MUTATIONAL ANALYSIS OF THE *ARX* GENE AND WHOLE EXOME SEQUENCING IN EPILEPTIC ENCEPHALOPATHY PATIENTS

by

Hande Özünlü

B.S., Molecular Biology and Genetics, Istanbul Technical University, 2012

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

Graduate Program in Molecular Biology and Genetics
Boğaziçi University

2019

MUTATIONAL ANALYSIS OF THE *ARX* GENE AND WHOLE EXOME SEQUENCING IN EPILEPTIC ENCEPHALOPATHY PATIENTS

APPROVED BY:

Prof. Dr. Esra Battaloğlu (Thesis Supervisor)	……
Prof. Dr. Hande Çağlayan (Thesis Co-Supervisor)	……
Assoc. Prof. İbrahim Yaman	……
Assist. Prof. Özlem Yalçın Çapan	……

DATE OF APPROVAL: 23.07.2019

To my beloved family …

ACKNOWLEDGEMENTS

Before anyone else, I would like to thank to my former supervisor Prof. Dr. Hande Çağlayan for her valuable insights and directions that have been my guidance during this M.Sc. thesis. To gain experience at various branches, and to expand my scientific knowledge beyond what I could imagine are great opportunities she provided. I am therefore grateful to her for being a mentor in academic life. Also, I would like to thank to my current supervisor Prof. Dr. Esra Battaloğlu for accepting me for the last term of my thesis journey and to guide me with her knowledge even it was a short mentoring period. I would also like to thank to Assoc. Prof. İbrahim Yaman and Assist. Prof. Özlem Yalçin Çapan for evaluating my thesis and guiding me to improve it. In addition, I would like to thank Elif Everest for being a wonderful friend, for her great support, for sharing her knowledge and helping me every time I needed. I would also like to thank Nalan Yildiz for her friendship and support during my master's programme and İlker Karacan for sharing his own experiences and knowledge for my thesis project. I am also grateful for Aslı Gündoğdu and Suna Usluer for sharing their knowledge and helping me during my thesis. Lastly, I am deeply grateful to my family for giving me the chance to experience this master's programme, for their endless love and support.

ABSTRACT

MUTATIONAL ANALYSIS OF THE ARX GENE AND WHOLE EXOME SEQUENCING IN EPILEPTIC ENCEPHALOPATHY PATIENTS

The term epileptic encephalopathy (EE) is used to describe epileptic syndromes with generally early onset, within the first year of life, refractory to drug treatment and poor developmental outcome. The genetics of epileptic encephalopathy disorders can be monogenic or complex. *ARX* gene mutations were observed in epileptic encephalopathy phenotypes such as Ohtahara and West Syndrome. The most common mutation seen in *ARX* gene is the polyalanine tract expansion, which is due to duplication of a 24 base-pair polyalanine tract that increases the alanine number to 20. *ARX* is a homeobox-containing gene encoding for Homeobox protein ARX which has roles in normal brain development. The polyalanine expansion in the *ARX* gene results in intellectual disability. Epileptic encephalopathy diagnosis cannot be solely made based on clinical criteria, and *ARX* gene mutations and repeat expansions are considered as an important parameter for both diagnosis and treatment. The aim of this study was to explore the genetic basis of epileptic encephalopathy. Genetic causes of EE are heterogeneous; hence, whole exome sequencing (WES) is necessary to reveal the causative variants. However, repeat expansions cannot be detected by WES technique. Therefore, initially *ARX* gene mutations and repeat expansions in epilepsy patients with EE were analysed by Agilent and Sanger sequencing in patients with *ARX*-related phenotypes. In addition, WES was performed in 3 patients with EE and without *ARX* mutations to reveal pathological variants, and candidate variants were determined. For *ARX* repeat analysis, exon 2.1 part with the repeat region of 16 patients that present *ARX*-related phenotype was screened using Agilent system, revealing no significant repeat number increase in the patients. Other exon regions were analysed by PCR and Sanger sequencing for the 16 patients. Analyses of Sanger sequencing data using FinchTV software revealed no mutations in the exons of *ARX* gene. WES was then performed in 3 out of 16 EE patients with *ARX*-related disease phenotype. After family segregation analysis, *CUL4B* and *MAP2* variants were selected as the strongest candidates responsible for EE phenotype in two patients.

ÖZET

EPİLEPTİK ENCEFALOPATİ HASTALARINDA ARX GENİNİN TÜM EKZOM DİZİLENMESİ VE MUTASYON ANALİZİ

Epileptik ensefalopati terimi; genellikle erken başlangıçlı (doğumdan sonraki ilk yıl içerisinde), tedaviye cevap vermeyen ve zayıf gelişim ile sonuçlanan epileptik sendromları tanımlamaktadır. EE, monogenik veya kompleks kalıtımlı olabilmektedir. *ARX*, "homeobox"-içeren bir gen olup, Homeobox protein ARX'ı kodlamaktadır ve normal beyin gelişiminde görev almaktadır. *ARX* gen mutasyonları; Ohtarahara ve West sendromu gibi EE fenotiplerinde rapor edilmiştir. En sık rastlanan *ARX* mutasyonu, 24 baz-çiftinden oluşan ve alanin sayısını 20'ye çıkaran polialenin bölge genişlemesidir. Bu polialenin genişlemesi, ayrıca zeka geriliğine yol açmaktadır. EE tanısı yalnızca klinik kriterlere göre yapılamamakta olup, *ARX* gen mutasyonlarının ve tekrar artışlarının hem tanı hem de tedavi için önemli parametreler olduğu düşünülmektedir. Bu çalışmanın amacı, epileptik ensefalopatinin genetik temelinin araştırılmasıdır. EE'nin genetik sebepleri heterojen olduğundan, hastalıktan sorumlu varyantların saptanabilmesi için tüm ekzom dizileme yöntemi gerekmektedir. Fakat, tekrar artışları bu yöntemle tespit edilememektedir. Bu nedenle, ilk olarak *ARX* gen mutasyonları taşıyabileceği düşünülen fenotipik özelliklere sahip EE hastalarında *ARX* mutasyonları ve tekrar artışları Agilent ve Sanger dizileme yöntemleriyle analiz edilmiştir. Ek olarak, *ARX* mutasyonu bulunmayan 3 EE hastasında ekzom dizileme yapılarak patolojik aday varyantlar tanımlanmıştır. *ARX* tekrar analizleri için tekrar bölgesini içeren ekzon 2.1 bölgesi, seçilen 16 EE hastasında Agilent sistemi kullanılarak analiz edilmiş, fakat anlamlı ölçüde baz artışı gözlenmemiştir. Diğer ekzon bölgeleri, PZR ve Sanger dizileme ile bu 16 hastada analiz edilmiş, ancak verilerin FinchTV yazılımı ile analizleri sonucunda *ARX* gen mutasyonu tespit edilmemiştir. Bunun üzerine, *ARX*-ilişkili hastalık fenotipine sahip 16 hastadan 3'ünde tüm ekzom dizileme gerçekleştirilmiştir. Ailesel segregasyon analizlerinin ardından *CUL4B* ve *MAP2* gen varyantları 2 hastada EE için en güçlü aday varyantlar olarak belirlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS /ABBREVIATIONS

A	Adenine
ARX	Aristaless related homeobox
bp	base pair
C	Cytosine
chr	Chromosome
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
EE	Epileptic Encephalopathy
EEG	Electroencephalography
EtBr	Ethidium Bromide
G	Guanine
GABA	gamma-aminobutyric acid
GWAS	Genome Wide Association Studies
ILAE	International League Against Epilepsy
LGS	Lennox-Gastaut Syndrome
$MgCl_2$	Magnesium Cloride
MGI	Mouse Genome Informatics
OH	Ohtahara Syndrome
OMIM	Online Mendelian Inheritance in Man
PA	Polyalanine tracks
PCR	Polymerase Chain Reaction
SB	Suppression-Burst
SCN	Sodium Channel
SNP	Single Nucleotide Polymorphism
T	Thymine
TBE	Tris-Borate-EDTA
UV	Ultraviole
V	Voltage
WES	Whole Exome Sequencing

1. INTRODUCTION

1.1. Epilepsy

Epilepsy is a group of neurological disorders characterized by unprovoked, spontaneous, recurrent epileptic seizures. An epileptic seizure is defined by International League Against Epilepsy (ILAE) as “a transient occurrence of signs and/or symptoms due to abnormal excessive and synchronous neuronal activity in the brain” [1]. The prevalence of epilepsy is 1% of the general population [2]. Environmental effects such as dementia, stroke, brain tumour can contribute to epilepsy phenotype, whereas genetic abnormalities such as causing ion channel deficiencies are the main causes of epileptic seizures/phenotype [3]. Proper classification of epilepsies is crucial for accurate diagnosis and related proper clinical consultation, and for development of novel antiepileptic therapies. According to ILAE classification of the epilepsies [4], the classification depends on three levels: seizure type, epilepsy type, and epilepsy syndrome (Figure 1.1). When available, all three information should be interpreted together, as well as aetiology of the specific epilepsy case. Seizure types are classified into three: focal onset, generalized onset, and unknown onset, with subcategories of motor, non-motor, with retained or impaired awareness for focal seizures [5]. The second level classification, epilepsy type, has four categories: generalised epilepsy, focal epilepsies, combined generalised and focal epilepsy, and unknown category [6, 4]. Seizure and epilepsy types are determined based on electroencephalography (EEG), neuroimaging, and other indicators of aetiology of the epilepsy including structural, genetic, infectious, metabolic, and immune aetiologies. The last level of epilepsy classification is epilepsy syndrome diagnosis; which combines all the information available to make a definite diagnosis, including also information on such as age of onset, seizure triggers, and prognosis [7].

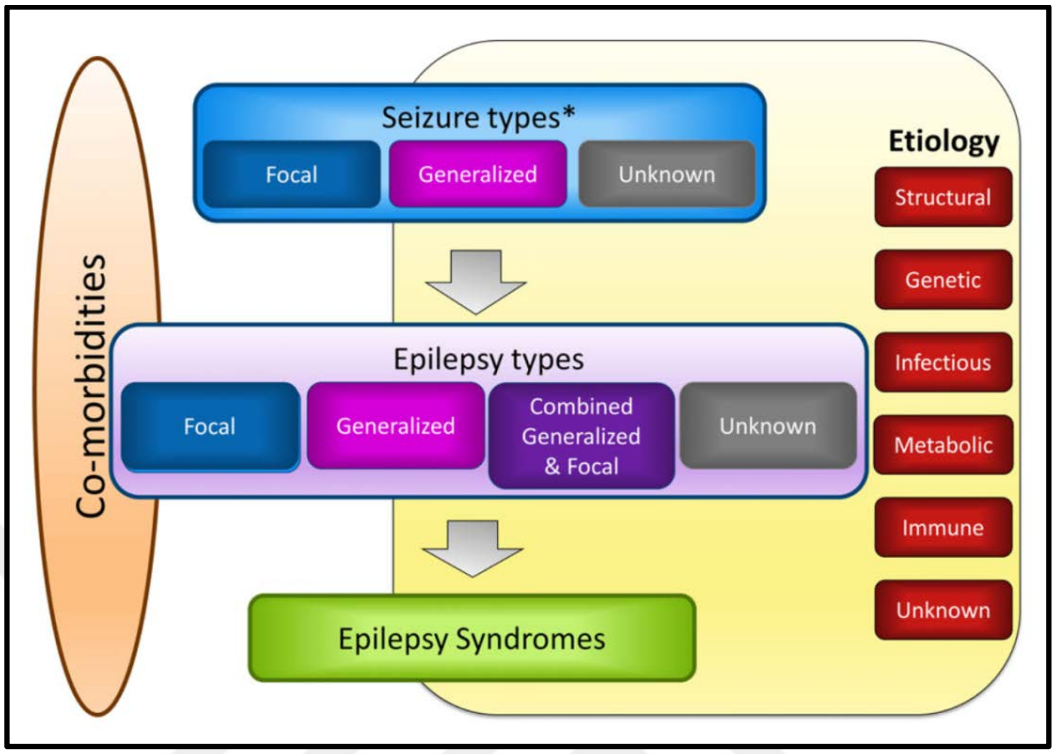

Figure 1.1. ILAE classification of the epilepsies [4].

1.2. Epileptic Encephalopathy

Epileptic encephalopathy (EE) is a severe epilepsy syndrome characterized by cerebral disturbance due to epileptic seizures [8]. The terminology of “encephalo” means brain, and “pathy” means disorder. Epileptic encephalopathy together refers to brain disorder characterized by epileptic seizures. EE is a heterogeneous group of epilepsy syndromes with progressive cognitive and behavioural disturbances. The term epileptic encephalopathy is used to describe epileptic syndromes with generally early onset, within the first year of life, refractory to drug treatment and poor developmental outcome [8]. In March 2017, ILAE defined a new term as “developmental and epileptic encephalopathy, DEE”; because the mutations that cause epileptic encephalopathies also cause development dysfunctions such as psychomotor impairment, language developmental delay or mental development may slow down [9].

1.2.1. Epilepsies Associated with Developmental and Epileptic Encephalopathy

Epileptic encephalopathy is associated with many epilepsy syndromes such as Dravet, Ohtahara, West, Lennox-Gastaut Syndrome, Early Myoclonic Encephalopathy and Landau-Kleffner Syndrome. Dravet Syndrome is one of the common epileptic encephalopathy disorder which is defined by neurological development disturbance. The common cause of Dravet Syndrome is Sodium channel, voltage-gated, type I, alpha subunit (*SCN1A*) mutations that are defined as channelopathy. The onset is generally in the first years of life, mostly while at 5 or 8 months [10]. Ohtahara Syndrome (OS) is considered as one of the earliest epileptic encephalopathies because the onset is generally neonatal. The syndrome is also called as Early Infantile Epileptic Encephalopathy with burst suppression since it has a unique EEG pattern which is named as the suppression-burst pattern (SB) [11]. The pattern is defined as “high voltage bursts alternating with almost flat suppression phases”. OS has frequent seizures, and EEG abnormalities are severe and continuous, which result in mental dysfunctions [12]. Genes related with OS are Aristaless Related Homeobox gene (*ARX*), *CDKL5*, *SLC25A22* and *STXBP1*genes [13]. West Syndrome (infantile spasms), is a severe epileptic disorder which is described by myoclonic tonic spasms (seizures) that occur during the first year of life. West Syndrome develops mainly due to *ARX* mutations that result in loss of interneurons in mouse models [14]. Lennox-Gastaut Syndrome (LGS) is a severe epileptic encephalopathy syndrome with developmental delays and behavioural and cognitive dysfunctions. The onset is generally between 3 to 8 years. Lennox-Gastaut Syndrome has several seizure types and abnormal EEG [15]. Thirty percent of LGS causes are unknown [16]. Early Myoclonic Encephalopathy is the other early-onset epileptic encephalopathy which generally develops in new-borns. It has specific seizure types and EEG abnormalities that differ from Ohtahara. Identified associated genes so far are *ERBB4, SIK1, SLC25A22* and a GABAA receptor subunit gene *GABRB2* which has a relation with GABAergic inhibition dysfunction in the brain [17]. Landau-Kleffner Syndrome is a rare epileptic encephalopathy with onset of 3 to 9 years. The seizures commonly result in language regression. 100% recovery can be achieved in some patients. *GRIN2A* mutations have been associated with Landau-Kleffner Syndrome in recent studies [18].

1.3. Genetics of Epileptic Encephalopathy

The genetics of epileptic encephalopathy disorders can be monogenic or complex inheritance. Genetic diagnosis approaches have increased over the past 10 years. The genetic background of diseases as well as epileptic disorders can be studied with next-generation sequencing techniques such as whole exome sequencing (WES) in combination with traditional approaches such as linkage analysis [19]. All studies in combination indicate that there is a complex heterogeneity in epileptic encephalopathy where different gene mutations can cause the same phenotype [19]. There are 71 genes associated with early infantile EE in Online Mendelian Inheritance in Man (OMIM) database [20]. Most commonly observed mutations are present on the voltage-gated sodium channels which are encoded by SCN family and controls the electrical communication between neurons and potassium, calcium channel genes. Gamma amino butyric acid (GABA) receptor family mutations also results in GABAergic interneuron deficiency [21].

ARX is neither a channel protein-coding gene nor a GABA-related gene, yet the *ARX* gene mutations were observed in epileptic encephalopathy phenotypes such as Ohtahara, West and Lennox Gastaut Syndrome [19].

1.3.1. ARX Gene

ARX gene is aristaless related X-linked homeobox DNA binding transcription factor. It is mainly synthesized in the forebrain playing an important role in embryogenesis of central nervous system [22] and essential for forebrain development [23]. It is the first identified gene in epileptic encephalopathy patients [24]. *ARX* gene is located on Xp22 and is a small gene which contains 5 exons (Fig.1.2).

Figure 1.2. *ARX* gene exons [29].

The mutations of *ARX* gene are associated with disorders such as infantile spasms (WEST syndrome), Partington syndrome, X-linked myoclonic epilepsy and X-linked mental retardation [23]. ARX gene plays an important role in neuronal development and GABAergic interneuron migration during development stage [25]. GABAergic neurons are the chief inhibitory neurons in the brain and control the electrical activity by secreting inhibitory neurotransmitters. The impairment of *ARX* gene results in dysfunction of GABAergic interneuron signalling and causes epileptic phenotypes and can also cause brain malformation phenotypes such as lissencephaly or agenesis of corpus callosum as per the mutation type and location in *ARX* gene [26].

The most common mutation seen in *ARX* gene is the polyalanine tract expansion, which is due to duplication of 24 base-pair polyalanine tract and increases the alanine number to 20 [27]. These polyalanine expansion in the *ARX* gene results in intellectual disability and the *ARX* was demonstrated as the most frequent reason for these phenotypic disorders [26].

ARX gene has 4 polyalanine tracts (PA) in the second exon region. The most common mutant polyalanine tracts are c.428-451dup in second PA tract which expands alanine residues from 12 to 20 and c.333-334ins(GCG)7 in first polyalanine tract leading to 23 alanine instead 16, and they both are associated with WEST syndrome [28].

There is a correlation reported between the ARX mutations and disorders. Missense, nonsense mutation in homeodomain and frameshift mutations mostly lead to more severe, malformative types such as XLAG or PROUD syndrome, whereas the non-malformative types such as Ohtahara and West are due to mostly poly-alanine tract expansions, as shown in Figure 1.3. As per the Figure 1.3, severity of diseases increase from bottom to top and the mutations were coloured same with their related phenotypes.

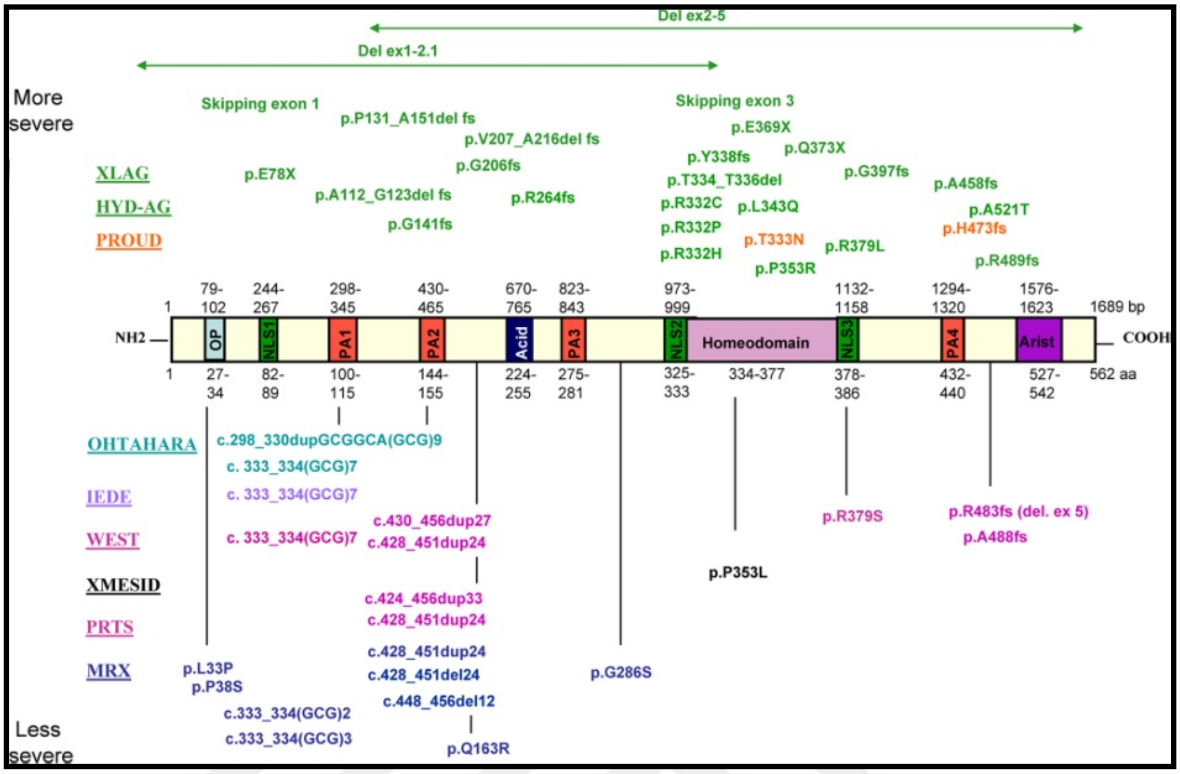

Figure 1.3. Phenotype correlations of *ARX* mutations [30].

1.4. Aim of the Study

Epileptic encephalopathy diagnosis cannot solely be made based on clinical criteria, and *ARX* gene mutations and repeat expansions are considered as an important parameter for both diagnosis and treatment. Current diagnosis relies on whole exome sequencing. The aim of this study is to initially analyse *ARX* gene for point mutations and repeat expansions in epileptic encephalopathy patients. To this end, exon regions of *ARX* gene were screened by Sanger sequencing in 16 epilepsy patients with *ARX*-related phenotypes, except for the first part of the exon2, containing the poly alanine tracts of *ARX*. Due to high GC region in polyalanine tracks in this exon, it is difficult to screen the *ARX* mutations via regular sequencing or with whole exome sequencing. Therefore, GC-rich part of the *ARX* exon2 were screened via Agilent technologies in 16 selected patients for poly alanine expansion screening. Consequently, whole exome sequencing was performed in 3 patients with epileptic encephalopathy and without *ARX* mutations to reveal pathological variants, and candidate variants were determined.

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. DNA Samples

Blood samples of EE patients were provided by clinicians in neurology or child neurology departments of various Medical Schools in Turkey. DNA was previously extracted using standard protocols.

2.1.2. Equipment

Table 2.1. The list of the equipment which are used in this thesis project.

Equipment	**Company/Country**
5415C Centrifuge	Eppendorf Germany
Microcentrifuge	Beckmann Coulter/USA
Vortex	Labnet Turkey
Nanodrop	Thermo Scientific/USA
T100 Thermal Cycler	Bio-Rad/USA
Agarose gel electrophoresis system	Wealtec/USA
Weighing scale	Precisa/Switzerland
Heat block	Techne /England
Refrigerator	+4^{o} C Medicool, Sanyo,/Japan & Arçelik 4042T/Turkey
ND 1000, NanoDrop	Thermo Scientific/USA
Freezer -20	Arcelik/Turkey
Microwave	DBK/Turkey
Bioanalyzer	Agilent/USA
GelDoc Documentation System	Bio-Rad/USA
MAC-601 Autoclave	Eyela/Japan

2.1.3 Chemicals, kits, and buffers

Table 2.2. The list of the chemicals, kits, and buffers which are used in this thesis project.

Chemicals, kits, and buffers	Company/Country
Ethanol pure	Riedel de Haen/Germany
Taq DNA polymerase kit	Bioline/England
Nuclease-free water	Labnet Turkey
Agilent DNA 1000 Assay Kit	Thermo Scientific/USA
Agarose	Bio-Rad/USA
DNA ladder	Wealtec/USA
6X DNA Loading Dye	Fermentas/Lithuania
Ethidium Bromide	Sigma/Germany
Qiaquick Gel Extraction Kit solutions	Qiagen/Germany
TE Buffer (20mM Tris-HCl pH:8.0, 1mM Na_2EDTA(pH8.0)	Sigma/Germany
10X TBE Buffer (0.89 M Tris-Base, 0.89 M Boric Acid, 20mM Na_2EDTA pH:8.3)	Bio-Rad/USA

2.2. Methods

2.2.1. Clinical information of selected patients

ARX polyalanine expansion phenotype are linked with WEST and OHTAHARA syndrome. In this study we selected 16 male EE patients among 91 that present WEST and OHTAHARA syndromes. All 16 patients have EE diagnosis and general features were as follows: development before seizure was abnormal, growth retardation areas are same among patients such as fine-gross motor skills and social skills. Abnormal symptoms in MRI like lissencephaly and schizencephaly was present in most patients. Patients those have normal MRI findings showed abnormalities in EEG, have epileptic spasms and therefore diagnosed as EE for certain.

After multiple experiments that showed individual absence of *ARX* mutations, three male patients were selected for whole exome sequencing among the first 16 male patients. Those 3 patients had epileptic infantile spasms and had mental retardation, which were characteristics of WEST syndrome. Patients also show burst suppressions in their EEG

findings, which again confirms the EE diagnosis commonly seen in OS. Patients 26EE86 and 28EE92 showed also autistic features.

2.2.2. Mutational analysis of the *ARX* gene

2.2.2.1. ARX Repeat Expansion Analysis. DNA samples from selected 16 EE patients were analysed using Agilent DNA 1000 kit with Agilent 2100 bioanalyzer software. First of all, chip priming station was prepared with a syringe and the chip was inserted in the station. Gel-Dye mix preparation was held using DNA dye concentrate and DNA gel matrix. 9.0 µl gel-dye mix was loaded into the specific well on the chip after the mix reached at room temperature for 30 min. Plunger was positioned at 1 ml and then the chip priming station was closed. Timer was set to 60 seconds. After 60 seconds, the plunger was released with the clip release mechanism. 9.0 µl gel-dye mix was again pipetted into the specific wells shown in the Agilent DNA 1000 Assay protocol. 5 µl DNA marker was placed into the specific well marked with the ladder symbol and into each of the 12 sample wells. After this process, 1 µl ladder was loaded into the ladder symbol well. PCR products were loaded into the 12 sample wells or 1 µl deionized water to unused wells. The chip was placed into IKA vortex mixer provided by the Agilent 1000Assay kit and vortexed at 2400rpm. After vortex process, the chip was placed into Agilent 2100 Bioanalyzer. The analysis was performed using the 2100 Expert Software.

2.2.2.2. Detection of *ARX* Exon Mutations. *Polymerase Chain Reaction (PCR)*: PCR procedure was performed in order to amplify all the exon regions of *ARX* gene to detect any possible mutations rather than repeat expansions. Specific exon primers were designed using Primer3 [31] and NCBI Primer-Blast [32]. All primers used in the study are given in Table 2.1. PCR conditions are shown in Table 2.2. Optimizations of the PCR amplification of the primers for specific regions were performed within framework of this thesis. For all amplicons 80ng genomic DNA was used as template. Genomic DNA concentrations were measured by NanoDrop spectrophotometer. The PCR kit buffer contained 1X buffer, 1,5 mM $MgCl_2$, 10 mM dNTP mix. 0,10 mM of each primer and 0,2 µl Taq DNA polymerase were added into final reaction mixture which had a total volume of 25 µl.

Table 2.3. PCR primer sequences for *ARX* exons.

Amplified Region	Primers	Product Size
ARX Exon1	Forward 5'- CCCTTAGTAAGTGCCTGACGG -3' Forward 5'- GCCCACAGGCCGACG -3'	557
ARX Exon2.1	Forward 5'- TCCTCCGGGTGCGTGA -3' Forward 5'- CAAGGCGTCGAAGTCTGGT -3'	555
ARX Exon2.2	Forward 5'- CCTACGCGCATACCTGGTG -3' Forward 5'- CTGGGACACGCTCAAGATCAG -3'	622
ARX Exon 3	Forward 5'- CAGGGACTCTAAAGCCCTTGTT-3' Forward 5'- TGTAGACTTGAAGCCACCCTTG -3'	589
ARX Exon 4	Forward 5'- CAGTCTCTGAGGTTGGGTCAAA-3' Forward 5'- ACGCGTCCGAAAACAACCTG -3'	617
ARX Exon 5	Forward 5'- CTGCGCTCTCTCAGTGCC -3' Forward 5'- CCTCGGGGAATATCTGGACTC -3'	600

Table 2.4. PCR conditions used to amplify *ARX* exons and the regions of the candidate gene variants.

Step	Temperature	Time	Cycle number
Initial denaturation	95ºC	2 min	1
Denaturation	95ºC	30 sec	35
Annealing	Primer-specific	30 sec	35
Extension	72ºC	1 min	35
Final extension	72ºC	7 min	1
Hold	4ºC	∞	-

Agarose Gel Electrophoresis: PCR products of *ARX* exon regions (25 µl) were loaded on a 2% agarose gel, which is prepared with 0,5X TBE buffer, and run at 120V for 40 minutes. Bands were identified under UV light and displayed with GelDoc software. Bands were cut using a clean, sharp scalpel, placing into 1.5 ml Eppendorf tubes for gel extraction and purification to obtain high-quality results from Sanger sequencing.

Gel Extraction: PCR products of *ARX* exon regions were isolated from the gel pieces using QIAquick gel extraction kit (QIAGEN):

All centrifuge process was performed at 13,000rpm.

- Gel slices were weighed in a colourless tube. 3 volumes of Buffer QX1 was added to 1 volume for each gel slice. Gel (100 mg ~ 100 ml).

- Gel slices were incubated on heat block at 50ºC for 10 min. Tubes were vortexed every 2-3 min during the incubation for better dissolving.
- After the colour of the mixture turned yellow, 1 gel volume of isopropanol was added for each and tubes were mixed gently.
- QIAquick spin columns were places in the provided 2ml collection tubes.
- All sample were applied to the QIAquick column and centrifuged for 1 min.
- The flow was discarded. 0.5 ml Buffer QXI was added, and the tubes were centrifuged for 1 min.
- For washing process, 0.75 ml Buffer PE was added to QIAquick column and centrifuged for 1 min. After discarding the flow through, additional 1 min centrifugation was performed.
- The QIAquick columns were placed into clean 1.5 ml microfuge tubes. To elute DNA, 30 ml elution buffer was added to the center of the QIAquick column and centrifuged 1 min.
- After the gel extraction process, all obtained samples were sent to Macrogen for Sanger Sequencing.

Sanger Sequencing: An adequate amount of purified PCR products of *ARX* exon regions (25 µl) were sent to Macrogen Inc. The results were obtained online as ABI document and analysed in Genious and Finch TV version 1.4.0, chromatogram analysing programmes.

2.2.3. Whole Exome Sequencing

2.2.3.1. Candidate Variant Selection. DNA samples of EE patients were sent to BGI genomic co., China for exome sequencing. The results obtained with Illumina sequencer and were processed from bam data into vcf files. The vcf files were transcribed into excel sheet with wANNOVAR software, which was used for variant annotation. Variant type, genomic location, gene, effects on possible amino acid sequence information were obtained from the results. Furthermore, the frequencies on several populations were observed by 1000G and ExAC databases. *In silico* functional effects were observed by the results as per the SIFT, PolyPhen, MutationTaster databases. The exonic and splicing region variants which have less than %1 frequency in the population databases were chosen from the results obtained. After all steps and filtering methods were performed, selected genes/variants were searched

in the literature and various databases for conservation rates of variant locations among species, the most expressed tissues, molecular pathways and disease interactions. The databases that were used during the analyses are as follows: GeneCards Human Gene Database [33], UniProt [34], OMIM [35], UCSC Genome Browser [36], dbSNP [37], Ensembl [38], Mouse Genome Informatics [39], ClinVar [40], Expression Atlas [41].

2.2.3.2. Family Segregation Analyses. After selecting the candidate variants for each patient, family segregation analyses were performed. To this end, regions containing the candidate variants were amplified by PCR as described in section 2.2.2, using specific primer sets (Table 2.3). PCR products were run on an agarose gel, purified, and sent for Sanger sequencing as described above (in section 2.2.2).

Table 2.5. Primer sequences for candidate variant regions.

Amplified Region	Primers	Product Size
PCDHA12	Forward 5'-AGTTAGCGAGTTGGTACCGC-3' Reverse 5'-GTGCTGATCTCGCCAGTGTA-3'	171
CACNA1I	Forward 5'-ATTTGCTTTTCAGAGCTGTGCC-3' Reverse 5'-TTCTAGGTGGACAGACACACAA-3'	554
SHANK3	Forward 5'-CCTACAGCACCTTGCTCTTCC-3' Forward 5'- ATGTGCAGGACACACAGTCG -3'	166
CUL4B	Forward 5'- ATGTCCCCCAGACCTCTTAAC-3' Reverse 5'- TGGCTTTGCTGGATAATGCTA-3'	300
ADGRV1	Forward 5'- GAGTGTGGGAGGTGACAGTT -3' Forward 5'-ACTTGGAAATGAGCTGTCATACAA-3'	349
GRIN2B	Forward 5'- CTCCCTGCAGCCCCTTTTTA -3' Forward 5'- GCTCTTTCTTTGTGTTTGCAGG-3'	471
SYN3	Forward 5'- CCTACACCCCAGCTCTATCCT -3' Forward 5'- ATTTGTGTACAGCCATATCTATCCT -3'	411
PLXNB2	Forward 5'- TGGATGAACTGTAGGGGCCG -3' Forward 5'- AGGAATTCACAGGTAGGGACCG -3'	438
MAP2	Forward 5'- CAAAAGTGTATGGAGAGAAAAGGGA -3' Forward 5'- GGCTATGGAATCTAGGCAAGACA -3'	149

3. RESULTS

3.1. *ARX* Repeat Expansion Analysis

For *ARX* repeat analysis, the part of exon-2 containing the repeat region of 16 patients that present *ARX*-related phenotype and the same healthy control in each analysis was screened using Agilent system, revealing no significant repeat number increase in the patients. Agilent analysis results are given in Figure 3.1. Other exon regions were analysed by PCR and Sanger sequencing for the 16 patients.

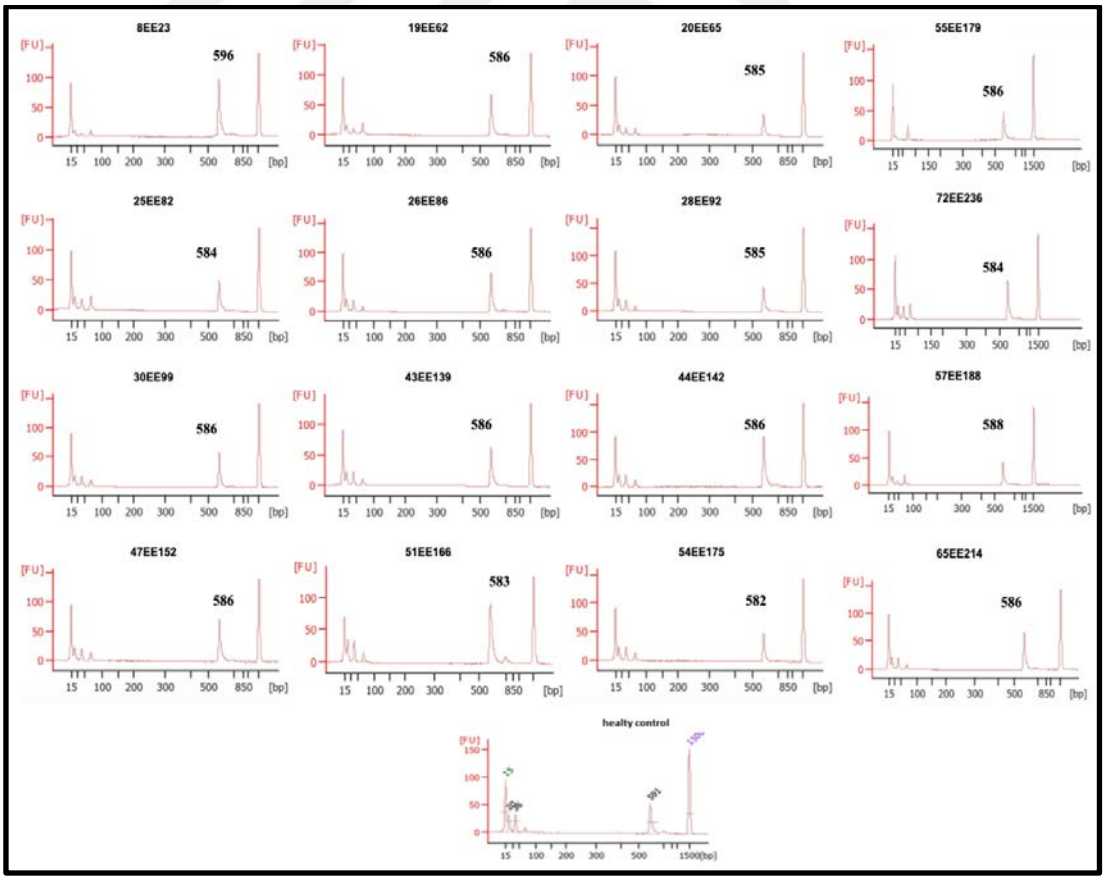

Figure 3.1. Agilent results of 16 EE patients and a healthy control. Patient codes and exon base-pair numbers are indicated on each graph.

3.2. Detection of *ARX* Exon Mutations

Possible mutations on each exon (exon-1, -2, -3, -4, and -5) of *ARX* gene was analysed by Sanger sequencing for 16 patients. *ARX* exon-2 part containing the repeat region (555 bp) was referred as exon 2.1 and the remaining part of the exon-2 (622 bp) was referred as exon 2.2 which were analysed separately. PCR optimisation was performed using DNA of a healthy volunteer. PCR product sizes and agarose gel results of each PCR optimization are given in Table 3.1 and Figure 3.2 respectively. Gel images of *ARX* PCR products for all 16 patients were given in Figure 3.3, Figure 3.4, Figure 3.5, Figure 3.6 and Figure 3.7. Analyses of Sanger sequencing data using FinchTV software revealed no mutations in the exons of *ARX* gene. Sanger results showing correct region sequencing and showing no *ARX* exon mutations are demonstrated in one example for patient 26EE86 in Figures 3.8.

Table 3.1. PCR product sizes of *ARX* exon regions.

***ARX* exon regions**	**Size (base pair)**
1	557
2.1	555
2.2	622
3	589
4	617
5	600

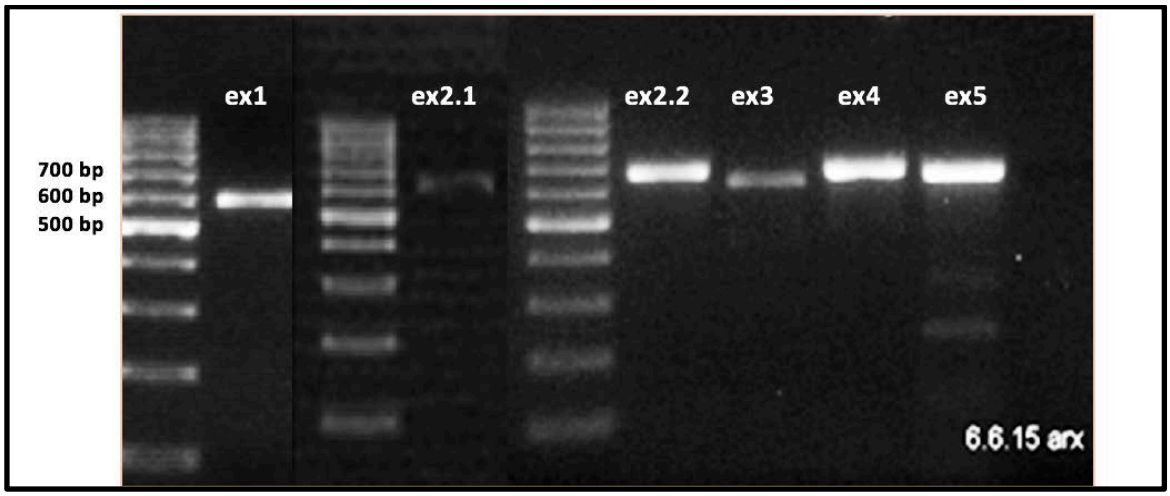

Figure 3.2. PCR products of amplified *ARX* exon regions.

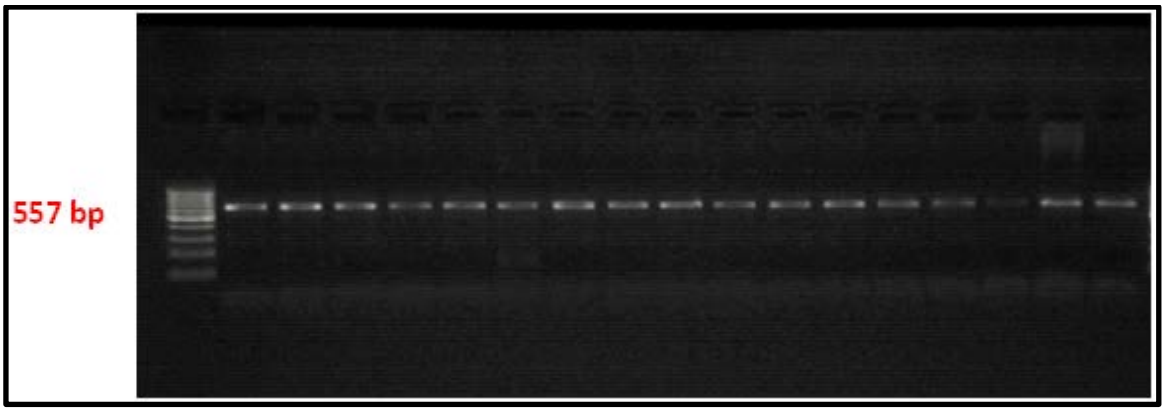

Figure 3.3. PCR products of amplified *ARX* exon 1 regions for 16 patients

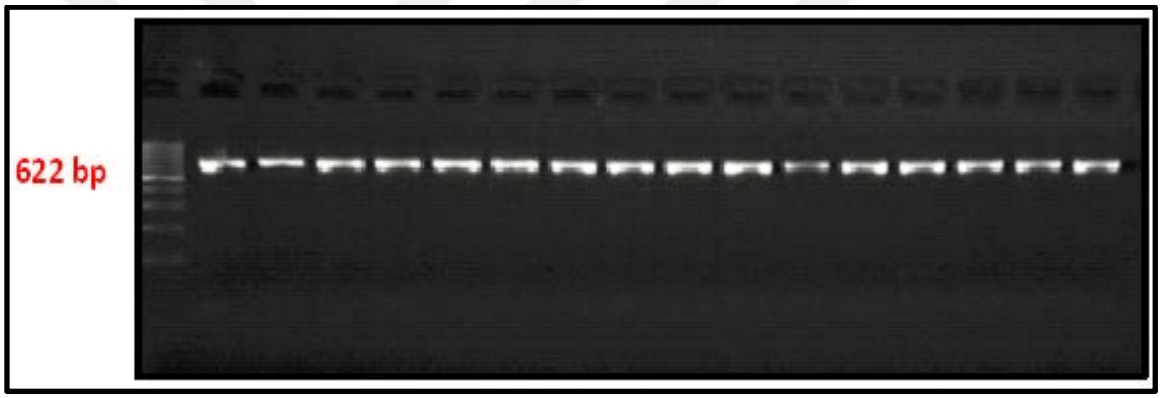

Figure 3.4. PCR products of amplified *ARX* exon 2.2 regions for 16 patients.

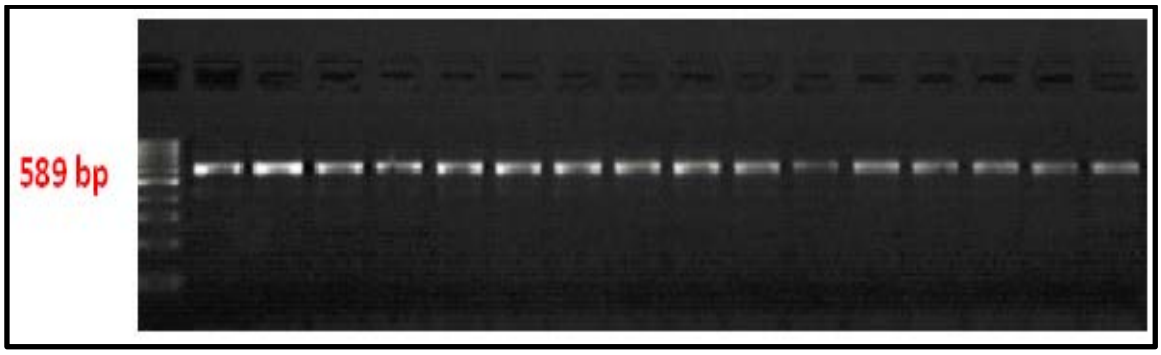

Figure 3.5. PCR products of amplified *ARX* exon 3 regions for 16 patients.

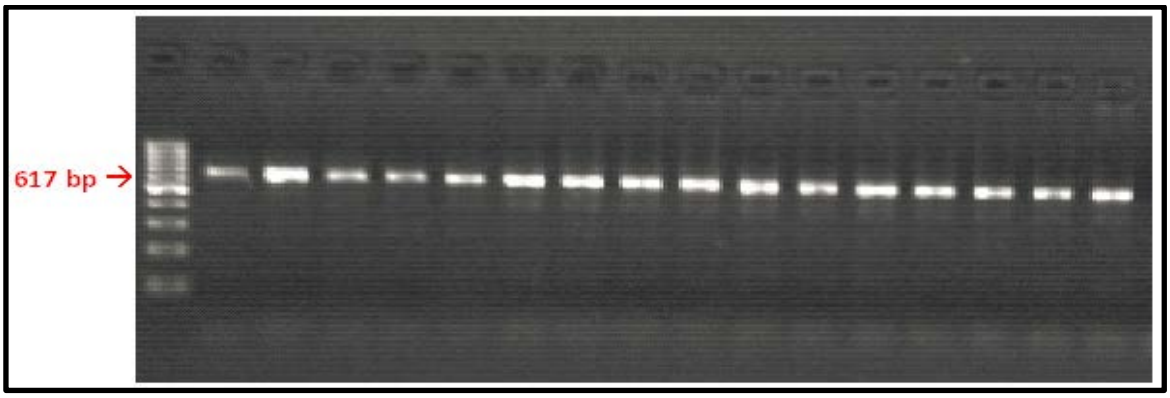

Figure 3.6. PCR products of amplified *ARX* exon 4 regions for 16 patients.

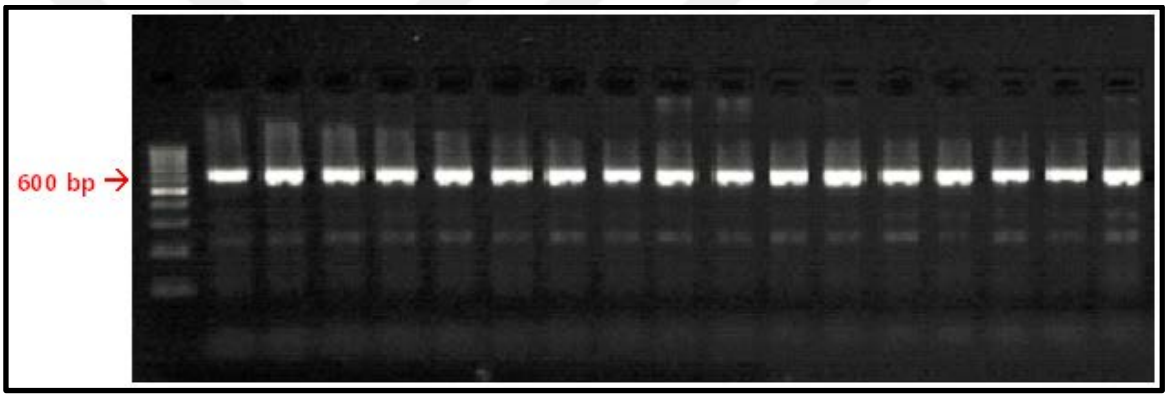

Figure 3.7. PCR products of amplified *ARX* exon 5 regions for 16 patients.

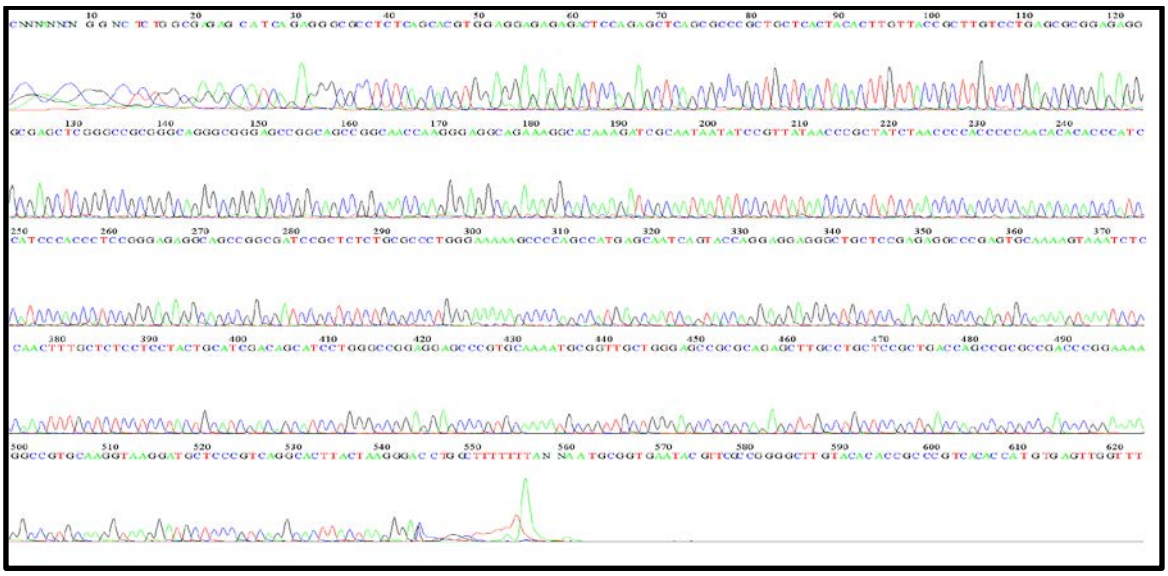

Figure 3.8. Example Sanger Chromatogram for *ARX* exon1 of patient 26EE86.

3.3. Whole Exome Sequencing in *ARX* Gene Mutation-Negative Patients

Upon revealing that there is no mutation on *ARX* gene, WES was performed in 3 EE patients with *ARX*-related disease phenotype without *ARX* gene mutations. A total of 3.760 variants, 2.542 of them being single nucleotide change and 1.218 being insertion/deletion (indel) variations, were detected by WES. Table 3.2 and Table 3.3 show numbers and types of detected variations.

Table 3.2. Number and types of detected SNPs by WES.

Samples	Synonymous	Missense	Stopgain	Stoploss	Startloss	Splicing
26EE86	10012	9117	67	30	12	58
28EE92	9860	9125	55	27	21	63
30EE99	9684	8851	71	31	15	54
Overall	15714	14658	126	38	28	88

Table 3.3. Number and types of detected indels by WES.

Samples	Frameshift	Non-frameshift Insertion	Non-frameshift Deletion	Stoploss	Startloss	Splicing
26EE86	211	114	93	1	4	47
28EE92	222	114	106	3	4	44
30EE99	219	109	92	6	3	49
Overall	323	172	165	7	5	52

In order to determine candidate variants; firstly, synonymous and intronic variants were excluded from the list of each patient's WES data. Both homozygous and heterozygous variants were selected. Among the variants that are located on exons/exon-intron junctions, and that are stop-loss/start-loss; the ones with a European population frequency of less than 1% were selected (1000 Genomes and ExAC Browser). Among those; more conserved variants among species (USCS Genome Browser), and variants with higher probability of being pathogenic (PolyPhen-2 & MutationTaster) were selected. In addition, variants that have been previously shown to be associated with epilepsy and having high expression levels in the brain were further chosen, resulting in a total of 9 candidate variants on 9 different genes (Table 3.4).

Table 3.4. Candidate pathological variants selected from WES data.

Patient ID	Gene	Chr	Start	End	Het/ Hom	Change	Read Num-ber	Read Quality	Pathog-enicity
26EE86	*PCDHA12*	chr5	140256870	140256870	Het	c.G1813T, p.A605S	30	Passed	T, D, D
	CACNA1I	chr22	40045583	40045583	Het	c.G1540A, p.D514N	30	Passed	T, D, D
	SHANK3	chr22	51117800	51117800	Het	c.G829A, p.G277R	67	Passed	T, B, N
	CUL4B	chrX	119668348	119668348	Hom	c.G2254A, p.A752T	68	Passed	., B, D
28EE92	*ADGRV1*	chr5	90040949	90040949	Het	c.C10636A, p.P3546T	87	Passed	T, D, D
	GRIN2B	chr12	14018972	14018972	Het	c.T171A, p.D57E	73	Passed	T, B, D
	SYN3	chr22	32914312	32914312	Het	c.G1328A, p.R443H	36	Passed	T, D, D
	PLXNB2	chr22	50719237	50719237	Het	c.C3929T, p.P1310L	69	Passed	T, B, N
	MAP2	chr2	210558947	210558947	Hom	c.A2053G, p.R685G	92	Passed	D, D, D
30EE99	*SHANK3*	chr22	51117800	51117800	Het	c.G829A, p.G277R	66	Passed	T, B, N

Pathogenicity prediction is based on SIFT, Polyphen, and MutationTaster, respectively.
B: benign, D: damaging, N: neutral, T: tolerable, .: unpredicted.

For validation and further selection of candidate pathogenic rare variants associated with EE, regions on which the listed variants are located were amplified, and Sanger sequencing was performed. The segregation analysis included all 3 patients and their families for the final candidate variant selection.

3.3.1. Variant Validation and Family Segregation for Patient 26EE86

As seen in Table 3.5, variants on *PCDHA12* and *CACNA1I* genes were eliminated, since the patient has one unaffected parent with the same variants for each one (Figure 3.9 and 3.10). In case of *SHANK3* variant, it was revealed that this was a false positive result of the WES; since Sanger sequencing of 26EE86 resulted in homozygous wild-type (Figure

3.11). However, segregation analysis of c.C2254T variant on *CUL4B* gene showed that the patient is homozygous for the variant, whereas the mother is heterozygous and the father is homozygous wild-type (Figure 3.12). Therefore, this variant was selected as the candidate variant. Family segregation for 26EE86 variants were presented in a family pedigree as shown in Figure 3.13.

Table 3.5. Sanger sequencing results of candidate variants for patient 26EE86 and parents.

Gene name / **Patient name**	***PCDHA12* (c.G1813T) (p.A605S)**	***CACNA1I* (c.G1540A) (p.D514N)**	***SHANK3* (c.G829A) (p.G277R)**	***CUL4B* (c.C2254T) (p.A752T)**
26EE86	Heterozygous	Heterozygous	Homozygous wild-type	Homozygous variant
26EE86 mother	Homozygous wild-type	Heterozygous	Homozygous wild-type	Heterozygous
26EE86 father	Heterozygous	Homozygous wild-type	Homozygous wild-type	Homozygous wild-type

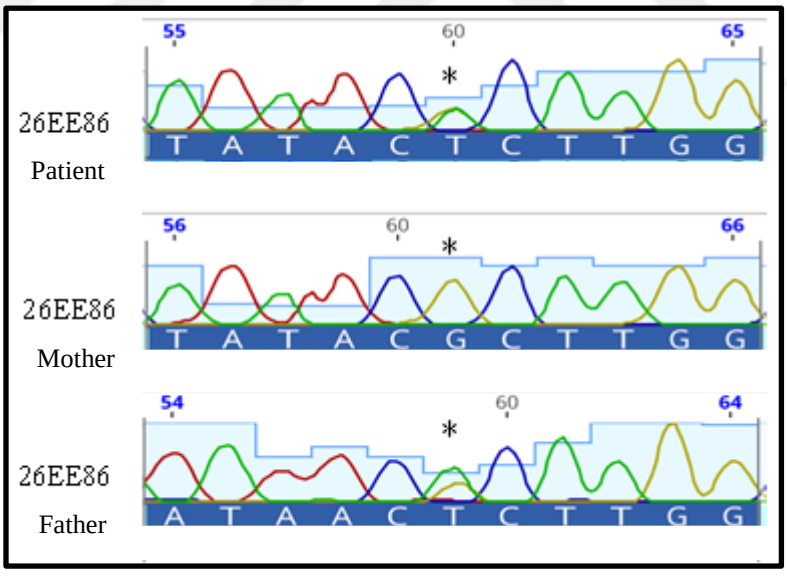

Figure 3.9. Sanger chromatograms of the *PCDHA12* variant for 26EE86 family

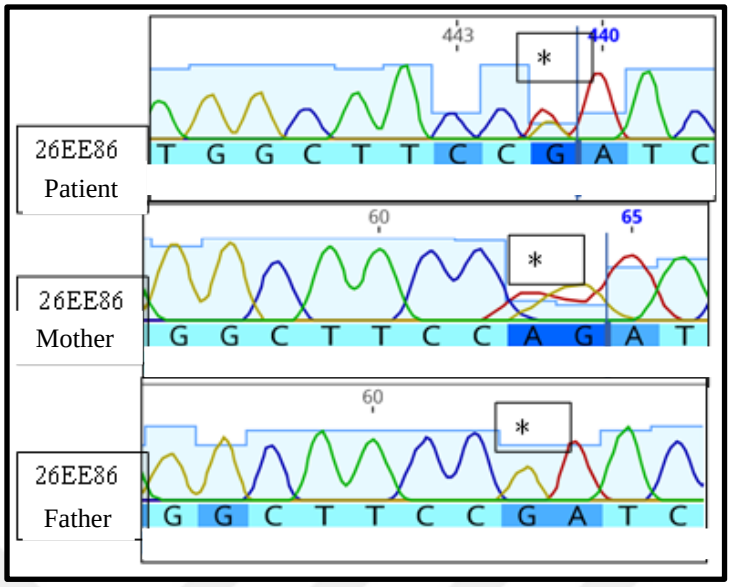

Figure 3.10. Sanger chromatograms of the *CACNA1I* variant for 26EE86 family.

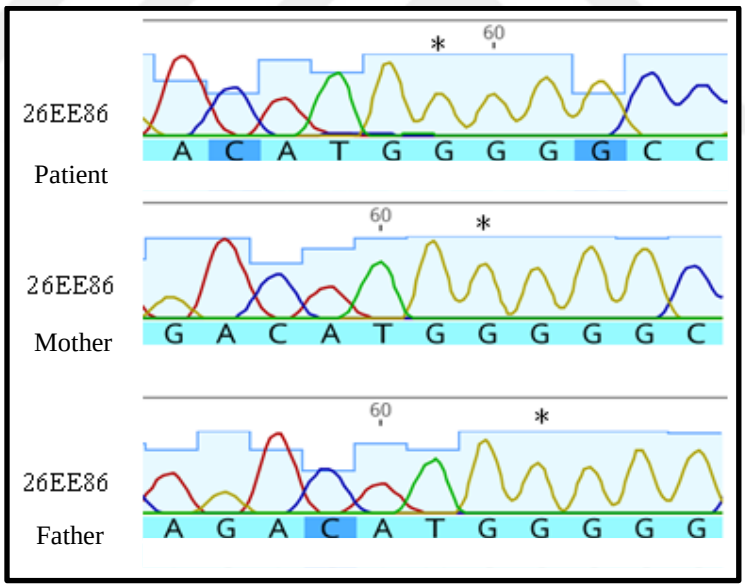

Figure 3.11. Sanger chromatograms of the *SHANK3* variant for 26EE86 family.

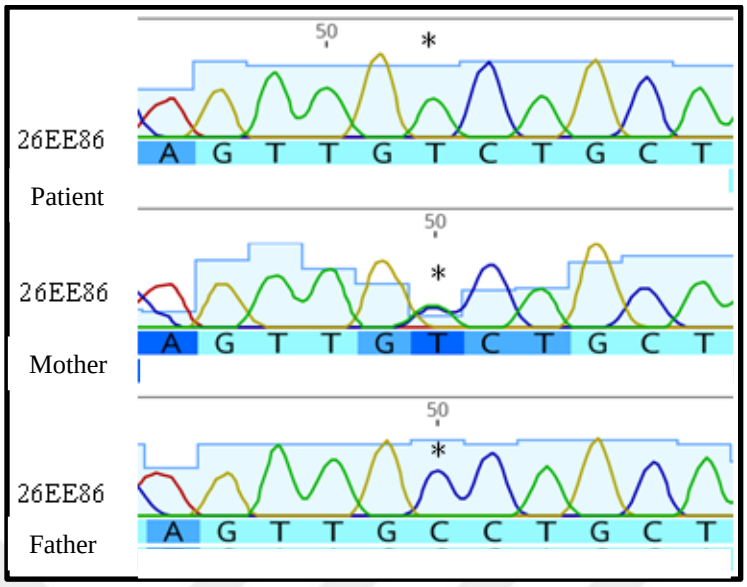

Figure 3.12. Sanger chromatograms of the *CUL4B* variant for 26EE86 family.

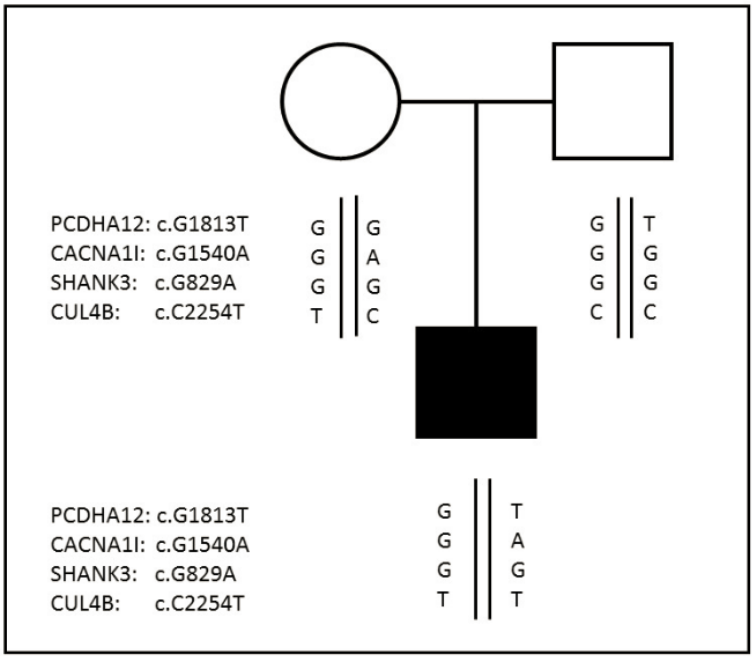

Figure 3.13. Family pedigree for 26EE86 family

3.3.2. Variant Validation and Family Segregation for Patient 28EE92

As seen in Table 3.6, variants on *ADGRV1, GRIN2B, SYN3,* and *PLXNB2* genes were eliminated for patient 28EE92, since the patient has one unaffected parent with the same variants for each one (Figure 3.14 -3.17 and Figure 3.19). However, segregation analysis for

c.A2053G variant on *MAP2* gene showed that the patient was homozygous for that variant, whereas both unaffected parents were heterozygous (Figure 3.18). Therefore, this variant was selected as the strongest candidate for this patient.

Table 3.6. Sanger sequencing results of candidate variants for patient 28EE92 and parents.

Gene name / **Patient name**	***ADGRV1* (c.C10636A) (p.P3546T)**	***GRIN2B* (c.T171A) (p.D57E)**	***SYN3* (c.G1328A) (p.R443H)**	***PLXNB2* (c.C3929T) (p.P1310L)**	***MAP2* (c.A2053G) (p.R685G)**
28EE92	Heterozygous	Heterozygous	Heterozygous	Heterozygous	Homozygous variant
28EE92 mother	Heterozygous	Homozygous wild-type	Homozygous wild-type	Homozygous wild-type	Heterozygous
28EE92 father	Homozygous wild-type	Homozygous variant	Heterozygous	Heterozygous	Heterozygous

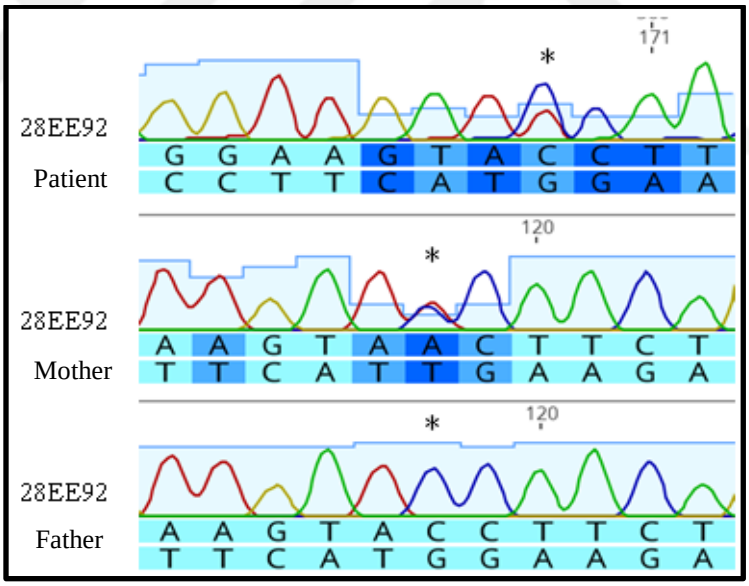

Figure 3.14. Sanger chromatograms of the *ADGRV1* variant for 28EE92 family.

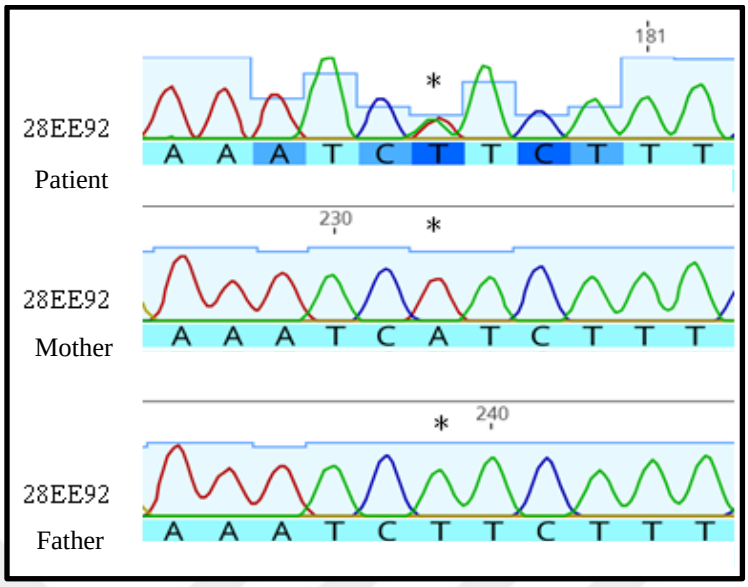

Figure 3.15. Sanger chromatograms of the *GRIN2B* variant for 28EE92 family.

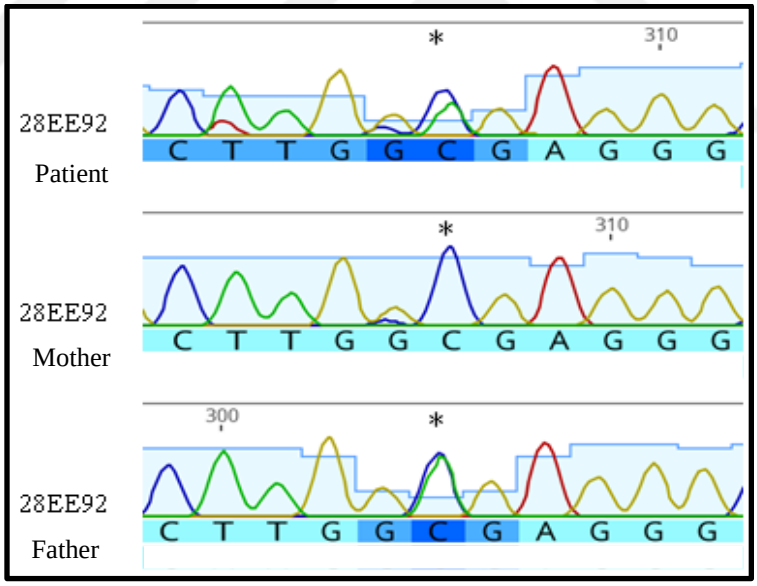

Figure 3.16. Sanger chromatograms of the *SYN3* variant for 28EE92 family.

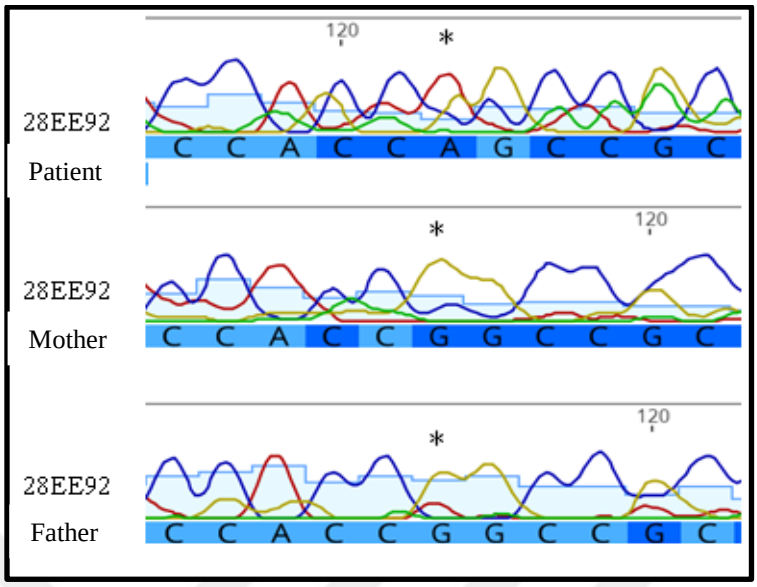

Figure 3.17. Sanger chromatograms of the *PLXNB2* variant for 28EE92 family.

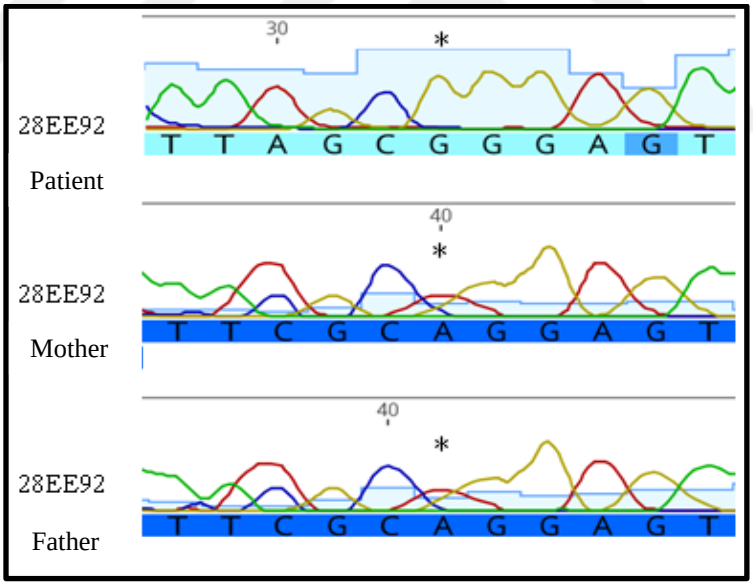

Figure 3.18. Sanger chromatograms of the *MAP2* variant for 28EE92 family.

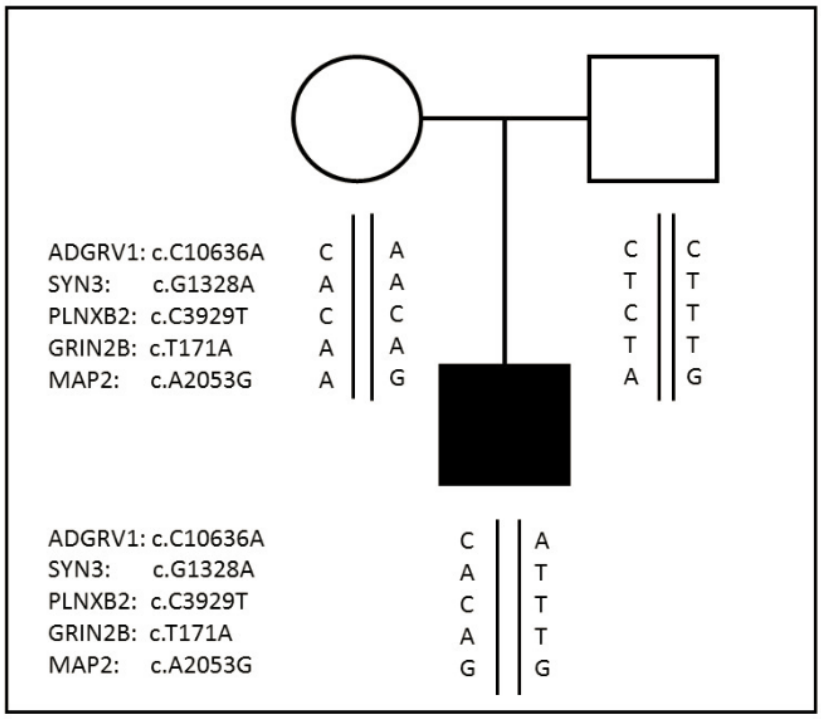

Figure 3.19. Family pedigree for 28EE92 family

3.3.3. Variant Validation and Family Segregation for Patient 30EE99

In case of patient 30EE99, the only candidate variant, *SHANK3* variant, was a false positive result of the WES same as in patient 26EE86. Sanger sequencing of patient 30EE99 resulted in homozygous wild-type sequence; therefore, this variant was eliminated (Table 3.7, Figure 3.20 and Figure 3.21)

Table 3.7. Sanger sequencing results of candidate variants for patient 30EE99.

Gene name / Patient name	SHANK3 (c.G829A) (p.G277R)
30EE99	Homozygous wild-type
30EE99s' Mother	Homozygous wild-type
30EE99s' Father	Homozygous wild-type

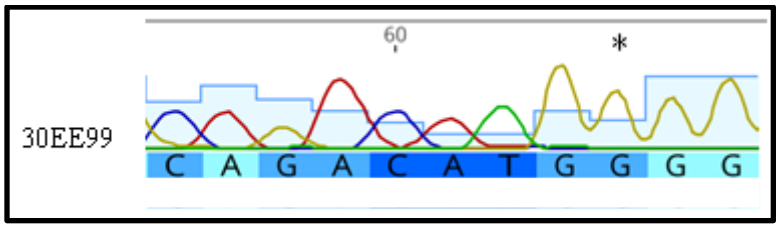

Figure 3.20. Sanger chromatograms of the *SHANK3* variant for 30EE99 patient.

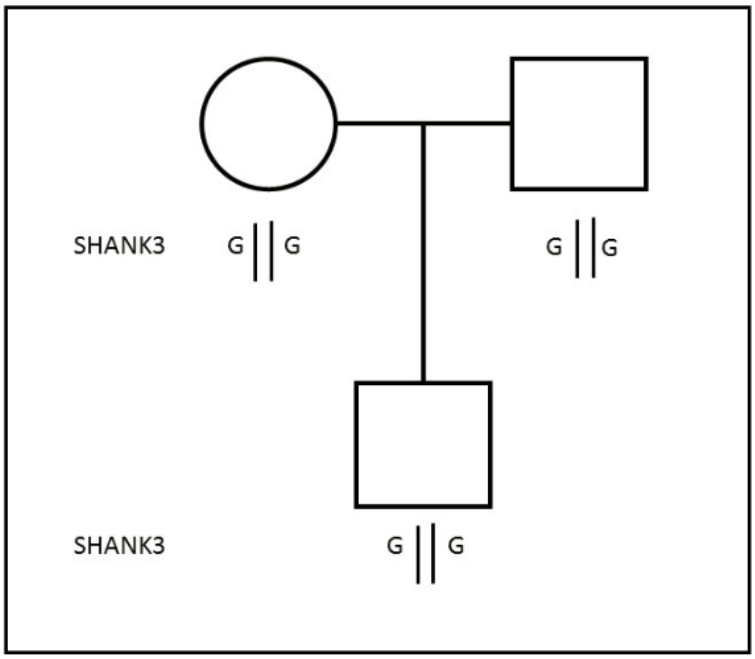

Figure 3.21. Family pedigree for 30EE99 family.

4. DISCUSSION

In this study, genetic basis of epileptic encephalopathy in selected patients with a common phenotype indicating WEST syndrome was explored by Agilent technology and whole-exome sequencing in the *ARX* gene. No *ARX* mutations were detected in 16 EE patients by the Agilent system, indicating existence of pathological mutations on other genes responsible for EE development. Exome sequencing of 3 selected patients presenting similar disease phenotype was performed, and 2 candidate variants that may be involved in the disease pathophysiology were detected.

After exome sequencing analyses, 9 candidate pathogenic variants in 3 patients were selected based on gene function and variant properties. *PCDHA12* (Protocadherin Alpha 12) gene encodes for a cell-adhesion protein that involves in maintenance of neuronal connections in brain, and the protein is highly expressed in the brain and abundant in cerebrospinal fluid. *CACNA1I* (Calcium Voltage-Gated Channel Subunit Alpha1 I) encodes for alpha subunit of a voltage-gated calcium channel, and the gene was previously associated with epilepsy [42]. *SHANK3* (SH3 And Multiple Ankyrin Repeat Domains 3) encodes for a scaffold protein that has roles in connecting neurotransmitter receptors and ion channels to actin cytoskeleton and in synapse formation. Mutations on this gene were associated with schizophrenia [43] and autism spectrum disorder [44]. Moreover, *Shank3*-mutant mice shows various nervous system phenotypes including environmentally induced seizures. *CUL4B* (Cullin 4B) is a cullin family member coding gene, which functions as E3 ubiquitin ligase. *CUL4B* mutations causes mental retardation [45], therefore the gene was selected as a candidate since similar phenotypes were observed in the studied patients. The receptor encoded by *ADGRV1* (Adhesion G Protein-Coupled Receptor V1) binds calcium in the central nervous system and previously associated to familial febrile seizures [46]. *GRIN2B* (Glutamate Ionotropic Receptor NMDA Type Subunit 2B) gene encodes for a subunit of N-methyl-D-aspartate receptor ion channel, which is an agonist for glutamate binding site. *GRIN2B* mutations are responsible for early infantile EE and mental retardation [47-49]. *SYN3* (Synapsin III) encodes for a neuronal phosphoprotein which may have roles in neurotransmitter release regulation and synaptogenesis. Syn3-mutant mice shows abnormal

excitatory and inhibitory postsynaptic currents. Moreover, *SYN3* was previously suggested to be linked with seizure disorders (ClinVar). *PLXNB2* (Plexin B2) is involved in axon guidance and cell migration, and has not been associated with a disease phenotype yet. Plxnb2 mutations in mice cause several nervous system phenotypes, mostly including abnormal brain development. *MAP2* (Microtubule Associated Protein 2) is essential in neurogenesis due to its roles in microtubule dynamics. *MAP2* was reported to be involved in many neurological conditions including Alzheimer's disease [50], Hirschsprung disease [51], prion disease [52], seizure-related brain damage [53].

After familial segregation analyses by Sanger sequencing of the variants, 7 of the variants were eliminated. In patient 26EE86 SHANK3 variant was false positive and *PCDHA12* and *CACNA11* variants in heterozygous state were eliminated since they were inherited from healthy parents. *CUL4B* variant however, was homozygous in the patient, heterozygous in the mother, and homozygous wild type in the father. One possibility for this situation is that there may be a deletion in the patient's genome affected the wild-type allele inherited from the father. This may lead to homozygous genotype result of the variant even though the patient has only one allele inherited from the mother. Other possibility is that the second allele in the patient could have resulted from a de novo mutation which is quite unlikely. Since the epileptic encephalopathies are complex disorders which genetic bases are not very certain, there is always a chance to find novel de novo mutation in an individual. These possibilities can be further investigated with whole SNP sequencing and with copy number variation analysis, which can identify result if it is due to a deletion or not. Lastly, this situation can also be seen if there is a different biological father. In order to increase our knowledge for this *CUL4B* variant, patient, mother and father should be examined for other genes in order to make the father data more reliable, since we could not ask the father to take a paternity test. Nevertheless, we argue that the homozygous condition of CUL4B variant in the patient deserves to be treated as the most likely candidate pathological variant.

CUL4B variant (c.G2254A) is located on a highly conserved base among species (Figure 4.1) as per UCSC Genome Browser. According to the database, when the dark blue line approaches to 4.88, it reveals the high conservation among species whereas if it becomes closer to -4.5, it means that the variant is less conserved among species. In Figure 4.1, our variant showed high conservation in the light of the database information. The variant has

not been reported in the literature, therefore no Rs number was found for the variant in dbSNP database. *CUL4B* gene is located in Xq24. Its product Cullin 4B is a scaffold protein and functions as E3 ubiquitin ligase in response to DNA damage, cell cycle progression, cell growth, and neuron projection development. *CUL4B* is expressed in a wide range of tissue types including the brain. Phenotypes presented in Cul4b-mutant mice include nervous system abnormalities such as increased susceptibility to pharmacologically induced seizures, abnormal dendrite and dendritic spine morphology, and decreased neuron number. It is well-known that *CUL4B* mutations cause recessively inherited X-linked mental retardation. However, *CUL4B* mutations have also been previously reported in individuals experiencing seizure events [54]. In a study by Tarpey *et al.*, 250 families with X-linked mental retardation were screened and *CUL4B* mutations were detected in 8 of the families [55]. During the adolescence of the patients, a specific phenotype including aggressive outbursts, seizures, relative macrocephaly, central obesity, hypogonadism, pes cavus, and tremor were observed. In another study, a novel variant (c.2107A-->T, p.703K-->X) in exon 18 of *CUL4B* gene was detected in three brothers with X-linked mental retardation [56]. The patients presented severe mental retardation, speech impairment, hyperactivity, seizures, intention tremor, inguinal hernia, small feet, and craniofacial dysmorphism. Similarly, Ravn *et al.* reported a 14-year-old monozygotic twin pair with a 28 kilobase-length deletion in *CUL4B* gene, presenting moderate mental retardation, short stature, truncal obesity, characteristic facial dysmorphism, tremor, seizures, and relative macrocephaly [57]. In 2012, a group successfully produces Cul4b-mutant mice with a deletion. They reported a decrease in the GABAergic interneurons and dendritic deficits in hippocampal neurons. They present decrease in inhibitory regulation and abnormal dendritic properties in hippocampal neurons, together resulting to epileptic susceptibility and spatial learning deficits. Since hippocampus is responsible for learning and memory formation, it can be concluded that *CUL4B* mutations in hippocampal neurons are linked to mental retardation [62]. In our study, *CUL4B* missense variant causes conversion of alanine to threonine in the 752. amino acid. Cullin 4B has a nuclear localization; therefore, contains a nuclear localization signal (55-58 amino acids) and a chain, consisting of a total of 913 amino acids. The detected variant is not thought to change localization of the protein and there has been no specific modification detected at the 752. amino acid. Even though cannot be accurately concluded, the amino acid change may alter interaction pattern of Cullin 4B with it's one or more specific interaction partners during the scaffolding process for ubiquitination. Cullin 4B has hundreds of potential interaction

partners. Failure of ubiquitination, thus degradation of target proteins may lead to disruption in several cellular regulations that may result in high ion channel number or activity causing the seizures. Also, other ubiquitin mutations were reported in the literature linked with epileptic syndromes. In 2014, a group studied with the E3 ubiquitin ligase complex CRL4A. Ion channels are targeted by CRL4A for ubiquitination and they are stored in endoplasmic reticulum, prevented their localization in cell membrane. It was shown that inactivation of the E3 ubiquitin ligase complex releases the ion channels to the cell membrane, therefore increasing the channel activity. The group worked with Crl4a-mutant mice and reported seizure induction sensitivity, which were caused by the high expression of ion channels on the cell surface [63].

In our study, the patient carrying this variant has West syndrome showing mental regression and EEG abnormalities. The developmental delay areas of the patient are psychomotor learning and language. Consistently, the patient also presents autistic features. As per the correlation between 26EE86 patient phenotype and the *CUL4B* genetical and molecular background from literature, we suggested that this variant can lead to West syndrome phenotypes, and therefore selected as the novel candidate variant for this study.

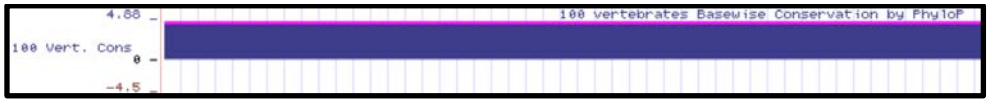

Figure 4.1. Vertebrate conservation rate of *CUL4B* variant

The patient carrying the *MAP2* variant (c.A2053G) is homozygous for this location, whereas both parents are heterozygous, showing a typical recessive inheritance. The variant is located on a conserved base among vertebrates (Figure 4.2), as the dark line approaches to 4.88; but not as much as the variant in *CUL4B.* The variant has not been reported in the literature, therefore no Rs number was found for the variant in dbSNP database. The cytogenetic location of *MAP2* gene is 2q34, MAP-2 protein mainly has roles in microtubule stabilization, and its related biological processes include axonogenesis, neuronal development, and synaptic vesicle transport. *MAP2* is overexpressed in the brain. Map2-mutant mice show nervous system phenotypes such as abnormal hippocampus morphology and abnormal Purkinje cell morphology. *MAP2* mutations have not been associated to any

disease with Mendelian-inheritance in the literature; however, the gene and its protein product have been implicated in many pathophysiological conditions. In a study of Ballough et.al. [53], seizure induction in the rat brain showed extreme MAP-2 reduction in the most damaged brain regions. Moreover, decreased MAP-2 levels in dorsolateral prefrontal cortex brain regions of two individuals with autism was reported by Mukaetova-Ladinska *et al.* [58]. In a recent study, a 4.5-year-old patient with epilepsy, mild developmental delay, and behavioural abnormalities was described and copy number variations were screened by an SNP array analysis [59]. The analysis revealed a 1.5 Mb *de novo* microdeletion on chromosome 2 (2q34). *UNC80*, *LANCL1*, and *MAP2* were determined as the candidate genes located on this region, and the authors suggested that *MAP2* was the strongest candidate due to its cellular functions. In another study by Tang *et al.*, relationship between MAP-2 and epilepsy was investigated using epileptic rats [60]. Mossy fiber sprouting is growth of granule cell axons into their dendritic field, creating recurrent excitatory circuits, which can explain the increased excitability in the epileptic brain [61]. Tang *et al.*, showed that overexpression of MAP-2 plays an important role in mossy fiber sprouting in epileptic mice. In 2016, a group reported a correlation between hippocampal dendrites and MAP-2 levels in status epilepticus, Map-2 induction resulted in destabilization of dendrite microtubules, lead to microglia accumulation in hippocampal region. This status epilepticus induced reduction in dendrite branching has been reported to be parallel with epileptogenesis and hippocampal-related cognitive defects [64]. Our patient carrying the *MAP2* variant has arginine to glycine conversion in the 685. amino acid which has no specific modification at that location detected so far. Even though 685. amino acid is not in a known location in terms of specific domain, activity, or binding sites of MAP-2, the variant converts a positively charged amino acid into an uncharged amino acid, which in turn may affect MAP-2 binding to microtubules or other potential interacting partners. It is well-known that MAP-2 stabilizes microtubules; therefore, amino acid change-based reduction in MAP-2 binding to microtubules may lead to increased microtubule depolymerisation [65]. This may result in increased neurite formation and increased excitability as suggested in the study of Danzer, et al [61]. Disruption in the microtubule organisation may result in transport dysregulation of numerous numbers of proteins including ion channels which may be yet another reason of dysregulation in neuronal electrical activity. Our patient has not diagnoses with a definite EE syndrome but in EEG, burst suppression was seen as in Ohtahara phenotype. The patient also has mental retardation and neurological developmental delay. In our patient's MRI,

cortical atrophy was detected, which can be explained by abnormal neural cell development. Therefore, even though association of *MAP2* to epilepsy has not as well-established as *CUL4B* in the literature, it is likely that *MAP2* has roles in neuronal processes that may be related to pathophysiological mechanisms in epilepsy, since patient 28EE92 show similar phenotypes as per *MAP2* mutation phenotypes from literature, therefore selected as a candidate novel variant for this study.

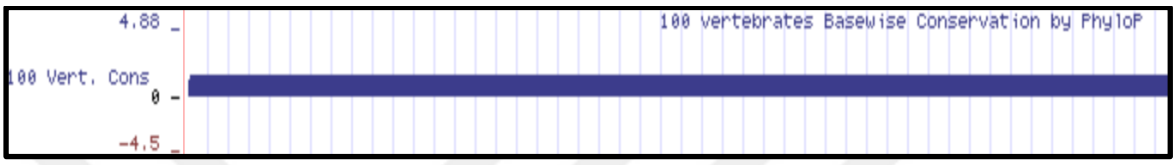

Figure 4.2. Vertebrate conservation rate of *MAP2* variant.

In conclusion, we have identified 2 novel candidate variants for epileptic encephalopathy using WES. Prevalence of less than %1 among population were selected from WES data, but in order to detect homozygous recessive variants, WES result can be filtered also for less than %5 prevalence in ExAc and 1000G databases for further examination for this study. Although the mentioned gene, variant, and protein properties indicate that *CUL4B* and *MAP2* are promising candidate genes and have previously been associated with phenotypes including epileptic features; these specific variants should be detected in a larger patient group for further evidence. In order to increase our inheritance knowledge, the segregation analysis for the selected variants should be performed in a large sample group in a specific family. Moreover, molecular effects of these variants should be identified in cell culture models by *in vivo* and *in vitro* experiments. Specific phenotypic features that may be caused by these variants should be studied in validated animal models for epilepsy to increase our knowledge in epileptic encephalopathy one step further.

REFERENCES

1. Fisher, Robert S., *et al.* "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)." *Epilepsia* 46.4 (2005): 470-472.

2. Ali, Niaz, and Muhammad Nabi. "The prevalence, incidence and etiology of epilepsy." *International Journal of Clinical and Experimental Neurology* 2.2 (2014): 29-39.

3. Annegers, John F., Walter A. Rocca, and W. Allen Hauser. "Causes of epilepsy: contributions of the Rochester epidemiology project." *Mayo Clinic Proceedings*. Vol. 71. No. 6. Elsevier, 1996.

4. Scheffer, Ingrid E., *et al.* "ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology." *Epilepsia* 58.4 (2017): 512-521.

5. Fisher, Robert S., *et al.* "Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology." *Epilepsia* 58.4 (2017): 522-530.

6. Fisher, Robert S., *et al.* "ILAE official report: a practical clinical definition of epilepsy." *Epilepsia* 55.4 (2014): 475-482.

7. CoCaTotILA, Epilepsy. "Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy." *Epilepsia* 30.4 (1989): 389-99.

8. Parisi, Pasquale, *et al.* "“Epileptic encephalopathy” of infancy and childhood: electro-clinical pictures and recent understandings." *Current neuropharmacology* 8.4 (2010): 409-421.

9. Fisher, Robert S. "The New Classification of Seizures by the International League Against Epilepsy 2017." *Current neurology and neuroscience reports*17.6 (2017): 48.

10. Dravet, Charlotte. "The core Dravet syndrome phenotype." *Epilepsia* 52.s2 (2011): 3-9.

11. Ohtahara, Shunsuke and Yamatogi, Yasuko. “Chapter 12: Neonatal Epilepsies with Suppression–Burst Pattern”. *Pediatric Epilepsy* (2002): ISBN 978-0-07-149621-6.

12. Ohtahara, Shunsuke, and Yasuko Yamatogi. "Ohtahara syndrome: with special reference to its developmental aspects for differentiating from early myoclonic encephalopathy." *Epilepsy research* 70 (2006): 58-67.

13. Pavone, Piero, *et al.* "Ohtahara syndrome with emphasis on recent genetic discovery." *Brain and Development* 34.6 (2012): 459-468.

14. Wheless, James W., *et al.* "Infantile spasms (West syndrome): update and resources for pediatricians and providers to share with parents." *BMC pediatrics* 12.1 (2012): 108.

15. Markand, Omkar N. "Lennox-Gastaut syndrome (childhood epileptic encephalopathy)." *Journal of Clinical Neurophysiology* 20.6 (2003): 426-441.

16. Isojarvi, Jouko, *et al.* "Clobazam-treated patients with Lennox-Gastaut syndrome experienced fewer seizure-related injuries than placebo patients during trial OV-1012." *Epilepsia* 57.6 (2016).

17. Ishii, Atsushi, *et al.* "A de novo missense mutation of GABRB2 causes early myoclonic encephalopathy." *Journal of medical genetics* (2016): jmedgenet-2016.

18. Conroy, Judith, *et al.* "Towards the identification of a genetic basis for Landau-Kleffner syndrome." *Epilepsia* 55.6 (2014): 858-865.

19. Gonsales, Marina C., *et al.* "Recent developments in the genetics of childhood epileptic encephalopathies: impact in clinical practice." *Arquivos de neuro-psiquiatria* 73.11 (2015): 946-958.

20. Online Mendelian Inheritance in Man, An Online Catalog of Human Genes and Genetic Disorders, https://www.omim.org/, accessed at May 2019.

21. Escayg, Andrew, and Alan L. Goldin. "Sodium channel SCN1A and epilepsy: mutations and mechanisms." *Epilepsia* 51.9 (2010): 1650-1658.

22. Absoud, Michael, *et al.* "A novel ARX phenotype: rapid neurodegeneration with Ohtahara syndrome and a dyskinetic movement disorder." *Developmental Medicine & Child Neurology* 52.3 (2010): 305-307.

23. Gécz, Jozef, Desiree Cloosterman, and Michael Partington. "ARX: a gene for all seasons." *Current opinion in genetics & development* 16.3 (2006): 308-316.

24. Helbig, Ingo (Kiel). "The OMIM epileptic encephalopathy genes – a 2014 review". http://epilepsygenetics.net/2014/05/04/the-omim-epileptic-encephalopathy-genes-a-2014-review/ accessed at May 2019.

25. Ekşioğlu, Yaman Z., Amanda W. Pong, and Masanori Takeoka. "A novel mutation in the aristaless domain of the ARX gene leads to Ohtahara syndrome, global developmental delay, and ambiguous genitalia in males and neuropsychiatric disorders in females." *Epilepsia* 52.5 (2011): 984-992.

26. Marques, Isabel, *et al.* "Unraveling the pathogenesis of ARX polyalanine tract variants using a clinical and molecular interfacing approach." *Molecular genetics & genomic medicine* 3.3 (2015): 203-214.

27. Suri, Mohnish. "The phenotypic spectrum of ARX mutations." *Developmental medicine and child neurology* 47.2 (2005): 133-137.

28. Cossée, Mireille, *et al.* "ARX polyalanine expansions are highly implicated in familial cases of mental retardation with infantile epilepsy and/or hand dystonia." *American Journal of Medical Genetics Part A* 155.1 (2011): 98-105.

29. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/genome/gdv/browser/?context=gene&acc=170302, accessed at May 2019.

30. Friocourt, Gaëlle, and John G. Parnavelas. "Mutations in ARX result in several defects involving GABAergic neurons." *Frontiers in cellular neuroscience* 4 (2010).

31. Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386. http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi, accessed at May 2019.

32. Madden T. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. In: McEntyre J, Ostell J, editors. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002-. Chapter 16. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch, accessed at May 2019.

33. Genecards Human Gene Database-https://www.genecards.org/, assessed at May 2019.

34. Uniprot , https://www.uniprot.org/ assessed at May 2019.

35. Online Mendelien Inheritance in Man (OMIM), https://www.omim.org/, assessed at May 2019.

36. UCSC Genome Browser, https://genome.ucsc.edu/, assessed at May 2019.

37. dbSNP, https://www.ncbi.nlm.nih.gov/snp/ assessed at May 2019.

38. Ensembl, https://www.ensembl.org/index.html, assessed at May 2019.

39. Mouse Genome Informatics (MGI), http://www.informatics.jax.org/, assessed at May 2019.

40. Clinvar, https://www.ncbi.nlm.nih.gov/clinvar/, assessed at May 2019.

41. Expression Atlas, https://www.ebi.ac.uk/gxa/home, assessed at May 2019.

42. Rajakulendran S, Graves TD, Labrum RW, Kotzadimitriou D, Eunson L, Davis MB, *et al.* Genetic and functional characterisation of the P/Q calcium channel in episodic ataxia with epilepsy. J Physiol (2010) 588:1905–13. doi:10.1113/jphysiol.2009.186437

43. Gauthier, J., Champagne, N., Lafreniere, R. G., Xiong, L., Spiegelman, D., Brustein, E., Lapointe, M., Peng, H., Cote, M., Noreau, A., Hamdan, F. F., Addington, A. M., and 18 others. De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. Proc. Nat. Acad. Sci. 107: 7863-7868, 2010.

44. Soorya, L., Kolevzon, A., Zweifach, J., Lim, T., Dobry, Y., Schwartz, L., Frank, Y., Wang, A. T., Cai, G., Parkhomenko, E., Halpern, D., Grodberg, D., Angarita, B., Willner, J. P., Yang, A., Canitano, R., Chaplin, W., Betancur, C., Buxbaum, J. D. Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. Molec. Autism 4: 18, 2013.

45. Ravn, K., Lindquist, S. G., Nielsen, K., Dahm, T. L., Tumer, Z. Deletion of CUL4B leads to concordant phenotype in a monozygotic twin pair. (Letter) Clin. Genet. 82: 292-294, 2012

46. Nakayama, J., Fu, Y.-H., Clark, A. M., Nakahara, S., Hamano, K., Iwasaki, N., Matsui, A., Arinami, T., Ptacek, L. J. A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. Ann. Neurol. 52: 654-657, 2002.

47. Epi4K Consortium and Epilepsy Phenome/Genome Project. De novo mutations in epileptic encephalopathies. Nature 501: 217-221, 2013.

48. Lemke, J. R., Hendrickx, R., Geider, K., Laube, B., Schwake, M., Harvey, R. J., James, V. M., Pepler, A., Steiner, I., Hortnagel, K., Neidhardt, J., Ruf, S., Wolff, M., Bartholdi, D., Caraballo, R., Platzer, K., Suls, A., De Jonghe, P., Biskup, S., Weckhuysen, S. GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. Ann. Neurol. 75: 147-154, 2014.

49. Platzer, K., Yuan, H., Schutz, H., Winschel, A., Chen, W., Hu, C., Kusumoto, H., O'Heyne, H., Helbig, K. L., Tang, S., Willing, M. C., Tinkle, B. T., and 63 others. GRIN2B encephalopathy: novel findings on phenotype, variant clustering, functional consequences and treatment aspects.

50. Yen, S. H., Dickson, D. W., Crowe, A., Butler, M., & Shelanski, M. L. (1987). Alzheimer's neurofibrillary tangles contain unique epitopes and epitopes in common with the heat-stable microtubule associated proteins tau and MAP2. The American journal of pathology, 126(1), 81-91.

51. Yang, W. I., & Oh, J. T. (2013). Calretinin and microtubule-associated protein-2 (MAP-2) immunohistochemistry in the diagnosis of Hirschsprung's disease. Journal of pediatric surgery, 48(10), 2112-2117.

52. Zhang, J., & Dong, X. P. (2012). Dysfunction of microtubule-associated proteins of MAP2/tau family in Prion disease. Prion, 6(4), 334-338.

53. Ballough, G. P. H., Martin, L. J., Cann, F. J., Graham, J. S., Smith, C. D., Kling, C. E., ... & Filbert, M. G. (1995). Microtubule-associated protein 2 (MAP-2): a sensitive

marker of seizure-related brain damage. Journal of neuroscience methods, 61(1-2), 23-32.

54. Jie Wang, Zhi-Jian Lina, Liu Liu, Hai-Qing Xu, Yi-Wu Shi, Yong-Hong Yi, Na He, & Wei-Ping Liao (2017). Epilepsy-associated genes. Seizure 44, 11-20.

55. Tarpey, P. S., Raymond, F. L., O'Meara, S., Edkins, S., Teague, J., Butler, A., ... & Barthorpe, S. (2007). Mutations in CUL4B, which encodes a ubiquitin E3 ligase subunit, cause an X-linked mental retardation syndrome associated with aggressive outbursts, seizures, relative macrocephaly, central obesity, hypogonadism, pes cavus, and tremor. The American Journal of Human Genetics, 80(2), 345-352.

56. Badura-Stronka, M., Jamsheer, A., Materna-Kiryluk, A., Sowińska, A., Kiryluk, K., Budny, B., & Latos-Bieleńska, A. (2010). A novel nonsense mutation in CUL4B gene in three brothers with X-linked mental retardation syndrome. Clinical genetics, 77(2), 141-144.

57. Ravn, K., Lindquist, S. G., Nielsen, K., Dahm, T. L., & Tümer, Z. (2012). Deletion of CUL4B leads to concordant phenotype in a monozygotic twin pair. Clinical genetics, 82(3), 292-294.

58. Mukaetova-Ladinska, E. B., Arnold, H., Jaros, E., Perry, R., & Perry, E. (2004). Depletion of MAP2 expression and laminar cytoarchitectonic changes in dorsolateral prefrontal cortex in adult autistic individuals. Neuropathology and applied neurobiology, 30(6), 615-623.

59. Westphal, D. S., Andres, S., Makowski, C., Meitinger, T., & Hoefele, J. (2018). MAP2–A Candidate Gene for Epilepsy, Developmental Delay and Behavioral Abnormalities in a Patient With Microdeletion 2q34. Frontiers in genetics, 9, 99.

60. Tang, L., Lu, Y., Zheng, W., & Li, Y. (2014). Overexpression of MAP-2 via formation of microtubules plays an important role in the sprouting of mossy fibers in epileptic rats. Journal of Molecular Neuroscience, 53(1), 103-108.

61. Danzer, S. (2017). Mossy fiber sprouting in the epileptic brain: taking on the Lernaean Hydra. Epilepsy currents, 17(1), 50-51.

62. Chen, C. Y., Tsai, M. S., Lin, C. Y., Yu, I. S., Chen, Y. T., Lin, S. R., ... & Lin, S. W. (2012). Rescue of the genetically engineered Cul4b mutant mouse as a potential model for human X-linked mental retardation. Human molecular genetics, 21(19), 4270-4285.

63. Liu, J., Ye, J., Zou, X., Xu, Z., Feng, Y., Zou, X., ... & Cang, Y. (2014). CRL4A CRBN E3 ubiquitin ligase restricts BK channel activity and prevents epileptogenesis. Nature communications, 5, 3924.

64. Schartz, N. D., Herr, S. A., Madsen, L., Butts, S. J., Torres, C., Mendez, L. B., & Brewster, A. L. (2016). Spatiotemporal profile of Map2 and microglial changes in the hippocampal CA1 region following pilocarpine-induced status epilepticus. Scientific reports, 6, 24988.

65. Bloom, G. S., & Vallee, R. B. (1983). Association of microtubule-associated protein 2 (MAP 2) with microtubules and intermediate filaments in cultured brain cells. The Journal of cell biology, 96(6), 1523-1531.