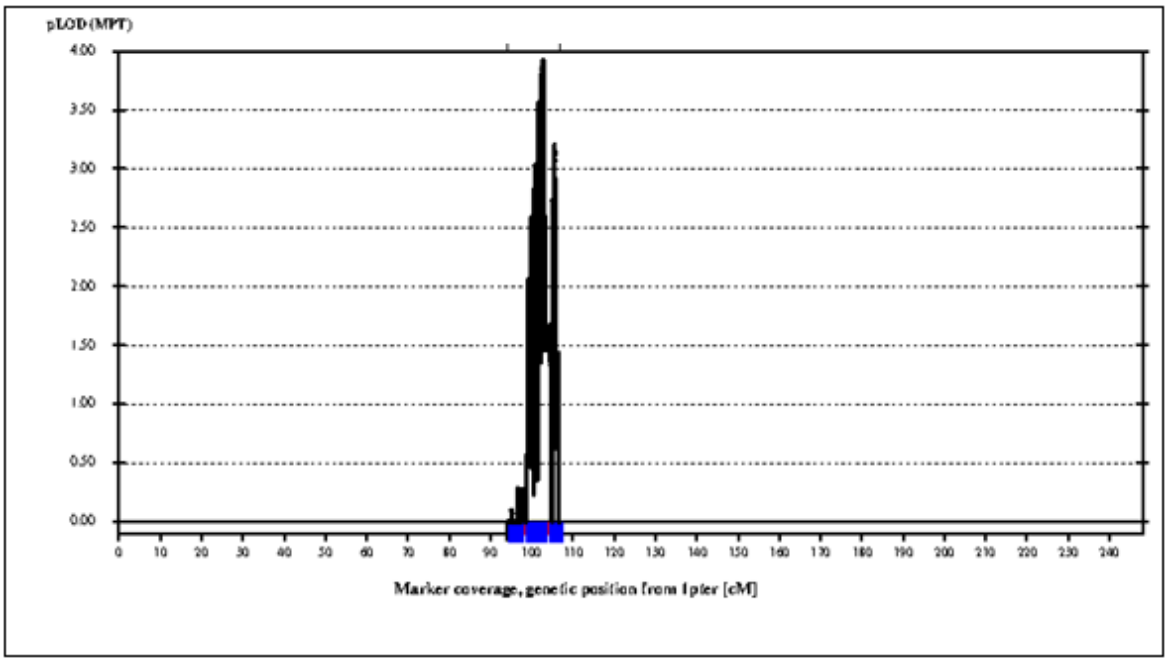


Figure 5.44. Result for parametric multipoint linkage analysis on chromosome 1 with AD and 90% penetrance model and all markers in range and 75 markers per group.

Table 5.17. Gene content of the linkage region obtained with AD linkage analysis on chromosome 1, analyzed by Genedistiller program.

LOD	**First SNP Last SNP**	**Genetic location (cM)**	**Physical Location (bp)**	**Av. LOD**	**All Genes**	**Protein Coding**	**miRNA**	**Epilepsy Remark in OMIM Report**	**Nervous System Phenotype**
>2	rs236285 rs338466	94cM-109	94132788 - 109486196	2,85	116	47	MiR137	CNN3 DPYD NTNG1	*BCAR3* *GLCM* *ABCA4* *F3* *PTBF2* *LPPR4* *S1PR1* *COL11A* *VAV3*
>3	rs7543296 rs12137571	101-107	101348100 - 107430479	3,35	34	13		No	*S1PR1* *COL11A*

5.3.3.2. Whole Genome Linkage Analysis Assessed by Autosomal Recessive Inheritance Model. In the result of linkage analysis with Autosomal Recessive (AD) inheritance and 70% penetrance model, in 1 marker/cM resolution and in sets of 100 markers, a significant LOD score above 2 was found on a region on chromosome 2 (.Figure 5.45). The analysis was repeated for chromosome 2 with 0.1cM resolution and sets of 100 markers. However the maximum LOD score achieved was only 2.25 (Figure 5.46). But when this region was analyzed again including every genotyped marker in sets of 100 markers and 70% penetrance, for a region from 85cM to 100 cM, LOD score above 3 was obtained for 3 regions that contained a few genes (Figure 5.47). However when whole region with LOD score above 2 was included, the linkage region spanned a 10cM genomic locus with more than 200 genes. Detailed analysis of linkage locus is listed on Table 5.18.

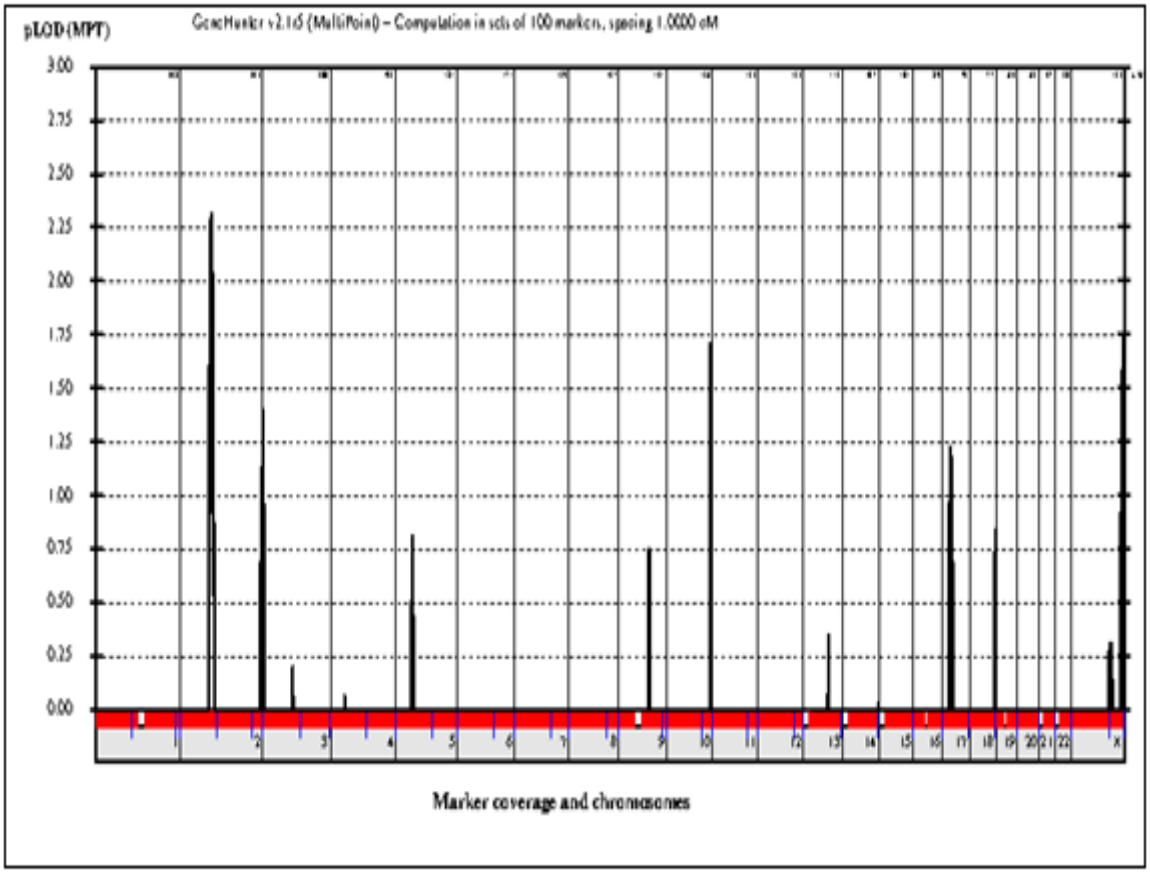


Figure 5.45. Result for parametric multipoint whole genome linkage analysis with AR and 90% penetrance model and in 1 marker/cM resolution and 100 markers per group.

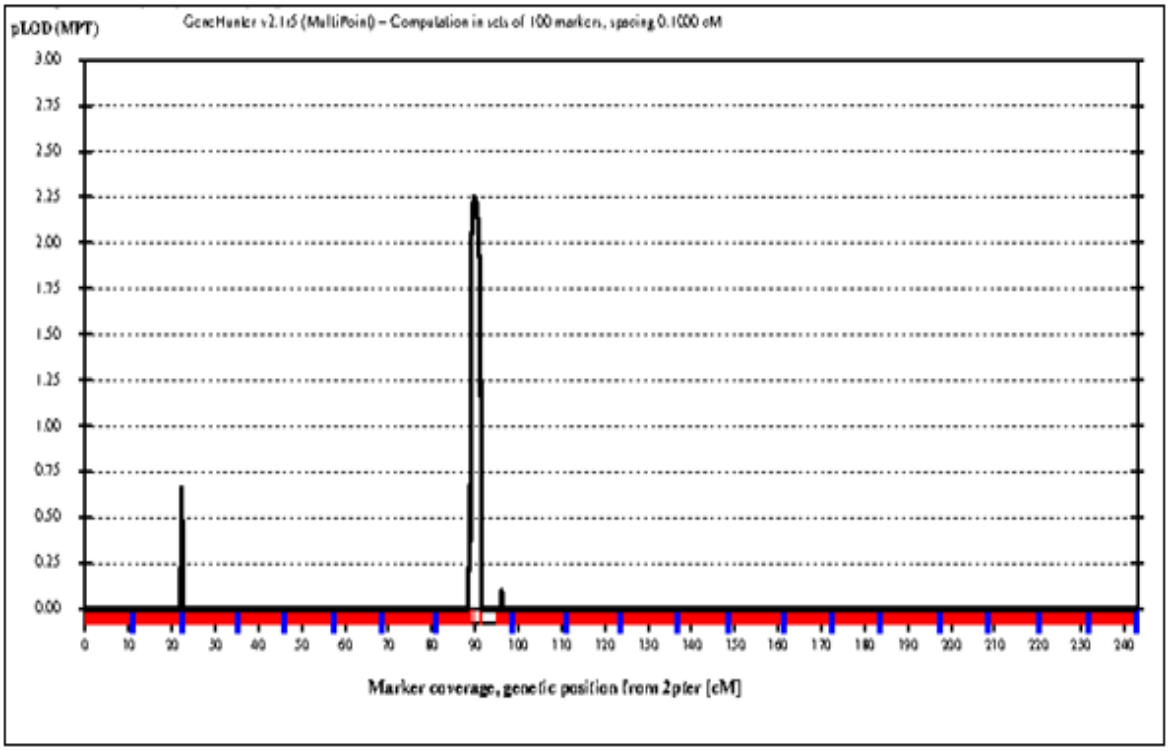


Figure 5.46. Result for parametric multipoint linkage analysis on chromosome 2 with AD and 90% penetrance model and in 0.1 marker/cM resolution and 100 markers per group.

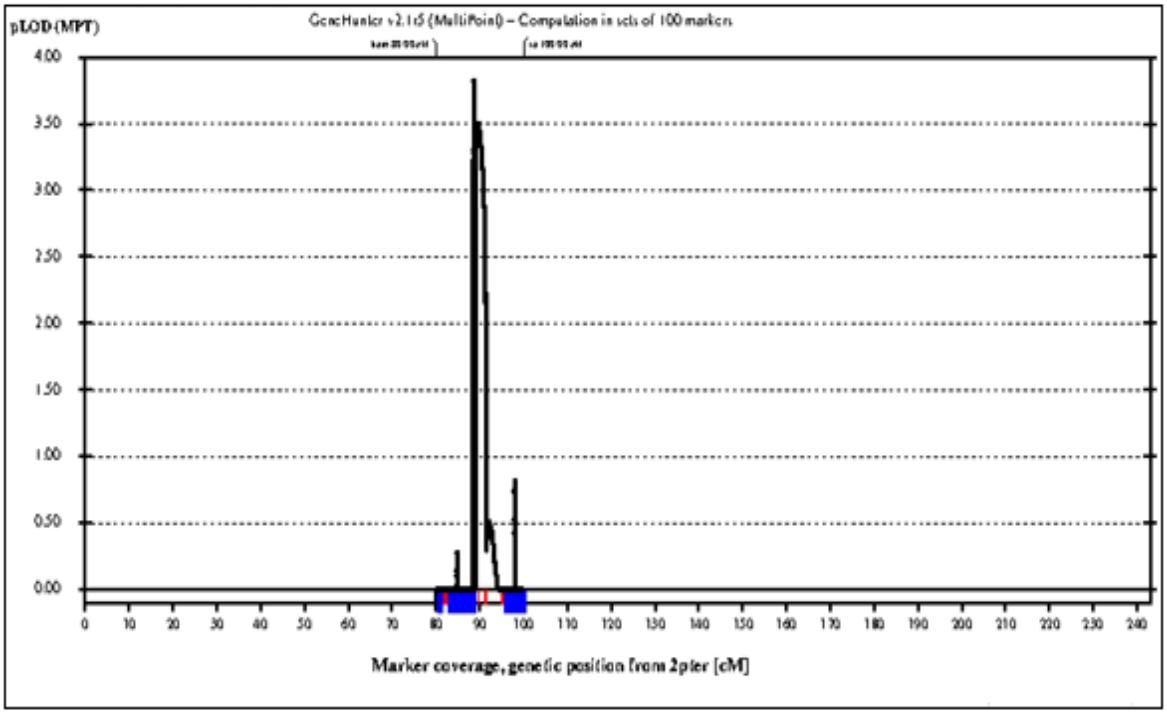


Figure 5.47. Result for parametric multipoint whole genome linkage analysis with AD and 70% penetrance model using every marker on range and 100 markers per group.

Table 5.18. Genetic content of linkage region obtained with AR linkage analysis on chromosome 2, analyzed by Genedistiller program.

Locus Number	First SNP Last SNP	Genetic Location (cM)	Physical Location (mB)	Average LOD Score	Genes				
					All	Protein-Coding	Other	Epilepsy	Nervous System Phenotype
1	rs4444527 rs867014	88.65-88.7	88.8 – 88.9	3.57	2	*EIF2AK3* *AC104134.2*			
2	rs17838437 rs1640007	89.1-90.67	89.1 – 91.8	3.45	93	*IGK* *LOC101060017* *LOC101060169*	50 (IGK Locus)		
3	rs10184457 rs10169419	97.95-98.02	98.5 – 98.6	3.05		*TMEM131*			
Whole Region		88-98	88mB-98mB		240	73	51	RPIA KCNIP3	*EIF2AK3* *KCNIP3* *SEMA4C* *CNGA3* *INPP4A*

5.3.3.3. Whole Genome Homozygosity Mapping. Besides linkage analysis since the inheritance was thought to be autosomal recessive, homozygosity mapping was also conducted using homozygositymapper.org web site. In the result of the analysis there were several regions showing homozygosity at high levels and there were more than 100 genes in those homozygosity stretches (Figure 5.48 and 5.50). However, when we compared these genes with whole exome sequencing data, the only gene mutation in the homozygosity region was on Semaphorin4C (*SEMA4C*) gene.

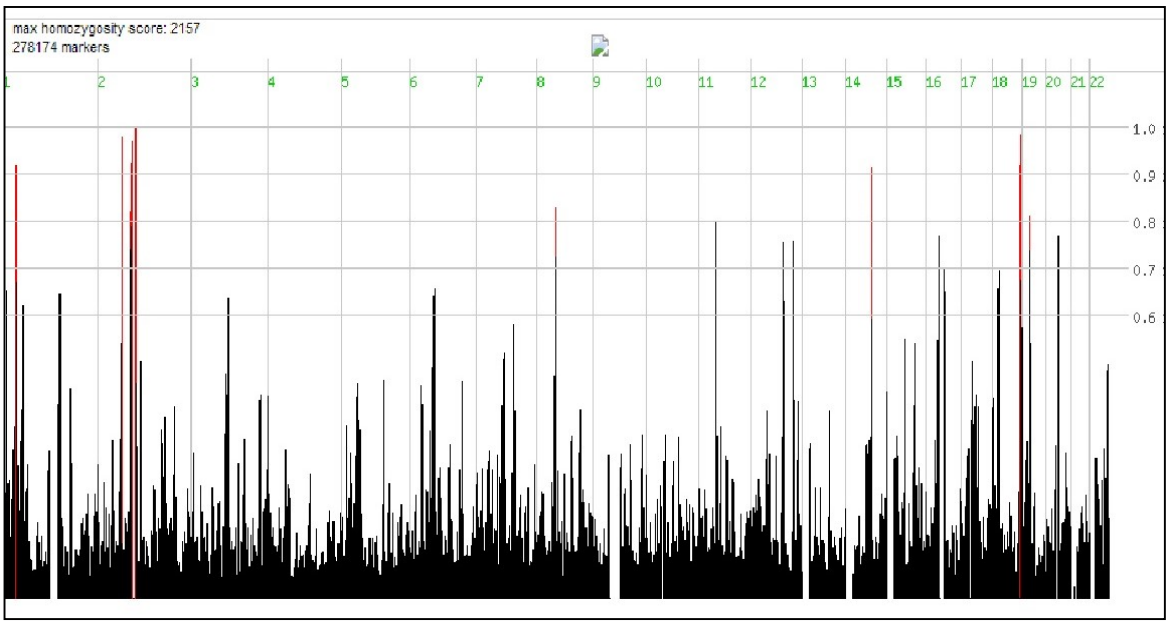


Figure 5.48. Genome-wide homozygosity mapping for GEFS samples.

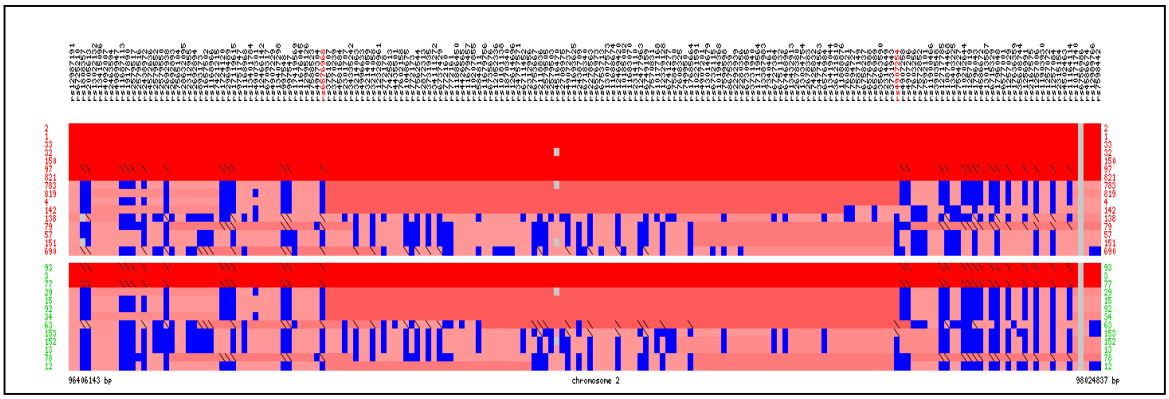


Figure 5.49. Homozygosity map in linkage region. Red samples are affected and green samples are controls. Homozygous regions are shown in red and darker the red color larger the homozygosity stretch.

5.3.4. Co-analysis of Linkage and WES Data

The regions identified by whole genome linkage analysis were compared with the whole exome sequencing data to identify putative variations. Variants with low Minor Allele Frequency (MAF) were selected for validation regardless of sharing between patients. Depending on this criteria 6 variants were selected and validated with Sanger sequencing (Table 5.19-Table 5.20). There were no false positive among these variants. Selected variants were also analyzed in the core family and for individuals selected for linkage analysis, if required (Table 5.21). It was observed that variant in *DPYD* gene was present only in one sample (GEFS3) and was not further analyzed. Variants in *AGL*, *COL11A1* and *INPP4A* genes did not fit the inheritance pattern of analyzed region and were omitted from segregation analysis, which was conducted only for *SEMA4C* and *ABCD3* variants. In the result of segregation analysis, *ABCD3* variant was observed only in one healthy and one affected individual other than the core family. *SEMA4C* variant was present in 4 healthy and 7 affected individual, however two point LOD score of this variant was insignificant both with AR (0.14, 70% penetrance) and AD (1.02, 70% Penetrance) inheritance models.

Moreover WES data of samples GEFS2 and GEFS3 were also analyzed with Integrated Genomic Viewer (IGV) tool to check whether all target exons in linkage region were sequenced or not. Table 5.22 shows IGV analysis results. According to this analysis

14 regions had low coverage. Most of the uncovered region contains untranslated regions of first exons. However microRNA gene mir137 was not covered at all.

Table 5.19. Selected WES variants in linkage regions (Chr: Chromosome, Loc: Location of variant on chromosome, Ref: Reference Allele, Obs: Observed allele).

Chr	Loc.	Ref	Obs	Het/ hom	Function	Gene	Annotation
1	94946011	C	T	het	exonic; splicing	*ABCD3*	NM_002858.3: c.685-9T>C
1	97770920	C	T	het	exonic	*DPYD*	NM_000110: c. 2194 G>A p .V732I
1	100376325	G	A	het	exonic	*AGL*	NM_000645: c.G3707A:p.R1236H
1	103491461	T	G	het	exonic	*COL11A1*	NM_080629: c. 828 A>C p.K276N
2	99193628	C	T	hm	splicing	*INPP4A*	NM_001134224.1 c.2801+22 C>T
2	97527073	C	T	hom	exonic	*SEMA4C*	NM_017789 :c. 1792G> A p.E598K

Table 5.20. Annotation and evaluation of selected WES variants in linkage region (TUB: presence of variant in TUBITAK database; P: Present, NP: Not Present, PP: PolyPHEN2; MT:Mutation Taster, MAF: Minor Allele Frequency).

Gene	Annotation	TUB	MAF	dbSNP ID	SIFT	PP	PhyloP	MT
ABCD3	NM_002858.3 c.685-9T>C	NP	0.0016	rs563660571				
DPYD	NM_000110 c. 2194 G>A p.V732I	P	0.04	rs1801160	0.01	0.905	0.999	0.967
AGL	NM_000645: c.G3707A:p.R1236H	P	0.01	rs12043139	0.07	0.154	0.956	0.999
COL11A1	NM_080629: c. 828 A>C p.K276N	P	0.08	rs12731843	0.49	0.061	0.784	4.0E-6
INPP4A	NM_001134224.1 c.2801+22 C>T	P	0.0026	rs372486077				
SEMA4C	NM_017789 :c. 1792G> A p.E598K	NP	0.0004	rs140700361	0.47	0.001	0.998	0.078

Table 5.21. Genotyping results for selected variants. (Affected individuals are shown in bald).

Patient ID	*SEMA4C* p.E598K	*ABCD3* c.685-9T>C	*DPYD* p.V732I	*AGL* p.R1236H	*COL11A1* p.K276N	*INPP4A* c.2801+22C>T
1	AA	TC	GG	AA	AC	TT
2	AA	TC	GG	AA	AC	TT
3	AA	TT	GA	AG	AA	TT
4	GA	TC	GG	GG	AA	TC
12	GG	TT				
13	GA	TC				
15	GA	TT				
29		TT				
32	AA	TT				
33	AA	TT				
34	GA					
57	GA	TC				
63	GG	TT				
77	GG	TT				
78	GG	TT				
79	GG	TT				
92	GA					
93	GG	TT				
97	GG	TT				
138		TT				
142	GA	TT				
150	AA	TT				
151		TT				
152	GA	TT				
154	GG	TT				
690	GG	TT				
783	GA	TT				

Table 5.22. IGV analysis result in linkage region.

Chromosome	Gene	Exon	Problem
2	*eIF2AK3*	Exon1	Not covered
2	*SEMA4C*	Exon1	Not covered
2	*CNGA3*	Exon1	Not covered
2	*CNGA3*	Exon6	Low Coverage
2	*INPP4A*	Exon1	Not covered
2	*INPP4A*	Exon 25	1 unreported variant
2	*TMEM131*	Exon1	Not covered
1	*EXTL2*	Exon1	Not covered
1	*Mir137HG*	Exon1	Not covered
1	*MiR137*		Not covered
1	*S1PR1*	Exon1	Not covered
1	*RNPC3*	Exon 4	Low Coverage
1	*RNPC3*	Exon 7	Low Coverage
1	*RNPC3*	Exon 10	Low Coverage

5.3.5. Micro RNA 137 (*miR137*) Analysis

When linkage region was evaluated for coverage in WES analysis, it was seen that miR137 (gene ID: 31523) was not covered for any of the samples, along with other genes. Due to its implication in epilepsy and other psychiatric phenotypes micro RNA sequence of *miR137* was amplified by PCR and sequence was analyzed by Sanger sequencing. There were no sequence alterations in *miR137* gene. However, copy number of a VNTR (rs58335419, -/CGCTGCCGCTGCTAC) which is normally 3 in the reference genome was 3 and 5 in the index patient. The VNTR is 6 nucleotides upstream of the exon and inside a CpG island. When the wild type and mutant sequence was analyzed with ESE Finder (Release 3.0) it was seen that VNTR sequence harbored a binding site for SR proteins (Figure 5.50). *miR137* locus was amplified for reliable DNA samples of the large pedigree and the copy number of rs58335419 estimated using DNA1000 chips in Agilent Bioanalyzer. The result is shown in large pedigree in Appendix C.

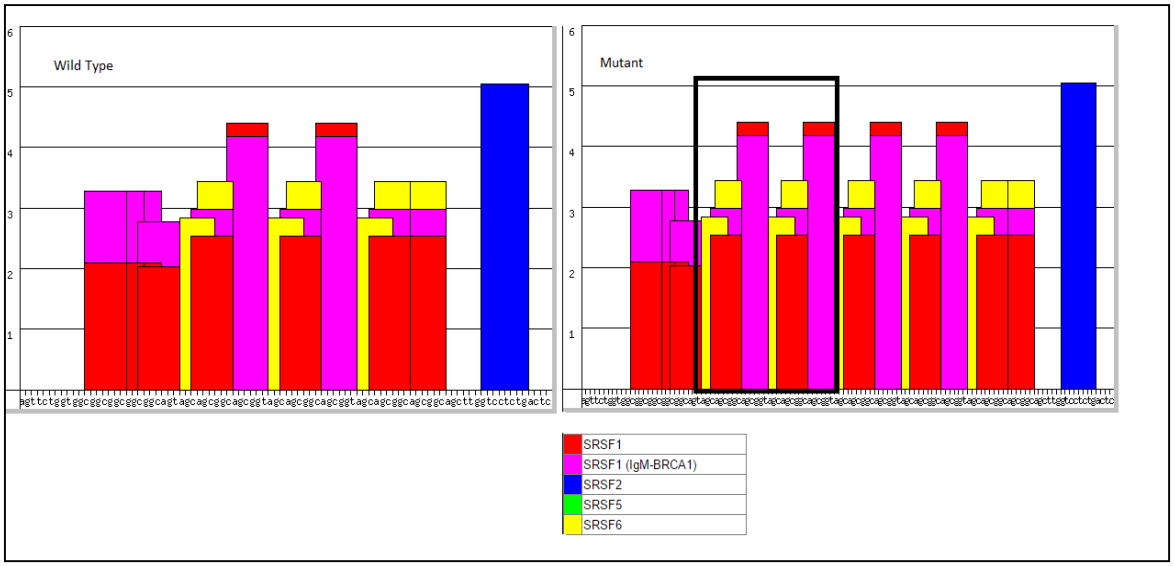


Figure 5.50. ESE Finder result for miR137 variant. Left pane shows the wild type and the right pane shows the mutant sequence. New SR proteins introduced by increasing copy number of VNTR is inside the box. (Wild type sequence: 3copy VNTR, mutant sequence: 5 copy VNTR).

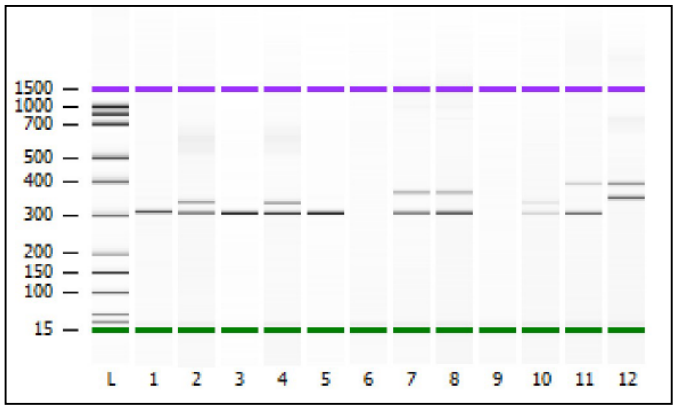


Figure 5.51. Representative gel image produced by Agilent Bioanalyzer. Every lane represents a sample. First lane is molecular weight marker and band sizes are indicated on the left. Green and purple bars represent sample size markers.

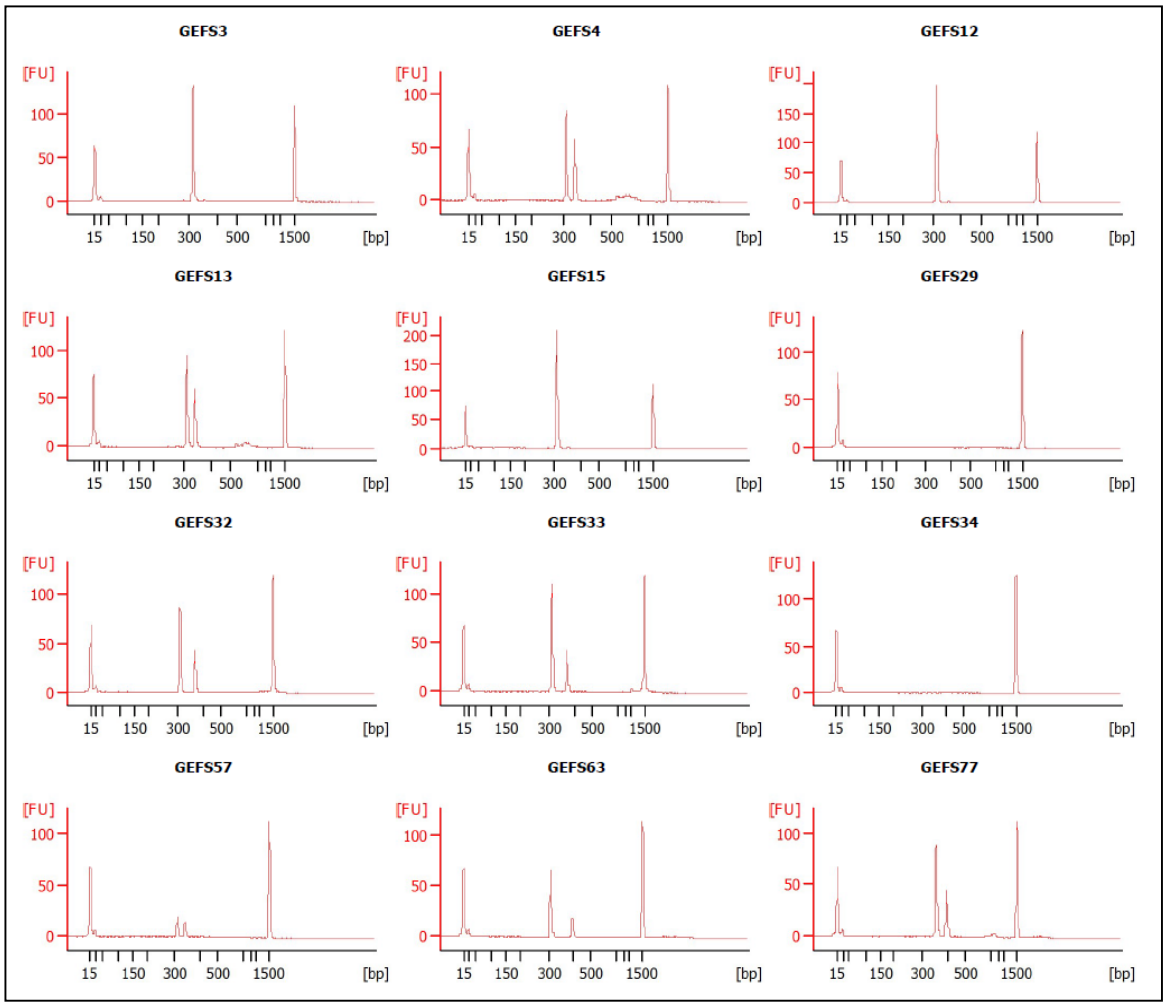


Figure 5.52. Representative electrophoregraph produced by Agilent Bioanalyzer. Each graph represents a PCR product.

5.3.5.1. Statistical Analysis of miR137 VNTR. Inheritance of rs58335419 VNTR in the whole family was analyzed by SuperLink program for two-point LOD score (AD inheritance, 70% penetrance) but resulted without any significant linkage to any repeat number (Figure 5.53). This was evident in the pedigree analysis where there were no repeat allele inherited together with the disease phenotype. However regardless of expansion number, affected individuals had the phenotype that included epilepsy, depression, anxiety or head ache.

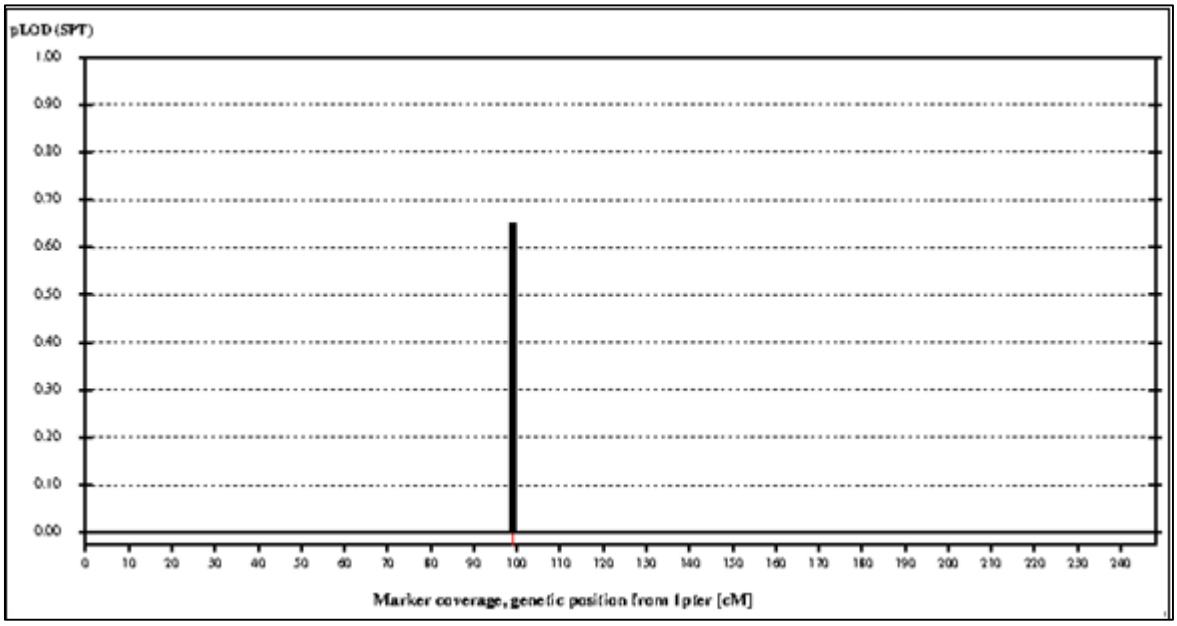


Figure 5.53. Two point LOD score analysis for *mir137* VNTR in GEFS pedigree.

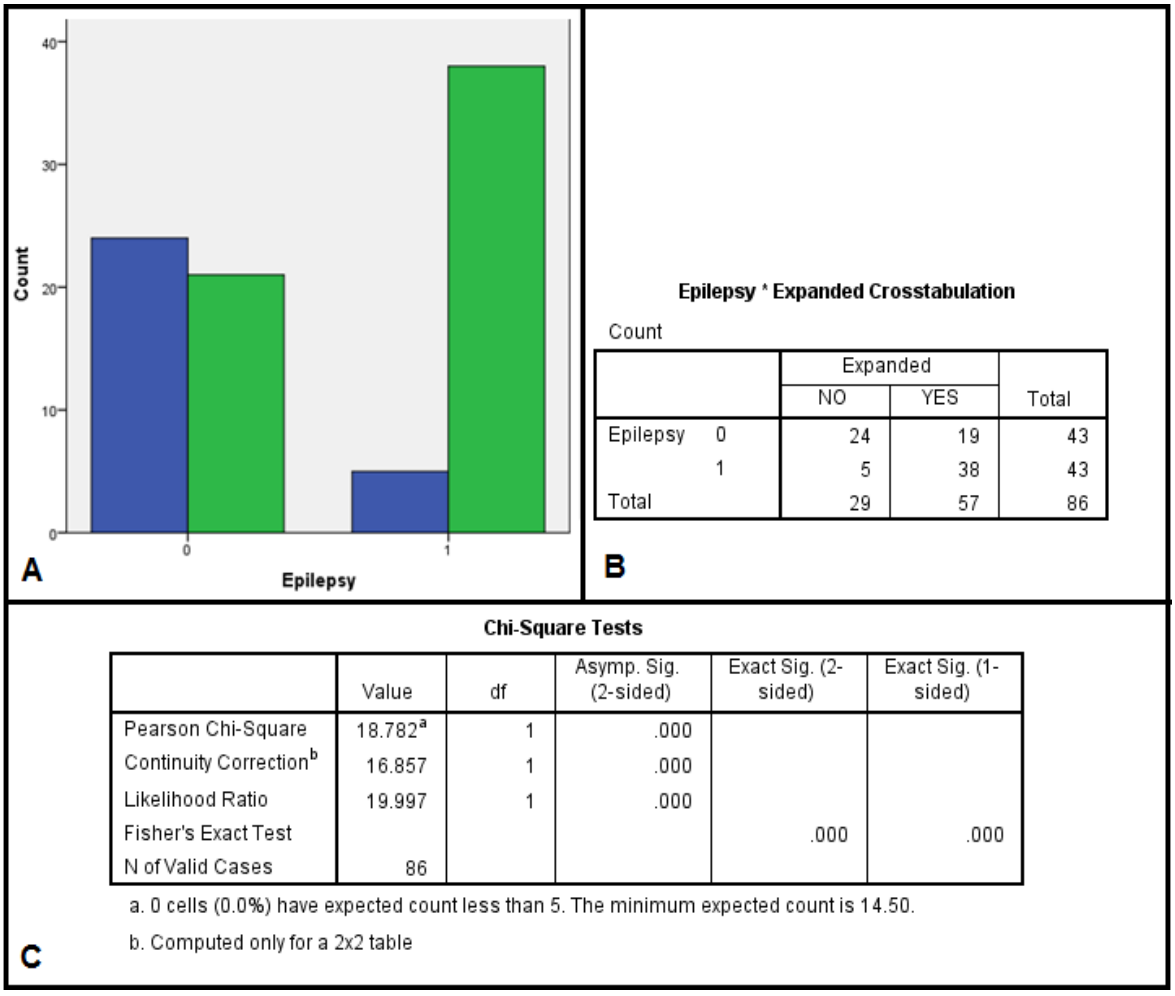
In order to analyze whether expansion of rs58335419 VNTR was associated with any disease phenotype, 86 individuals that were examined both by a neurologist and also by a psychiatrist were included into the analysis. The genetic data converted into nominal as individuals over 4 copy accepted as “expanded” and also VNTR copy numbers were added together to obtain a total expansion score.

In order to analyze whether there was a significant difference between patients with/without epilepsy phenotype in regard of *miR137* expansion status, chi-square test is applied. Bar chart in Figure 5.54 shows that individuals with epilepsy phenotype have *miR137* VNTR expanded more than the individuals without epilepsy.

The significance of association between *miR137* VNTR expansion and epilepsy status was analyzed by cross-tabulation option of SPSS software. "The cross tabulation table is the basic technique for examining the relationship between two categorical (nominal or ordinal) variables, possibly controlling for additional layering variables." Chi-Square test for epilepsy phenotype showed significant difference between individuals with or without epilepsy phenotype (Figure 5.54, Panel A). SPSS cross tabulation option conducted four different chi-square tests. Though all of them applied in situations where groups of nominal values were compared to each other Fisher's Exact Test was the most reliable for conditions with low sample number. However, all four tests showed significant difference between epilepsy patients and unaffected individuals in regard of *miR137* VNTR expansion.

Additionally in order to check whether there was a difference between total expansion score between individuals with/without epilepsy, independent samples –t test was applied. Box-plot shows the difference between group means (Figure 5.55). Table 5.23 shows the results for t-test. Group statistics show there is a difference between mean total expansion between individuals with or without epilepsy. Independent samples-t test check whether this difference was random or not. T- Test assumes equal variance between groups and test it using Levene's test for equality of variances. This test showed there is not a significant difference between group variances. Lastly, significance of t-test was 0.000, which meant that there was a significant difference of total expansion score between individuals with or without epilepsy.

In our cohort epilepsy patients showed four distinct phenotypes, namely IGE and partial epilepsy with motor, sensory and psychiatric implications. Histogram in Figure 5.56 shows the number of each epilepsy type in our cohort and box-plot in Figure 5.57 shows the total expansion difference between epilepsy types. In these graphs, red bar represents individuals with IGE phenotype, yellow bar represents partial motor epilepsy, blue bar represent partial epilepsy with sensory features, purple bar represents the partial epilepsy with psychiatric features, green bar represents individuals with unclassified epilepsy and beige bar represents unaffected individuals. In order to check whether there is any difference between epilepsy types in regard to *miR137* VNTR total expansion one-way ANOVA test was applied, however, it did not show a significant difference (Table 5.24).



Epilepsy * Expanded Crosstabulation

Count

		Expanded		Total
		NO	YES	
Epilepsy	0	24	19	43
	1	5	38	43
Total		29	57	86

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	18.782[a]	1	.000		
Continuity Correction[b]	16.857	1	.000		
Likelihood Ratio	19.997	1	.000		
Fisher's Exact Test				.000	.000
N of Valid Cases	86				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 14.50.

b. Computed only for a 2x2 table

Figure 5.54. Chi-square test results for miR137 expansion status between individuals with or without epilepsy. Panel A: Bar graph showing *mir137* expansion status between individuals with (1) and without (0) epilepsy (Blue=Not Expanded; Green=Expanded). Panel B: Epilepsy vs Expanded cross tabulation showing the number of individuals in corresponding situation. Panel C: Results for four types of chi-square test.

Table 5.23. Independent samples t-test results for the difference between individuals with or without epilepsy regarding miR137 VNTR expansion.

Independent Samples Test									
	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	99% Confidence Interval of the Difference	
								Lower	Upper
Total_Expansion	.013	.909	3.983	84	.000	2.023	.508	.684	3.362

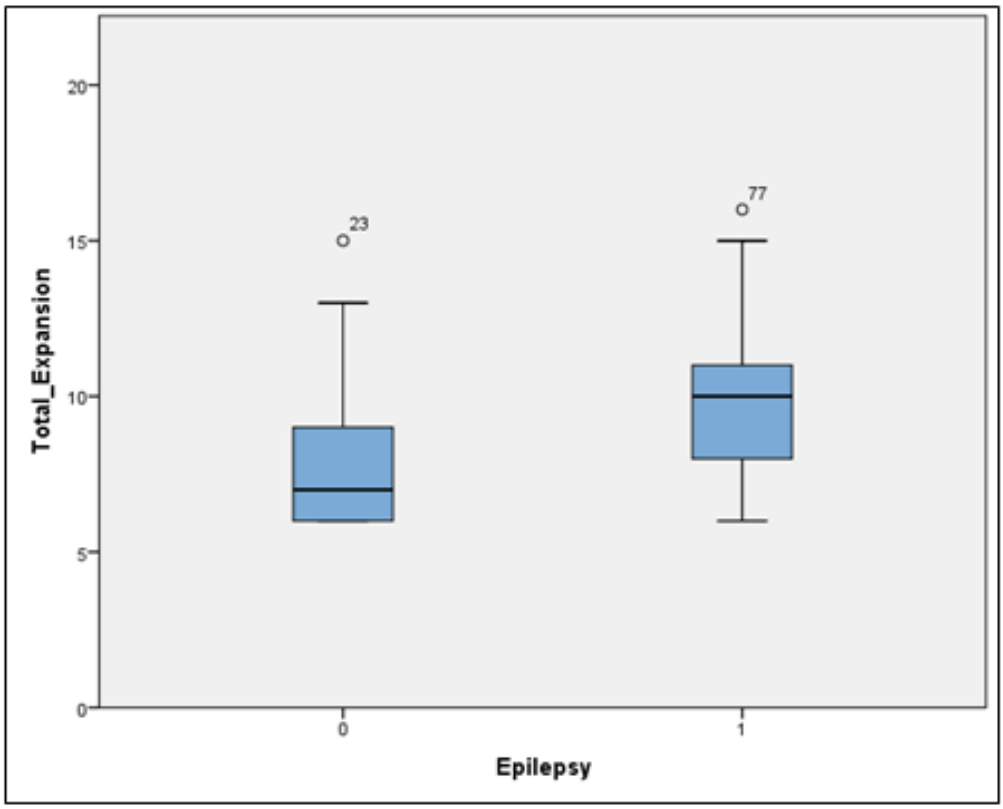


Figure 5.55. Box plot showing the difference between individuals with or without epilepsy phenotype. Center of the boxes show mean for each group, box borders represent one standard deviation and bars represents minimum and maximum.

Table 5.24. Test results for one-way ANOVA comparing total expansion score difference between different epilepsy groups.

ANOVA					
Total Expansion					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.033	4	1.008	.157	.959
Within Groups	243.595	38	6.410		
Total	247.628	42			

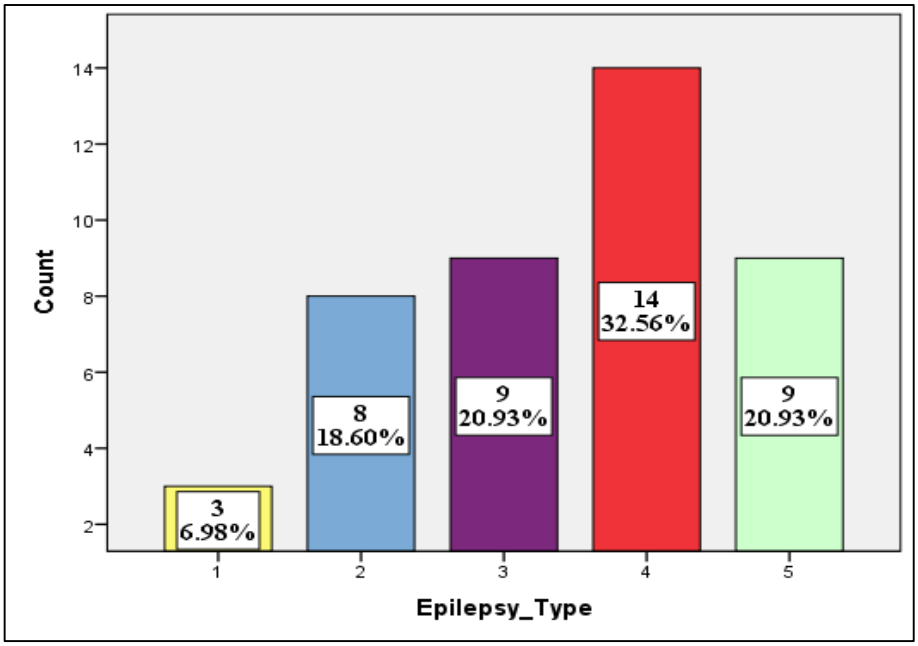


Figure 5.56. Histogram showing the count of different epilepsy types in the cohort.

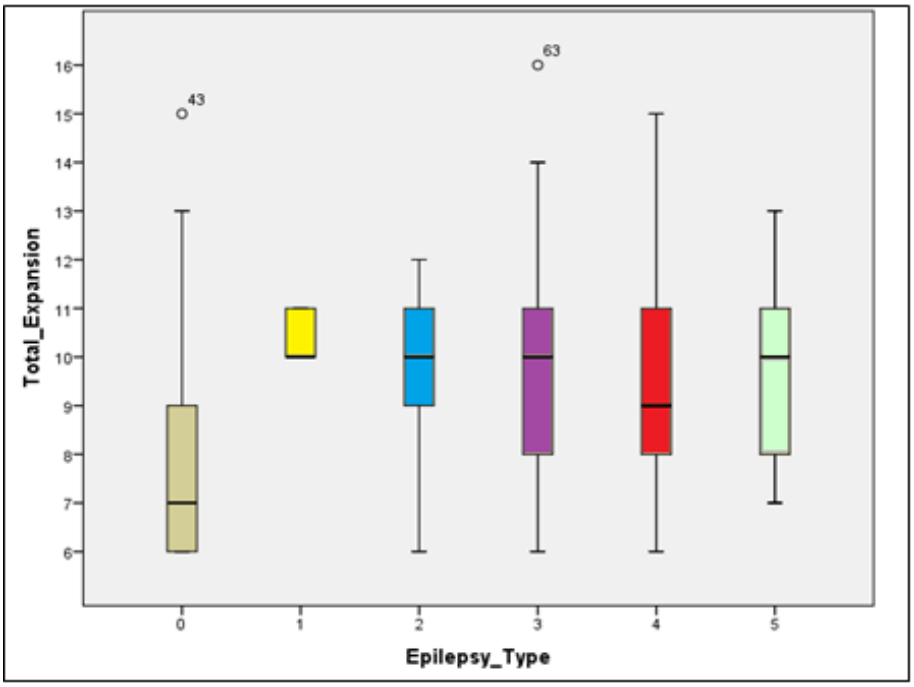


Figure 5.57. Box-plot showing mean *miR137* VNTR total expansion score of different epilepsy types in the cohort.

<u>5.3.5.2. Phenotype Expansion.</u> Other than epilepsy, depression, anxiety and headache was frequent in the pedigree. Figure 5.59 shows the distribution of these phenotypes in the

analyzed cohort. In further analysis; epilepsy, depression, anxiety and headache were accepted as continuum of a phenotype. Affection status was analyzed into two categories, affected individuals carry at least one of the aforementioned phenotype and healthy individuals have none.

Additionally, an affection score was computed by counting the number of phenotypes, such that if an individual had epilepsy together with depression his/her score was 2, if an individual only had epilepsy or migraine or depression or epilepsy his/her score was 1. So there were 5 different categories form 0 to 5 (Figure 5.59). According to this analysis, 20 people had no phenotype and 66 individuals had at least one of the aforementioned phenotypes though most of them had only one phenotype (n=37).

The relationship between affection status (expanded affection) and expansion status was analyzed by chi-square tests of SPSS program. Figure 5.60A shows the histogram of affected individuals who had *miR137* expansion. 49 affected individuals had a *miR137* expansion on the other hand only 8 unaffected individual had *miR137* expansion. Fisher's exact test also confirmed this difference between affected and unaffected individuals (p=0,007) though not as significant as difference between individual with/without epilepsy only (Figure 5.60B).

Figure 5.61 compares total *miR137* VNTR expansion score between individuals with no phenotype and individuals with different affection scores. Individuals without any affection score had lower mean expansion score than other groups. On the other hand, individuals who had all of four phenotypes (Affection Score =4) had the maximum mean expansion score. Independent samples t-test was also applied to test whether there were any difference in total expansion score between affected and unaffected individuals. Table 5.25 shows the test result. Though sample sizes were not equal between affected and unaffected individuals, Levene's test showed equal distribution of variation between groups, which showed that the t-test was reliable to use in this analysis. T-test showed a significant difference with p value of 0.016, but similar to chi-square test, it was not as strong as difference between individuals with/without epilepsy.

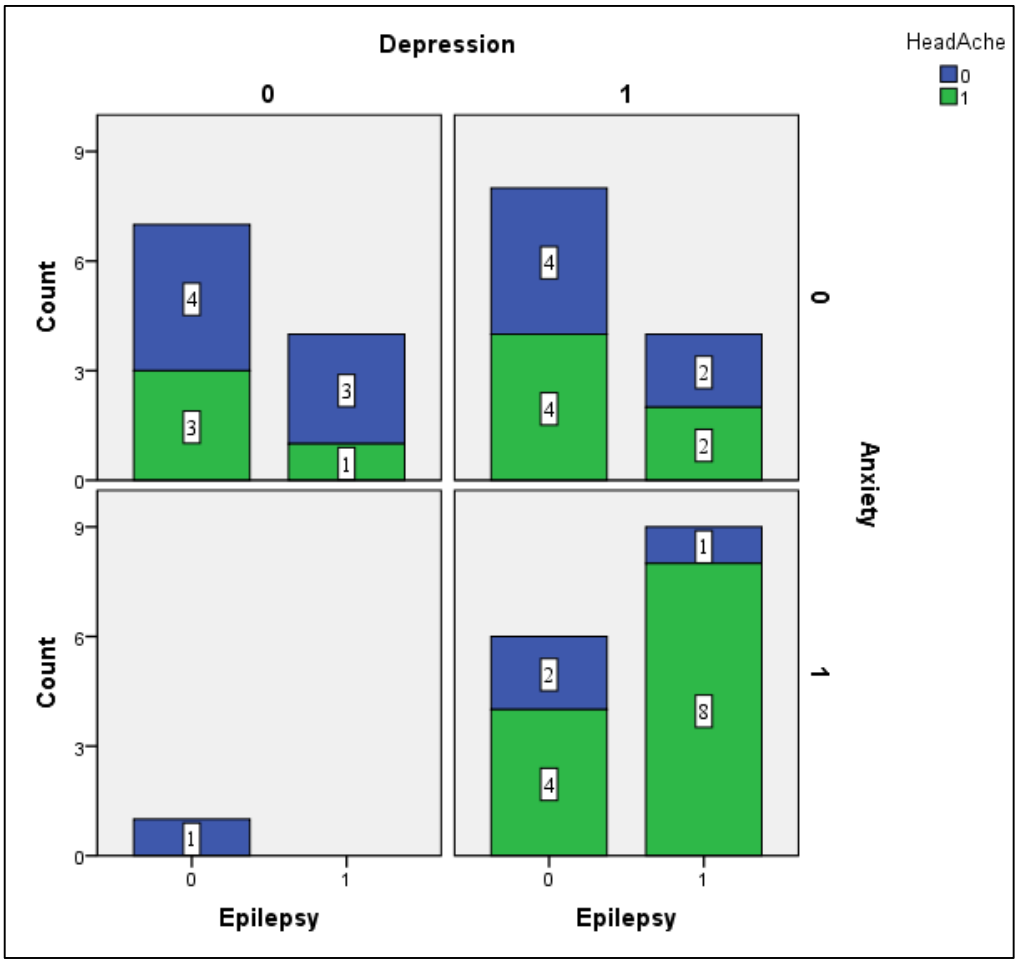


Figure 5.58. Stacked bar graph showing the count of different phenotypes in different combinations.

Table 5.25. Independent samples t-test comparing total expansion score of *miR137* VNTR between affected and unaffected individuals.

Independent Samples Test									
	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	99% Confidence Interval of the Difference	
								Lower	Upper
Total Expansion	2.634	.108	2.460	84	.016	1.558	.633	-.111	3.226

One way –ANOVA test was conducted to check whether there was any difference between affection score groups regarding *miR137* expansion and showed that there was not a significant difference between different affection scores (Table 5.26).

Table 5.26. One-way ANOVA result comparing total expansion score of *miR137* VNTR between different total affection score groups.

ANOVA					
Total Expansion					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	34.521	3	11.507	1.698	.177
Within Groups	420.100	62	6.776		
Total	454.621	65			

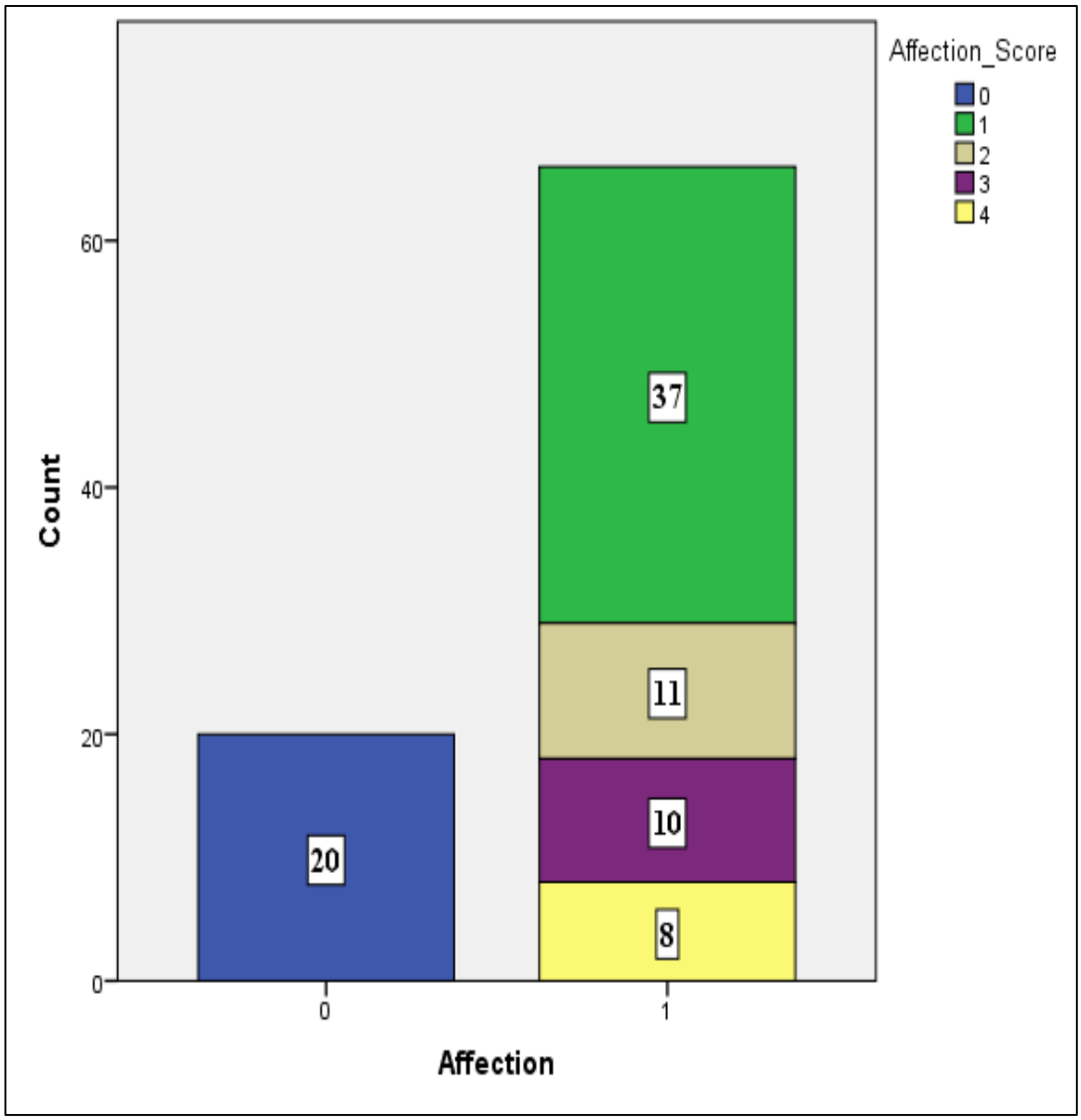
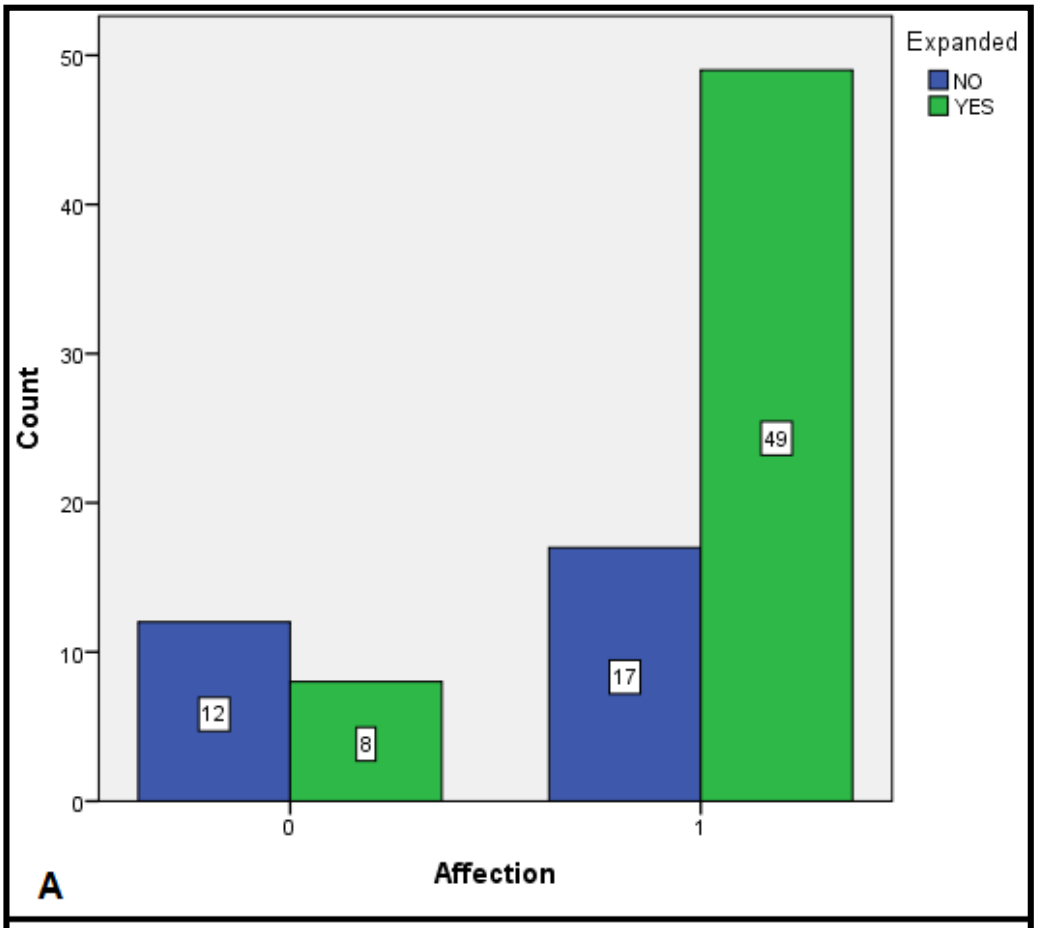


Figure 5.59. Stacked bar graph showing the count of different affection cohorts in each affected group and number of unaffected individuals.



B

	Value	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Fisher's Exact Test		.007	.006
N of Valid Cases	86		

Figure 5.60. Fisher's exact test results for affection regarding *miR137* VNTR expansion. (A.) Histogram showing the count of individuals affected or not affected who had *miR137* VNTR expansion or not. (B). Result for Fisher's exact test comparing expansion status between affected and unaffected individuals.

In conclusion of the statistical tests, individuals with epilepsy phenotype had significantly more *miR137* VNTR expansion shown both by Fisher's exact test ($p<0.001$) and also by independent samples t-test ($p<0.001$). However there was not a significant difference between epilepsy subgroups. When depression, anxiety and headache were assumed as expansion of the epilepsy phenotype and individuals were accepted as affected if they had either of the phenotypes, a significant difference between affected and unaffected individuals in regard to *miR137* VNTR expansion was observed both with

Fisher's exact test (0.07) and independent samples t-test (0.016). But the difference was not as significant as difference between individuals with or without epilepsy. Additionally there was not a significant difference between individuals with different total affection score, which was count of phenotypes observed in one individual.

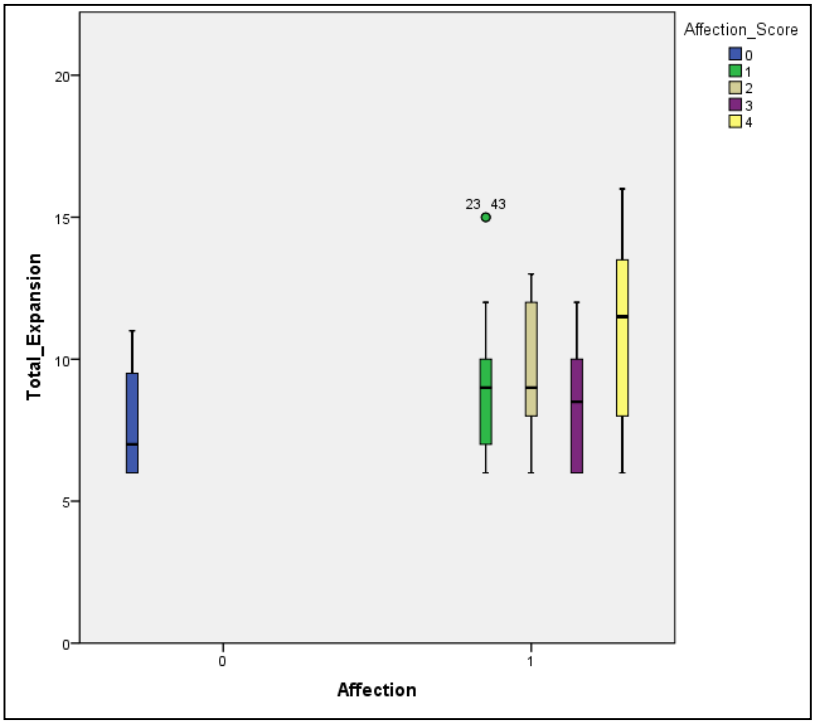


Figure 5.61. Box plot of total *miR137* VNTR expansion score between individuals with no phenotype and individuals with different affection scores.

5.3.6. Whole Genome CNV Analysis

5.3.6.1. Copy Number Analysis from SNP Data. Together with whole genome SNP genotyping, using Illumina CytoSNP chips, whole genome copy number variation of samples were also evaluated. There were no chromosomal aberration above the threshold level (Figure 5.62), but when the thresholds were not taken into consideration there were 4 copy number aberrations in more than one sample. However when the facts that these aberrations were observed in both heathy and affected individuals and that they were reported as possibly population CNVs in DECIPHER database were taken into consideration, they were rather evaluated as benign. Results are listed in Table 5.27.

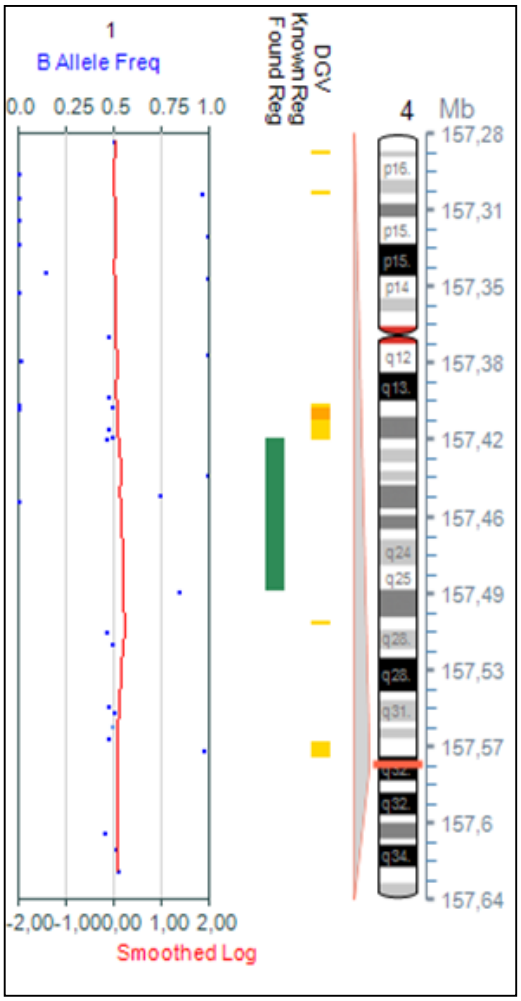


Figure 5.62. CNV analysis results by Illumina Karyostudio program showing a copy number gain region.

Table 5.27. Whole genome CNV analysis results for GEFS samples.

Sample	Chr	Start	Stop	Size	Copy#	Gene
4	12	7993316	8086083	92767	1	*SLC2A14, NANOGP1, SLC2A3*
2	12	8000336	8129708	129372	1	*SLC2A14, NANOGP1, SLC2A3*
12	12	8000336	8129708	129372	1	*SLC2A14, NANOGP1, SLC2A3*
92	Y	2929037	6581824	3652787	3	*PCDH11Y*
4	Y	3801503	6616258	2814755	3	*PCDH11Y*
2	Y	4663804	4925716	261912	3	*PCDH11Y*
33	17	44169808	44350293	180485	3	*KANSL1* Region

Table 5.27. Whole genome CNV analysis results for GEFS samples (Cont.).

Sample	Chr	Start	Stop	Size	Copy#	Gene
34	17	44169808	44350293	180485	3	*KANSL1* Region
15	17	44169808	44350293	180485	3	*KANSL1* Region
690	17	44215896	44345063	129167	3	*KANSL1* Region
63	17	44222183	44345063	12280	3	*KANSL1* Region
1	9	40215520	42374011	2158491	3	*SPATA31A3* *ZNF658* *SPATA31A4* *SPATA31A5* *ANKRD20A2*
138	9	40215520	42374011	2158491	3	*SPATA31A3* *ZNF658* *SPATA31A4* *SPATA31A5* *ANKRD20A2*
151	9	40215520	42374011	2158491	3	*SPATA31A3* *ZNF658* *SPATA31A4* *SPATA31A5* *ANKRD20A2*
152	9	40215520	42374011	2158491	3	*SPATA31A3* *ZNF658* *SPATA31A4* *SPATA31A5* *ANKRD20A2*

5.3.6.2. Array CGH Results. Additionally, samples GEFS1 and GEFS4 were also analyzed with Agilent ISCA 8X60K array. With default analysis method no alteration was observed in sample GEFS4, however there were one deletion and one duplication regions in GEFS1. The whole gene region of *CATSPER2* gene was deleted (Figure 5.63, left). *CATSPER2* is a sperm specific calcium channel. CNVs on the gene region was observed in population, but suspected to be more frequently observed in epilepsy patients (Nicholl *et al*., 2013).

The other gene number variation was the duplication on the imprinted and maternally expressed non-coding RNA gene *H19*, which is implicated in Beckwith-Wiedemann syndrome, Silver-Russel syndrome and Wilms Tumor type2 (OMIM) (Figure 5.63, right). Pathogenicity of this variant needs to be further investigated.

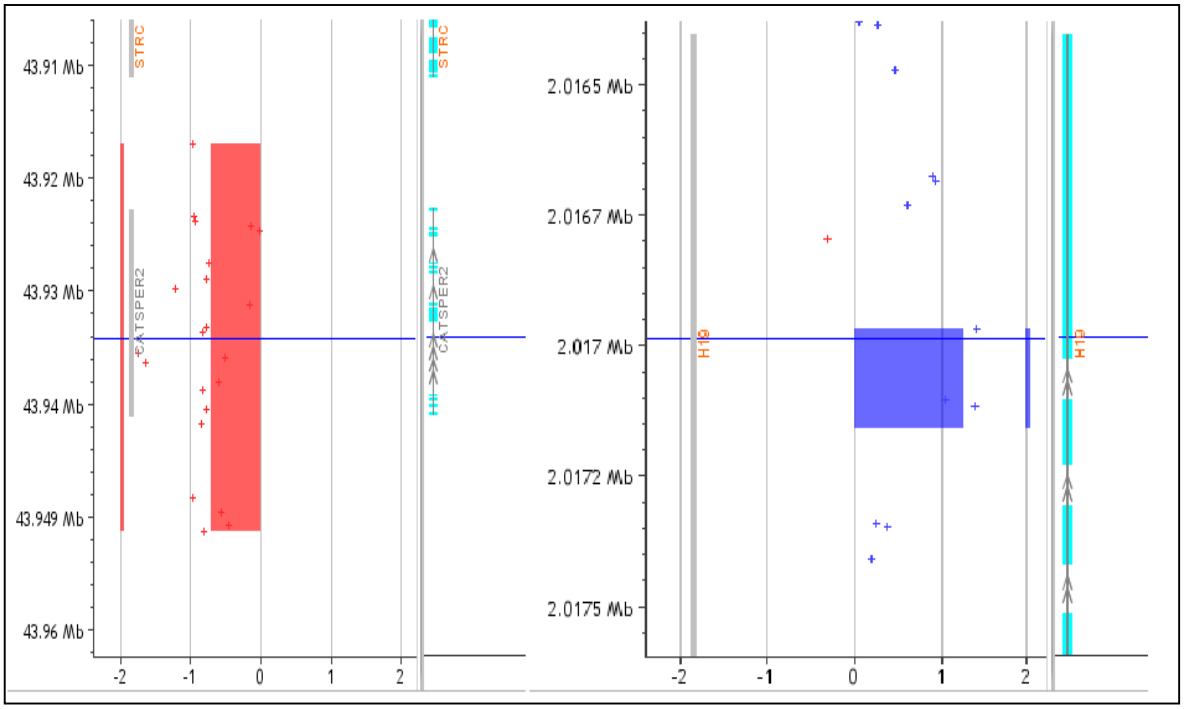


Figure 5.63. Array CGH results for patient GEFS1, showing copy number loss in *CATSPER2* gene and copy number gain on *H19* gene.

6. DISCUSSION

6.1. Genetic Analysis of 4BF Family

BFNS, BFNIS and BFIS are the three similar epilepsy types that occur in the first year of life, the age of onset of seizures being the major clinical distinctive feature. Genetically, they are caused by mutations in different genes with overlaps to some degree. Among these the association of the *SCN2A* gene mutations with the BFNS phenotype is the most common (Heron *et al* 2002). The *KCNQ2* gene mutations are generally observed in BFINS phenotypes however, there are also reports of *KCNQ2* mutations in BFIS and BFNS. On the other hand majority of BFIS patients have a truncating *PRRT2* mutation (Zara *et al*., 2013).

In the first part of this study genome-wide linkage analysis of the multigenerational and multiplex BFIS family with index patient negative for mutations in *KCNQ2* and *SCN2A* genes resulted in a haplotype segregating with the disease with a significant LOD score of 3.48 in a 15 Mb region on the chromosomal segment 19p12-q13 that has previously been reported as candidate linkage region for BFIS (Guipponi *et al*., 1997). The mutational analysis of the most plausible positional candidate gene, the *SCN1B* gene, identified a synonymous nucleotide exchange (NM_001037.3 c.492T>C; NP_001028.1 p.T164T). The mutation was inherited together with the disease phenotype except that there were three healthy individuals in the family who carried the mutation most probably due to incomplete penetrance. One homozygous affected member did not have a more severe phenotype.

Since there was no any other strong candidate in the large linkage region whole exome sequencing of the index (4BF14) and the affected individual who was homozygous for the *SCN1B* variant (4BF17) was conducted. WES data of the linkage region identified Sanger validated *ZNF792*, and *GRAMD1A* variants shared by both patients and their co-segregation with the risk haplotype implicated that both variants resided in the risk haplotype. *ZNF792* is a zinc finger protein and may have a role in transcription regulation but not much is actually known about its function. According to the Expression Atlas

database of EMBL-EBI it is not expressed in the brain and the effect of the variant was shown to be benign by in silico testing. The other segregating variant was a splice site variant in *GRAMD1A* gene that codes for GRAM domain-containing protein 1A, a seven pass transmembrane protein (SwissProt). According to the Expression Atlas database of EMBL-EBI it has low expression in the brain and is a low frequency SNP in dbSNP (rs558298045). Both variants in *SCN1B* and *GRAMD1A* genes were not detected in at least 300 Turkish population controls. Another variant on linkage region was in *KCTD15* gene which codes for the Potassium Channel Tetramerization Domain Containing protein. This protein has function in neural crest formation. Though gene appears as a good candidate, the frameshift mutation in *KCTD15* gene was not shared by the two patients analyzed by WES, and segregated by a lower LOD score of 2.02 in the family. All data obtained after linkage and WES analysis leaves the *SCN1B* gene as the most plausible candidate gene.

On the other hand, WES data of patient 4BF14 showed the common variant, p.R217PfsX8, in the *PRRT2* gene at 16q11.2, which was associated with the BFIS phenotype (Heron *et al*., 2012). However, the variant had a reduced penetrance since 6 unaffected individuals carried the mutation and two affected individuals did not carry the mutation and should be considered as phenocopies. Reported penetrance of *PRRT2* mutations is around 70%, which implies a modifying effect on the phenotypic variance of BFIS. Many reports of unaffected carriers of *PRRT2* p.R217PfsX8 mutation was attributed to diagnostic difficulties (Nobile and Striano, 2014). However, the pedigree and phenotype information in our family was very carefully taken. Moreover, affected individuals in the family were generally hospitalized due to long cluster of seizures.

The structural analysis of the genome of patient 4BF18 by a custom CGH array that specifically covered the linkage region and BFIS associated genes in exon resolution and rest of the genome with less resolution, a *KCNQ2* amplification was detected. No other exonic structural variant was seen in the linkage region. In total 3 healthy and 6 affected individuals were selected specifically from individuals who show obscure inheritance of the *PRRT2* mutation and analyzed for *KCNQ2* duplication by qPCR analysis. A possibly heterozygous duplication of the *KCNQ2* gene was observed in all affected individuals and also in two unaffected individuals. Pathogenicity of *KCNQ2* deletions were well established as a major cause of the BFIS phenotype (Zara *et al*., 2013). However, there was

only one report on *KCNQ2* duplication, which is a partial duplication encompassing a few exons (Heron, *et al*., 2007). Reported pathological copy number gains in DECIPHER and ISCA databases for the *KCNQ2* gene, all encompassed multiple genes besides *KCNQ2*. The effect of *KCNQ2* duplications on the phenotype needs to be further explored.

The most plausible candidate is the *SCN1B* synonymous variant (NM_001037.3 c.492T>C; NP_001028.1 p.T164T) was shown to have a possible effect on splicing by the tools ESE Finder and PESX. The cDNA sequence obtained from total blood mRNA of the index patient (4BF14) and his parents (4BF15, 4BF16) had the normal sequence without an effect on splicing. This fact might be due to NMD of the mutant allele however, the presence of both wild type and mutant sequences after Sanger sequencing eliminated the possibility of NMD.

However, the most striking feature of BFIS and the two other similar syndromes is their time dependent occurrence. They have specific age of onset and patients become totally disease free after age of one or two. On the other hand, it has been shown by Kazen-Gillepsie *et al*. that *SCN1B* splice variants were expressed at different periods of human development. β1B variant (which is not a transmembrane protein) is dominant during fetal development and replaced by β1 after birth (Kazen-Gillespie *et al*., 2000). It might be speculated that tissue specific factors may be shifting the balance of splicing regulators. Moreover, all of the reported pathological *SCN1B* variants implicated in other epilepsy phenotypes were observed on the same gene region shared by the two splice variants (Brackenbury and Isom, 2011). The synonymous change detected in 4BF family, however, was on the B1 variant. These data are depicted on Figure 6.1. ın this Figure panel A shows the genomic structure of *SCN1B* gene. Exons are shown as green boxes (dark green is coding region and light green is non-coding region) and introns are shown as lines. Panel B1 and B2 shows mRNA structure of two splice variants (β1B in panel B1 and β1 in panel B2). In both of these panels, blue bars represent mRNA, green bars represent exons and yellow bars represent coding region. Light blue arrow shows region shared by two splice variant. Light green bars show Ig-like C2-type domain and red bar shows transmembrane domain. Yellow star on exon 4 shows the place of the c.492T>C mutation.

The *in-vitro* effect of the variant was shown using the system developed by Kishore *et al*. by constructing a splicing reporter plasmid; inserting *SCN1B* exon 4 into the pSpliceExpress vector and expressing the plasmid with native and variant *SCN1B* exon 4 in neuroblastoma cells. The total mRNA produced by the reporter vector when amplified by two step RT-PCR showed proper splicing with both the wild type and variant exon 4 but inhibition of splicing of exon 4 in the variant seemed to be removed. Exon 4 of *SCN1B* codes for the transmembrane domain of the protein. During early development, in neuronal cells, the loss of splicing inhibition in B1 variant may interfere with the balance between the two variants. Technically the effectiveness of the minigene system testing for splicing defects was shown by assessing the function of a well-established *GABRG2* splice site variant during MS thesis study by Öyküm Kaplan. This *in vitro* system is promising for assessment of developmental regulation of *SCN1B* gene splicing.

The large amounts of data generated by current technologies may actually have significant contributions to the understanding of basic genetic concepts such as low penetrance and modifier effects, however, the pedigree in Figure B.1 in Appendix B, showing the genotypes of all 3 variants in the 3 genes did not give an obvious clue as to the modifier effects of *KCNQ2* duplication or the *PRRT2* common mutation. The synergistic effects of mutations in these genes should be investigated in larger family cohorts.

Figure 6.1. Genomic and mRNA structures of two splice variants of the *SCN1B* gene.

Considering the results obtained by whole genome linkage and exome sequencing analysis and *in vitro* expression analysis, it was concluded that the *SCN1B* splice site variant was the most plausible genetic cause of BFIS in the family. This study provided the first genetic evidence of *SCN1B* mutation implication in the BFIS phenotype and showed the importance of synonymous changes that were generally ignored. In addition, the presence of mutations in 3 plausible genes namely *SCN1B* synonymous change affecting splicing efficiency, *KCNQ2* duplication and a common *PRRT2* mutation with variable degrees of penetrance in the family clearly reveal the oligogenic nature of the disease phenotype.

6.2 Mutational Analysis of Other BFIS and BFNS Families

6.2.1. BFIS Families

Beside large BFIS family 11 other smaller families with putative BFIS phenotype was also included in the study. These families were characterized by infantile onset benign epilepsy however, disease onset at day 45 in index case of 6BF45 is rather early for BFIS phenotype and for case 12BF74, and disease onset is age of 3 years, which is considered as late. Other than these cases, index patients have disease onset between 3 months and 9 months which is characteristic for BFIS. Characteristically BFIS patients are seizure free 1 or 2 years after disease onset however, for cases 9BF59 and 14 BF88 seizures continue until age of 3.5 and 11 years, respectively. Though this is not an expected situation for BFIS patients, this might be a rare sub-phenotype.

6 index cases with BFIS phenotype were analyzed for an alteration in *SCN1B* gene resulting in no novel variation. But previously reported SNPs are observed in heterozygous state for each patient except for 7BF49. Though they are previously reported minor allele frequency (MAF) of these SNPs are below 3% and they are considered as rare variants. Additionally, SIFT analysis resulted as damaging for three of the SNPs. They are considered as benign in ClinVar however, contribution of these alleles to the disease phenotype needs to be further investigated.

5BF is a relatively larger family with 11 individuals in three generations, 7 healthy and 4 affected individuals. Simulation analysis with Autosomal Dominant and Autosomal Recessive disease phenotypes with 99% penetrance produced 2.09 and 1,79 maximum LOD score for the family, yielding 4 and 2 loci reaching maximum LOD score, respectively. There were several protein coding genes in these regions.

On the other hand after discovery of large implication of *PRRT2* gene on BFIS phenotype, this family is screened for BFIS mutation together with 24 index patients in 16 different families either with BFIS or BFNS phenotype. In the result of the analysis, a two nucleotide deletion (c.215-216 delCA) is observed in heterozygous state in all affected individuals together with a healthy member of the family. This was a surprising result since there was not a significant linkage to *PRRT2* locus. But this may be explained by unaffected individual who carry the mutation. On the other hand, low penetrance is a common attribute of PRRT2 mutations.

Shortly after discovery of *PRRT2* mutation in Paroxysmal Dyskinesia patients(Chen *et al*., 2011; Wang *et al*., 2011) it was reported as major gene for BFIS phenotype with 70 to 80% mutation rate in different populations (de Vries *et al*., 2012; Heron *et al*., 2012; Schubert *et al*., 2012; Zara *et al*., 2013). However, on the contrary in our cohort of BFIS and BFNS patient mutation rate is very low; only 2 families (4BF and 5BF) out of 16 have *PRRT2* mutations. Additionally, this variation shows incomplete penetrance in our families. Taking these facts into account, though repeatedly reported, implication of *PRRT2* in BFIS phenotype need to be further evaluated.

6.2.2. Mutational Analysis of BFNS Patients

Blood samples of 5 patients are provided by Prof. Dilşad Türkdoğan upon observation of a common clinical phenotype, namely apneic seizures triggered by breast feeding originating from temporal lobe. However, the main difference in cohort is age of onset. Three of patients have disease onset in first 10 days and two other has later onset.

The reported clinical onset resembles benign epilepsy syndromes of first year so all implicated genes namely *KCNQ2*, *SCN2A* and *PRRT2* were screened in these patients and

all were negative for point mutations. In addition, the copy number alterations in these genes were investigated by a custom design CGH array, resulting in *KCNQ2* gene duplication for four patients and *PRRT2* gene duplication in all. *KCNQ2* gene duplication is confirmed also by qRT-PCR for three of four patient and duplication found to be false positive for one patient. On the other hand, *PRRT2* duplication found to be false positive for all patients.

When we analyze patients with *KCNQ2* duplication we see that all three patients have disease onset in the first 10 days of life that strikes out as a common feature for patients from independent families and support the pathogenicity of the duplication. When family members were also analyzed by qRT-PCR and it was observed that both reportedly healthy parents of affected indexes also have *KCNQ2* duplications which may be explained by incomplete penetrance as discussed for the large BFIS family, pathogenic consequence of *KCNQ2* duplication may be a part of an oligogenic system that needs to be investigated through WES analysis.

6.3. GEFS+ Study

Index patients of GEFS+ study belongs to large kindred with a degree of inherited epilepsy phenotype. Consanguineous marriage is frequent in the cohort. There are several different types of epilepsy including Idiopathic Generalized epilepsy (IGE) and partial epilepsies with motor, sensory or psychiatric implications. The symptoms in patients do not fit into any classification proposed by the International League Against Epilepsy (ILEA) but most similar to GEFS+, which is characterized by large families segregating different types of epilepsy.

Sampling, clinical evaluation and pedigree drawing was complicated with large number of patients and conducted in three visits to village. All visits include several neurologists equipped with a mobile EEG machine, and geneticists. A psychiatrist joined the crew at last visit and evaluated individuals with psychiatric problems. However, clinical evaluation of index patient and his family is conducted in Yeditepe University Hospital.

Blood samples were obtained from affected individuals and family members during first visit and DNA was extracted from collected blood samples by Seda Salar. However no indexing was done during sample collection and bloods were labeled only with patient names which caused sample confusion. Before last visit a large pedigree was drawn using family relation data obtained in previous visits and individuals were indexed with unique ID numbers. Resampling of individuals who segregate disease phenotype was done in last visit using patient ID numbers and confirming family relations on the pedigree drawn. Though not all of targeted individuals were sampled, genomic DNA was isolated from 73 blood samples collected. Reliable DNA samples from previous visits were also used in analysis.

Mutational analysis was carried on GEFS+ index patient for *GABRG2* (Özlem Yalçın Çapan), *SCN1A* (Seda Salar), *SCN1B* and *SCN2A* (M.Aslı Kayserili) genes, by PCR amplification of coding region and exon-intron boundaries followed by direct DNA sequencing. *GABRG2, SCN1A, SCN1B* and *SCN2A* genes were all negative for point mutations. Flowingly Whole Exome Sequencing was conducted for patient GEFS4 (index) and his daughter (GEFS1) to identify any common epilepsy or channel related gene mutations which yielded 112 putative variations. Six variations were selected based on the function and pathologic effect of the mutation and analyzed in other family members. However, two variants observed only in two individuals (*ATP1A* and *BDKRB1*) and four other variants were not inherited together with disease phenotype. Additionally, WES data for individuals GEFS2 and GEFS3 were also obtained 2 years after initial WES. These data were used for the evaluation of the linkage region.

Since no plausible causative variants were identified in WES, Whole Genome Linkage analysis was used to pinpoint disease associated loci. In order to focus on more stringent phenotype 15 individuals from a pedigree showing IGE inheritance was selected and genotyped. However, after analysis with AD and AR models no significant LOD score (above 3) was obtained. In order to improve the power of the pedigree 14 additional family members were also genotyped and analyzed using different inheritance models (AD or AR), penetrance values (70, 80 0r 100%) and also different phenotype assumptions, such as individuals with pure FS are accepted as affected or not. Before linkage analysis Alohomora and PedCheck programs were used to check family relationship and confirms

relation between parents and offspring and markers showing Mendelian inconsistencies were removed from analysis.

In the result of AD linkage analysis a significant LOD score around 3 was found on chromosome 1, a 15MB region on 1p21.3-13.3, containing 47 protein coding protein. Though it was a rather large region interestingly there were 4 rare variants on the region identified in WES data. On the other hand AR whole genome linkage analysis produced slightly significant LOD score above 2 on chromosome 2p11.2-q11.2 and when the region was analyzed in detail LOD score increased to 3,5 on three loci. 10MB region contains centromere and 73 protein coding genes. Also this region includes Immunoglobulin kappa (IGK) locus with several genes. There were 2 rare variants identified in WES data for this region.

Moreover, in order to find homozygous regions shared by affected individuals, hommozygositymapper.org tool was used. In the result of the analysis there were several regions with significant homozygosity scores. On the other hand, especially affected individuals have long stretches of homozygous regions shared by non-affected individuals. Due to this fact 20% of all genotyped SNPs were non-informative in the pedigree. The only variant in the homozygous region was on *SEMA4C* gene which was also present on the AR linkage region.

In the result of linkage analysis 6 variants are selected for further analysis. Among these 4 were exonic and 2 were intronic. All of the selected variants were rare SNPs but two variants *ABCD3* c.685-9T>C and *SEMA4C* p.E598K variants were not present in TUBITAK IBGAM database which consisted of more than 300 human exomes.

ABCD3, ATP-Binding Cassette, Subfamily 3, Member 3, encodes a peroxisomal membrane transporter functioning in transport of branched-chain fatty acids and C27 bile acids. A homozygous truncating mutation was reported for a Turkish female infant with bile acid synthesis defect resulting progressive liver failure due to peroxisomal dysfunction. Her parents were unaffected (OMIM). On the other hand, segregation analysis showed that this variant was present only in one healthy and one affected individual other than the core family. Taking all these fact into account, this variant was accepted as a rare

variant in the family.

SEMA4C p.E598K was another variant present in the AR linkage region but not in LOD score 3 segment. *SEMA4C* gene codes for Semaphorin 4C protein, which belongs to semaphoring family, transmembrane or secreted proteins that appear to function in growth cone guidance of neurons. *SEMA4C* is widely expressed in adult brain and spinal cord (OMIM). On the other hand, identified variant was not observed in Turkish population and had very low MAF (0.0004). When these functions were taken into account this variant was a good candidate, however it did not segregate with the disease phenotype. However, the modifier effect of this variant still needs to be evaluated.

Other variants in the linkage region were either observed only in one individual (*DPYD*), or also observed in healthy member of core family in homozygous state (*INPP4A*) or though present in healthy or affected members, not present in index case (*AGL, COL11A1*). Because of these facts, they were only genotyped in the core family.

Variants in, *AGL, COL11A1* and *INPP4A* genes were observed only in 2nd WES data obtained for individuals 2 and 3 but not in the initial WES data obtained for individuals 1 and 4 though they were present in Sanger sequencing in results. They were considered as false negative for initial WES data and indicated how unreliable this data was. In order to check whether genes implicated in neuronal phenotypes are covered by WES, we analyzed linkage region using IGV tool manually and observed that the region listed in Table 5.22, were not covered in WES analysis.

Among the regions that were not covered in WES analysis one micro RNA gene, *miR137*, was especially important due to its implication in neuropsychiatric disease. So specific primers were designed and miRNA region and surrounding intronic region were analyzed by Sanger sequencing. In the result of analysis no sequence variation was observed however, there was an expansion in a VNTR (rs58335419, -/CGCTGCCGCTGCTAC) just 6 nucleotide before the miRNA coding exon. The VNTR normally found in 3 copies in the reference genome was present in 3 and 5 copies in the index patient. The VNTR was genotyped in all available genomic DNA samples (n=109). The VNTR length were between 3 copies to 12 copies but there were not a significant copy

number association with disease phenotype (LOD score 0.65).

The reason of obtaining no significant LOD score on a linkage region can be attributed to several factors. The main factor may be attributed to computer analysis limitations. In order to get a significant LOD score, 29 individuals are genotyped and BIT score of family was 40, which was difficult for a program to analyze. Due to this fact the pedigree was investigated in four parts and the total LOD score was calculated. Additionally, due to program limitations not all of the genotyped SNPs were used to calculate genome-wide LOD score but rather SNPs were selected in 1cM intervals. In the result we obtained an average LOD score calculated for 4 different family, though LOD score for individual families were variable. When haplotypes with different resolutions are analyzed it is observed the disease haplotype is not inherited similarly in all pedigrees.

On the other hand SNP markers have only two alleles whereas VNTR markers may have several. So ıt is plausible to have difference between multipoint parametric linkage analysis with SNP markers and single point parametric linkage analysis with a VNTR marker.

Analyzed cohort belong to a cultural isolate founded by five families, where endogamy was prevalent. There are advantages and disadvantages of working with a population isolate. Prevalent consanguineous marriages increase the length of LD blocks shared as IBD among affected individuals and they can be detected by less number of SNPs. On the other hand this limits the resolution of the SNP based linkage analyses (Hatzikotoulas, Gilly, and Zeggini, 2014; Sheffield, Stone, and Carmi, 1998). During this study 20% of all genotyped SNPs were uninformative which limits the power around the uninformative genomic stretches. Another disadvantage of working with isolated populations is rare SNPs in other populations might reach significant minor allele frequency just due to presence of this SNP in the inherited haplotype (Sheffield *et al*., 1998).

Another challenge of working isolated groups is the common environmental stimulus they are exposed to. Isolated populations have same nutritional, lifestyle and cultural habits. This may increase the phenocopy rate in the cohort (Hatzikotoulas *et al*., 2014).

Since family based LOD score analysis was uninformative chi square test was applied to analyze association of *miR137* VNTR with disease phenotype. Sampling on field poses diagnostic challenges and contribute to the phenotypic variance in the cohort. Though, affected individuals are diagnosed by a group of neurologist and a psychiatrist, some individuals are sampled due to their family relations and diagnosed as unaffected without examination. In order to prevent conflict samples evaluated both by a neurologist and psychiatrist is included in the statistical analysis (n=86).

In first part of statistical analysis chi square tests were applied and individuals with or without epilepsy phenotype is compared if they have *miR137* expansion or not. We used SPSS Cross Tabulation option which calculated four different chi-square tests and all of them yielded significant difference between epilepsy patients and non-epileptic patients. According to this data having a *miR137* expansion can be considered as a risk factor for epilepsy.

Besides epilepsy depression, anxiety and headache were also prevalent in the cohort. There are reports in the literature where epilepsy is inherited together psychiatric disorders. Especially work by Angelicheva *et al.* (2009) on isolated population of Bulgarian gypsies reports presence of temporal lobe epilepsy together with mood disorders, learning difficulty and alcohol/substance dependence (Angelicheva *et al*., 2009). Taking these facts into account disease phenotype is expanded by including individuals with headache, depression and anxiety as affected. Again we applied chi-square test and compared affected and unaffected individuals in regard of *miR137* VNTR expansion (expanded or non-expanded). In the result all three chi-square tests show significant difference between groups but not as strong as in the epilepsy case ($p<0.01$). This might be attributed to discrepancy between the numbers of affected and unaffected individuals and also to diagnostic difficulties of psychiatric symptoms in isolated communities.

6.3.1. MicroRNA137

miR137 is a microRNA that was first discovered in melanoma cells, (Bemis *et al*. (2008) followed by several studies on cancer cell lines as well as nervous system

development schemes. It is expressed in embryonic and adult brains and enriched in synaptosomes. Its involvement in neurogenesis is investigated in several setups. In a recent study, overexpression of *miR137* was shown to decrease maturation and increase proliferation while antagonism of miR-137 in adult neural stem cells increased neuronal differentiation and reduced proliferation (Szulwach *et al*., 2010). However, in different regions of the brain effect of *miR137* overexpression is vice versa suggesting a fine sub-region *miR137* expression balance is required for neural development (Wright, *et al.*, 2013). Interestingly, homozygous deletion of *miR137* is lethal in mice but mice with heterozygous deletion have the same phenotype and *miR137* as wild type suggesting allelic compensation by single gene (Crowley *et al*., 2014).

In a large GWAS study on schizophrenia patients conducted on 2011 a SNP (rs1625579) on an intron of *miR137* gene had the most significant association score ($P = 1.6 \times 10^{-11}$) and other four loci with genome-wide significant score was on putative targets of *miR137* (Ripke *et al*., 2011). Same finding was repeated in another independent GWAS in schizophrenia patients with different population origin (Ripke *et al*., 2013). Furthermore, several known schizophrenia genes were shown as targets of *miR137* both with electronic tools and with *in vitro* experiments (Wright *et al*., 2013).

Green *et al*. suggested that gene discovery in schizophrenia is hampered by broad definition of the disorder and grouped patients according to their cognitive function. They reported a subgroup of schizophrenia patients who were more likely to be unemployed, had an earlier illness onset, and greater severity of functional disability and negative symptoms. Furthermore, these patients had the risk haplotype on Sz associated SNP on *miR137* (Green *et al*., 2012). This data supports the findings by Willemsen *et al.* that chromosome 1p21.3 microdeletions comprising *DPYD* and *miR137* are associated with intellectual disability in 5 different patients with mild to moderate intellectual disability and also autism spectrum behavior disturbances. They show reduced *miR137* expression in lymphoblastoid cell lines of patients, and also high expression of *miR137* in hippocampus, occipital cortex, and frontal cortex in human post-mortem tissue, as well as in the synaptosomal fractions in mouse brain preparations (Willemsen *et al*., 2011).

Though association of *miR137* with Sz and ID was established many studies failed to find a causative mutation on *miR137* gene. However, in a recent study by Strazisar *et al*. analyzed 13 microRNA genes for sequence variants in patients with Sz and Bipolar disorder. They identified a single nucleotide change 4nt before miRNA sequence and also they analyze VNTR sequence. They found that increase in VNTR as 8 copies decreased the *miR137* expression 0.63 fold in SH-SY5Y cells. Additionally, they analyzed gene expression between SH-SY5Y cell lines transduced by different *miR137* constructs and found dysregulation in genes implicated in synapse organization and biogenesis, transmission of nerve impulse, synaptic transmission and regulation of blood pressure (Strazisar *et al*., 2014).

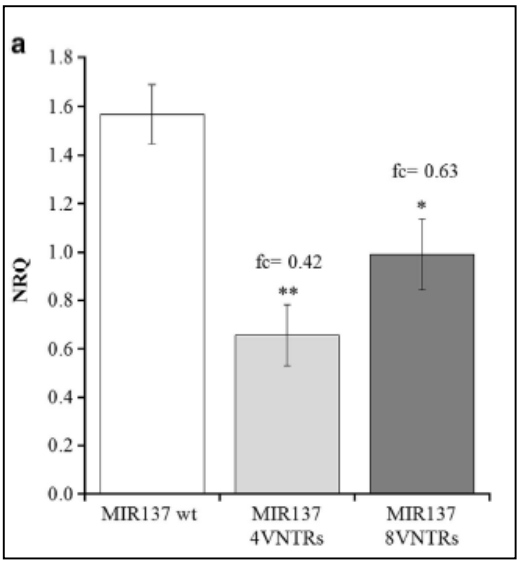


Figure 6.2. Comparison of the mature mRNA levels in SH-SY5Y cell lines transduced with 4copies and 8 copies *miR137* VNTR (Adapted from Strazisar *et al*, 2014).

The association of *miR137* with epilepsy phenotype is not thoroughly investigated though functional analysis including *miR137* overexpression or silencing studies point out epilepsy related genes as listed in the Figure 6.3 (Strazisar *et al.,* 2014) and also in other studies (Collins *et al*., 2014). Furthermore, expression of *miR137* is regulated by the *MECP2* gene and upon deletion of this gene *miR137* expression increases together with an altered methylation pattern in *miR137* promoter region (Szulwach *et al*., 2010). MeCP2 codes for methyl-CpG-binding protein 2, which binds methylated CpGs, is a chromatin-associated protein that can both activate and repress transcription, required for maturation

of neurons and is developmentally regulated (summary by Swanberg and Nagarajan 2009). Mutations in *MECP2* can cause Rett syndrome, mental retardation, or encephalopathy, and have been implicated in autism susceptibility (OMIM).

Results of the GSEA showing the top ranked gene sets and the genes contributing to the core enrichment

GSEA analysis	*C5 enriched*	*Gene set size*	*NES*	*FDR*	*NOM P-value*	*CE genes*
mir-137: 4- and 8VNTR versus 3VNTR	Synapse organization and biogenesis	23	1.834	0.248	< 0.001	ACHE, **NLGN1, PCDHB10, PCDHB13, PCDHB2, PCDHB3, PCDHB5**
	Transmission of nerve impulse	186	1.834	0.159	< 0.001	**ACCN1, AKAP5**, AMIGO1, **APOE, BAIAP3, CACNA1B, CALY, CARTPT, CBLN1**, CD9, **CHRNE, CPLX1, DLG4, DRD2, DRD4, DTNA, GAD1, GLRB, GPR176, GRIA2, GRIK2, GRIN2A, GRM1, HAP1, HRH3, HTR3A, HTR3B, KCNA1, KCNMB1**, KCNMB3, **KCNMB4, KCNN3, KCNQ3, KCNQ5, KLK8, NLGN1, NPBWR2, NQO1, PCDHB10, PCDHB13, PCDHB2, PCDHB3, PCDHB5, PLP1, PMP22, RAB3A, SCN1B, SLC1A1, SLC1A2, SLC6A5, SST, SYN1, SYN2, SYT5**
	Synaptic transmission	171	1.834	0.212	< 0.001	**ACCN1, AKAP5, APOE, BAIAP3, CACNA1B, CALY, CARTPT, CBLN1, CHRNE, CPLX1, DLG4, DRD2, DRD4, DTNA, GAD1, GLRB, GPR176, GRIA2, GRIK2, GRIN2A, GRM1, HAP1, HRH3, HTR3A, HTR3B, KCNA1, KCNMB1, KCNMB4, KCNN3, KCNQ3, KCNQ5, KLK8, NLGN1, NPBWR2, NQO1, PCDHB10, PCDHB13, PCDHB2, PCDHB3, PCDHB5, PLP1, PMP22, RAB3A, SCN1B, SLC1A1, SLC1A2, SLC6A5, SST, SYN1, SYN2, SYT5**
	Regulation of blood pressure	22	1.828	0.137	0.002	ACE, AGT, CALCA, **CARTPT**, GCGR, NPPA, PTGS1, REN

Abbreviations: C5 enriched, name of the top ranked gene sets when based on C5 gene ontology set collections; CE genes, genes that contributed to the enrichment of the specific gene set; FDR: false discovery rate; Gene set size, number of the genes in the gene set; GSEA, gene set enrichment analysis; NES, normalized enrichment score; NOM *P*-value, nominal *P*-value; wt, wild-type. The gene sets met significance threshold of FDR < 0.05 and NOM *P*-value < 0.001 or FDR < 0.25 and NOM *P*-value < 0.01 (when no gene sets met the more stringent threshold). The first column defines the compared microarray data sets. In bold are the genes that contributed to gene set enrichments when comparing microarray results originating from cells transduced with different constructs.

Figure 6.3. Gene Set Enrichment Analysis results for the genes expression of which altered upon transduction of SH-SY5Y cells with 4 or 8 copies of *miR137* VNTR, compared to cells transduced with wild type *miR137* (Adapted from Stratizar *et al.*, 2014).

The association of *miR137* with epilepsy phenotype is not thoroughly investigated though functional analysis including *miR137* overexpression or silencing studies point out epilepsy related genes as listed in the Figure 6.3 (Strazisar *et al.,* 2014) and also in other studies (Collins *et al*., 2014). Furthermore, expression of *miR137* is regulated by the *MECP2* gene and upon deletion of this gene *miR137* expression increases together with an altered methylation pattern in *miR137* promoter region (Szulwach *et al*., 2010). MeCP2 codes for methyl-CpG-binding protein 2, which binds methylated CpGs, is a chromatin-associated protein that can both activate and repress transcription, required for maturation of neurons and is developmentally regulated (summary by Swanberg and Nagarajan 2009). Mutations in *MECP2* can cause Rett syndrome, mental retardation, or encephalopathy, and have been implicated in autism susceptibility (OMIM).

Though the patients analyzed in this study are diagnosed as GEFS+, the exact diagnosis was not established due to phenotypic variance in individuals. Beside IGE

phenotype, especially seizures in partial epilepsy patients are characterized by delusions and hallucinations. Depression is common in individuals with or without epilepsy. Besides, the daughter of the index patient is the only person intelligence testing was applied and she was found to have mild Intellectual Disability. Literature about *miR137* explains features observed especially in partial epilepsy patients which resemble Sz phenotype. On the other hand, literature suggest implication of *miR137* in epilepsy as risk factor it should be further evaluated. Furthermore, VNTR sequence expansion in *miR137* might define a new syndrome with borderline schizophrenia, intellectual disability, seizures and mood disorder.

APPENDIX A: SEQUENCING RESULTS FOR *LGI4* GENE

Table A.1. Genotyping result for *LGI4* gene for the index patient of BFIS pedigree (NM_139284.2 Relative to CDS start.).

Exon	Known SNP in analyzed region	Position	SNP Genotype	Patient Genotype	New Alteration
Exon 1	no	c.-16		G/T	NM_139284.2 Ht. c.-16 G>T In 5' UTR
Exon 2	rs35849647	c.185	G>A	G/G	No
Exon 3	No				No
Exon 4	No				No
Exon 5	rs1673007	c.456	T>C	C/C	
Exon 6	rs1688001	c.495-33	G>A	A/A	
Exon 7	No				No
Exon 8	rs1687998	c.834	G>C	C/C	
	rs77635806	c.960	C>G	C/C	
	rs77357560	c.1180	G>A	G/G	
	rs1319969	c.1203	A>G	G/G	
	rs11617066 1	c.1317	T>G	T/T	
	rs8105771	1299+20	C>G	C/C	
	rs937087	c.1299+52	T>C	C/C	
Exon 9	rs36102542	c.1395_1396	delGCinsAT	GC	
	rs12610234	:c.1395	G>A	GG	
	rs3826989	c.1683	C>T	C/C	
	rs11084800	c.1684	G>A	A/A	
	rs1048690	c.1754	C>T	T/T	

APPENDIX B: SEGREGATION OF *SCN1B, PRRT2* AND *KCNQ2* GENE VARIANTS IN 4BF FAMILY

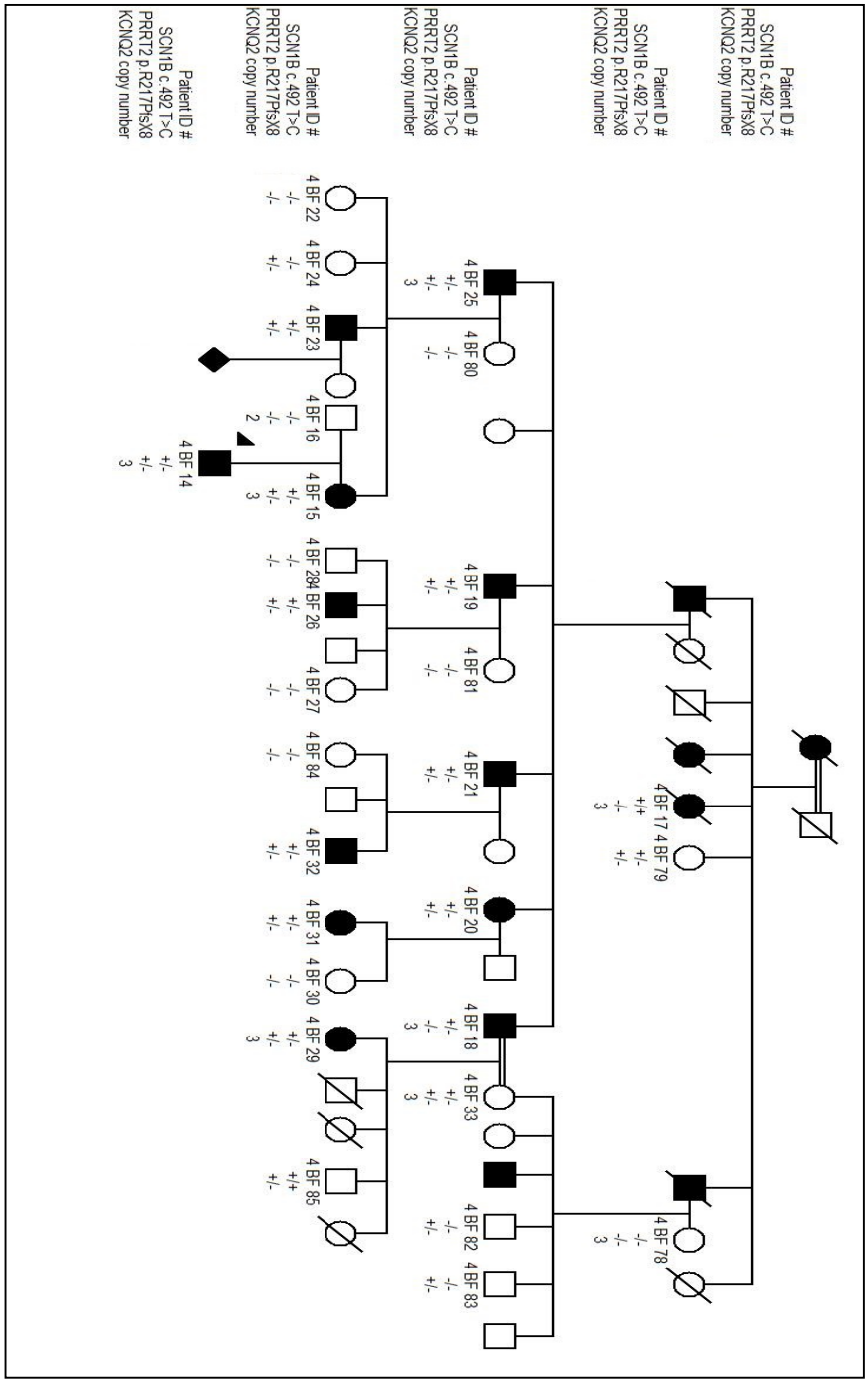


Figure B.1. Segregation of *SCN1B*, *KCNQ2* and *PRRT2* gene alterations in 4BF pedigree. Affected individuals are depicted as black. For each genotype "+" represent mutant and "-" represents wild type allele. For *KCNQ2* copy number gain phenotype, "2" represents 2 copies and "3" represents 3 copies.

APPENDIX C: GEFS+ PEDIGREE SHOWING MIR137 INHERITANCE

In order to facilitate reading, pedigree is printed larger and attached to next page

REFERENCES

Adams, D. R., M. Sincan, K. Fuentes Fajardo, J. C. Mullikin, T. M. Pierson, C. Toro, C. F. Boerkoel, C. J. Tifft, W. a. Gahl, and T. C. Markello, 2012, "Analysis Of DNA Sequence Variants Detected By High-throughput Sequencing", *Human Mutation*, Vol. *33*, No.4, pp.599–608.

Altshuler, D., M. J. Daly, and E. S. Lander, 2008, "Genetic Mapping In Human Disease.", *Science (New York, N.Y.)*, Vol. *322*, No.5903, pp.881–888.

Angelicheva, D., I. Tournev, V. Guergueltcheva, V. Mihaylova, D. N. Azmanov, B. Morar, M. Radionova, S. J. Smith, D. Zlatareva, J. M. Stevens, R. Kaneva, V. Bojinova, K. Carter, M. Brown, A. Jablensky, L. Kalaydjieva, and J. W. Sander, 2009, "Partial Epilepsy Syndrome In A Gypsy Family Linked To 5q31.3-q32", *Epilepsia*, Vol. *50*, No.7, pp.1679–1688.

Audenaert, D., L. Claes, B. Ceulemans, a Löfgren, C. Van Broeckhoven, and P. De Jonghe, 2003, "A Deletion In SCN1B Is Associated With Febrile Seizures And Early-onset Absence Epilepsy.", *Neurology*, Vol. *61*, No.6, pp.854–856.

Audenaert, D., C. Van Broeckhoven, and P. De Jonghe, 2006, "Genes And Loci Involved In Febrile Seizures And Related Epilepsy Syndromes", *Human Mutation*. Vol. 42, No.3, pp. 453-460.

Bamshad, M. J., S. B. Ng, A. W. Bigham, H. K. Tabor, M. J. Emond, D. a. Nickerson, and J. Shendure, 2011, "Exome Sequencing As A Tool For Mendelian Disease Gene Discovery", *Nature Reviews Genetics*, Vol. *12*, No.11, pp.745–755.

Bean, B. P., 2007, "The Action Potential In Mammalian Central Neurons", *Nature Review of Neuroscience*, Vol. *8*, No.6, pp.451–465.

Belhedi, N., F. Bena, A. Mrabet, M. Guipponi, C. B. Souissi, H. K. Mrabet, A. B. Elgaaied, A. Malafosse, and A. Salzmann, 2013, "A New Locus On Chromosome 22q13.31 Linked To Recessive Genetic Epilepsy With Febrile Seizures Plus (GEFS+) In A Tunisian Consanguineous Family.", *BMC Genetics*, Vol. *14*, No.1, pp.93.

Bennett, L. B., E. S. Roach, and a M. Bowcock, 2000, "A Locus For Paroxysmal Kinesigenic Dyskinesia Maps To Human Chromosome 16.", *Neurology*, Vol. *54*, No.1, pp.125–130.

Bermingham, J. R., H. Shearin, J. Pennington, J. O'Moore, M. Jaegle, S. Driegen, A. van Zon, A. Darbas, E. Ozkaynak, E. J. Ryu, J. Milbrandt, and D. Meijer, 2006, "The Claw Paw Mutation Reveals A Role For Lgi4 In Peripheral Nerve Development.", *Nature Neuroscience*, Vol. *9*, No.1, pp.76–84.

Biervert, C., and O. K. Steinlein, 1999, "Structural And Mutational Analysis Of KCNQ2,

The Major Gene Locus For Benign Familial Neonatal Convulsions", *Human Genetics*, Vol. *104*, No.3, pp.234–240.

Brackenbury, W. J., and L. L. Isom, 2011, "Na+ Channel B Subunits: Overachievers Of The Ion Channel Family", *Frontiers in Pharmacology*, Vol. *SEP*, .

Brodie, M. J., S. C. Schachter, and P. Kwan, 2012, "*Fast Facts : Epilepsy*". Health Press Limited.

Caraballo, R. H., R. O. Cersósimo, H. Amartino, P. Szepetowski, and N. Fejerman, 2002, "Benign Familial Infantile Seizures: Further Delineation Of The Syndrome.", *Journal of Child Neurology*, Vol. *17*, No.9, pp.696–699.

Caraballo, R., S. Pavek, a Lemainque, M. Gastaldi, B. Echenne, J. Motte, P. Genton, R. Cersósimo, V. Humbertclaude, N. Fejerman, a P. Monaco, M. G. Lathrop, J. Rochette, and P. Szepetowski, 2001, "Linkage Of Benign Familial Infantile Convulsions To Chromosome 16p12-q12 Suggests Allelism To The Infantile Convulsions And Choreoathetosis Syndrome.", *American Journal of Human Genetics*, Vol. *68*, No.3, pp.788–794.

Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang, and A. R. Krainer, 2003, "ESEfinder: A Web Resource To Identify Exonic Splicing Enhancers", *Nucleic Acids Research*, Vol. *31*, No.13, pp.3568–3571.

Chen, C., R. E. Westenbroek, X. Xu, C. a Edwards, D. R. Sorenson, Y. Chen, … L. L. Isom, 2004, "Mice Lacking Sodium Channel Beta1 Subunits Display Defects In Neuronal Excitability, Sodium Channel Expression, And Nodal Architecture.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. *24*, No.16, pp.4030–4042.

Chen, W.-J., Y. Lin, Z.-Q. Xiong, W. Wei, W. Ni, G.-H. Tan, S.-L. Guo, J. He, Y.-F. Chen, Q.-J. Zhang, H.-F. Li, Y. Lin, S.-X. Murong, J. Xu, N. Wang, and Z.-Y. Wu, 2011, "Exome Sequencing Identifies Truncating Mutations In PRRT2 That Cause Paroxysmal Kinesigenic Dyskinesia", *Nature Genetics*, Vol. *43*, No.12, pp.1252–1255.

Choi, M., U. I. Scholl, W. Ji, T. Liu, I. R. Tikhonova, P. Zumbo, A. Nayir, A. Bakkaloğlu, S. Ozen, S. Sanjad, C. Nelson-Williams, A. Farhi, S. Mane, and R. P. Lifton, 2009, "Genetic Diagnosis By Whole Exome Capture And Massively Parallel DNA Sequencing.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. *106*, No.45, pp.19096–19101.

Collins, a L., Y. Kim, R. J. Bloom, S. N. Kelada, P. Sethupathy, and P. F. Sullivan, 2014, "Transcriptional Targets Of The Schizophrenia Risk Gene MIR137.", *Translational Psychiatry*, Vol. *4*, No.7, pp.e404.

Crowley, J. J., A. L. Collins, R. J. Lee, R. J. Nonneman, M. S. Farrell, N. Ancalade, J. W. Mugford, K. L. Agster, V. D. Nikolova, S. S. Moy, and P. F. Sullivan, 2014, "Disruption Of The MiR-137 Primary Transcript Results In Early Embryonic

Lethality In Mouse", *Biological Psychiatry*.

De Vries, B., P. M. C. Callenbach, J. T. Kamphorst, C. M. Weller, S. C. Koelewijn, R. T. Houten, I. F. M. de Coo, O. F. Brouwer, and a. M. J. M. van den Maagdenberg, 2012, "PRRT2 Mutation Causes Benign Familial Infantile Convulsions", *Neurology*.

Dibbens, L. M., L. a. Harkin, M. Richards, B. L. Hodgson, a. L. Clarke, S. Petrou, I. E. Scheffer, S. F. Berkovic, and J. C. Mulley, 2009, "The Role Of Neuronal GABAA Receptor Subunit Mutations In Idiopathic Generalized Epilepsies", *Neuroscience Letters*, Vol. *453*, No.3, pp.162–165.

Eisenstein, M., 2014, "Unrestrained Excitement", *Nature OUTLOOK*, Vol. *511*, No.7508, pp.S4 S5.

Epstein, M. P., C. D. Veal, R. C. Trembath, J. N. W. N. Barker, C. Li, and G. a Satten, 2005, "Genetic Association Analysis Using Data From Triads And Unrelated Subjects.", *American Journal of Human Genetics*, Vol. *76*, No.4, pp.592–608.

Fendri-Kriaa, N., F. Kammoun, I. H. Salem, C. Kifagi, E. Mkaouar-Rebai, I. Hsairi, a. Rebai, C. Triki, and F. Fakhfakh, 2011, "New Mutation C.374C>T And A Putative Disease-associated Haplotype Within SCN1B Gene In Tunisian Families With Febrile Seizures", *European Journal of Neurology*, Vol. *18*, No.5, pp.695–702.

Fukata, Y., H. Adesnik, T. Iwanaga, D. S. Bredt, R. a Nicoll, and M. Fukata, 2006, "Epilepsy-related Ligand/receptor Complex LGI1 And ADAM22 Regulate Synaptic Transmission.", *Science (New York, N.Y.)*, Vol. *313*, No.5794, pp.1792–1795.

Fukuyama, Y., 1963, "Borderland Of Childhood Epilepsy-special Reference To Febrile Convulsions And So-called Infantile Convulsions", *Seishin Igaku*.

Green, M. J., M. J. Cairns, J. Wu, M. Dragovic, a Jablensky, P. a Tooney, R. J. Scott, and V. J. Carr, 2012, "Genome-wide Supported Variant MIR137 And Severe Negative Symptoms Predict Membership Of An Impaired Cognitive Subtype Of Schizophrenia", *Molecular Psychiatry*.

Guipponi, M., F. Rivier, F. Vigevano, C. Beck, A. Crespel, B. Echenne, P. Lucchini, R. Sebastianelli, M. Baldy-Moulinier, and A. Malafosse, 1997, "Linkage Mapping Of Benign Familial Infantile Convulsions (BFIC) To Chromosome 19q", *Human Molecular Genetics*, Vol. *6*, No.3, pp.473–477.

Gutman, G. a., 2005, "International Union Of Pharmacology. LIII. Nomenclature And Molecular Relationships Of Voltage-Gated Potassium Channels", *Pharmacological Reviews*, Vol. *57*, No.4, pp.473–508.

Haitin, Y., and B. Attali, 2008, "The C-terminus Of Kv7 Channels: A Multifunctional Module.", *The Journal of Physiology*, Vol. *586*, No.7, pp.1803–1810.

Hatzikotoulas, K., a. Gilly, and E. Zeggini, 2014, "Using Population Isolates In Genetic Association Studies", *Briefings in Functional Genomics*, Vol. *13*, No.5, pp.1–7.

Helbig, I., 2014, "New Technologies In Molecular Genetics: The Impact On Epilepsy Research", In O. K. Stenlein (Ed.), *Genetics of Epilepsy* (p. 253). Elsevier B.V.

Helbig, I., and D. H. Lowenstein, 2013, "Genetics Of The Epilepsies: Where Are We And Where Are We Going?", *Current Opinion in Neurology*, Vol. *26*, No.2, pp.179–85.

Heron, S. E., K. Cox, B. E. Grinton, S. M. Zuberi, S. Kivity, Z. Afawi, R. Straussberg, S. F. Berkovic, I. E. Scheffer, and J. C. Mulley, 2007, "Deletions Or Duplications In KCNQ2 Can Cause Benign Familial Neonatal Seizures.", *Journal of Medical Genetics*, Vol. *44*, No.12, pp.791–796.

Heron, S. E., K. M. Crossland, E. Andermann, H. a. Phillips, A. J. Hall, A. Bleasel, M. Shevell, S. Mercho, M. H. Seni, M. C. Guiot, J. C. Mulley, S. F. Berkovic, and I. E. Scheffer, 2002, "Sodium-channel Defects In Benign Familial Neonatal-infantile Seizures", *Lancet*, Vol. *360*, No.9336, pp.851–852.

Heron, S. E., B. E. Grinton, S. Kivity, Z. Afawi, S. M. Zuberi, J. N. Hughes, … L. M. Dibbens, 2012, "PRRT2 Mutations Cause Benign Familial Infantile Epilepsy And Infantile Convulsions With Choreoathetosis Syndrome", *American Journal of Human Genetics*, Vol. *90*, No.1, pp.152–160.

Herranz-Pérez, V., F. E. Olucha-Bordonau, J. M. Morante-Redolat, and J. Pérez-Tur, 2010, "Regional Distribution Of The Leucine-rich Glioma Inactivated (LGI) Gene Family Transcripts In The Adult Mouse Brain", *Brain Research*, Vol. *1307*, pp.177–194.

Hillman, S. C., S. Pretlove, A. Coomarasamy, D. J. McMullan, E. V. Davison, E. R. Maher, and M. D. Kilby, 2010, "Additional Information From Array Comparative Genomic Hybridisation (array CGH) Technology Over Conventional Karyotyping In Prenatal Diagnosis - A Systematic Review And Meta-analysis.", *Ultrasound in Obstetrics & Gynecology : The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, No.July 2010, pp.6–14.

Ishii, A., B. Zhang, S. Kaneko, and S. Hirose, 2010, "Positive Association Between Benign Familial Infantile Convulsions And LGI4", *Brain and Development*, Vol. *32*, No.7, pp.538–543.

Jegla, T. J., C. M. Zmasek, S. Batalov, and S. K. Nayak, 2009, "Evolution Of The Human Ion Channel Set.", *Combinatorial Chemistry & High Throughput Screening*, Vol. *12*, No.1, pp.2–23.

Jensen, C. S., H. B. Rasmussen, and H. Misonou, 2011, "Neuronal Trafficking Of Voltage-gated Potassium Channels", *Molecular and Cellular Neuroscience*, Vol. *48*, No.4, pp.288–297.

Kaplan, R. E., and D. J. Lacey, 1983, "Benign Familial Neonatal-infantile Seizures.", *American Journal of Medical Genetics*, Vol. *16*, No.4, pp.595–599.

Kazen-Gillespie, K. a., D. S. Ragsdale, M. R. D'Andreall, L. N. Mattei, K. E. Rogers, and L. L. Isom, 2000, "Cloning, Localization, And Functional Expression Of Sodium

Channel B1A Subunits", *Journal of Biological Chemistry*, Vol. *275*, No.2, pp.1079–1088.

Kishore, S., A. Khanna, and S. Stamm, 2008, "Rapid Generation Of Splicing Reporters With PSpliceExpress", *Gene*, Vol. *427*, No.1-2, pp.104–110.

Klassen, T., C. Davis, A. Goldman, D. Burgess, T. Chen, D. Wheeler, J. McPherson, T. Bourquin, L. Lewis, D. Villasana, M. Morgan, D. Muzny, R. Gibbs, and J. Noebels, 2011, "Exome Sequencing Of Ion Channel Genes Reveals Complex Profiles Confounding Personal Risk Assessment In Epilepsy", *Cell*, Vol. *145*, No.7, pp.1036–1048.

Korff, C. M., and D. R. Nordli, 2006, "Epilepsy Syndromes In Infancy", *Pediatric Neurology*, Vol. *34*, No.4, pp.253–263.

Kruglyak, L., M. J. Daly, M. P. Reeve-Daly, and E. S. Lander, 1996, "Parametric And Nonparametric Linkage Analysis: A Unified Multipoint Approach.", *American Journal of Human Genetics*, Vol. *58*, No.6, pp.1347–1363.

Lee, H. Y., Y. Huang, N. Bruneau, P. Roll, E. D. O. Roberson, M. Hermann, L. J. Ptáček, 2012, "Mutations In The Gene PRRT2 Cause Paroxysmal Kinesigenic Dyskinesia With Infantile Convulsions", *Cell Reports*, Vol. *1*, No.1, pp.2–12.

Li, M., F. Niu, X. Zhu, X. Wu, N. Shen, X. Peng, and Y. Liu, 2015, "PRRT2 Mutant Leads To Dysfunction Of Glutamate Signaling", *International Journal of Molecular Sciences*, Vol. *16*, No.5, pp.9134–9151.

Li, N., H. Li, H. Jiang, L. Shen, X. Yan, J. Guo, Y. Song, Q. Yang, Y. Wang, X. Li, R. Xiang, X. Zi, X. Long, Z. Hu, Q. Pan, K. Xia, and B. Tang, 2010, "Mutation Detection In Candidate Genes For Benign Familial Infantile Seizures On A Novel Locus.", *The International Journal of Neuroscience*, Vol. *120*, No.3, pp.217–221.

Lindner, T. H., and K. Hoffmann, 2005, "EasyLINKAGE: A PERL Script For Easy And Automated Two-/multi-point Linkage Analyses", *Bioinformatics*, Vol. *21*, No.3, pp.405–407.

Malacarne, M., E. Gennaro, F. Madia, S. Pozzi, D. Vacca, B. Barone, … F. Zara, 2001, "Benign Familial Infantile Convulsions: Mapping Of A Novel Locus On Chromosome 2q24 And Evidence For Genetic Heterogeneity.", *American Journal of Human Genetics*, Vol. *68*, No.6, pp.1521–1526.

Malafosse, a, C. Beck, H. Bellet, M. Di Capua, O. Dulac, B. Echenne, L. Fusco, P. Lucchini, S. Ricci, and R. Sebastianelli, 1994, "Benign Infantile Familial Convulsions Are Not An Allelic Form Of The Benign Familial Neonatal Convulsions Gene.", *Annals of Neurology*, Vol. *35*, No.4, pp.479–482.

Maljevic, S., and H. Lerche, 2014, "Potassium Channel Genes And Benign Familial Neonatal Epilepsy.", In *Progress in brain research*.

Maljevic, S., T. V Wuttke, and H. Lerche, 2008, "Nervous System KV7 Disorders: Breakdown Of A Subthreshold Brake.", *The Journal of Physiology*, Vol. *586*, No.7, pp.1791–1801.

McMullin, D., 2012, "Epi4K: Gene Discovery In 4, 000 Genomes", *Epilepsia*, Vol. *53*, No.8, pp.1457–1467.

Miller, S. a., D. D. Dykes, and H. F. Polesky, 1988, "A Simple Salting Out Procedure For Extracting DNA From Human Nucleated Cells", *Nucleic Acids Research*, Vol. *16*, No.3, pp.1215.

Nakayama, J., 2009, "Progress In Searching For The Febrile Seizure Susceptibility Genes", *Brain and Development*, Vol. *31*, No.5, pp.359–365.

Narain, C., 2014, "Complex Expressions", *Nature OUTLOOK*, Vol. *511*, No.7508, pp.S8 S9.

Nicholl, J., W. Waters, S. Suwalski, S. Brown, Y. Hull, M. G. Harbord, … J. C. Mulley, 2013, "Epilepsy With Cognitive Deficit And Autism Spectrum Disorders: Prospective Diagnosis By Array CGH", *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, Vol. *162*, No.1, pp.24–35.

Nobile, C., and P. Striano, 2014, "PRRT2: A Major Cause Of Infantile Epilepsy And Other Paroxysmal Disorders Of Childhood.", *Progress in Brain Research*, Vol. *213*, pp.141.

O'Connell, J. R., and D. E. Weeks, 1998, "PedCheck: A Program For Identification Of Genotype Incompatibilities In Linkage Analysis.", *American Journal of Human Genetics*, Vol. *63*, No.1, pp.259–266.

Otto, J. F., N. A. Singh, E. J. Dahle, M. F. Leppert, C. M. Pappas, T. H. Pruess, K. S. Wilcox, and H. S. White, 2009, "Electroconvulsive Seizure Thresholds And Kindling Acquisition Rates Are Altered In Mouse Models Of Human KCNQ2 And KCNQ3 Mutations For Benign Familial Neonatal Convulsions", *Epilepsia*, Vol. *50*, No.7, pp.1752–1759.

Patino, G. a, L. R. F. Claes, L. F. Lopez-Santiago, E. a Slat, R. S. R. Dondeti, C. Chen, H. a O'Malley, C. B. B. Gray, H. Miyazaki, N. Nukina, F. Oyama, P. De Jonghe, and L. L. Isom, 2009, "A Functional Null Mutation Of SCN1B In A Patient With Dravet Syndrome.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. *29*, No.34, pp.10764–10778.

Patino, G. a., W. J. Brackenbury, Y. Bao, L. F. Lopez-Santiago, H. a. O'Malley, C. Chen, J. D. Calhoun, R. G. Lafreniere, P. Cossette, G. a. Rouleau, and L. L. Isom, 2011, "Voltage-Gated Na+ Channel Beta 1B : A Secreted Cell Adhesion Molecule Involved In Human Epilepsy", Vol. *31*, No.41, pp.14577–14591.

Patino, G. a., and L. L. Isom, 2010, "Electrophysiology And Beyond: Multiple Roles Of Na+ Channel B Subunits In Development And Disease", *Neuroscience Letters*, Vol. *486*, No.2, pp.53–59.

Pinkel, D., R. Segraves, D. Sudar, S. Clark, I. Poole, D. Kowbel, C. Collins, W. L. Kuo, C. Chen, Y. Zhai, S. H. Dairkee, B. M. Ljung, J. W. Gray, and D. G. Albertson, 1998, "High Resolution Analysis Of DNA Copy Number Variation Using Comparative Genomic Hybridization To Microarrays.", *Nature Genetics*, Vol. *20*, No.2, pp.207–11.

Qin, N., M. R. D'Andrea, M. Lou Lubin, N. Shafaee, E. E. Codd, and A. M. Correa, 2003, "Molecular Cloning And Functional Expression Of The Human Sodium Channel ??1B Subunit, A Novel Splicing Variant Of The ??1 Subunit", *European Journal of Biochemistry*, Vol. *270*, No.23, pp.4762–4770.

Reid, C. A., S. F. Berkovic, and S. Petrou, 2009, "Mechanisms Of Human Inherited Epilepsies", *Progress in Neurobiology*, Vol. *87*, No.1, pp.41–57.

Ripke, S., C. O'Dushlaine, K. Chambert, J. L. Moran, A. K. Kähler, S. Akterin, … P. F. Sullivan, 2013, "Genome-wide Association Analysis Identifies 13 New Risk Loci For Schizophrenia.", *Nature Genetics*, Vol. *45*, No.10, pp.1150–9.

Ripke, S., A. R. Sanders, K. S. Kendler, D. F. Levinson, P. Sklar, P. A. Holmans, … P. V. Gejman, 2011, "Genome-wide Association Study Identifies Five New Schizophrenia Loci [Letter]", Vol. *43*, No.10.

Robbins, J., G. M. Passmore, F. C. Abogadie, J. M. Reilly, and D. A. Brown, 2013, "Effects Of KCNQ2 Gene Truncation On M-Type Kv7 Potassium Currents", *PLoS ONE*, Vol. *8*, No.8.

Rüschendorf, F., and P. Nürnberg, 2005, "ALOHOMORA: A Tool For Linkage Analyisi Using 10K SNP Array Data", *Bioinformatics*, Vol. *21*, No.9, pp.2123–2125.

Savage, N., 2014, "Epidemiology: The Complexities Of Epilepsy", *Nature*, Vol. *511*, No.7508, pp.S2–S3.

Scheffer, I. E., and S. F. Berkovic, 1997, "Generalized Epilepsy With Febrile Seizures Plus. A Genetic Disorder With Heterogeneous Clinical Phenotypes", *Brain*, Vol. *120*, No.3, pp.479–490.

Scheffer, I. E., L. A. Harkin, L. M. Dibbens, J. C. Mulley, and S. F. Berkovic, 2005, "Neonatal Epilepsy Syndromes And Generalized Epilepsy With Febrile Seizures Plus (GEFS+)", In *Epilepsia* (Vol. 46, pp. 41–47). http://doi.org/10.1111/j.1528-1167.2005.00358.x

Scheffer, I. E., L. a. Harkin, B. E. Grinton, L. M. Dibbens, S. J. Turner, M. a. Zielinski, R. Xu, G. Jackson, J. Adams, M. Connellan, S. Petrou, R. M. Wellard, R. S. Briellmann, R. H. Wallace, J. C. Mulley, and S. F. Berkovic, 2007, "Temporal Lobe Epilepsy And GEFS+ Phenotypes Associated With SCN1B Mutations", *Brain*, Vol. *130*, No.1, pp.100–109.

Schubert, J., R. Paravidino, F. Becker, A. Berger, N. Bebek, A. Bianchi, … Y. G. Weber, 2012, "PRRT2 Mutations Are The Major Cause Of Benign Familial Infantile

Seizures", *Human Mutation*, Vol. *33*, No.10, pp.1439–1443.

Seelow, D., J. M. Schwarz, and M. Schuelke, 2008, "GeneDistiller--distilling Candidate Genes From Linkage Intervals.", *PloS One*, Vol. *3*, No.12, pp.e3874.

Sheffield, V. C., E. M. Stone, and R. Carmi, 1998, "Use Of Isolated Inbred Human Populations For Identification Of Disease Genes", *Trends in Genetics*, Vol. *14*, No.10, pp.391–396.

Shieh, C. C., M. Coghlan, J. P. Sullivan, and M. Gopalakrishnan, 2000, "Potassium Channels: Molecular Defects, Diseases, And Therapeutic Opportunities.", *Pharmacological Reviews*, Vol. *52*, No.4, pp.557–594.

Singh, N. A., J. F. Otto, E. J. Dahle, C. Pappas, J. D. Leslie, A. Vilaythong, J. L. Noebels, H. S. White, K. S. Wilcox, and M. F. Leppert, 2008, "Mouse Models Of Human KCNQ2 And KCNQ3 Mutations For Benign Familial Neonatal Convulsions Show Seizures And Neuronal Plasticity Without Synaptic Reorganization.", *The Journal of Physiology*, Vol. *586*, No.14, pp.3405–3423.

Siva, N., 2008, "1000 Genomes Project.", *Nature Biotechnology*, Vol. *26*, No.3, pp.256.

Smith, P., C. Zhang, and J. Wang, 2006, "An Increased Specificity Score Matrix For The Prediction Of SF2/ASF-specific Exonic Splicing Enhancers", *Human Molecular ….*

Sobel, E., J. C. Papp, and K. Lange, 2002, "Detection And Integration Of Genotyping Errors In Statistical Genetics.", *American Journal of Human Genetics*, Vol. *70*, No.2, pp.496–508.

Specchio, N., and F. Vigevano, 2006, "The Spectrum Of Benign Infantile Seizures.", *Epilepsy Research*, Vol. *70 Suppl 1*, pp.S156–67.

Stafstrom, C., 2008, "Traffic Jam At The Sodium Channel", *Epilepsy Currents*.

Steinlein, O., 2014, "Mechanisms Underlying Epilepsies Associated With Sodium Channel Mutations.", In *Genetics of Epilepsy*. Elsevier.

Strazisar, M., S. Cammaerts, K. van der Ven, D. a Forero, a-S. Lenaerts, a Nordin, L. Almeida-Souza, G. Genovese, V. Timmerman, a Liekens, P. De Rijk, R. Adolfsson, P. Callaerts, and J. Del-Favero, 2014, "MIR137 Variants Identified In Psychiatric Patients Affect Synaptogenesis And Neuronal Transmission Gene Sets.", *Molecular Psychiatry*, No.October 2013, pp.1–10.

Striano, P., L. Bordo, M. L. Lispi, N. Specchio, C. Minetti, F. Vigevano, and F. Zara, 2006, "A Novel SCN2A Mutation In Family With Benign Familial Infantile Seizures.", *Epilepsia*, Vol. *47*, No.1, pp.218–20.

Swanberg, S. E., R. P. Nagarajan, S. Peddada, D. H. Yasui, and J. M. Lasalle, 2009, "Reciprocal Co-regulation Of EGR2 And MECP2 Is Disrupted In Rett Syndrome And Autism", *Human Molecular Genetics*, Vol. *18*, No.3, pp.525–534.

Szulwach, K. E., X. Li, R. D. Smrt, Y. Li, Y. Luo, L. Lin, N. J. Santistevan, W. Li, X. Zhao, and P. Jin, 2010, "Cross Talk Between MicroRNA And Epigenetic Regulation In Adult Neurogenesis", *Journal of Cell Biology*, Vol. *189*, No.1, pp.127–141.

Tomita, H. a, S. Nagamitsu, K. Wakui, Y. Fukushima, K. Yamada, M. Sadamatsu, … N. Niikawa, 1999, "Paroxysmal Kinesigenic Choreoathetosis Locus Maps To Chromosome 16p11.2-q12.1.", *American Journal of Human Genetics*, Vol. *65*, No.6, pp.1688–1697.

Tzingounis, A. V, and R. A. Nicoll, 2008, "Contribution Of KCNQ2 And KCNQ3 To The Medium And Slow Afterhyperpolarization Currents.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. *105*, No.50, pp.19974–19979.

Vigevano, F., 2005, "Benign Familial Infantile Seizures.", *Brain & Development*, Vol. *27*, No.3, pp.172–7.

Vigevano, F., L. Fusco, M. Di Capua, S. Ricci, R. Sebastianelli, and P. Lucchini, 1992, "Benign Infantile Familial Convulsions", *European Journal of Pediatrics*, Vol. *151*, No.8, pp.608–612.

Wallace, R. H., D. W. Wang, R. Singh, I. E. Scheffer, A. L. George, H. A. Phillips, K. Saar, A. Reis, E. W. Johnson, G. R. Sutherland, S. F. Berkovic, and J. C. Mulley, 1998, "Febrile Seizures And Generalized Epilepsy Associated With A Mutation In The Na+-channel Beta1 Subunit Gene SCN1B.", *Nature Genetics*, Vol. *19*, No.4, pp.366–70.

Wallace, R., I. Scheffer, and G. Parasivam, 2002, "Generalized Epilepsy With Febrile Seizures Plus: Mutation Of The Sodium Channel Subunit SCN1B", *Neurology*.

Wang, J. L., L. Cao, X. H. Li, Z. M. Hu, J. Da Li, J. G. Zhang, … B. S. Tang, 2011, "Identification Of PRRT2 As The Causative Gene Of Paroxysmal Kinesigenic Dyskinesias", *Brain*, Vol. *134*, No.12, pp.3490–3498.

Watanabe, H., D. Darbar, D. W. Kaiser, K. Jiramongkolchai, S. Chopra, B. S. Donahue, P. J. Kannankeril, and D. M. Roden, 2009, "Mutations In Sodium Channel B1- And B2-Subunits Associated With Atrial Fibrillation", *Circulation: Arrhythmia and Electrophysiology*, Vol. *2*, No.3, pp.268–275.

Watanabe, H., E. Nagata, A. Kosakai, M. Nakamura, M. Yokoyama, K. Tanaka, and H. Sasai, 2000, "Disruption Of The Epilepsy KCNQ2 Gene Results In Neural Hyperexcitability", *Journal of Neurochemistry*, Vol. *75*, No.1, pp.28–33.

Weber, Y. G., and H. Lerche, 2008, "Genetic Mechanisms In Idiopathic Epilepsies", *Developmental Medicine and Child Neurology*, Vol. *50*, No.9, pp.648–654.

Weckhuysen, S., S. Mandelstam, A. Suls, D. Audenaert, T. Deconinck, L. R. F. Claes, … P. De Jonghe, 2012, "KCNQ2 Encephalopathy: Emerging Phenotype Of A Neonatal Epileptic Encephalopathy", *Annals of Neurology*, Vol. *71*, No.1, pp.15–25.

Willemsen, M. H., a. Valles, L. a. M. H. Kirkels, M. Mastebroek, N. Olde Loohuis, a. Kos, … T. Kleefstra, 2011, "Chromosome 1p21.3 Microdeletions Comprising DPYD And MIR137 Are Associated With Intellectual Disability", *Journal of Medical Genetics*, Vol. *48*, No.12, pp.810–818.

Wright, C., J. a Turner, V. D. Calhoun, and N. Perrone-Bizzozero, 2013, "Potential Impact Of MiR-137 And Its Targets In Schizophrenia.", *Frontiers in Genetics*, Vol. *4*, No.April, pp.58.

Zara, F., N. Specchio, P. Striano, A. Robbiano, E. Gennaro, R. Paravidino, C. Minetti, 2013, "Genetic Testing In Benign Familial Epilepsies Of The First Year Of Life: Clinical And Diagnostic Significance", *Epilepsia*, Vol. *54*, No.3, pp.425–436.

Zhou, X., A. Ma, X. Liu, C. Huang, Y. Zhang, R. Shi, S. Mao, T. Geng, and S. Li, 2006, "Infantile Seizures And Other Epileptic Phenotypes In A Chinese Family With A Missense Mutation Of KCNQ2.", *European Journal of Pediatrics*, Vol. *165*, No.10, pp.691–5.