T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

NEW GENOMIC APPROACHES TO EXPLORE THE NEUROGENETIC DISEASE BURDEN OF CONSANGUINEOUS MARRIAGES IN TURKEY

ELMASNUR YILMAZ

MOLECULAR BIOLOGY AND GENETICS

Master of Science

December 2018

IZMIR

T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

NEW GENOMIC APPROACHES TO EXPLORE THE NEUROGENETIC DISEASE BURDEN OF CONSANGUINEOUS MARRIAGES IN TURKEY

MOLECULAR BIOLOGY AND GENETICS

Master of Science

ELMASNUR YILMAZ

Thesis Advisor: Asst. Prof. Dr. Yavuz OKTAY

This study was supported by TUBITAK (The Scientific and Technological Research Council of Turkey) project No. 216S771

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
İZMİR ULUSLARARASI BİYOTIP VE GENOM
ENSTİTÜSÜ MÜDÜRLÜĞÜ

iBG-Enstitü

YÜKSEK LİSANS TEZ SAVUNMA SINAVI TUTANAĞI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı Moleküler Biyoloji ve Genetik Yüksek Lisans Programı 2016850012 numaralı öğrencisi Elmasnur Yılmaz, "Türkiye'deki akraba evliliklerine bağlı nörogenetik hastalık yükünün araştırılmasında yeni genomik yaklaşımlar" konulu Yüksek Lisans tezini 28/12/2018 tarihinde yapılan savunma sınavı sonucunda başarılı olmuştur.

Dr.Öğr.Üyesi Yavuz OKTAY

BAŞKAN

JÜRİ ÜYESİ

Prof. Dr. Şermin GENÇ

JÜRİ ÜYESİ

Dr. Öğr. Üyesi
Cihangir YANDIM

İzmir Uluslararası Biyotıp ve Genom Enstitüsü
Dokuz Eylül Üniversitesi
Sağlık Yerleşkesi No:58/1
Balçova 35340 - İzmir / TÜRKİYE
Tel : (232) 277 63 53
Faks : (232) 412 86 59
e-posta : ibg@deu.edu.tr

https://www.ibg.deu.edu.tr

İzmir önceklerin şehri

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor Asst. Prof. Dr. Yavuz Oktay for allowing me to be a member of his research group. By giving me this opportunity, he has definitely changed my life for good. I am deeply grateful to him allowing me to explore different projects and ideas, and also being kind and patient with me. I benefited greatly from his intelligence, open-mindedness, and enthusiasm for science. His advice has helped me grow as a researcher and an individual.

I would also like to thank Prof. Dr. Semra Hız for assisting CONSEQUITUR project. I appreciate her valuable ideas and support throughout this study.

This thesis has been a piece of CONSEQUITUR project and would not have been possible without the work of many people, for which I am deeply grateful. I want to thank Dr. Ana Topf who helped me and provided me a lot of suggestions and advice on data analysis. I also wish to express my appreciation to Prof. Dr. Rita Horvath and Prof. Dr. Hanns Lochmüller. It has been a great privilege to work with them and learn from such brilliant and thoughtful people. In addition, I would also like to thank to other collaborators Prof. Dr. Serdal Güngör, Dr. Bilge Özgör, Prof. Dr. Ahmet Yaramış. Most notably, without close collaboration with these clinicians, this research would not have been possible.

I also gratefully thank my thesis advisory committee members Prof. Dr. Şermin Genç and Asst. Prof. Dr. Cihangir Yandım for sparing their time and for their important feedback.

I want to express my gratitude to Assoc. Prof. Dr. Güneş Özhan. She has been supportive since the first day I began working in IBG. It was a great privilege and honour to know her. I will always be extremely grateful for her friendship, empathy, and support. I would also like to thank all the members of Özhan Lab who always helpful to me. Their friendships have supported me along the way. I wish you all the best! But I will particularly mention Yağmur Azbazdar, Esra Katkat, Mehmet Ergüven for bringing fun and colour in my life during these years.

The Özhan Lab alumni also deserve acknowledgement for helping me towards my studies in many ways. Thank you İsmail Küçükçaylak, Betül Haykır, Ansarullah for all the good times, insight, knowledge and inspiration. I have no doubt that our paths, scientific and otherwise, will cross many times again.

Last but not the least, I must express my very profound gratitude to my family for first fostering my interest in science and then patiently putting up with my extreme enthusiasm for science throughout the years of my education. I am deeply grateful to my parents for their love, prayers, caring and sacrifices for educating and preparing me for my future. Their warmth presence is all I need, and this work is dedicated to them.

This thesis is a part of ongoing project called CONSEQUITUR which is supported by MRC-RCUK and TUBITAK (The Scientific and Technological Research Council of Turkey) under Katip Çelebi-Newton Fund Programme.

So long, and thanks for all the fish!

Elmasnur YILMAZ

December 2018

Izmir

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

ACC	Agenesis of the Corpus Callosum
AD	Autosamal Dominant
AR	Automasal Resessive
CADD	Combined Annotation Dependent Depletion
CGH	Comparative Genomic Hybridization
CMV	Cytomegalovirus
CNV	Copy Number Variation
FCD	Focal Cortical Dysplasia
gDNA	genomic DNA
HW	Hardy-Weinberg
IBD	Identical by Descent
LoF	Loss of Function
MAF	Minor Allele Frequency
MCD	Malformation of Cortical Development
MRI	Magnetic Resonance Imaging
mtDNA	Mitochondrial DNA
NGS	Next-Generation Sequencing
OMIM	Online Mendelian Inheritance in Man
PMG	Polymicrogyria
ROH	Runs of Homozygosity
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variation
WES	Whole Exome Sequencing

NEW GENOMIC APPROACHES TO EXPLORE THE NEUROGENETIC DISEASE BURDEN OF CONSANGUINEOUS MARRIAGES IN TURKEY

ABSTRACT

Consanguineous marriages are common in Turkey and carry an increased risk of genetic conditions with autosomal recessive inheritance. The nervous system and muscle are affected by these conditions leading to severe disability or premature death. While each condition is individually rare, there are several hundred genetically defined neurogenetic disease entities, collectively resulting in significant health and economic burden. A definitive molecular diagnosis has not been achieved for most of the children with neurogenetic conditions in Turkey, as infrastructure and access to sophisticated diagnostic options such as MRI, comprehensive metabolic testing, muscle biopsies and genetic testing are limited.

The advance of next-generation sequencing now offers the opportunity to understand the genetic causes of childhood neurogenetic disorders better, ultimately providing a definite diagnosis for many families, and in some cases effective treatments.

In this study, children with undiagnosed neurogenetic disorders born to consanguineous parents were recruited to the study. Disease-causing variants were identified through systematic deep phenotyping of neurogenetic patients from consanguineous families, combined with whole exome sequencing (WES). Variants were analyzed through RD-Connect Genome-Phenome Analysis Platform (GPAP) and assessed for predicted deleterious effects with an integration of data from several bioinformatics tools. WES provided a molecular diagnosis for 62 families out of 138 (45%).

This study will help to develop new prevention and treatment strategies, improved outcomes for affected families, and to an essential body of genomic data relevant to the population and increased genomics research capacity within the Turkish medical community.

Key words: consanguineous marriage, neurogenetic diseases, recessive inheritance, whole exome sequencing, brain malformations, disease-causing variant

TÜRKİYE'DEKİ AKRABA EVLİLİKLERİNE BAĞLI NÖROGENETİK HASTALIK YÜKÜNÜN ARAŞTIRILMASINDA YENİ GENOMİK YAKLAŞIMLAR

ÖZET

Akraba evlilikleri Türkiye'de yaygındır ve otozomal resesif kalıtımla artan genetik hastalık riski taşırlar. Bu durumların birçoğu, sinir sistemi ve kasları etkiler, bu da ciddi sakatlık veya erken ölümlere yol açar. Her bir nörogenetik vaka tek başına nadir olmasına rağmen, toplu bir şekilde ciddi bir sağlık ve ekonomik yük ile sonuçlanan yüzlerce nörogenetik vaka ile karşılaşılmaktadır. Türkiye'de nörogenetik rahatsızlığı olan çocukların çoğunda MRG, kapsamlı metabolik testler, kas biyopsileri ve genetik testler gibi tanı seçenekleri kullanılmış olmasına rağmen çoğunda kesin bir moleküler tanıya ulaşılamamıştır.

Yeni nesil dizilemenin ilerlemesi, çocukluk çağındaki nörogenetik bozuklukların genetik nedenlerini daha iyi anlama, sonuçta birçok aile için kesin teşhis ve bazı durumlarda etkili tedavi fırsatı sunmaktadır. Genom dizilimi gün geçtikçe ekonomik olarak daha uygun hale gelmektedir. Yaygın olarak görülen Mendel hastalıklarına ve akraba evliliği sonucunda ortaya çıkmış nörogenetik hastalıklara yol açan nadir veya özel varyantların tanımlanması için de uygun hale gelmiştir.

Bu çalışmada akraba evliliği sonucu ortaya çıkan nörogenetik hastalıklara neden olan varyantlar, akraba olan ailelerden gelen nörogenetik hastaların sistematik derin fenotiplemesi, tüm ekzom sekanslama ve uluslararası biyoenformatik platformlarındaki verilerin entegrasyonu ile tanımlandı. Tüm ekzom sekanslama ile çalışmaya katılan 138 aileden 62 aile sekanslama sonucunda tanılandı.

Bu çalışma, önleme ve tedavi stratejilerinin geliştirilmesine, etkilenen aileler için iyileştirilmiş sonuçlara ve toplumla ilgili önemli bir genomik veri bilgi birikimine ve Türk tıp camiasında genom araştırmalarının artmasına yol açacaktır.

Anahtar kelimeler: akraba evliliği, nörogenetik hastalıklar, resesif geçişli hastalıklar, tüm ekzom dizilime, beyin malformasyonarı, mutasyonlar

1. INTRODUCTION

Consanguineous marriages carry an increased risk of genetic conditions due to stretches of homozygosity in the genome and result in significant health and economic burden. Many of these conditions are difficult to diagnose despite using comprehensive investigative methods. Next-generation sequencing offers the opportunity to understand the genetic causes of childhood neurogenetic disorders better. The main aim of this study was to obtain an in-depth understanding of the genetic causes of childhood neurogenetic disorders in Turkey.

The main objectives in this study:

1. Identify new disease genes and causal disease variant that underlie neurogenetic diseases (focusing on brain malformations) found in consanguineous families in Turkey

2. Identify modifier genes that affect the penetrance and expressivity of these neurogenetic diseases

Factors modifying expressivity and penetrance are still mostly unknown in many neurogenetic conditions. With the help of this study, it is expected that a significant proportion of cases will show differences that cannot be fully explained by simple Mendelian inheritance

3. Increase the proportion of families with a genetic diagnosis to enable provision of reproductive counselling and disease-specific therapies

As a result of this study, many patients were able to get diagnosed with precise disease-causing defects. Furthermore, many neurogenetic conditions have effective treatments available which can only be offered once the genetic diagnosis is known. Therefore, it is expected to provide a treatment for several patients when a certain diagnosis is completed.

4. Explain genotype-phenotype relations in detail through systematic deep phenotyping of all affected individuals

Precise phenotypic characterization of the neurogenetic conditions affecting consanguineous families in Turkey is crucial not only for gene discovery but also for extending knowledge on genotype-phenotype correlations as the foundation for further research into the social and economic burden of these conditions and appropriate healthcare provider.

5. Improve the biological understanding of recessive neurological disorders in children

Functional *in vitro* and *in vivo* studies of the many genes, pathways and proteins provided a better understanding of the underlying molecular and cellular pathophysiology. These contributions will help to identify new therapeutic targets and strategies in the future.

6. Create a comprehensive body of knowledge and data on a large cohort of Turkish consanguineous families accessible to authorized researchers worldwide as the basis for future research

Data collected with this project will form the comprehensive genomic and phenotypic cohort of Turkish consanguineous families characterized to date and it will contribute to the Turkish genomic variome, and identification of disease-associated variants.

7. Create a lasting legacy for Turkish genomic research and healthcare provision through knowledge transfer, training and education.

This thesis is a part of ongoing project called CONSEQUITUR. Patients were recruited from three centers in Turkey (Izmir, Diyarbakır, Malatya) in this study. The child with suspected neurogenetic disorders born into consanguineous marriages were included in the study. Consent forms were collected from all volunteers (index patients, parents, siblings) who agreed to participate in the study, and the use of all genome sequencing and data. Patients who agreed to participate in the study were phenotyped according to standardized phenotyping protocols by clinicians. Diseases included in the study are neonatal, infantile or childhood-related and primarily affect the brain or muscles, mental retardation (with or without epilepsy), cortical and other brain malformations, leukodystrophies, ataxias, myopathies and muscular dystrophies,

inherited neuropathies are congenital myasthenic syndromes and mitochondrial disorders. Inclusion criteria was being accessible for in-depth phenotyping and blood sampling studies for both index patient and family members.

Even though there were many good results, the project is still contiuning and there will be better results in the near future.

2. LITERATURE REVIEW

2.1. Consanguineous Marriages

in clinical perspective, a marriage between two persons associated with or closer to the second cousin is described as a consanguineous marriage. Marriages between third cousins, in other words; marriages between people who are more distantly related can also be included on some consanguinity reports. Even though this distinction in definition leads to change in the total consanguinity rate, the mean inbreeding coefficient is not significantly altered due to the lower inbreeding coefficient. Consanguineous marriages, also called endogamous marriages, often leads to an unequal distribution of founder mutations among populations [1].

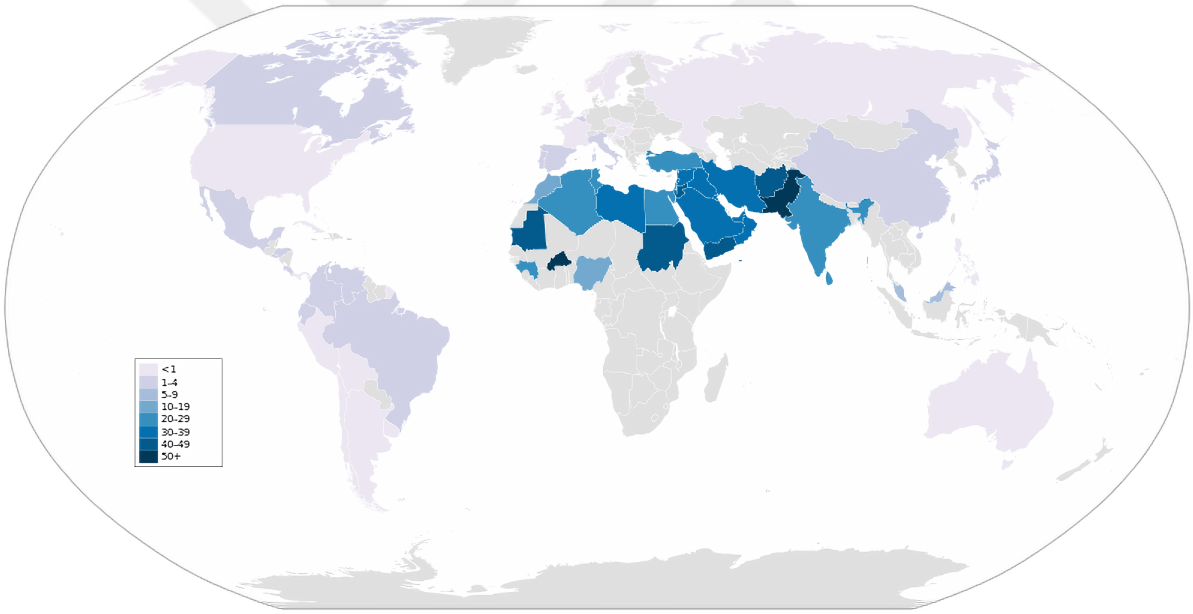

Figure- 1 A world map of the current consanguinity rates
(Retrieved December 10, 2018, from http:/www.consang.net)

The consanguinity prevalence and its rates show differences globally within and between populations based on geography, culture, ethnicity,and religion (Figure- 1) [1].

The availability of consanguineous relatives at similar ages, the similarity of social and economical conditions, and physical characteristics between relatives and

traditions for consanguineous marriage are some of the factors that affect consanguineous marriage possibility [2].

There are many different socio-economic reasons for consanguineous marriages around different regions of the world and sometimes consanguinity rate can be up to 70% [3]. Although consanguinity is essential for genetic research, many populations with high consanguinity rate are not investigated thoroughly in genetic and sociological aspects [4], [5]. In oder to benefit from these consangunious populations, big sequencing projects are needed. As a result, consanguinity studies can provide a great body of knowledge on human genetics [6].

Several ancient populations of founders showing little evidence of bottlenecks suggest that the tradition of consanguinity has led to an increased burden of recessive disease [7].

Increased expression of rare autosomal recessive genetic disorders is reported as the main effect of consanguinity. Therefore, it is believed that populations with high consanguinity rate offer a unique chance to identify genes inherited from recessively in neurogenetic diseases [8].

Mostly due to autosomal recessive diseases, the prevalence of congenital anomalies is predicted to be two, three times higher than the non-consanguineous population risk in the offspring of first cousin marriages [8].

The inbreeding coefficient symbolized by F is measured by the degree of inbreeding. F also indicates the likelihood of two identical alleles at a specific locus [9]. The extensive data set analysis with many parent–child trios demonstrates that descendants from first–cousin marriages show higher F values than descendants from non-consanguine marriages. In other words, extended homozygosity runs (ROHs) are seen in populations with high rates of consanguinity [7]. The increased length of ROHs enables the identification of homozygous loss of function variants. Therefore, potentially disease-causing variants in the consanguineous populations are more advantageous to solve recessive conditions [1].

Rare (<0.05) and very rare variants (<0.01) are enriched in longer ROHs. Conversely, common variants (all-frequency <0.05) are seen in shorter ROHs.

Therefore, it can be said that the longer ROHs are the consequence of consanguineous marriages that occurred recently [7].

i. Genetic Changes

Gene mutations can be classified as inherited mutations and acquired (or somatic) mutations in two main groups. Hereditary mutations in nearly each cell of the body throughout the life of a person are inherited from the parent. These mutations are also known as germline mutations, as they occur in the parent's egg or sperm cells, also known as germline mutations. If the DNA has a mutation in germ cells, the offspring who from this fertilized egg will contain the mutation in each of its cells [10].

Acquired (or somatic) mutations occur through a person's lifetime and are found solely in certain cells. Environmental factors can cause these changes, or somatic mutations can arise from a DNA copy error made during cell division. This kind of mutations in somatic cells are not inherited [11]. Genetic changes referred as *de novo* mutation may be either hereditary or somatic. The mutation occurs in egg or sperm cells in some cases, but is not present in any of the other cells [11].

Furthermore, somatic mutations occurring at an early embryonic stage in a single cell can lead to mosaicism. These genetic changes do not exist in the cells of any parent or fertilized egg but occur when the embryo increases the number of the cell. Cells with the altered gene prompt to many cells carrying mutation as cells dividing during growth and development, while other cells will not. Mosaicism can cause a disease but it depends on the mutation and affected cell number [11].

ii. Polymorphisms

Genetic changes in >1% of the population are referred to as polymorphisms (also called variation). They are common enough to be regarded as regular changes in the DNA. Polymorphisms account for many typical differences in the colors of the eyes, hair and blood. While many polymorphisms do not adversely affect a person's health, some of these differences can influence the risk of certain disorders [12].

The function of each cell is directly correlated with thousands of proteins to

function errorless. In many cases, mutations interrupt these proteins to function correctly [11]. By changing the instructions of a gene to make a protein, a variation may lead to the protein to malfunction or to be completely missing. If that specific variation alters a protein which is critical in the body, normal development can be disrupted.

The frequency of disease-causing (pathogenic) variants is expected to be lower than polymorphisms that are expected as regular changes. Therefore, pathogenic (disease-causing) variations are not common in the general population.

A genetic disorder can be described as a condition caused by one or more disease-causing variations (mutation). When genes essential for the development of these changes occur, the early stage of embryo development is often distorted. Because these mutations are serious, some are not consistent with life [13].

Mitochondria are rod-shaped organelles within the cells that generate energy in eukaryotic cells. Even though most DNA is packaged in nucleus, mitochondria contain 37 genes in their own DNA known as mitochondrial DNA or mtDNA. Hereditary changes in mitochondrial DNA may also cause growth, development and function problems of the body systems in some cases.

It is well-known that gene changes can alter the function of a protein in the body and cause health problems. It is now clear that changes in non-coding DNA can also cause to illness. Many regions of non-coding DNA have an effect on gene activity, determining when and where certain genes are turned on or off.

The non-coding DNA may affect health and development as the same genetic changes arise in the genes or change the chromosome structure. The changes in points, insertions, deletions and translocation include changes in points. These changes include DNA mutations that are not encoded can not be acquired or inherited by a parent during one's lifetime.

Non-coding regions of DNA and their role still mostly remains unknown. Therefore, it's very difficult to assess the effects of these genetic changes in non-coding DNA on certain genes and health conditions. The attempt to study of roles

and impact of genetic changes in non-coding DNA are increasing.

Mutations in non-coding DNA have been linked to developmental disorders [14]. Many cancer types were also associated with non-coding DNA mutations. Other regulatory elements, such as promoters, insulators, and silencer can disrupt these mutations in addition to enhancers. The disease is also resultes by mutations in areas providing guidance on the production of functional RNA molecules, such as RNA transport, microRNAs and long non-coding RNAs [15].

The discovery of functional genomic variants with pathogenic mutations represents an important step towards understanding phenotype physiopathology mechanism and paves the way for developing therapeutic interventions. In recent years, developments in genome sequencing technologies and the copy number variants (CNVs) identification by adapting comparative genomic hybridization (CGH) methods has provided unique opportunities for identification of a significant number of genes in consanguine families by associations between genotypes and phenotypes [8]. It is also possible to design a single nucleotide variant (SNV) array to complete this process much more efficiently [6].

Recent genetic advances can offer opportunities for comprehensive studies and research in highly consanguine populations is expected to reduce this gap to identify coding or regulatory factors, including those in complex diseases such and cancer, which affect human traits and diseases in general, autism, diabetes.

Homozygosity rate is high in the descendants of consanguine marriages. Autozygous stretches are parts of the genome where two alleles at a locus originate from a common ancestor by way of nonrandom mating in their genomes. Therefore, autozygosity is likely to lead to homozygous loss of function mutations resulted in gene inactivation [16], [17]. The study of offspring born to consanguineous marriages with clinical phenotypes helps to identify causal disease mutations. However, the majority of human genes currently has not been linked to any disorder. Furthermore, function of most of these genes is still unknown. The main reason behind that is the homozygous loss of function variants which are not seen in high rate in outbred populations. In addition to this, numerous genes do not lead a separate or defined

phenotype. However, the study of consanguineous populations can accelerate the characterization new genes and their function [6].

The human genome has a variability of about four billion DNA sequences according to Human Variome Project (www.humanvariomeproject.org/)[18]. The characterization of these variants is an important step to be considered in genetic studies. Likewise, reverse genetic studies performed in model organisms by inactivation of gene of interest and then examining its effect on certain phenotype, function of genes should be observed when they are entirely 'knocked out' to get a better understanding about gene function. Knockout studies by use of model organisms are preferred oftenly in order to get better understanding of the human genome and the genes taking a place in certain biological pathways. But model organism studies may be limited due to differences in the underlying mechanism. Also, some so-called 'orphan' genes may not even have homologues in commonly used model organisms. For these reasons, genes of interest obtained from knockouts studies cannot be used in a human model directly. The studies in consanguineous marriages are great alternatives to such model organism studies. Therefore, random sampling from a consanguineous population makes enable natural human knockouts to be identified. For this reason, consanguineous population studies are referred as a study of 'quasi-reverse genetics' aiming correlation genotype to phenotype [6].

Even they do not have distinct clinical phenotypes, offsprings born into consanguineous marriages should be analyzed as well to determine which genes contain homozygous loss of function mutations after clarifying the function of these genes, as cohort studies can be carried out to investigate any long-term effects. Molecular studies should also be conducted to observe changes in the expression of other genes. The mutated gene can be deleted; as a result, it could become a pseudogene. In addition to these, 'Knock out' can have protective effects in some genes on certain disorders or diseases.

The analysis of consanguineous populations helps to get some perspective on a gene-centred approach instead of a traditional disorder-based approach. There are >7500 Mendelian disorders exist and molecular basis of many of these diseases is

unknown. More than half of these Mendelian disorders are caused by autosomal dominant mutations. Autosomal recessive and X-linked inheritance lead to the rest of these diseases. Therefore, the effects of homozygous loss-of-function mutations in more than 10,000 genes have not been observed considering of 20,000 genes [6].

The rate of mutations discovered in autosomal dominant disorder mutations is much higher than the mutations identified in autosomal recessive disorders. It is explicitly indicated in the Hardy-Weinberg (H-W) equation showing the rate of heterozygotes (i.e., 2pq) in an outbred population will be higher than the causal variant homozygotes (2pq > q2, 2q for very low q). For this reason, it is expected to identify more autosomal dominant mutations in disease phenotypes [19].

The increased risk of expressing an autosomal recessive disorder in the offspring born into consanguineous marriage is inversely proportional to the disease allele frequency in the gene pool [20].

2.2. Autozygosity Mapping

Autozygosity mapping is a widely used powerful tool. It is a method for the identification of genetic loci in consanguine families for recessive disorders. The identification of causal mutations in many families is required to be convincing evidence [21].

There are homozygous regions in the human genome that have significant stretches or 'runs' of homozygosity (ROH). The length of these ROHs depends on the degree of paternal ancestry shared by individuals who are descended from consanguineous marriages or isolated groups. It is based on that individuals with a recessive disease are likely to have homozygous regions around the disease locus because of their identity-by-descent. Consequently, the ROH analysis shared by affected persons in the same family often helps to identify the gene that causes the disease.

- Haplotype Map

Genetic variants are often inherited together in the DNA segments known as

haplotypes. These ancient genomic segments are inherited as separate, generations-long, genetically transforming units. Because haplotypes are frequent in most people, genetic differences can be deciphered, which make some more susceptible than others to disease. Today, the patterns of common genome haplotypes are mapped by an International consultation co-ordinatedby national health institutes. HapMap is a major contributor in this global effort to create new data and to develop new analytical methods for the study of Haplotypes in the Broad Institute which plays a key role in generating new data and creating new analytical methods for studying haplotype information [22].

Single nucleotide polymorphisms (SNPs) are the sequence variants and they can be found in within a haplotype. In many cases, the genetic diversity of the entire block is characterized by redundancies in a particular haplotype. The main objective of the HapMap project is to identify sets of SNPs or tags that provide this capability of prediction. With the number of tags identified, the number of SNPs needed for genome study will decline. As a result, efforts to determine risk genes can be made cost-effectively for certain diseases [23].

2.3. Whole-Exome Sequencing (WES)

Next-generation sequencing provides the human genome can be sequenced fully and multiple genes can be analyzed [24]. Compared to traditional homozygosity mapped using previously reported SNPs, WES has added an advantage in identifying the disease-causing variantd. Therefore, WES can identify both the loci of the applicant and the genetic defect itself in a single step instead of using two different methods at nucleotide level. As there are many constraints, further bioinformatics development is needed in order to overcome these obstacles [21].

Whole- exome(WES) sequencing has been demonstrated an effective way of identifying pathogenic variants, especially in highly heterogeneous genetic conditions. WES covers more than 95% of the regions that harbour the majority of the genetic variants associated with human disease phenotypes. It makes possible to focus on the most relevant portion of the genome (the coding regions) and facilitates the discovery and validation of common and rare variants [25].

Sanger sequencing was the gold standard in molecular diagnosis of Mendelian diseases and continues to be the first choice to confirm a suspected diagnosis allowing accurate genetic counselling [23]. However, Sanger sequencing gene-by-gene is not the most economical or effective approach to genetic heterogeneity diseases. Classical approaches to unraveling several disease-causing genes have studied in monogenic disorders. Due to the cost and the time needed to screen every mutation by Sanger sequencing, the presence of many genes in a candidate genomic region is a limiting factor.

Mendelian diseases are caused by pathogenic variants in genes which comply with Gregor Mendel's original biological heritage laws [26]. The disease-causing gene can be automal, allosomal. In addition to these, genes can be inherited to the next generations as dominant or recessive. Even though inherited diseases are considered to be rare alone, these diseases can affect approximately 7.9 million children (born with a serious genetic congenital disability) collectively [27].

Marriage of a close family increases the likelihood that two deleterious copies of a recessive gene will be inherited by a descendant. Therefore, children from unbreathed couples are more likely to develop recessive autosomal disorders [28].

Furthermore, since alleles are parts of haplotypes, the affected descendant will not only have two identical copies of the ancestral allele, but also the surrounding haplotype segment will also be homozygous. The child will have this homozygous segment, so-called homozygotic runs (ROH) [29].

In general, these homozygous segments which are identical by descent (IBD) are longer in cases of consanguinity. However, ROH may be detected in geographically isolated populations and historical bottleneck events in the absence of known recent inbreeding [30].

Parental ancestry degree and age affect the length of the ROH. Recent inbreeding/parental consanguinity events tend to have longer ROH as there are fewer recombination events that interrupt IBD segments. In contrast, older ROHs are usually much shorter, as the homozygous stretches have been divided over

generations by repeated meioses, except genomic regions with lower recombination rates [29]. The mapping of homozygosity (also known as autozygosity mapping) is the identification of homozygotic regions in the genome. This strategy is powerful to link new genes to diseases [31], [32]. As mentioned above, affected persons are likely to have two IBD alleles on markers near the disease locus and are therefore homozygous for these markers. This method is based on the search for ROH, which is shared by affected persons of the same family. However, it requires sophisticated techniques such as the use of numerous microsatellite markers or high-density single nucleotide polymorphism (SNP) genotyping to scan the genome for blocks of homozygosity, although efficient and straightforward [33]. For most autozygosity mapping projects, there are many analysis software has been carried out using a recessive disease model. In general, haplotypes are manually inspected for homozygous regions that are shared by all affected persons and can be deduced to be IBD if there are genotypes from parents or other close relatives. However, in practice, conventional linkage analysis methods for large data sets in complex consanguineous families are often difficult due to the time and computational power required, affected individuals tend to have larger homozygous blocks (mean 4.4 Mb) and contain dozens or hundreds of genes [34].

NGS, and in particular whole-exome sequencing (WES), has led to considerable progress in the study of genetic diseases and genetic discovery. However, this technology still has certain limitations [35]. First, some variations in non–coding regions suspected to be deleterious cannot be detected by WES. Genetic and phenotypic heterogeneity in affected individuals makes it difficult to interpret exome sequencing [36]. The sequencing errors associated with poor capture efficiency and misalignment of repetitive regions can cause to misinterpretation of the resulting data. Therefore, it requires Sanger sequencing to validate the causative variants of the candidate. Choosing candidate variants involves a series of functional and deleterious effect prediction and allele frequency filtering [2]. This step is crucial because it limits the analysis of data and therefore influences the results. For example, nearly half the variants can be excluded because they are synonymous. However, it is well-known fact that many synonymous mutations can cause to disease phenotype

even though synonymous variants are not generally considered harmful [37].

Furthermore, during the data analysis, variants are usually filtered by frequency which is typically set by excluding variants with a low allele frequency above a certain threshold (more than 1% in most cases). Frequency filter should be used in regard to the disease prevalence proportion. The frequency of damaging alleles in populational variant databases can also be higher than the threshold set in some cases because in recessive disorders carriers do not show any signs of the disease. This means that this variant is incorrectly excluded during filtering. Studies conducted recently have showed the power and effectiveness of WES combined with homozygosity mapping in an attempt to identify causative genes in Mendelian disorders.

The combination of these two methods is favourable for discovering the causal variant regardless of the gene in question. The homozygosity map allows narrowing down of the target data set at first. After that, evaluation of these genes at the level of nucleotide make possible to identify candidate causal variants [24].

There are many tools available to perform ROH analysis. For instance, HomozygosityMapper [38] provides an online service to users for homozygosity mapping by analyzing NGS data interactively. Furthermore, PLINK [39] and GERMLINE [40] initially build up to analyze SNP array data, are ROH detection tools based on sliding-window algorithms [41]. A small frame of the sequencing data are measured in a sliding window analysis. At each new position, the window gradually moves throughout the region of interest and calculation is performed based on the reported statistics. Therefore, genome is scanned by a moving, fixed size window and differences or variations in genetic markers across the genome can be detected. This type of analysis reveals variation patterns across a genomic segment [42]. However, the sliding-windows approaches cannot be used easily with short/medium ROH sizes. To solve this issue, a new algorithm, H3M2, was proposed [43]. That is capable of detecting smaller ROH. AgileGenotyper and HomSI are some of other tools that can be used for the graphical visualization of ROHs.

Neurogenetic disorders are a large and heterogeneous disease group caused by

changes in one or more genes have a role on neuroectoderm and its derivatives [44]. This group of diseases can cause abnormalities on brain, muscle, nerves, and spinal cord.

2.4. Cortical developmental malformations (CDMs)

2.4.1. Brain Development

Human brain cortex development is a complex dynamic process. In developmental process, the differentiation of neuroepithelial cells begins at a gestation of about 26 days in the subventricular layer. The larger pyramid cells are born first and differentiate early to serve as targets for nervous system migration [45]. In the forebrain, ventricular and subventricular areas lining the cerebral cavity during the first stage, stem cells proliferate and differentiate genuinely into young neurons or glial cells [44], [46]. In the second stage, cortical neurons migrate away from their place of origin: most cells migrate from the periventric region to the pial surface along the radial glial fibres, where each successive generation passes through and settles inside the cortical plate in an inner pattern. When neurons reach their final destination, they stop migrating to specific 'architectural' patterns that guide cells in the cerebral cortex to the right position. This third phase involves the final organization of the typical six cortex layers (Figure- 2) associated with synaptogenesis and apoptosis [47].

The development of the brain cortex of humans is an orchestrated process involving neural progenitors in periventric germ zones, cell proliferation with symmetric and asymmetric mites, followed in six highly ordered, functionally specialized levels by the migration of post-mitotic neurons into their final destinations, which require precision of multiple genetic and epigenetic interaction depending on regulation of cellular and molecular mechanisms. The distruption of these mechanisms leads to several of neurodevelopmental phenotypes depending on the spatial and temporal effects of the disorder [48], [49].

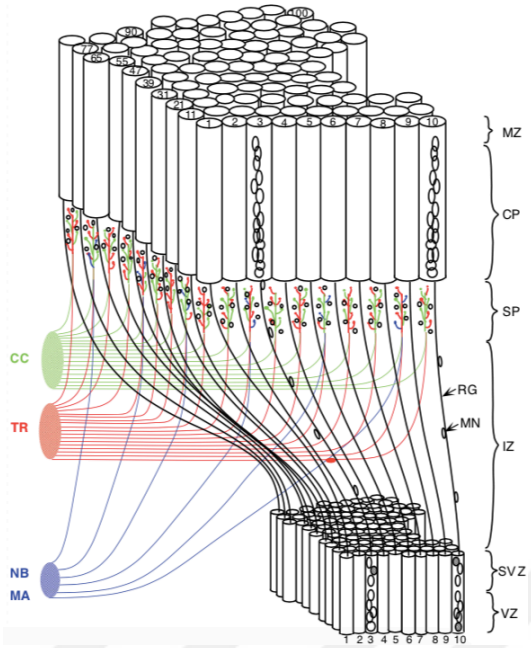

Figure- 2 Organization of the six cortical layers

(VZ: ventricular zone, SVZ: subventricular zone, IZ: intermediate zone, SP: subplate, CP: cortical plate, MZ: margizal zone, CC: contralateral corticocortical, TR: thalamic radiation, MN: migrating neurons, RG: radial glial cells, and subcortical centers; NB: nucleus basalis, MA: monoamine)

Cortical malformations are often has recognized with severe consequences, including mental delay, epilepsy, paralysis and blindness. This recognition is mainly due to improved magnetic resonance imaging (MRI) resolution, which enables the distribution and depth of cortical sulci, cortical thickness, the boundaries between grey and white matter and signal intensity variations to be assessed. Abnormalities of any or all of these characteristics can be observed in various malformations of cortical developmental (MCDs), which can be limited to discrete cortical areas or can be diffuse [44].

A classification scheme was developed that categorizes MCD into three major groups that summarize the main developmental steps as cell proliferation, neuronal migration or post-migration cortical organization and connectivity malformations. However, new evidence suggests that MCD is much more heterogeneous than suggested by this classification. To date, over 100 genes have been linked to one or more MCD types. The biological pathways include multi-step cell-cycle regulation (especially mitosis and cell division), apoptosis, cell-fate specification, cytoskeletal

structure and function, neuronal migration and basement-membrane function and many inborn metabolism errors. A subset of MCD genes is associated with post-zygotic (i.e. mosaic) mutations, especially those associated with megalencephaly. Genetic testing requires accurate imaging and family distribution assessments, if any, and may be straightforward in some disorders, but requires a complex diagnostic algorithm in other disorders. Due to the substantial genotypic and phenotypic heterogeneity in most of these genes, a comprehensive analysis of clinical, imaging and genetic data is required to define these disorders correctly. Exome sequencing and ultra-high MRI fields change the classification of these disorders rapidly [44], [50].

Neuronal development can be separated into three main processes: neurogenesis, neuronal migration and cortical organization and circuit formation after migration. Therefore, cortical developmental malformations can be classified into three major groups according to the main underlying mechanism (Figure-3) [49], [51].

As is to be expected a number of conditions fall outside these three groupings and are placed in a miscellaneous group: those associated with other diseases or yet to be classified. The classification present below is a summary with examples adapted from the one proposed by Barkovich et al. in 1996 and updated in 2001, 2005 and 2012 [51]. The full classification is extensive, contains numerous rare and specific abnormalities beyond the scope of this or other related articles and will undoubtedly evolve over the coming years.

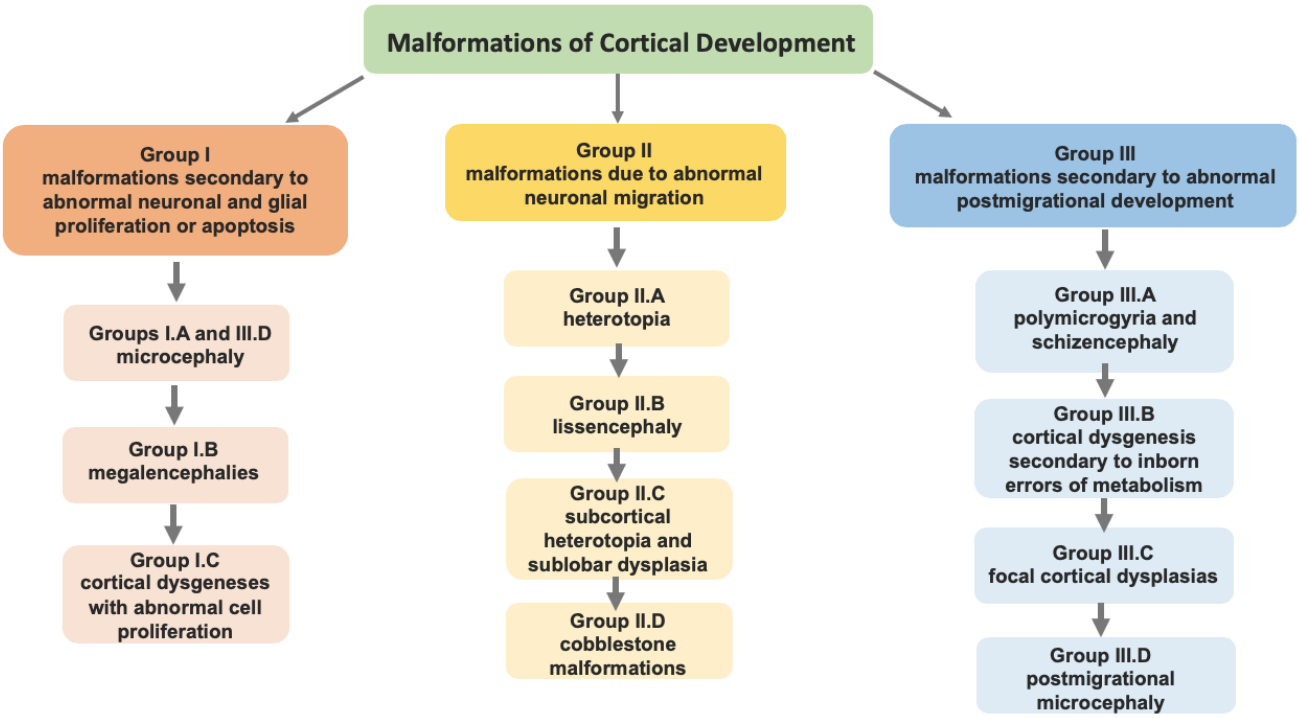

Figure- 3 Malformations of Cortical Development

The knowledge of the genetic and molecular basis of these malformations is progressing quickly with often reported new genes, which shows that the various cortical malformations are secondary to anomalies interfering with different stages of development. In the same experiment, screening of several genes has also the valuable advantage of improving new genotype-phenotype correlations and identifying possible genetic causes of MCD which are still not diagnosed by molecules. However, the next sequence analysis can cover many variants with sometimes difficult interpretation, depending on the number of genes screened in each experiment. In order to overcome misinterpretations of the pathogenic potential of variants can be avoided by assessing their interpretation in the context of each patient's phenotype [44], [50].

The malformations of cortical development classification should be ideally based on the knowledge of biological pathways which are not available at currently. Recently identified wide range of malformations with mutations in *WDR62, DYNC1H1*, and *TUBG1* genes, in fact, demonstrate the fading boundaries of subsequent cortical

organization, migration, or neuronal proliferation disorders [50].

Detailed descriptions of MCD are given in clinical papers and are of great conceptual significance. For the most common of these disorders, these problems have been dealt with both from a general and a more specific phenotypeAlthough the MCD classification has advanced considerably during the past decade, only a few categories, including lissencephaly, polymicrogyria, schizencephaemia, focal cortical dysplasia (FCD) and nodular heterotopia periventricular are used in practice. But new evidence suggests that MCD is much heterogeneous, particularly in situations where irregular or galvanised cortical surfaces have been classified as polymicrogyria, although typical curvilinear microsulcans have not been seen [50].

Cortical malformations are a number of structural brain disorders that are often devastating, and reflect neuronal growth, migration or organization. The use of traditional mapping methods has been particularly challenging to discover genes in these syndromes, where only one member of family members has been the most common. Recent development in next generation sequencing technologies can drastically increase the gene discovery in cortical malformations.

2.4.2. Group I: Malformations secondary to abnormal neuronal and glial proliferation or apoptosis

Microcephaly is the most common physical finding in intellectually disabled children in general. Historically, different and often confusing terms have been used to describe different types of microcephalus, such as 'primary microcephalus' and 'microcephalus vera' but the most useful classification of microcephalus from an etiological and clinical point of view is congenital and postnatal. Congenital microcephaly involves an early and severe brain developmental defect, most commonly the mitotic microtubule spindle apparatus and centrosome, as well as DNA damage response and repair proteins. Many of these syndromes include global growth dysregulation with a delay in prenatal or postnatal growth and cancer predisposition in some cases. It is therefore prudent for a child with microcephaly to be assessed thoroughly as soon as possible by a geneticist. The majority of congenital

microcephalus syndromes have an autosomal recessive pattern. Postnatal microcephaly, on the other hand, is much more genetically and clinically heterogeneous, often a challenge for geneticists and neurologists to diagnose. These syndromes have a wide range of neurodevelopmental and neurobehavioral characteristics that are often underrecognized in many functional areas such as pain tolerance, mood, food behaviour, gastrointestinal motility, hand function, breathing rhythms. A thorough understanding of these characteristics is essential for accurate diagnosis and genetic testing [52].

Brain overgrowth disorders, highlighting the exciting recent discovery of hemi-megalencephalic genetic causes; severe overgrowth of the brain often linked to intractable early epilepsy and deep developmental delays. The identification of mosaic mutations in the PI3K–AKT pathway in hemimegalencephaly and other focal megalencephaly phenotypes corresponds to the PTEN pathway activation disorders and increases the exciting potential for future pathway–based therapies [52].

The abnormally large brain is referred to as megalencephaly [44] that exceeds the age and gender [46]. Megalencephaly has mostly been classified as a brain-size disorder, but recent studies have shown that normal cortex megalencephaly can be a result of mutations in the PI3K-AKT-mTOR pathway genes [52]. Usually, patients were born with large heads and head can continue to grow in the first few months of life [45].

Hemimegalencephaly and FCD constitute a spectrum of cortical development malformations with common neuropathological characteristics. The former is mainly defined by macroscopic enlargement of (more or less) one hemisphere, whereas FCD is defined primarily by histopathology [44]. As currently classified, FCD includes a wide range of cortical malformations with variable characteristics, including neuronal microscopic heterotopia, dyslamination and abnormal cell types. The FCD was divided into three main types and nine subtypes based on the histopathological characteristic [50].

Another example of a defective outgrowth of neuritis in humans is the failure of cortical axons to cross the midline leading to the absence or agenesis of the corpus

callosum. ACC can be found in an impressive range of cortical dysplasia, leading to several gene defects. However, ACC discusses the importance of the gene involved in neurite growth when present and is, therefore, a valuable indication for the selection of candidate genes [53].

2.4.3. Group II: Malformations due to abnormal neuronal migration

The most common cortical brain malformations, beginning with lissencephaly, subcortical band heterotopia and a growing family of tubulin-related malformations (tubulinopathies). In addition to *LIS1*, the most common cause of post-predominant lissencephaly, more than ten genes are now associated with various patterns of lissencephaly, including *DCX, ARX, RELN, VLDLR* and many tubules. There have been exciting advances in lissencephaly research in recent years, as animal therapeutic studies have shown that phenotypes can be rescued in Lis1 and Dcx1 mutant animal models [52].

i. Heterotopia

There are three main groups of heterotopias: periventric (usually nodular: PNH), subcortical and leptomeningeal (glioneuronal heterotopia found over the brain surface), only the first 2 of which can be detected through imaging. PNH is by far the most common [44]. While several hypotheses have been proposed, the exact mechanism for aberration is not defined yet. These include damage to radial glial fibers, premature glial cell conversion into astrocytes or lack of specific neuroblast or radial glial cell surface molecules that perturb normal migrations [54]. Heterotopias are often found to be isolated defects which can only lead to epilepsy.

The high- resolution imaging showed that displaced neurons or heterotopies are more prevalent than was once thought, and that up to 25% of childhood epilepsy is due to heterotopia [53].

The result of heterotopia is often thought to be interference with neuronal migration mechanisms that cause serious global neurological impairment during brain development. These malformations have been associated with mutations in the *LIS1, ACTB, ACTG1, ARX, CDK5, DCX, KIF2A, KIF5C, RELN, TUBG1, VLDLR* genes. Studies

more recently have also shown a relationship between lissencephaly, with or without associated microcephaly, corpus callosum dysgenesis and cerebellar hypoplasia, and sometimes a morphological pattern consistent with polymicrogyria with multiple gene mutations (*TUBA1A, TUBA8, TUBB, TUBB2B, TUBB3 and DYNC1H1*) that regulates the synthesis and function of microtubules and centrosomes. MCD affecting only neuron subsets, such as mild subcortical band heterotopia and periventricular heterotopia, has been linked to gene abnormalities of *DCX, FLN1A and ARFGEF2* and causes neurological and cognitive impairment from severe to mild deficits [44], [53].

ii. Lissencephaly

Lisbonencephaly refers to the external appearance of the cerebral cortex in disorders in which a neuronal migration aberration causes the cortical surface to be relatively smooth. The full spectrum includes agyria and pachygyria, the true abnormality being the large heterotopic neuron under the cortical surface. The lack of cortical attraction because of incorrect axon paths results in abnormality to develop gyri or sulci in this disorder [45], [53].

iii. Cobblestone Complex (Lissencephaly)

The cobblestone complex involves the migration of young heterotopic neurons into the outer basement membrane across the marginal zone-future layer I. This may be due to changes in the extracellular matrix, a hypothesis supported in Fukuyama 's congenital muscular dystrophy by the identification of fukutin. The neuron movement into leptomeninges can be destructive to the subarachnoid space and lead to an impaired absorption of the cerebrospinal fluid causing hydrocephalus or ventricular expansion. At first glance, polymicrogyrias can hardly distinguish the magnetic resonance imaging properties of cobblestone lissencephaly [53].

2.4.4. Group III: Malformations secondary to abnormal postmigrational development

The abnormality of neural cells is known in neuronal position disorders involving brain syndromes of cortical malformation. Cell movement can be distrupted in many respects by movement in the ventricular area from its origins.

A disturbed cell proliferation, destination determination and programmed death

of the cell can alter the primary neurogenesis or number of cells. If certain cells do not distinguish or if their birth times are not correct, the destiny and position of other cells in the region can be altered. Ventricular cell migration may be reduced through interferences with mechanical motors and cytoskeletal cell dynamics. In addition, the molecular signals initiating motion can guide the migration of the cell and inform it that it has reached the final position. The position of the cellular body further developing and refining neurite/axonal projection and synapses, partly working on neuronal survival and related glia [53].

Molelecules that affect neuronal migration are highly interesting recently, as their actions may not be limited to cell movements. Through neurogenesis, early and late migration, axon projection and guidance, all events organizing the brain structure are expected to overlap. Therefore, caution should be taken until the function of the responsible gene product is established that malformation is classified as a neuronal migration syndrome.

For moving and organizing cells in the developing cortex, at least four requirements must be met. Cells must receive 'go' signals at first. Adhesive and contractile elements must then be coordinated to ensure strong cell front adhesion and weak adhesion at the back to net cell motion, as cytoskeletal elements contract. Thirdly, the direction of movement must be determined by the signal. Finally, cells must be informed of 'stop' signals when they reach their final destination. Genetic mutations can cause many problems by disrupting neuronal migration. Early indications of mutations in the *reeler* and Kallmann syndrome suggested that extracellular matrix molecular molecules with selective brain expression are the unique features of brain neuronal migration. The incredible insights from the changes to human life is the distinctive character of the intracellular cytoskeletal machinery of migratory neurons. To date, more than a dozen neuronal migration molecules have been reported in the brain. At least half of these were first identified in clinical human studies.

i. Polymicrogyria

Polymicrogyria (PMG) is still a complex malformation that is often misunderstood

and often misdiagnosed. The word polymicrogyria defines the abnormally small quantities of gyri that produce a lumpy, irregular cortical surface. Polymicrogyria is an abnormal late-cortical structure that is permanently associated with an abnormal neuronal move. PMG is emphasized with its heterogeneous genetic and non-genetic aetiology and related syndromes [44].

While polymicrogyria occurs as a single malformation most often, it can be associated with several other brain defects such as microcephaly, megalencephaly, heterotopias of gray matter, ventriculomegaly, pellucidal abnormalities of the septum, corpus callosum, brainstem, and cerebellum. The edges of the cortex may seem to fuse or remain at a distance when schizencephaly is present [55], [52].

ii. Schizencephaly

A unillateral or bilateral split within the cerebral cortex extending through the cortex to the lateral ventricular surface is a feature of schizencephaly (split-brain). This may appear as a large cleft (open lip) or as a narrow groove (closed lip). Gen mutations that can interfere with neuronal migration are increasingly common. The pieces of this complex puzzle are demonstrated by studies stimulated with human clínical syndromes and animal brain malformation models and fascinating patterns. These genes have many functions and have limitations between primary neurogenesis, survival of cells, migration, extension of the rite, axon-pathology and synaptogenesis. This basic knowledge helps to develop strategies to repair and regenerate brains affected by ageing, trauma, malfunction, stroke, and immunity [53].

3. MATERIALS AND METHODS

3.1. Type of the Study

This study consisted of a group of inherited Mendelian neurogenetic disorders. Study involves obtaining blood sapmles, extraction and analysis of DNA, collection of associated phenotypic data, and functional studies in cell culture.

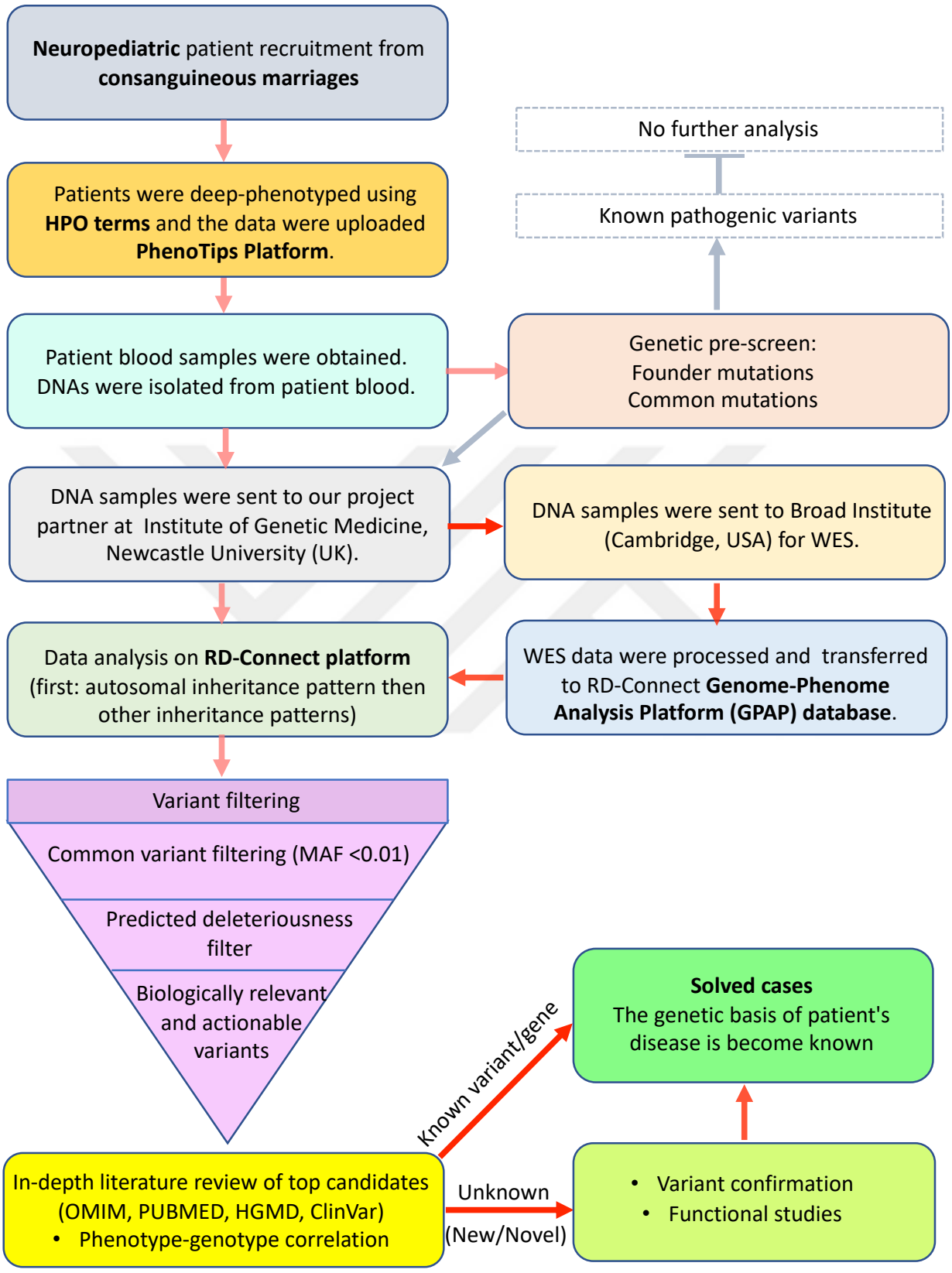

Figure- 4 The approach of the study

3.1.1. *Phenotyping of families and biobanking of samples*

i. Children from three provinces in Turkey (Izmir, Diyarbakır, Malatya) with suspected neurogenetic disorders born into consanguineous marriage are included to this study.

ii. Diseases included in the study are neonatal, infantile or childhood-related and primarily affect the brain or muscles, mental retardation (with or without epilepsy), cortical and other brain malformations, leukodystrophies, ataxias, myopathies and muscular dystrophies, inherited neuropathies are congenital myasthenic syndromes and mitochondrial disorders.

iii. Phenotyping of patients who agreed to participate in the study was performed according to the standardized phenotyping protocol. In-depth phenotyping of patients included in the study was performed by clinicians and these phenotype data forms were transferred to the PhenoTips database to be coded through the PhenoTips software in order to make them computer-readable and matchable within the dataset of this project.

iv. Consent forms were collected from all volunteers (index patients, parents) on the use of all exom sequencing and data.

v. Inclusion criteria for this study was being accessible for in-depth phenotyping and blood sampling studies for index patient and family members.

vi. Index patients were deep phenotyped using standardized data collection forms including detailed clinical and laboratory examinations, e.g. neurophysiology and MRI.

vii. All participants were asked to contribute blood samples for DNA and a small skin biopsy for fibroblast culture for some.

viii. Blood samples were collected in EDTA tubes. All names of volunteers were kept confidential and the samples were shared by IDs and barcodes assigned to them. The assigned IDs were followed by the letter F (family) followed by 3 letters indicating the province. FMAL for Malatya, FIZM for İzmir and FDIY IDs for Diyarbakır were used and these IDs were followed with numbers indicating the family number.

ix. Samples from all 3 clinical sites were coded and processed, and biobanked in IBG.
x. DNA samples were quality-controlled and provided for sequencing at the Broad Institute in batches.
xi. All samples went undergo quantification, purity and integrity assays, and genotyping with a panel of SNPs, which allows immediate confirmation of sample identity and rapid flagging of any sequence data discordant with genotyping data.

Patients were classified according to following disease groups;

a. Intellectual disability with epilepsy
b. Intellectual disability without epilepsy
c. Cortical and other brain malformations
d. Leukodystrophy
e. Spinocerebellar ataxias
f. Spastic paraplegia
g. CMT and hereditary neuropathies
h. Muscular dystrophies and myopathies
i. Mitochondrial disorders
j. Cerebellar hypoplasia or atrophy
k. Other phenotypes

3.1.2. *Whole-Exome Sequencing (WES) and Variant Confirmation*

i. Whole exome sequencing (WES) was performed on affected patients, and both parents by the MIT Broad Institute and the Harvard Genomics Platform (Cambridge, MA, USA) using > 250 ng DNA.
ii. Illumina Exome Capture Kit (38 Mb target) is used at the Broad Institute of MIT and Harvard.
iii. Sequencing data were processed at the Centro Nacional de Análisis Genómico (CNAG), Barcelona, and data analyzes carried out on the RD-Connect Genome-Phenome Analysis Platform (https:/platform.rd-connect.eu/genomics) using

standard filtering criteria for rare diseases, including Minor Allelelele Frequency (MAF) < 0.01, Variant Effect Predictor (VEP)=mod / high.

iv. Variants with sub-optimal quality scores were excluded.

v. The remaining variants were compared to the list of reported human gene mutations in the database.

vi. Variants were prioritized with a lower allele frequency of less than 0.1% for obviously protein-altering variants (missense, nonsense, splice sites, and insertions or deletions) and were evaluated for conservation and predicted deleterious effects by several bioinformatics algorithms (CADD, SIFT, Mutation Taster, and PolyPhen-2).

vii. Several variants have been molecularly modelled to further detect harmful protein effects. Gene functions, their pathways and their clinical potential have been examined through a comprehensive literature review and patient phenotyping.

viii. Family segregation patterns have been evaulated to identify expected inheritance modes and to consider the similarity of identified phenotypes to those described in previous reports.

All putative causative alleles were subject to extensive literature and database search and the results were discussed by PIs and doctors from three centers in Turkey during teleconference sessions. Due to ambiguous records in databases or literature, this review sometimes led to the reclassification of the variant status. Physicians with experience in medical genetics have carried out clinical family characterisation.

The study was endorsed by Dokuz Eylül University Ethical Review Board, with the approval number 2016/03-01. Each patient and family members were included in the study only after their written consent was obtained and data analysis has been carried out from November 2017 to the present day. Human Phenotype Ontology has been used to record patient's phenotype. We recruited 192 families with at least 1 offspring with suspected neurogenetic diseases from which the parents are consanguineous.

3.2. Time and Place of the Study

This multi-national study was mainly conducted in IBG between September 2016 to until today. However, patients recruited to this study were from three centers in Turkey: Izmir, Malatya and Diyarbakır. Patient blood samples were obtained in these three centers. After that, samples were coded and processed, and stored in IBG.

Their phenotypic data was uploaded to the PhenoTips Platform (https://platform.rd-connect.eu/phenotips) (Online).

DNA samples of patients were sent to Institute of Genetic Medicine, Newcastle University (Newcastle upon Tyne, UK) in oder to send and perfom WES in Broad Institute (Cambridge, MA, USA).

Sequencing data were processed at the Centro Nacional de Análisis Genómico (CNAG) (Barcelona) and data analyzes carried out on the RD-Connect Genome-Phenome Analysis Platform (https:/platform.rd-connect.eu/genomics) (Online)

Some functional studies were carried out at Newcastle University, Center for Life (Newcastle upon Tyne, UK)

3.3. Study groups of the Study

This study was practised with collaboration of different disciplines (Figure-4).

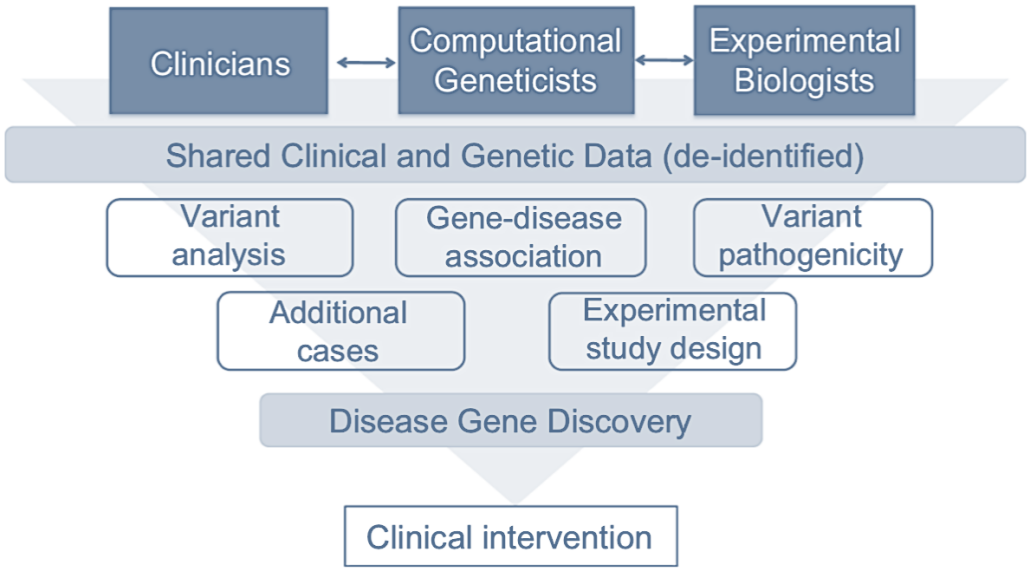

Figure- 5 Workflow of the Study Groups

Adapted from "An integrated clinical program and crowdsourcing strategy for genomic sequencing and Mendelian disease gene discovery", by Haghighi et al., 2018, *NPJ genomic medicine*, 3(1), 21[56].

3.4. Materials and Method

3.4.1. Manual DNA Isolation

A. Materials

To extract DNA from patient blood samples, following materials were used.

a. Reagents

- Tris-HCL
- Sucrose MgCl2
- Triton X-100
- NaOH
- EDTA NaCI
- SDS
- Sodium perchlorate

b. Equipment and Consumables

- 50 mL sterile centrifuge tubes
- 15 mL sterile centrifuge tubes
- 1.5 mL sterile Eppendorf tubes
- 500 mL and 2 L glass bottles
- pH meter
- Disposable plastic pipette tips

B. Methods

Buffers used in manual DNA isolation protocol were prepared as follows.

i. *Preparation of Buffer A*

To prepare 1 L of Buffer A, following reagents were weighted and added to a glass bottle:

- 10 mM Tris-HCL 1.21 g
- Sucrose 109.55 g
- 5mM $MgCI_2$ 1.02 g
- Distilled H_2O (dH_2O) 650 mL

After that, initial pH was measured (calibrate with pH 7 and pH 10 standards). The buffer pH was adjusted to 8.0 with 40% NaOH (mix carefully between each pH measurement). Next, the solution was autoclaved. Once autoclaved, Triton X-100 (10 mL for 1 L) was added to the solution and mixed well by inverting gently. The buffer A was stored at 4°C.

ii. *Preparation of Buffer B*

To prepare 500 mL of Buffer B, following reagents were weighted and added to a glass bottle:

- 400 mM Tris-HCL 15.25 mL
- 0.5 M EDTA 30 mL
- 150 mM NaCI 2.20 g
- Distilled H_2O (dH_2O) up to 250 mL

After reagents were added, initial pH was measured (calibrate with pH 7 and pH 10 standards). The buffer pH was adjusted to 8.0 with 40% NaOH (mix carefully

between each pH measurement). Next, the solution was autoclaved. Once autoclaved, 1% SDS (25 mL for 250 mL) was added to the solution and mixed well by inverting gently. The buffer B was stored at room temperature.

iii. *DNA Extraction*

Before initiating protocol, 50 mL aliquot of chloroform was chilled at -20°C and water bath was set to 65°C.

1. Cell preparation from whole blood (5-10ml)

50 mL sterile centrifuge tubes were labeled for each patient. Patient blood (1-2 mL) sample was thawed in the hand and poured from vacutainer into a 50 mL sterile centrifuge tube as soon as possible. 20 mL of Buffer A was added to the 50 mL centrifuge tube. Next, the tube was shaked for 4 minutes at room temperature (RT). After that, it was centrifuged for 10 minutes at 3,000 rpm. After the centrifuge, supernatant was discarded and cell pellet was resuspended in 10 mL Buffer A. 4 minutes shake and 10 minutes centrifuge steps were repeated with same parameters.

2. Cell Lysis

1 mL Buffer B (SDS completely dissolved) was added onto the cell pellet. After that, the tube containing cell pellet was vortexed and resuspend. The suspension was transferred into a 15 mL centrifuge tube. 250 µL sodiumperchlorate was added onto that suspension and shaked for 10 minutes at RT. Then, the suspension was incubated for 25 minutes at 65°C in water bath. The suspension was quickly vortexed every 5 minutes.

Before proceeding next step, centrifuge was set to 4°C.

3. Extraction

1 mL chloroform was added to the suspension. Then, the suspension was shaked for 10 minutes at RT. After that, the suspension was centrifuged for 10 minutes at 3,000 rpm (4°C). In the mean time, 15 mL centrifuge tubes with 3mL ethanol (98%), and 1.5 mL Epp tubes with 250 µL ethanol (70%) were prepared for each patient. The supernatant was transferred into a 15 mL centrifuge tube containing 3 mL ethanol (98%) and mixed gently by inversion until the DNA precipitate 'Medusa' (white, thread-like formation) was visible. The 'Medusa' was captured by using a P200 micropipette and transfered into a 1.5 mL Epp tube containing 250 µL ethanol (70%). The tube was

spun down for 2 minutes at 13,000 rpm using Eppendorf microcentrifuge. Then, ethanol was removed by pipette. Next, the pellet was left to dry.

The pellet was dissolved in 50 µL of TE buffer and left overnight at 4°C. Next day, the tubes containing DNA samples were heated at 54°C for 1 hour to make sure DNA was fully dissolved. Finally, the DNA concentration was measured by Thermo Scientific NanoDrop 2000.

4. Storage

The tube containing patient's DNA sample was labeled with patient ID, date and DNA concentration for each patient and tubes were stored at -20°C.

3.4.2. DNA isolation with a kit

In this study, DNA extraction could not be performed by manual DNA extraction protocol at times. To overcome this problem, MN Genomic DNA purification with Nucleospin Blood L kit (REF: 740954.20) was used now and then.

i. Before starting a NucleoSpin Blood protocol, the followings were prepared according to the kit manual.

• Wash Buffer BQ2: The indicated volume of ethanol (96–100%) (48 mL for this kit) was added to Wash Buffer BQ2 Concentrate (12 mL). The label of the bottle was labeled to indicate that ethanol was added. Wash Buffer BQ2 was stored at room temperature (18–25 °C).

• Proteinase K: The indicated volume of Proteinase Buffer PB (1.35 mL) was added to dissolve lyophilized Proteinase K (30 mg). Proteinase K solution can be stored at 20 °C for up to 6 months.

ii. Before starting a NucleoSpin Blood protocol, the following conditions for the experiment were made ready.

- An incubator or a water bath was set to 56 °C.
- Elution Buffer BE was heated to 70 °C.
- A centrifuge with a swing-out rotor and appropriate buckets capable of reaching 4,000–4,500 x g was required.

iii. DNA isolation

2mL blood sample was thawed and added into a 15 mL centrifuge tube (not provided with the kit). 150 µL Proteinase K was added onto that blood sample. 2 mL

Buffer BQ1 (if processing less than 2 mL blood, add one volume of Buffer BQ1) was added to the samples and then, the mixture was vortexed vigorously for 10 s. Vigorous mixing was especially important to obtain high yield and purity of DNA. After that, samples were incubated at 56 °C for 15 min. Later, the samples were cooled down to room temperature before proceeding with addition of ethanol. The lysate should become brownish during incubation with Buffer BQ1. The incubation time with Proteinase K was extended to 15 min and vortexed every 5 minutes during incubation to increase the yield from older blood samples.

To adjust DNA binding conditions, 2 mL ethanol (96–100 %) was added (if processing less than 2 mL blood, add 1 volume of ethanol) to each sample and mixed by inverting the tube 10 times. High local ethanol concentration must be avoided by immediate mixing after addition. The lysate can only be loaded onto the column if the lysate has cooled down to room temperature. Loading of hot lysate may lead to diminished yields.

In order to bind DNA to the column, one NucleoSpin Blood L Column was placed in a collection tube for each sample. Later, 3 mL lysate was loaded onto the column. The rims of the columns must be kept dry and clean. The tubes were closed with screw caps and centrifuged 3 min at 4,500 x g. If there was any, remaining lysate was loaded onto respective NucleoSpin Blood L Column. It was centrifuged 5min at 4,500xg. The flow-through was discarded and the column was placed back into the collection tube. Next, the collection tube was removed with the column carefully from the rotor to avoid that the flow-through comes in contact with the column outlet. Any spilled lysate from the collection tube was wiped off before placing the column back.

In order to wash silica membrane, 2 mL Buffer BQ2 was added onto the column. The column was centrifuged for 2 min at 4,500 x g. It was not necessary to discard the flow-through after the first washing step. In order to perform second wash, 2 mL Buffer BQ2 was added to the column. It was centrifuged for 10 min at 4,500 x g. The column was removed carefully from the rotor in order to avoid that the flow-through comes in contact with the column outlet. By extending centrifugation time to 10 minutes during this second washing step, residual ethanolic washing Buffer BQ2 was removed from the silica membrane of the NucleoSpin Blood L Column. In addition to

this, prolonged centrifugation time helped to dry silica membrane.

In order to elute highly pure DNA, the column was inserted into a new collection tube (15 mL) and 150 µL preheated Buffer BE (70 °C) was applied directly to the center of the silica membrane. The tubes were incubated at room temperature for 2 min. After that, they were centrifuged at 4,500 x g for 2 min. Later, the concentration and purity of DNA samples were measured by Thermo Scientific NanoDrop 2000.

3.4.3. DNA clean-up procedure

When the purity of DNA samples was lower than the expected, clean-up procedure was applied by using MN Genomic DNA Clean-up Kit (REF: 740230.50).

All kit components were stored at room temperature (18–25 °C). Storage at lower temperatures may cause precipitation of salts. When the precipitation was occured, the bottle was incubated for several minutes at about 30–40°C and mixed well until the precipitate was dissolved.

Before starting any NucleoSpin gDNA Clean-up protocol, Wash Buffer DW was prepared as following. The indicated volume of ethanol (96–100 %) (60 mL) was added to Buffer DW Concentrate (25 mL). The label of the bottle was marked to indicate that ethanol has been added. Buffer DW was stored at room temperature (18–25°C).

To adjust DNA binding conditions, 450 µL Binding Buffer DB was added to 150 µL DNA solution. Later, it was vortexed for 5 s. If sample volume was less than 150 µL, the rest was filled up with water to 150 µL. If more than 150 µL of sample had to be processed, Binding Buffer DB was increased proportionally. In order to do that, multiple loading steps were necessary.

In order to bind DNA to the column, a NucleoSpin gDNA Clean-up Column was placed in a Collection Tube (2 mL). Up to 700 µL sample solution was loaded onto the column. It was centrifuged for 30 s at 11,000 x g. The flow-through was discarded and the column was placed back into the collection tube.

To wash silica membrane, 700 µL Buffer DW was added to the NucleoSpin gDNA Clean-up Column. The lid was closed and vortexed for 2 s. After that, the column was centrifuged for 30 s at 11,000 x g. The flow-through was discarded and the column was placed back into the collection tube in first wash. For the second wash, 700 µL

Buffer DW was added to the NucleoSpin gDNA Clean-up Column. The lid was closed and the column was vortexed for 2 s. Then, the column was centrifuged for 30 s at 11,000 x g. The flow-through was discarded and the column was placed back into the collection tube.

In order to dry silica membrane, the column was centrifuged for 1min at 11,000xg and then, the collection tube was discarded.

In order to elute DNA, the NucleoSpin gDNA Clean-up Column was placed into a new 1.5 mL microcentrifuge tube (not provided). 50 µL Buffer DE was added to the column and then, the column was incubated for 1 min at room temperature (18–25 °C). After that, the lid was closed and the column centrifuged for 30 s at 11,000 x g. Heating the elution buffer to 70 °C increased the elution efficiency in at times.

3.4.4. Data analysis with RD-Connect platform

Data analysis was carried out by using RD-connect platorm. The results were uploaded to the European-funded consortium, the RD-Connect platform for further bioinformatics analysis. This consortium has been a pioneer in the sharing of omic data generated in rare disease studies and has established a central platform and regulation for the online analysis of genomic data. The data sets belonging to the participants can be analyzed by us online and safely.

During alignment, human genome version Hs37d5, which is a modified version of the GRCh37 was used as a reference genome developed during the 1000 genomes project, which improves the quality of variant calling when using short-read alignment.

For the current release the transcript set corresponds to that of Ensembl Version 75 (http://feb2014.archive.ensembl.org/index.html). The current release of BWA-mem (v0.7.8), and GATK Haplotype Caller (v3.6) softwares were used for alignment and variant calling to generate gVCFs.

Some variants filtered prior to being uploaded to the platform. It was performed minimal filtering of very poor-quality variants, removing any variants for which the read-depth (DP) was less than 8, or the genotype quality (GQ), as provided by GATK, was less than 20.

SNV Effect Prediction filters were implemented as OR (not AND). That means

when D is selected for mutation taster and D for SIFT then the platform will return variants flagged as D in either, or both.

Candidate compound-heterozygote mutations were also identified at the level of the transcript within the platform.

The platform was currently using annotation data from ExAC version 0.1, CADD version 1.0, and the first release of 1000GP frequencies.

The meaning of High, Moderate etc. in the Variant Class filter was provided by SnpEff as effect prediction tool.

3.4.5. Variant classification and annotation

Variant classification and annotation were again performed on the RD-Connect platform.

Tools used within the platform were explained below.

i. CADD Score

The Combined Annotation Dependent Depletion (CADD) tool was used within the platform to score the predicted deleteriousness effect of single nucleotide variants as well as insertion and deletions variants in the human genome by integrating multiple annotations including conservation and functional information into one metric. Phred-style CADD raw scores were displayed and variants with higher scores were accepted as more likely to be deleterious.

CADD provided a ranking rather than a prediction or default cut-off, with higher scores more likely to be deleterious. For convenience, scores above 30 as 'likely deleterious' were displayed and scores below as 'likely benign' within the platform. Variants with scores over 30 are predicted to be the 0.1% most deleterious possible substitutions in the human genome. CADD score was a combination of 63 different types of databases.

The CADD score was interpreted as follows:

- above >30, its highly pathogenic
- above >20 then its pathogenic.
- Between 15-20, likely pathogenic.
- Below <15, likely benign,
- Below <10, its benign

Stop-gain mutations have CADD score above 30 in general. Therefore, stop-gained mutations were considered highly pathogenic.

ii. SIFT

SIFT was used to predict whether a protein function based on the homology of sequences and physical similitaries between alternative versions of amino acids would probably have a beating on an amino acid substitution. The data for each substitution of amino acid were a score and a qualitative prediction (both' tolerated' and' deleterous'). The score was the normalized chance of tolerating changes in amino acid, which made it more likely that score near zero would be deleterious. From the quality prediction, substitution with a score less than 0.05 was referred to as 'deleterious' and all other substitutions were called 'tolerated'.

iii. PolyPhen

In PolyPhen-2 a series of PDB annotations, 3D structures (as available) and a number of other databases and tools including DSSP and ncoil systems were used to predict the effect of amino acid surrounding a protein structure and protein function. Each substitution for amino acid was provided with both a qualitative prediction score as 'probably harmful', 'possibly harmful', 'benign' or 'unknown'. The PolyPhen score represented the probability that a substitution was damaging, so values nearer to 1 was predicted to more likely be deleterious on the contrary to SIFT).

iv. Exomiser tool

Exomiser tool was used to prioritize SNVs by comparing the phenotype across species using human or a model organism phenotype in company with the inheritance pattern. In this study, HPO terms were selected according to patient clinical phenotype during deep phenotyping by considering recessive inheritance model at first. HPO terms associated with specific diseases were found by using the Human Phenotype Ontology (HPO) database.

v. Homozygosity mapping

Homozygosity mapping was an effective way in consanguine families to detect molecular defects by delineating genomic DNA stretches that were dentical by descent. Constant developments created opportunities to combine complete sequencing (WES) with homozygosity mapping. The parameters used during homozygosity mapping in

families with autosomal recessive (AR) mutations were optimized and explained by in Kancheva et al. (2016). According to Kancheva et al. (2016), GenomeComb package had been integrated into the WES-based homozygosity mapping (HOMWES) for filtering and mapping with optimized parameters in the genomic data analysis.

3.5. Data Collection Equipments

Clinical data of patients was conducted by using Phenotips forms which was prepared by clinicans considering selected clinical phenotypes and recessive inheritance model. Phenotips forms were prepared by using Human Phenotype Ontology (HPO) database which contained over 8000 terms. Because, HPO terms captured phenotypic similarity, it was believed that using HPO terms was useful to compare disease state and correlate phenotype to genotype to discover biologica function of genes.

3.6. Work Plan

Patient and their families' blood samples were collected from the beginning of the project. Consent forms and phenotips forms corresponding to these patients were gathered while getting blood samples. Patient phenotips forms were uploaded to the RD-Connect PhenoTips Platform (https://platform.rd-connect.eu/phenotips) according to used HPO terms while filling PhenoTips forms.

DNA isolation from blood was performed all along the project. DNA samples from all volunteers were sent to our UK partners on 09.03.2017, 31.05.2017, 11.07.2017, 11.08.2017, 01.11.2017 and 30.11.2017, 27.11.2018, 04.12.2018 for genotyping and quality control (QC-Quality check). After the samples were subjected to quantification, purity and integrity tests, they were sent to the Broad Institute (Cambridge, MA), one of the world's leading centers of genomics, for whole exome sequencing (WES).

Thereafter, sequencing of patients was completed and this exome sequence raw data was uploaded to the RD-Connect database. Sequence analysis of patients was performed by using RD-Connect Platform (https://platform.rd-connect.eu/genomics/).

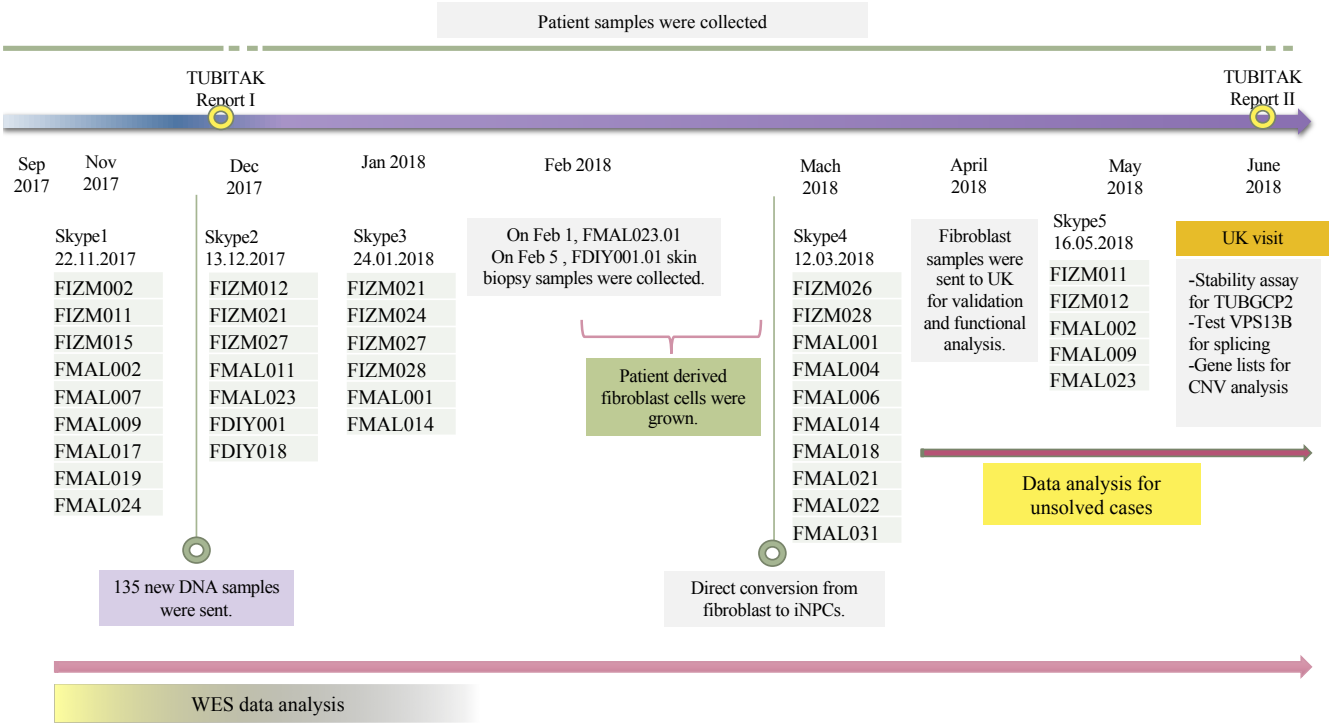

Figure- 6 Workflow of the study (September 2017- June 2018)

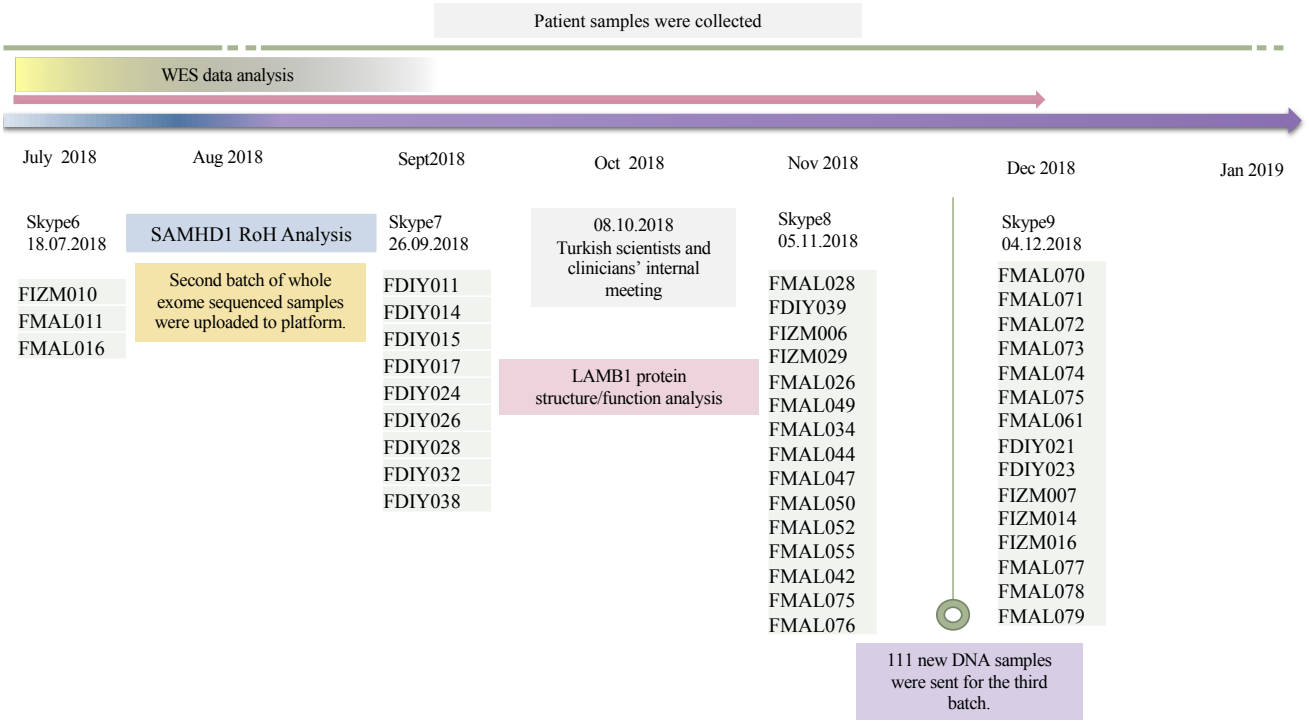

Figure- 7 Workflow of the study (July 2018- December 2018)

3.7. Evaluation of the Data

In the gene and causal variant discovery process, WES is preferred for highly penetrating Mendelian diseases. Because WES concentrates only on 1% of the genome, WES is limited to the only coding and splice-site variants. The exon capture step can also cause numerous technical bias during the procedure and restrict its use in the detection of copy-number variants and in less effective genomic regions. In addition to these challenges, high complexity and heterogeneity of the rare disease research data also limit the reseach. The process of data evaluation should be made in cooperation with different research groups and platforms in order to improve the data analysis process. In this study, RD-Connect platform was used to evaluate patient data. By the help of this analysis plaform, great numbers of patient data were analyzed in corcordance with different study groups by different researchers.

3.8. Limitations of the Study

Although the diagnostic yield in this study was already 60% with the currently limited knowledge, there are many candidates waiting to be confirmed. Because in the complex development of the central nervous system many genes and pathways with important functions remain unknown, suspected causal candidate genes can not be demonstrated without working properly. Also,

i. Some consanguineous families not willing to contribute to research project due to concerns about being disadvantaged through participation in the study.
ii. There were many known genes in the discovery cohort.
iii. Novel genes unlinked to any known function or pathway with a good score can not be confirmed without further studies.
iv. Some variants and genes were observed in only one family. There was not second family for validation.
v. There were many variants with unknown significance.

3.9. Ethical Committee Approval

Ethical Committee Protocol No: 302-SBKAEK

Decree No: 2016/03-01

Date: 11.02.2016

4. RESULTS

Patients were recruited from three centers in Turkey (Izmir, Diyarbakır, Malatya) in this study. The child with suspected neurogenetic disorders born into consanguineous marriages were included in the study. Consent forms were collected from all volunteers (index patients, parents, siblings) who agreed to participate in the study, and the use of all genome sequencing and data. Patients who agreed to participate in the study were phenotyped according to standardized standardized phenotyping protocols by clinicians. Diseases included in the study are neonatal, infantile or childhood-related and primarily affect the brain or muscles, mental retardation (with or without epilepsy), cortical and other brain malformations, leukodystrophies, ataxias, myopathies and muscular dystrophies, inherited neuropathies are congenital myasthenic syndromes and mitochondrial disorders. Inclusion criteria was being accessible for in-depth phenotyping and blood sampling studies for both index patient and family members.

To date, 192 families and a total of 854 people from these families were included in the study and blood samples were collected from EDTA tubes for all subjects and DNA isolation was performed for each of these samples for analysis. All volunteers' names were kept confidential and the samples were shared by IDs and barcodes assigned to them.

DNA isolation from blood was performed all along the project. DNA samples from all volunteers were sent to our UK partners on 09.03.2017, 31.05.2017, 11.07.2017, 11.08.2017, 01.11.2017 and 30.11.2017, 27.11.2018, 04.12.2018 for genotyping and quality control (QC-Quality check). After the samples were subjected to quantification, purity and integrity tests, they were sent to the Broad Institute (Cambridge, MA), one of the world's leading centers of genomics, for whole exome sequencing (WES). Whole exome sequencing (WES) was performed on affected patients, and both parents by the MIT Broad Institute and the Harvard Genomics Platform (Cambridge, MA, USA) using > 250 ng DNA. Illumina Exome Capture Kit (38 Mb target) was used.

Sequencing data were processed at the Centro Nacional de Análisis Genómico

(CNAG), Barcelona, and data analyses were carried out on the RD-Connect Genome-Phenome Analysis Platform (https:/platform.rd-connect.eu/genomics) (Figure-7) by using standard filtering criteria for rare diseases, including Minor Allelelele Frequency (MAF) < 0.01, Variant Effect Predictor (VEP)=mod / high.

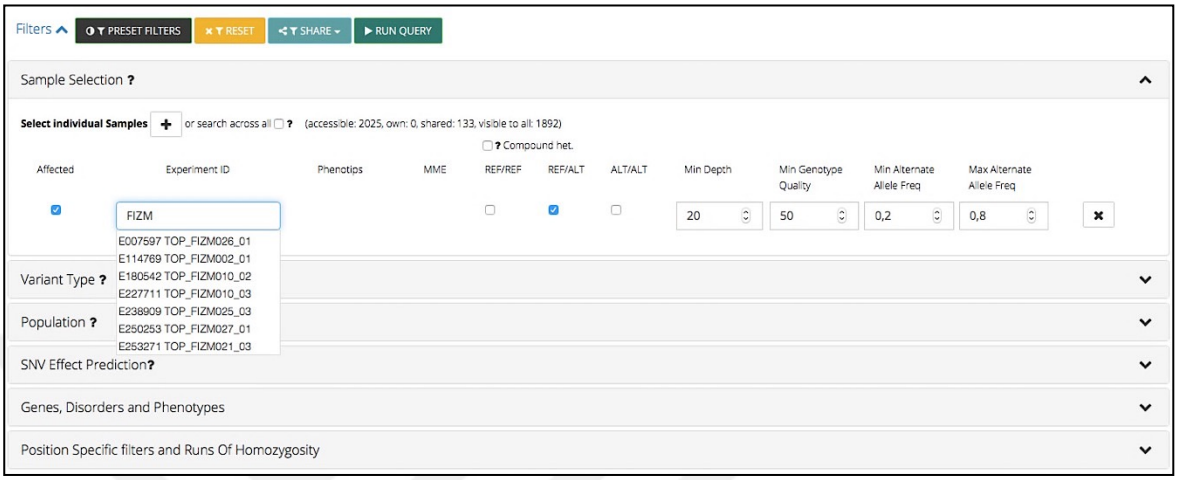

Figure- 8 RD-Connect Platform

4.1. Solved Cases

Four novel candidate genes were discovered in leukodystrophy, brain malformation, intellectual disability disease groups.

WES provided a molecular diagnosis for 62 families out of 138 (45%) (Table-1).

Table- 1 The number of families and people recruited to the study

Number	İzmir	Malatya	Diyarbakır	Total
Family	59	99	34	192
Person	202	491	161	854
WES	32	76	30	138
Solved	14	33	15	62
Ratio	44%	43%	50%	45%

In Izmir, 9 missense homozygout mutation, 1 X-linked stop gained, 1 X-linked missense, 1 farmeshift, 2 splice site, 1 compund heterozygous mutation were deteceted (Table-2).

Table- 2 Solved Izmir cases

Family ID	Disease Group	Gene	Variant
FIZM002	Brain Malformation	*PORCN*	X-linked dominant stop-gained (p.Arg95Ter)
FIZM006	Intellectual Disability	*WWOX*	Homozygous missense (novel)
FIZM007	CMT and Neuropathies	*DPAGT1*	Homozygous missense (p.Phe113Leu)
FIZM010	Brain Malformation	*TLK2*	Homozygous missense (p.Lys55Glu)
FIZM012	Muscular Dystrophies and Myopathies	*NDUFA12*	Homozygous frameshift (p.Glu41GlyfsTer10)
FIZM014	CMT and Neuropathies	*TBCD*	Homozygous missense (p.Leu300Phe)
FIZM015	Brain Malformation	*CASK*	X-linked missense (p.Gly637Asp)
FIZM016	Intellectual Disability	*ARX*	Splice donor (c.1448+1G>A)
FIZM020	Mitochondrial Disorders	*RMND1*	Homozygous missense (p.Val53Glu)
FIZM021	Brain Malformation	*ZRANB3* *GALK1*	Compound heterozygous missense (p.Leu911Phe + p.Gly449Glu) Homozygous missense (p.Arg37His)
FIZM024	Brain Malformation	*LAMB1*	Homozygous missense (novel)
FIZM027	Muscular Dystrophies and Myopathies	*CACNA1S*	Homozygous missense (p.Arg789His)
FIZM029	Intellectual Disability	*WWOX*	Homozygous missense(p.Gln230Pro)

In Malatya, 19 homozygous missense mutation, 7 farmeshift, 1 splice site, 1 compund heterozygous, 5 stop-gained, 1 *de novo* mutation were deteceted (Table-3).

Table- 3 Solved Malatya Cases

Family ID	Disease Group	Gene / Variant	Variant
FMAL002	Leukodystrophy	*TACO1*	Homozygous frameshift (p.His158ProfsTer8)
FMAL004	Muscular Dystrophies and Myopathies.	*COX6B1*	Homozygous splice donor variant (c.-12+2T>C)
FMAL007	Leukodystrophy	*CLP1*	Homozygous missense (p.Arg140His)
FMAL009	Spinocerebellar Ataxia	*PRKCG*	Homozygous missense (novel)
FMAL011	Brain Malformation	*OCLN* *MCCC2*	Homozygous frameshift mut. (p.Trp58PhefsTer10) Homozygous missense (p.Val339Met)
FMAL014	Muscular Dystrophies and Myopathies.	*COL12A1*	*De novo* missense variant (p.Gly1922Glu)
FMAL016	Brain Malformation	*KATNB1*	Homozygous missense (p.Arg34Trp)
FMAL017	Other	*SAMHD1*	Homozygous stop gained (p.Arg164Ter)
FMAL019	Leukodystrophy	*SAMHD1*	Homozygous stop gained (p.Arg164Ter)
FMAL023	Leukodystrophy	*TUBGCP2 (novel gene)*	Homozygous missense (p.Glu311Lys)
FMAL024	Muscular Dystrophies and Myopathies	*COLQ*	Homozygous stop gained (p.Trp138Ter)
FMAL026	Spastic Paraplegia	*ALS2*	Homozygous missense (p.Cys157Tyr)
FMAL028	Intellectual Disability	*WWOX*	Homozygous missense (novel)
FMAL031	Leukodystrophy	*STAMBP*	Homozygous missense (p.Tyr63Cys)
FMAL042	Spastic Paraplegia	*SACS*	Homozygous missense (p.Arg728Ter)
FMAL047	CMT and Neuropathies	*GDAP1*	Homozygous frameshift (p.Phe263LeufsTer22)
FMAL049	Spastic Paraplegia	*ALS2*	Homozygous frameshift (p.Val1525GlyfsTer17)
FMAL050	Muscular Dystrophies and Myopathies	*CLCN1*	Homozygous missense (p.Gly355Arg)
FMAL052	Muscular Dystrophies and Myopathies	*SGCA*	Homozygous missense (p.Arg284Cys)
FMAL053	Brain Malformation	*ATP8A2*	Stop-gained (p.Arg546Ter)
FMAL055	Spastic Parapleagia	*SPG11*	Homozygous frameshift (p.Glu1026ArgfsTer4)

FMAL061	Intellectual Disability	*STIL*	Homozygous stop-gained (p.Glu1213Ter)
FMAL070	Intellectual Disability	*AP3B2*	Homozygous frameshift (p.Pro974ArgfsTer5)
FMAL075	Intellectual Disability	*OTUD6B*	Homozygous missense (p.Arg127Trp)
FMAL076	Intellectual Disability	*PIGT*	Homozygous missense (p.Tyr215Cys)

In Diyarbakır, 10 homozygous missense mutation, 2 *De novo*, 1 homozygous frameshift, 1 synonmous, 1 stop-gained, 1 compund heterozygous mutation were deteceted (Table-4).

Table- 4 Solved Diyarbakır Cases

Family ID	**Disease Group**	**Gene**	**Variant**
FDIY001	Brain Malformation	*VPS13B*	Compound heterozygous (c.412+1G>T and c.7504+40A>T)
FDIY021	Brain Malformation	*HADH, SPP1*	Homozygous synonmous (p.Ala40=) Homozygous stop-gained (p.Gln105Ter)
FDIY023	Brain Malformation	*TAF1*	Homozygous missense (p.Ile1535Val)
FDIY018	Brain Malformation	*TCF20*	*De novo* missense (p.Gln794Pro)
FDIY024	Leukodystrophy	*COL4A1*	Homozygous missense (p.Gly1278Ser)
FDIY028	Brain Malformation	*P2RX7*	Homozygous missense (novel)
FDIY030	Brain Malformation	*TCC1*	Homozygous missense (p.Phe262Val)
FDIY039	Intellectual Disability	*WWOX*	Homozygous missense (novel)

4.2. Solved Brain Malformation Related Cases

FDIY001 *VPS13B* (OMIM 607817)

Studies have shown that the protein produced by the gene is a part of a Golgi apparatus that modifies newly manufactured proteins to carry out their functions. Studies suggest that losing the function of this protein interferes with Golgi's organization and impairs normal glycosylation.The protein VPS13B also appears to participate in the sorting and transport of proteins within the cell [57].

Studies show that there are numerous functions for the VPS13B protein within the body. The protein seems to have an important role to play in normal neuronal cell growth and development. It may also be part of adipocytes which are cells that store energy fats and play a role in the body's fat storage and distribution [58].

This gene is associated with Cohen syndrome. In individuals with Cohen syndrome more than 150 different mutations were found in the *VPS13B* gene. Persons with this condition typically have intellectual disability, eye problems and distinguishing facial characteristics. Obesity especially around the torso, not arms and legs (truncal obesity) is another common feature of the condition. The VPS13B protein most mutations that cause the Cohen syndrome cause a premature stop signal. This genetic change is believed to cause an abnormally short, non-functional version of the protein to be produced [59], [60]. 75% of the Finns have Cohen syndrome mutations that both copies of the VPS13B genedeleted. Deletion causes the VPS13B gene to have a premature stop signal.

In the Old Order Amish population, two common mutations occur. The first mutation adds a pair of bases, results in a premature stop signal. The second mutation changes the VPS13B protein to a single amino acid. In particular, this mutation replaces amino acid isoleucine at position 2820 (Ile2820Thr) with amino acid threonine [61],[57].

Table 5 summarize the information of the variant found in our patient.

Table- 5 *VPS13B* splice donor variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Aa change	Aa length	Exon	CDS Pos.
VPS 13B	ENST00000358544	splice_donor_variant	c.412+1G>T	c.412+1G>T		4022		
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
8	100108661	rs1057517295	G	T	G/T	G/T	G/G	HIGH
Clin Var	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
LP	26.4			D	NA	NA	NA	0.001189

Figure 9 illustates the VPS13B protein and its domains.

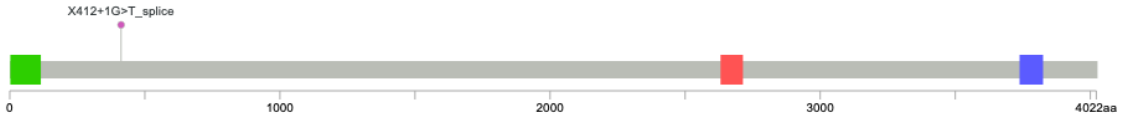

Figure- 9 VPS13B protein domains and the mutation

FDIY018 TCF20 (OMIM 603107)

TCF20 has more than 1,900 amino acids and contains a variety of feature areas. It is predicted to work as an activator or repressor according to its interaction with other factors [62].

Table 6 summarize the information of the variant found in our patient.

Table- 6 *TCF20 de novo* missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
TCF20	ENST00000359486	missense_variant	c.2381A>C	c.2381A>C	p.Gln794Pro	1960	1/5	2381/5883
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
22	42608931	.	T	G	T/G	T/T	T/T	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	24.4	D	D	D	NA	NA	NA	0.000131

Figure 10 illustates the TCF20 protein and its domain.

Figure- 10 TCF20 protein domains and the mutation

FIZM002 *PORCN* (OMIM 300651)

The *PORCN* gene is part of a family of proteins known as porcupine (Porc). While the precise function of the PORCN protein remains unknown, the Porc proteins are involved in transferring palmitoleic acid to Wnt proteins. Wnt proteins are involved in chemical body signals and play critical roles in development. The Porc family members are located in the reticular endoplasm. Palmitoleic acid transfer to wnt proteins makes it possible for these proteins to be released from the cells so as to regulate skin, bone and other structures [63].

Focal dermal hypoplasia is associated with this gene. Focal dermal hypoplasia has been observed with at least 29 mutations in the *PORCN* gene. These mutations may modify the structure of the protein, cause an abnormally shorter version of the protein to be produced or can delete the whole *PORCN* gene. The absence of any functional PORCN protein is apparent in all those mutations. Wnt proteins without the PORCN protein are believed to be unable to release from the cell. If Wnt proteins can not exit the cell they can not engage in the chemistry that is critical for normal development [64], [65].

Table 7 summarize the information of the variant found in our patient.

Table- 7 *PORCN de novo* stop-gained variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
PORCN	ENST00000326194	stop_gained	c.283C>T	c.283C>T	p.Arg95Ter	461	2/14	283/1386
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
X	48369829	.	C	T	C/T	C/C	C/C	HIGH
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	37.0			A	NA	NA	NA	0.000131

Figure 11 illustates the PORCN protein and its domain.

Figure- 11 PORCN protein domains and the mutation

FIZM010 *TLK2* (OMIM 608439)

This gene encodes a first identified nuclear serine / threonine kinase in Arabidopsis. The encoded protein is intended to operate by controlling histone levels H3/H4 in chromatin assemblies in the S phase of the cell cycle. This protein is linked to the repair of radiation caused by double-strand rupture of DNA damage. Chromosomes 10 and 17 contain pseudogenes of this gene. Alternate split results in several variants of transcripts [66], [67].

The chromatin assembly process involves serine/threonine protein kinase and probably DNA replication, transcription, repair and segregation of chromosomes. The chromatin assembly factors phosphorylate ASF1A and ASF1B. ASF1A phosphorylation prevents degradation through proteasome, improving the aseembly of chromatin [67].

A distinct neurodevelopmental phenotype characterized mainly by mild motor and language delay and facial dysmorphism, caused by heterozygous *de novo* or dominant mutations in the *TLK2* gene has recently been described. All cases reported carried

either truncating variants located throughout the gene, or missense changes principally located at the C-terminal end of the protein mostly resulting in haploinsufficiency of *TLK2*. Through whole exome sequencing, homozygous missense variant in *TLK2* was identified showing more severe symptoms than those previously described, including cerebellar vermis hypoplasia and West syndrome. Both parents are heterozygous for the pathogenic variant and clinically unaffected highlighting that recessive variants in *TLK2* can also be disease-causing and may act through a different pathomechanism.

TLK2 is associated with autosomal dominant 57 (MRD57), a type of mental retards, which is a disease which is clearly lower than the average general intellectual function associated with adaptive behavior impairments and which occurs during the developmental period. The delayed psychomotor development of the childhood or early childhood and several behavioral abnormalities are characterized by the MRD57 [MIM:618050] [68], [69].

Table 8 summarize the information of the variant found in our patient.

Table- 8 *TLK2* homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
TLK2	ENST00000326270	missense_variant	c.163A>G	c.163A>G	p.Lys55Glu	772	4/23	163/2319
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
17	60599574	rs774263147	A	G	G/G	A/G	A/G	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	26.9	D	D	D	0.000008	NA	0.000004	0.000527

Figure 12 illustates the TLK2 protein and its domain.

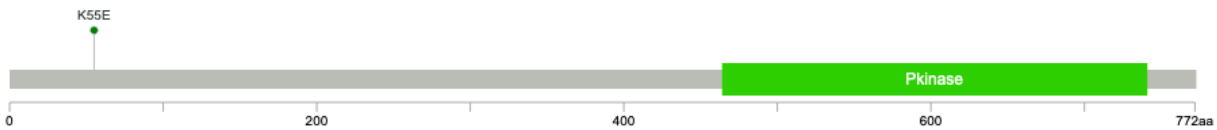

Figure- 12 TLK2 protein domain and the mutation

FIZM015 *CASK* (OMIM 300172)

CASK gene synthesize a protein called as calcium/calmodulin-dependent serine protein kinase (CASK). CASK protein is found primarily in neurons in the brain to help control expression of other genes involved in brain development. It also contributes to the regulation of the motion of chemical substances referred to as neurotransmitters and charged ions necessary for signals between neurons. Studies suggest that the CASK protein may also be interactive in order to promote nerve development that control the movement of the eye (the neural oculomotor network) in the protein produced from another gene (FRMD7) [70].

In people with CASK-related intellectual disabilities, over 35 mutations of *CASK* gene have been identified. This disorder affects the development of the brain and found in two primary forms as the severe form of pontinous and cerebellar hypoplasia (MICPCH) and a milder form called X-linked intellectual disability (XL-ID) with and without nystagmus. Nystagmus refers to quick, unintended movements in the back and forth of the eye. *CASK* gene mutations may disrupt the relationship of the protein CASK with the protein produced by the gene FRMD7, causing problems in the development of the neural oculomotor network and abnormal eye movements [71], [72].

Table 9 summarize the information of the variant found in our patient.

Table- 9 *CASK de novo* missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
CASK	ENST00000318588	missense_variant	c.1910G>A	c.1910G>A	p.Gly637Asp	921	21/27	1910/2766
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
X	41413101	.	C	T	C/T	C/C	C/C	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	31.0	D	D	D	NA	NA	NA	0.000131

Figure 13 illustates the CASK protein and its domains.

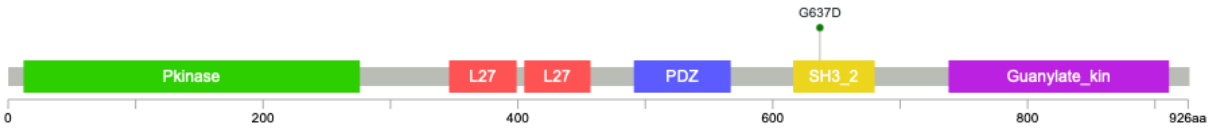

Figure- 13 CASK protein domains and the mutation

FIZM021 *GALK1* (OMIM 604313)

The *GALK1* gene is used to make the galactokinase enzyme 1. The body can process a simple sugar called galactose with this enzyme. Galactokinase 1 is responsible for one step in a chemical process in which galactose is converted into other molecules that the body can use. This enzyme modifies the galactose to form a similar galactose-1-phosphate molecule in particular. Galactose-1-phosphate is converted into another simple sugar called glucose by a number of additional steps, the main source of electricity for most cells. Galactose-1-phosphate can also be transformed into a form used to build proteins and fats that contain galactose. These modified proteins and fats play critical roles in chemical signaling, building cellular structures, transporting molecules, and producing energy [73].

In people with a type II galactosemia or galactokinase deficiency, more than 30 mutations in the GALK1 gene have been identified. The majority of these mutations change amino acids in galactokinase 1. Some mutations remove from GALK1 gene a small quantity of genetic material which results in this enzyme being stable or inactive.

Functional galactokinase 1 deficiency prevents the cells from processing the dietary galactose. Galactose and associated sugar known as galactitol can therefore accumulate in eye and resulted in cloudy and blurred vision [73], [74].

Table 10 summarize the information of the variant found in our patient.

Table- 10 *GALK1* homozygous missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
GALK1	ENST00000225614	missense_variant	c.110G>A	c.110G>A	p.Arg37His	392	1/9	110/1179
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
17	73761108	.	C	T	T/T	C/T	C/T	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	34.0	D	D	D	NA	NA	NA	0.000538

Figure 14 illustates the GALK1 protein and its domains.

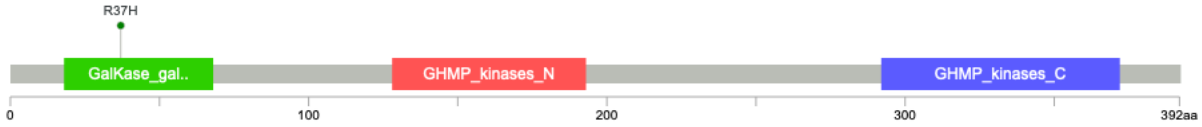

Figure- 14 GALK1 protein domains and the mutation

FIZM024 *LAMB1* (OMIM 150240)

Laminines, which are a family of extracellular glycoprotein matrix, are the main components of the cell membranes. A range of biological processes have been participated such as cell adhesion, differentiation, migration, signaling, outgrowth of neuritis and metastasis. Lamininin consists of three non-identical strings: alpha, beta and gamma laminin (formerly A, B1, and B2, respectively), forming the cross-structure of three short arms formed by a distinct chain, each of which consists of a large arm comprising all three strings. Every laminin chain is a separate gene encoded multidomain protein. There have been descriptions of several isoforms in each chain. In their order of discovery different heterotrimeric isoformes are combined in the alpha, beta and gamma chain.

While the biological functions of the different tissue chains and trimming molecules are largely unknown, some tissue distributions were shown to differ, reflecting probably different *in vivo* functions. The beta 1 chain has 7 different domains with other isomers in the beta chain. Domains I and II are separated from the alpha

domain in the helical C-terminal region, Domains III and V are replicated in EGF-like ways and domains IV and VI are configured globally.

Laminin beta 1 is found in the majority of tissues making cellular membranes and is one of the first 3 laminin 1, isolated from the tumor Engelbreth-Holm-Swarm (EHS). The metastases are identified and demonstrated by a sequence of cell attachment, chemo-taxis and binding to the betas 1 chain laminin receptor.

Laminin mediates cell attachment, migration and tissue organization via interactions with other extracellular matrix in embryonic development by bending cells through a high-affinity receptor. LAMB1 participates in the development of the brain cortex laminar architecture.

The ends of radial glial cells need to anchor on basement membrane/glia and as a physical barrier to migrating neurons. Radial glial cells play a central role in the development of the cerebral cortical system, in which they act as both the cerebral cortex 's proliferative unit and an eavesdrop for neurons migrating to the pial area [75], [76].

This gene is associated with Lissencephaly 5 (LIS5) which is autosomal recessive brain malformation characterized by cobblestone lisssencephaly, severe post-cortical region and subcortical band heterotopia. Serious delays in the development of psychomotors, hydrocephalus, and seizures are the feature found in affected persons [MIM:615191][75], [77].

Table 11 summarize the information of the variant found in our patient.

Table- 11 LAMB1 homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
LAMB1	ENST00000222399	missense_variant	c.4238G>A	c.4238G>A	p.Gly1413Glu	1786	28/34	4238/5361
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
7	107572773	rs757004190	C	T	T/T	C/T	C/T	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	34.0	D	D	D	0.000008	NA	0.000012	0.000905

Figure 15 illustates the LAMB1 protein and its domains.

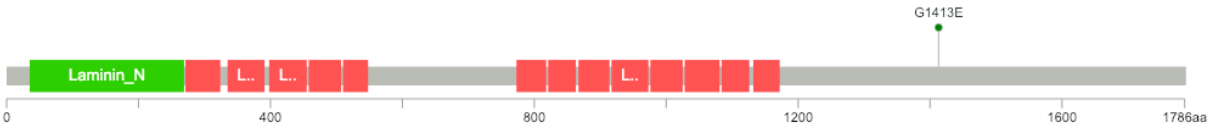

Figure- 15 LAMB1 protein domains and the mutation

FMAL007 *CLP1* (OMIM 608757)

This gene encodes a member of the family CLP1. The encoded protein is a multifunctional kinase that is part of the tRNA endonuclease splicing complex and part of the pre-mRNA cleavage complex II. This protein is involved in the maturation of tRNA, mRNA and siRNA. Pontocerebellar hypoplasia type 10 (PCH10) is associated with mutations of this gene. Alternatively, spliced transcript variants that encode different isoforms for this gene were found [78].

The pontocerebellar hypoplasia 10 (PCH10) is associated with this gene, which is a form of pontocerebellar hypoplasia, a condition that is characterized cerebellar structural defects that can be seen in the brain imaging. PCH10 has cortical dysgenesis characterized by simplified gyral patterns, cortical atrophy, mild or focused brain cerebellary loss, delayed myelination, progressive microcephaly, global growth and delays in development, serious intellectual disabilities and treatment-refractory convulsions. [MIM:615803][79].

Table 12 summarize the information of the variant found in our patient.

Table- 12 CLP1 homozygous missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
CLP1	ENST00000533682	missense_variant	c.419G>A	c.419G>A	p.Arg140His	425	2/3	419/1278
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
11	57427367	rs587777616	G	A	A/A	G/A	G/A	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
LP	29.3	D	D	D	0.000008	NA	0.000008	0.001200

Figure 16 illustates the CLP1 protein and its domains.

Figure- 16 CLP1 protein domains and the mutation

FMAL11 *OCLN* (OMIM 602876) and *MCCC2* (OMIM 609014)

OCLN

This gene encodes an integral protein membrane required for a tight inter-connection of the cytokine-regulated paracellular permeability barrier. Mutations in this gene, which is also known as pseudo-TORCH Syndrome, are seen as the cause of band-like calcifications with simplified gyrated and polymicrogyria (BLC-PMG). A number of variants of transcripts result in alternative splicing. An associated pseudogen can play the role of the paracelular permeability of the close crossover (TJ) in the q arm of a chromosome 5 1.5 Mb downstream. It may lead to adhesion when expressed in cells that do not have tight connections [80].

This gene is associated with band-like calcification with simplified gyration and polymicrogyria and Pseudo-TORCH syndrome 1. Pseudo-TORCH syndrome 1 is an automated recessive neurological disorder with features that mimic intrauterine TORCH disease in the absence of evidence of infection. Pseudo-TORCH syndrome 1. Congenital microcephaly, intracrane calcification and severe delays in development are present in individuals affected [MIM:251290] [80]–[82].

Table 13 summarize the information of the variant found in our patient.

Table- 13 *OCLN* frame-shift variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
OCLN	ENST00000355237	frameshift_variant	c.173_194delGGACCTCTCCTCCAGGAGTGAT	c.173_194delGGACCTCTCCTCCAGGAGTGAT	p.Trp58PhefsTer10	522	3/9	171/1569
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
5	68805087	rs797045841	AATGGACCTC..	A	A/A	AATGGACCTC..	AATGGACCTC..	HIGH
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
P	< 20				NA	NA	NA	0.000786

Figure 17 illustates the OCLN protein and its domains.

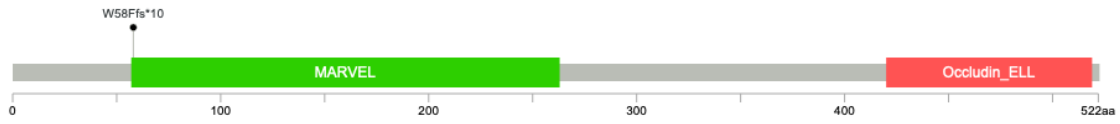

Figure- 17 OCLN protein domains and the mutation

The *MCCC2* gene sythesize the beta subunit of an enzyme called 3-methylcrotonoyl-CoA carboxylase or 3-MCC. Beta subunits join with larger alpha subunits made from the *MCCC1* gene; six of these pairings together form a functioning enzyme [83].

The 3-MCC enzyme in mitochondria is found in cells that produce energy. In breaking down proteins from the diet, this enzyme plays a key role. It is responsible specifically for the fourth step of leucine breakdown, an amino acid that is a component of many proteins. This step transforms the 3-methylcrotonyl-CoA molecule into a 3-methylglutaconyl-CoA molecule. Further chemical reactions transform 3-methylglutaconyl-CoA into energy molecules.

This gene is related to the carboxylase deficiency of 3-methylcrotonyl-CoA. In persons with a 3-methylcrotonyl-CoA carboxylase deficiency (also called 3-MCC deficiency), more than 40 mutations have been identified in the *MCCC 2* gene. Most of these mutations modify single amino acids in 3-MCC, seriously reducing enzyme activity. Additional mutations prevent any functional enzyme from developing. Leucine can not therefore be properly break down, and leucine processing by-products increase to toxic levels in the organism. These toxic substances can damage the brain and cause the symptoms and characteristics of 3-MCC deficiency [83], [84].

Table 14 summarize the information of the variant found in our patient.

Table- 14 *MCCC2* homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
MCCC2	ENST00000323375	missense_variant	c.901G>A	c.901G>A	p.Val301Met	525	10/16	901/1578
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
5	70936845	rs150591260	G	A	A/A	G/A	G/A	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
C	33.0	D	D	D	0.000675	NA	0.000796	0.003691

Figure 18 illustates the MCCC2 protein and its domain.

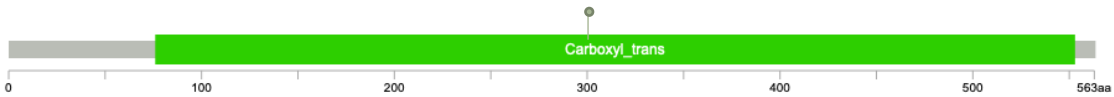

Figure- 18 MCCC2 protein domain and the mutation

FMAL017 and FMAL019 *SAMHD1* (OMIM 606754)

The *SAMHD1* gene provides instructions for making an enzyme that helps cut (cleave) molecules called deoxynucleoside triphosphates (dNTPs) into their deoxynucleoside and triphosphate components. The dNTP molecules are needed for the replication and maintenance of the genetic material (DNA) in the body's cells, and

the number of available dNTPs must be tightly controlled for these functions to proceed normally. Invading viruses such as the human immunodeficiency virus (HIV) also need dNTPs in order to replicate themselves. The SAMHD1 enzyme helps regulate the number of available dNTPs to both meet the needs of the body's cells and control viral infections [85], [86].

This gene is associated witlh Aicardi-Goutières syndrome At least 78 mutations of the *SAMHD1* gene, including severe brain dysfunction (encyphalopathy), skin losses, and other health problems, have been found in patients suffering from Aicardi-Goutières syndrome. Research suggests that mutations in this gene may lead to production of a SAMHD1 protein that does not properly regulate the number of available dNTPs. As a result, DNA maintenance is impaired, allowing DNA damage to accumulate in cells. Research suggests that this DNA damage may lead to cell signaling that inappropriately activates an immune response. This causes inflammatory damage to the brain, skin and others that lead to the typical features of Aicardi- Goutières Syndrome and the immune system attacks your own tissues and organ [87],[88].

Table 15 summarize the information of the variant found in our patient.

Table- 15 *SAMHD1* homozygous stop-gained variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
SAMHD1	ENST00000262878	stop_gained	c.490C>T	c.490C>T	p.Arg164Ter	626	4/16	490/1881
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
20	35563451	rs267607027	G	A	A/A	G/A	G/A	HIGH
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
P	39.0			A	0.000008	NA	0.000004	0.001049

Figure 19 illustates the SAMHD1 protein and its domains.

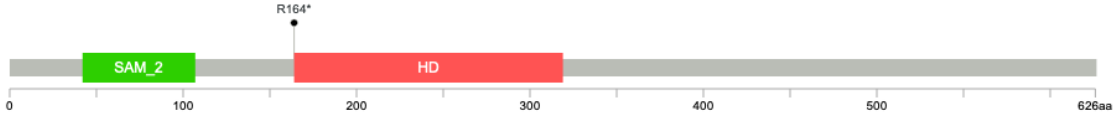

Figure- 19 SAMHD1 protein domains and the mutation

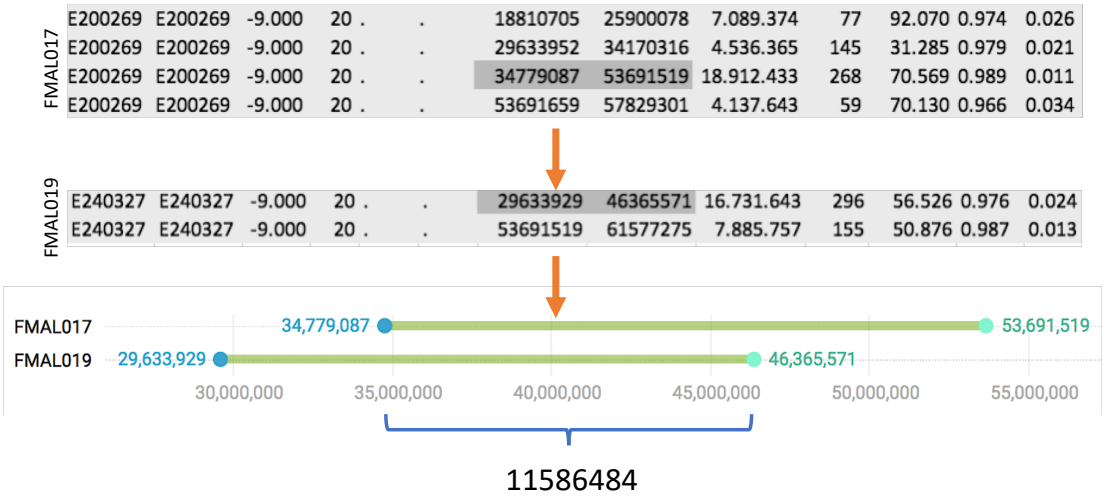

Figure- 20 Shared homozygous block (11 MB) in both FMAL017 and FMAL019 patients

FMAL016 *KATNB1* (OMIM 602703)

Katanin is the heterodimer of 60 kDa ATPase (p60 subunit A1) and 80 kDa (p80 subunit B1) accessory protein. The p60 subunit acts in a way that separates and disassembles the microtubules, while the enzyme to the centrosome is targeted by p80 subunit. Katanin is part of the AAA ATPase family.

KATNB1 participates in a complex that is ATP-dependent and can act on a centrosome-like basis in order to target the enzyme subunit of this complex. The microtubular separation may encourage the quick reorganization of cell microtubular arrays and the release of centrosome microtubules after nucleation. Microtubular release of mitotic spindle pole could allow the microtubule end to be depolymerized proximally to the spindle pole, which can cause the poleward movement of chromosomes. The release of microtubules into the cell body of neurons by microtubule-based motor proteins may be required for transport to neuronal processes. For axonal growth this transport is necessary.

This gene is linked with lissencephaly 6 with microcephaly, a subtype of lissencephaly, which has features like, a cortical developmental distortion with agyria, pachygyria, and disorganized neuronal lamination of the six-layer cortex. LIS6 includes corpus callosum hypoplasia, severe microcephaly and delayed development. [MIM:616212] [89].

Table 16 summarize the information of the variant found in our patient.

Table- 16 *KATNB1* homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
KATNB1	ENST00000379661	missense_variant	c.100C>T	c.100C>T	p.Arg34Trp	655	3/20	100/1968
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**			**Effect Impact**
16	57775658	rs371850267	C	T	T/T			MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	35.0	D	D	D	0.000065	NA	0.000037	0.000790

Figure 21 illustates the KATNB1 protein and its domains.

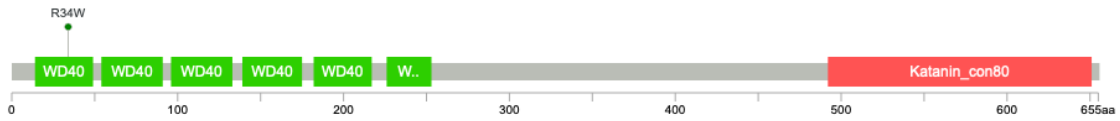

Figure- 21 KATNB1 protein domains and the mutation

FMAL023 *TUBGCP2* (OMIM 617817) (Novel Gene)

TUBGCP2 gene encodes the protein called gamma- tubulin complex component 2. It is part of the complex of gamma tubulins that is needed for microtubule nucleation at the centrosome.

The gamma-tubulin (TUBG; see 605785) ring complex is a microtubule-organizing center that acts as a template for polarized growth of microtubules essential for diverse cellular structures. The TUBG small complex, a component of the ring complex, consists of multimers of TUBGCP2, TUBGCP3 (617818), and TUBG and is the immediate template for growing microtubule ends[34].

Table 17 summarize the information of the variant found in our patient.

Table- 17 *TUBGCP2* homozygous missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
TUBGCP2	ENST00000252936	missense_variant	c.931G>A	c.931G>A	p.Glu311Lys	902	6/17	931/2709
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
10	135106636	.	C	T	T/T	C/T	C/T	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	34.0	D	D	D	NA	NA	NA	0.000920

Figure 22 illustates the TUBGCP2 protein and its domain.

Figure- 22 TUBGCP2 protein domain and the mutation

FMAL031 *STAMBP* (OMIM 606247)

The *STAMBP* synthesize a protein called a protein binding STAM. Even though its exact function is not well understood yet, this protein interacts within cells with large interrelated protein groups known as transport-needed endosomal sorting complexes (ESCRTs). ESCRTs help transportation of proteins into the cell, a process known as an endocytosis, from the outer cell membrane. In particular, the proteins that need to be disintegrated (degraded) or recycled by the cell are involved in the endocentoses of damaged or unnecessary. ESCRTs help to sort these proteins into so called lysosome-providing structures called multivicular bodies (MVB) [90].

STAM binding protein helps maintain the correct balance between protein production and (protein homeostasis) disruption that cells have to operate and survive, by associating it with ESCRTs. Studies show that multiple chemical signalling pathways within cells include the pathways necessary for the overall growth and development of new blood vibration (angiogenics) and the interaction between STAM binding Protein and ESCRTs [91].

Table 18 summarize the information of the variant found in our patient.

Table- 18 STAMBP homozygous missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
STAMBP	ENST00000339566	missense_variant	c.188A>G	c.188A>G	p.Tyr63Cys	424	2/10	188/1275
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
2	74058171	rs781694797	A	G	G/G	A/G	A/G	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	27.1	D	D	D	0.000074	NA	0.000049	0.000526

Figure 23 illustates the STAMBP protein and its domains.

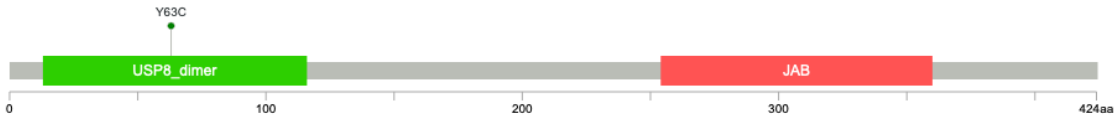

Figure- 23 STAMBP protein domain and the mutation

FDIY022 *DLAT* (OMIM 608770)

The *DLAT* gene synthesizes the E2 enzyme, which is part of a large group of proteins called pyruvate dehydrogenase complex (D2), it is also known as dihydrolipoamide acetyltransferase. Multiple copies of three enzymes, including E2, and various related protein are present in this complex. E2 is the core of the complex formed by the other proteins. The complex of pyruvate dehydrogenase plays a key role in translating energy from food into form that can be used by cells. This complex transforms a pyruvate molecule formed from a carbohydrate breakdown in a different molecule known as acetyl-CoA. One part of this chemical reaction is carried out by the E2 enzyme. Pyruvate conversion is essential to start with the series of adenotic triphosphate (ATP) chemical reactions, the main source of energy for cells [92].

In individuals with pyruvate dehydrogenase deficiency at least two mutations in the *DLAT* gene have been identified. Deficiency in pyruvate dehydrogenase is characterized by potentially life-threatening development, delays in development, and neurological problems of the lactic acidosis.

DLAT gene mutations lead, although the mechanism is unclear, to an abnormal E2 enzyme and decrease the pyruvate dehydrogenase complex. Pyruvate develops and translates into lactic acid by decreased activity in this complex, which leads to lactic acid. Moreover, cellular energy production has been reduced. The brain, which depends particularly on this energy, is seriously affected, leading to neurological problems with deficiency of pyruvate dehydrogenase [92][93].

Table 19 summarize the information of the variant found in our patient.

Table- 19 *DLAT* compound heterozygote missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
DLAT	ENST00000280346	missense_variant	c.1728C>A	c.1728C>A	p.Phe576Leu	647	13/14	1728/1944
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
11	111931812	rs119103240	C	A	A/A	C/A	C/A	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
P	29.1	D	D	A	NA	NA	0.000012	0.000766

Figure 24 illustates the DLAT protein and its domains.

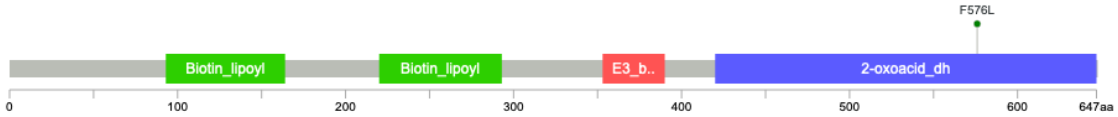

Figure- 24 DLAT protein domains and the mutation

FDIY024 *COL4A1* (OMIM 120130)

The *COL4A1* gene synthesize one of the type IV collagen components, which is a major flexible protein in the structure of numerous body tissues. This gene in particular produces the alpha1(IV) string type IV collagen. The chain combins a complete Alpha1-1-2 collagen type IV molecule with a second alpha1 chain type and an alternative alpha (IV) chain type called alpha2. Type IV collagen molecules are connected to form complex networks of proteins. These protein networks constitute the principal element of the cell membranes that separate thin sheet structures and

support them into many tissues.

Networks of type IV collagen alpha1-1-2 plays a significant role in virtually all tissues within the body's basement membranes, in particular those around the bird's blood vessels (vasculature). The IV network of collagen enables the basement membranes to interact with neighboring cells, which play a role in cell (migration), cell growth and divisions (proliferation), cell maturation (differentiation) and cell survival [94]–[96].

Table 20 summarize the information of the variant found in our patient.

Table- 20 *COL4A1* homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
COL4A1	ENST00000375820	missense_variant	c.3832G>A	c.3832G>A	p.Gly1278Ser	1669	43/52	3832/5010
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
13	110822020	rs757453900	C	T	T/T	C/T	C/T	MODERATE
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	28.2	D	D	D	0.000024	NA	0.000016	0.000786

Figure 25 illustates the COL4A1 protein and its domains.

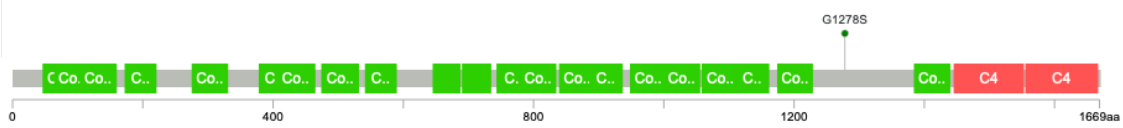

Figure- 25 COL4A1 protein domains and the mutation

FDIY028 *P2RX7* (OMIM 602566)

P2RX7 protein belongs to the ATP purinoceptor family. This receptor acts as a ligand-gated ion channel which forms membrane pores that are porous for large molecules and is responsible for the ATP-dependent macrophage lysis. Activation of this nuclear receptor by ATP in the cytoplasm can become a mechanism whereby cellular activity can be associated with changes in gene expression. Several alternative spliced variants, most of which fit nonsense-mediated decline (NMD) criteria, have

been identified [97].

Table 21 summarize the information of the variant found in our patient.

Table- 21 *P2RX7* homozygous missense variant

Gene Name	Transcript ID	Consequence	HGVS coding	Human Splicing Finder	Amino Acid change	Amino Acid length	Exon	CDS Position
P2RX7	ENST00000328963	missense_variant	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**index**	**Mother**	**Father**	**Effect Impact**
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
	32.0	D	D	D	NA	NA	[redacted]	[redacted]

Figure 26 illustates the P2RX7 protein and its domains.

Figure- 26 P2RX7 protein domain

FMAL053 *ATP8A2* (OMIM 605870)

The gene's protein is a member of the ATPase P4 protein group that is considered a part of a lipid fluctuation process, which involves the transmission of phospholipids from the exoplasmic leaflet in the cytosolic leaflet of the cell membrane, which aids asymmetries in membrane lipid production and maintenance. The E1 E2 ATPase, the haloacid dehalogenase-like (HAD) domain and multiple transmembrane domains are predicted to contain this protein. For the translocation of phosphatidylserine through membranes, associations between this protein and cell cycle control protein 50A are important.

Cerebellar ataxia and cognitive disabilities have been associated with mutations in this gene with a syndrome (CAMRQ4). In an individual with neurological dysfunction, a translocation breakpoint was also found within this gene. Multiple transcript variants encoder various isoforms result in alternative splintening [98].

P4-ATPase flippase catalytic component, which catalyzes ATP hydrolysis and carries aminophospholipids externally to the internal flux of various membrane and ensures asymmetric distribution of phospholipids. The formation of vesicles and lipid signaling molecules also appear to be involved in phospolipid translocation. In a lesser extent phosphatidyl serine (PS) and phosphatidylethanolamine (PE) is mainly contained in ATP8A2:TMEM30A flippase complex reconstituted with liposomes.

It is suitable for generating, maintaining and trafficking in vesicles in neuronal cells in photoreceptors disk membranes and neuronal axon membranes. This gene is involved in the regulation of neurite outgrowth; acts synergistically with TMEM30A and is required to perform a normal visual and auditory function and contributes to photoreceptor and internal ear survival of spiral ganglion cells.

Cerebel ataxia, mental delay and dysequilibrium syndrome 4 (CMARQ4) (MIM:615268) are associated with this gene. In patients with severe mental retardation and major hypotonia a chromosomal aberration disrupting ATP8A2 has also been found [99]–[101].

Table 22 summarize the information of the variant found in our patient.

Table- 22 *ATP8A2* homozygous missense variant

Gene Name	**Transcript ID**	**Consequence**	**HGVS coding**	**Human Splicing Finder**	**Amino Acid change**	**Amino Acid length**	**Exon**	**CDS Position**
ATP8A2	ENST00000255283	stop_gained	c.1636C>T	c.1636C>T	p.Arg546Ter	1123	19/35	1636/3372
Chr	**Position**	**dbSNP**	**Ref**	**Alt**	**Index**	**Mother**	**Father**	**Effect Impact**
13	26151250	rs755133567	C	T	T/T	C/T	C/T	HIGH
ClinVar	**CADD**	**SIFT**	**PP2**	**MT**	**ExAC**	**1000GP**	**gnomAD**	**Internal Freq**
LP	41.0			A	0.000008	NA	0.000008	0.000788

Figure 27 illustates the ATP8A2 protein and its domains.

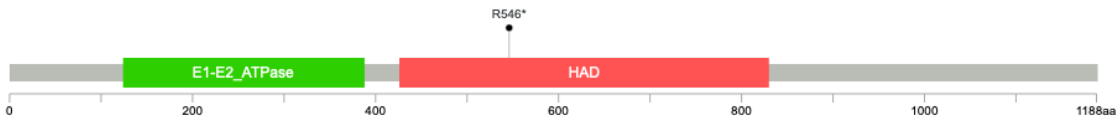

Figure- 27 ATP8A2 protein domains and the mutation

5. DISCUSSION

In this study, three paediatric neurology departments from different parts of Turkey (Izmir, Diyarbakir, Malatya) worked together in Izmir and Newcastle upon Tyne (UK). Important health-related issues for the Turkish population were the focus of the study.

It was a well-known fact that consanguineous marriages in Turkey was higly common. Consanguineous people shared pieces of their DNA because of their common ancestry. By doing this, they also shared some of the same genetic mutations if it was already occured in common ancestry. For this reason, the likelihood of recessive disorders was significantly higher in consanguineous families as compared to non-consanguineous families

More genes are active in the brain and nervous system than in any other part of the body. Recessive defects in these genes often lead to severe childhood disorders causing early death or severe disability including seizures, muscle weakness, breathing problems and severe learning difficulties. Thousands of children affected by neurogenetic conditions are born every year in Turkey in consanguineous families. A genetic diagnosis that would allow families to avoid having further affected children and, in some cases allow effective treatment to be prescribed is rarely achieved because the underlying causes are difficult to pinpoint and the tests required have until now been expensive and time-consuming.

Rather than looking at genes one by one, it has now become possible to read the entire genetic code of a child (the whole genome) in a single test and identify faulty genes by comparing this information with the general population and healthy relatives, in particular, the parents. This is especially effective in consanguineous families, as the affected child is expected to have inherited an identical mutation from both Father and Mother. Information obtained from consanguineous families has already been helpful in the identification of disease-causing genetic mutations and might also make it possible to find other genes that cause the primary disease more

or less severe.

The aim of this study was to build an important body of genomic data relevant to the Turkish population and increase genomics research capacity within the Turkish medical community by using exome sequencing. Exome sequencing allows early identification of treatable cases with high accuracy, even prior to the onset of severe or disease-specific symptoms. This research was also intended to lead development of prevention and treatment strategies and improved outcomes for affected families. In addition to these, this research was designed to generate sustainable resources for knowledge transfer and joint work, as well as creating economic opportunities for the diagnostic and healthcare market.

In terms of the first research aim, which focused at the to identify new disease genes and causal disease variant that underlie the neurogenetic diseases found in consanguineous families in Turkey, the findings of this study seem to discover new genes and variant that were never reported before.

One of the main aims of the study was to increase the proportion of families with a genetic diagnosis to enable provision of reproductive counselling and disease-specific therapies. The findings showed that WES provided a molecular diagnosis for 62 families out of 138 (45%). By the help of whole exome sequencing and WES data analysis, many patients were able to get diagnosed with precise disease-causing defects. Moreover, many neurogenetic conditions had effective treatments available which can only be offered once the genetic diagnosis is known. Therefore, theresults provided a treatment for several patients when a certain diagnosis is completed.

In the investigation of the other research question, explaining genotype-phenotype relations in detail through systematic deep phenotyping of all affected individuals, the body of knowledge on genotype-phenotype correlations and as the foundation for further research into the social and economic burden of these conditions and appropriate healthcare provider were extended.

The analysis of the WES data in relation to improve understanding the biological understanding of recessive neurological disorders in children, seems to help to provide a better understanding of the underlying molecular and cellular pathophysiology.

Many variants and genes were analyzed to discover underlying mechanism in regard to disease phenotype. These contributions will help to identify new therapeutic targets and strategies in the future.

One of the main concepts of this study was to form the comprehensive genomic and phenotypic cohort of Turkish consanguineous families characterized to date and contribute to the Turkish genomic variome, and identification of disease-associated variants. The findings of this study seem to lead new reseach questions and follow-up studies. With the help of this study, many genes and variants will be examined in the scope of disease inheritance pattern and disease mechanisms.

Our findings suggest that patients who had born into consangunious families with a suspected neurogenetic disease can also show other inheritance patterns apart from recessive inheritance pattern. Many variants were found as *de novo* and also compound heterozygous. Majority of the patients were found with homozygous, missense mutation. However, there were several cases in which patients had stop-gained and frameshift mutations too.

The human cerebral cortex development is a highly organized and complicated process. Any disruption in overlapping steps contributing to this process can lead to a large spectrum of disorders in development. Many of these disorders were recognized as brain malformations by the help of brain imaging studies. The classification of the various malformations of the cortical development has evolved to reflect their underlying processes. While these classifications summarize the main developental steps in brain formation, recent developments have challenged the established boundaries between these defined stages and suggest that the genes involved in many stages of development are genetically and functionally interdependent.

The recent developments in next-generation sequencing technologies has enabled the rapid identification of a wide range of genes and their mechanisms underlying brain malformation disorders. Further progress is needed but it is limited due to lack of availability of well-characterized and deep phenotyped patients and the ability to analyze gene function in detail.

The findings of this study are consistent with previous research. Results showed that there is considerable variation in underlying genetic basis of brain malformation anomalies. Many genes participate in wide range of biological processes were suggested as the reason for diseases related to brain malformation.

In this study, we applied whole exome sequencing (WES) to a cohort of 192 families in consanguine families 854 people congenital neurogenetic diseases. The genes and numerous rare disease-causing variants were identified by WES data analysis. Also, several novel genes not associated with that specific patient phenotype previously were suggested by considering their fit into known biological processes.

6. CONCLUSION

In this study, we examined the underlying reason of suspected neurogenetic diseases with neuropediatric patients born into consangunious marriges. We analyzed the patient WES data with through RD-Connect Genome-Phenome Analysis Platform. After examining of 854 people from 192 families in consanguine families, 45% of patient were diagnosed. Analysis concluded that the clinical usage of WES and WES analysis through RD-connect platform resulted in outcomes that were better than similr studies. Although previous findings indicated that WGS is more advantageous than WES, our study showed that a great number of patients were diagnosed with WES similar to those in major studies conducted by WGS previously. It is possible that outcomes would vary afterwards as the study continues.

The application of NGS technology produce many gene variations. The challenge is how these data can be used in clinical practice. The collection and characterisation of disease related mutations and their effects would improve the diagnosis of genetic disorders if the patient data could be transferred to clinical usage effectively.

When searching for disease-causing variants, the availability of the data belong to large families with many affected family member or numerous pedigrees with the same genetically homogeneous disorder is fundamentally important in the clinical practice. Therefore, multiple families with a similar condition can accelerate greatly finding the causal variants.

In the rapid era of NGS, WES and WGS are commonly used in clinical and diagnostic applications. Exome sequencing in affected family members with neurodevelopmental disorders born into consanguineous families have already provided a high diagnostic yield even though clinical and genetic heterogeneity of neurodevelopmental disease groups.

WES applications in medical research and how we understand the genetic mechanisms of the diseases. Genome sequencing allows early identification of

treatable cases with high accuracy, even before the onset of severe or disease-specific symptoms. Next generation platforms can significantly improve the efforts of gene discovery and diagnosis in cortical developmental malformations.

7. BIBLIOGRAPHY

[1] H. Hamamy, "Consanguineous marriages : Preconception consultation in primary health care settings.," *J. Community Genet.*, vol. 3, no. 3, pp. 185–92, Jul. 2012.

[2] H. Hamamy and S. Alwan, "The sociodemographic and economic correlates of consanguineous marriages in highly consanguineous populations," *Elsevier*, 2016.

[3] M. A. Hashmi, "Frequency of consanguinity and its effect on congenital malformation--a hospital based study.," *J. Pak. Med. Assoc.*, vol. 47, no. 3, pp. 75–78, 1997.

[4] A. H. Bittles and M. L. Black, "Consanguineous Marriage and Human Evolution," *Annu. Rev. Anthropol.*, vol. 39, no. 1, pp. 193–207, Oct. 2010.

[5] A. H. Bittles, "Endogamy, consanguinity and community genetics," *Journal of Genetics*, vol. 81, no. 3. pp. 91–98, 2002.

[6] A. M. Erzurumluoglu, H. A. Shihab, S. Rodriguez, T. R. Gaunt, and I. N. M. Day, "Importance of Genetic Studies in Consanguineous Populations for the Characterization of Novel Human Gene Functions," *Ann. Hum. Genet.*, vol. 80, no. 3, pp. 187–196, May 2016.

[7] E. M. Scott *et al.*, "Characterization of Greater Middle Eastern genetic variation for enhanced disease gene discovery," *Nat. Genet.*, vol. 48, no. 9, pp. 1071–1076, 2016.

[8] H. Hamamy, H., Antonarakis, S. E., Cavalli-Sforza, L. L., Temtamy, S., Romeo, G., Ten Kate, L. P., ... & Bathija, "Consanguineous marriages, pearls and perils: Geneva international consanguinity workshop report.," *nature.com*, vol. 13, pp. 841–847, 2011.

[9] F. C. Ceballos, P. K. Joshi, D. W. Clark, M. Ramsay, and J. F. Wilson, "Runs of homozygosity: Windows into population history and trait architecture," *Nature Reviews Genetics*. 2018.

[10] National Library of Medicine (US) Genetics Home Reference [Internet].Bethesda (MD): The Library, "What is a gene mutation and how do mutations occur?," *July 26*, 2016. .

[11] S. H. Tsang and T. Sharma, Eds., *Atlas of Inherited Retinal Diseases*, 1st ed., vol. 1085. Cham: Springer International Publishing, 2018.

[12] R. Karki, D. Pandya, R. C. Elston, and C. Ferlini, "Defining 'mutation' and 'polymorphism' in the era of personal genomics," *BMC Medical Genomics*, vol. 8, no. 1. p. 37, 15-Dec-2015.

[13] R. M. Cormack, D. L. Hartl, and A. G. Clark, "Principles of Population Genetics.," *Biometrics*, vol. 46, no. 2, p. 546, 1990.

[14] C. A. Scacheri and P. C. Scacheri, "Mutations in the noncoding genome," *Current Opinion in Pediatrics*. 2015.

[15] S. Chatterjee and N. Ahituv, "Gene Regulatory Elements, Major Drivers of Human Disease," *Annu. Rev. Genomics Hum. Genet.*, 2017.

[16] J. Muys, B. Blaumeiser, Y. Jacquemyn, and K. Janssens, "Prenatal homozygosity mapping detects a novel mutation in CHST3 in a fetus with skeletal dysplasia and joint dislocations," *Clin. Case Reports*, 2017.

[17] H. Chen, *Atlas of genetic diagnosis and counseling*. 2006.

[18] H. Z. Ring, P. Y. Kwok, and R. G. H. Cotton, "Human Variome Project: An international collaboration to catalogue human genetic variation," in *Pharmacogenomics*, 2006.

[19] O. Mayo, "A Century of Hardy–Weinberg Equilibrium," *Twin Res. Hum. Genet.*, vol. 11, no. 03, pp. 249–256, 2008.

[20] A. Bittles, "Consanguinity and its relevance to clinical genetics," *Clin. Genet.*, vol. 60, no. 2, pp. 89–98, Jul. 2008.

[21] M. S. Reuter *et al.*, "Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders," *JAMA Psychiatry*, vol. 74, no. 3, pp. 293–299, 2017.

[22] J. W. Belmont *et al.*, "The international HapMap project," *Nature*, 2003.

[23] L.-J. C. Wong, Ed., *Next Generation Sequencing Based Clinical Molecular Diagnosis of Human Genetic Disorders*. Cham: Springer International Publishing, 2017.

[24] F. S. Alkuraya, "The application of next-generation sequencing in the autozygosity mapping of human recessive diseases," *Human Genetics*, vol. 132, no. 11. pp. 1197–1211, 2013.

[25] G. Bademci *et al.*, "Identification of Copy Number Variants Through Whole-Exome Sequencing in Autosomal Recessive Nonsyndromic Hearing Loss," *Genet. Test. Mol. Biomarkers*, 2014.

[26] A. Gazzo *et al.*, "Understanding mutational effects in digenic diseases," *Nucleic Acids Res.*, 2017.

[27] WHO, "Genes and human disease," *Genomic Resource Centre*, 2016. .

[28] R. Hussain, A. H. Bittles, and S. Sullivan, "Consanguinity and early mortality in the muslim populations of India and Pakistan," *Am. J. Hum. Biol.*, vol. 13, no. 6, pp. 777–787, Nov. 2001.

[29] R. McQuillan, A. Leutenegger, and … R. A.-R., "Runs of homozygosity in European populations," *Elsevier*.

[30] J. Z. Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, "Genomic patterns of homozygosity in worldwide," *The American Journal of Human Genetics*. 2012.

[31] F. S. Alkuraya, "Homozygosity mapping: One more tool in the clinical geneticist's toolbox," *Genet. Med.*, vol. 12, no. 4, pp. 236–239, 2010.

[32] J. Goodship, H. Gill, J. Carter, A. Jackson, M. Splitt, and M. Wright, "Autozygosity mapping of a seckel syndrome locus to chromosome 3q22. 1-q24," *Am J Hum Genet*, vol. 67, no. 2, pp. 498–503, 2000.

[33] R. Calderón, C. L. Hernández, G. García-Varela, D. Masciarelli, and P. Cuesta, "Inbreeding in Southeastern Spain: The Impact of Geography and Demography on Marital Mobility and Marital Distance Patterns (1900–1969)," *Hum. Nat.*, 2018.

[34] E. G. Puffenberger *et al.*, "Genetic mapping and exome sequencing identify variants associated with five novel diseases," *PLoS One*, vol. 7, no. 1, p. e28936, Jan. 2012.

[35] Y. Saillour and J. Chelly, "Genetic Causes of Intellectual Disability: The Genes Controlling Cortical Development," in *Neuronal and Synaptic Dysfunction in Autism Spectrum Disorder and Intellectual Disability*, 2016.

[36] B. Rabbani, M. Tekin, and N. Mahdieh, "The promise of whole-exome sequencing in medical genetics," *Journal of Human Genetics*. 2014.

[37] Z. E. Sauna and C. Kimchi-Sarfaty, "Understanding the contribution of synonymous mutations to human disease," *Nature Reviews Genetics*, vol. 12, no. 10. pp. 683–691, 2011.

[38] D. Seelow and M. Schuelke, "HomozygosityMapper2012-bridging the gap between homozygosity mapping and deep sequencing," *Nucleic Acids Res.*, vol. 40, no. W1, 2012.

[39] S. Purcell *et al.*, "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *Am. J. Hum. Genet.*, vol. 81, no. 3, pp. 559–575, 2007.

[40] A. Gusev *et al.*, "Whole population, genome-wide mapping of hidden relatedness," *Genome Res.*, 2009.

[41] T. Pippucci, A. Magi, A. Gialluisi, and G. Romeo, "Detection of runs of homozygosity from whole exome sequencing data: State of the art and perspectives for clinical, population and epidemiological studies," *Hum. Hered.*, 2014.

[42] J. Oliveira, R. Pereira, R. Santos, and M. Sousa, "Homozygosity Mapping using Whole-Exome Sequencing: A Valuable Approach for Pathogenic Variant Identification in Genetic Diseases," in *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, 2017.

[43] A. Magi *et al.*, "H3M2: detection of runs of homozygosity from whole-exome sequencing data," *Bioinformatics*, 2014.

[44] E. Parrini, V. Conti, W. B. Dobyns, and R. Guerrini, "Genetic basis of brain malformations," *Molecular Syndromology*. 2016.

[45] G. D. Clark, "The classification of cortical dysplasias through molecular genetics," *Brain Dev.*, no. 26(6), pp. 351–362, 2004.

[46] R. Guerrini and E. Parrini, "Malformations of cortical development: Genetic aspects," in *Neonatology: A Practical Approach to Neonatal Diseases*, 2012.

[47] E. Parrini, P. Sgadò, and R. Guerrini, "Single Gene Mutations Causing Epileptogenic Malformations of the Cerebral Cortex," 2009.

[48] S. S. Sanders, "Whole-exome sequencing: A powerful technique for identifying novel genes of complex disorders," *Clinical Genetics*, vol. 79, no. 2. pp. 132–133, Feb-2011.

[49] E. Karaca *et al.*, "Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease," *Neuron*, 2015.

[50] R. Guerrini and W. B. Dobyns, "Malformations of cortical development: Clinical features and genetic causes," *The Lancet Neurology*, vol. 13, no. 7. pp. 710–726, 2014.

[51] A. J. Barkovich, R. Guerrini, R. I. Kuzniecky, G. D. Jackson, and W. B. Dobyns, "A developmental and genetic classification for malformations of cortical development: Update 2012," *Brain*, vol. 135, no. 5. pp. 1348–1369, 2012.

[52] G. M. Mirzaa and A. R. Paciorkowski, "Introduction: Brain malformations," *Am. J. Med. Genet. Part C Semin. Med. Genet.*, vol. 166, no. 2, pp. 117–123, Jun. 2014.

[53] M. E. Ross and C. A. Walsh, "Human Brain Malformations and Their Lessons for Neuronal Migration," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1041–1070, Mar. 2001.

[54] M. Raymond and F. Rousset, "An Exact Test for Population Differentiation," *Evolution (N. Y).*, vol. 49, no. 6, p. 1280, Dec. 1995.

[55] A. J. Barkovich and B. O. Kjos, "Gray matter heterotopias: MR characteristics and correlation with developmental and neurologic manifestations.," *Radiology*, vol. 182, no. 2, pp. 493–499, Feb. 1992.

[56] A. Haghighi *et al.*, "An integrated clinical program and crowdsourcing strategy for genomic sequencing and Mendelian disease gene discovery," *npj Genomic Medicine*. 2018.

[57] S. El Chehadeh, B. Aral, … N. G.-J. of medical, and undefined 2010, "Search for the best indicators for the presence of a VPS13B gene mutation and confirmation of diagnostic criteria in a series of 34 patients genotyped for," *jmg.bmj.com*.

[58] I. Ionita-Laza, M. Capanu, S. De Rubeis, K. McCallum, and J. D. Buxbaum, "Identification of Rare Causal Variants in Sequence-Based Studies: Methods and Applications to VPS13B, a Gene Involved in Cohen Syndrome and Autism," *PLoS Genet.*, vol. 10, no. 12, p. e1004729, Dec. 2014.

[59] M. A. Rafiq *et al.*, "Novel VPS13B Mutations in Three Large Pakistani Cohen Syndrome Families Suggests a Baloch Variant with Autistic-Like Features," *BMC Med. Genet.*, vol. 16, no. 1, p. 41, Dec. 2015.

[60] I. Balikova *et al.*, "Deletions in the *VPS13B* (*COH1*) gene as a cause of Cohen syndrome," *Hum. Mutat.*, vol. 30, no. 9, pp. E845–E854, Sep. 2009.

[61] J. Dastan *et al.*, "Exome sequencing identifies pathogenic variants of VPS13B in a patient with familial 16p11.2 duplication," *BMC Med. Genet.*, vol. 17, no. 1, p. 78, Dec. 2016.

[62] C. Babbs, D. Lloyd, … A. P.-J. of medical, and undefined 2014, "De novo and rare inherited mutations implicate the transcriptional coregulator TCF20/SPBP in autism spectrum disorder," *jmg.bmj.com*.

[63] X. Wang, V. Sutton, J. Peraza-Llanes, … Z. Y.-N., and undefined 2007, "Mutations in X-linked PORCN, a putative regulator of Wnt signaling, cause focal dermal hypoplasia," *nature.com*.

[64] P. Leoyklang, K. Suphapeetiporn, … S. W.-C., and undefined 2008, "Three novel mutations in the PORCN gene underlying focal dermal hypoplasia," *Wiley Online Libr.*

[65] S. Clements, J. Mellerio, … S. H.-B. J. of, and undefined 2009, "PORCN gene mutations and the protean nature of focal dermal hypoplasia," *Wiley Online Libr.*

[66] K. Jin-Ah, Y. Tan, X. Wang, X. Cao, and J. Veeraraghavan, "Genomic deregulation and therapeutic role of the cell-cycle kinase TLK2 in more aggressive breast cancers," 2015.

[67] J. Kim, M. Anurag, J. Veeraraghavan, … R. S.-M. C., and undefined 2016, "Amplification of TLK2 Induces Genomic Instability via Impairing the G2/M Checkpoint.," *AACR*.

[68] C. Jauset, S. S.-B.-… O. T. CELL, and undefined 2016, "Identification of

molecular determinants of Tousled-like kinase 2 (TLK2) activity," *AMER SOC CELL Biol. 8120 …*.

[69] M. Reijnders, K. Miller, M. Alvi, … J. G.-T. A. J. of, and undefined 2018, "De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder," *Elsevier*.

[70] L. Pisciotta, S. Uccella, T. Giacomini, … C. C.-, and undefined 2018, "Cask-Related Disorders: Clinical, Electroencephalographic and Neuroradiological Description of Four Genetically Confirmed Cases," *thieme-connect.com*.

[71] X. Bozarth, K. Foss, H. M.-A. J. of Medical, and undefined 2018, "A de novo in-frame deletion of CASK gene causes early onset infantile spasms and supratentorial cerebral malformation in a female patient," *Wiley Online Libr.*

[72] L. Pisciotta *et al.*, "Cask-Related Disorders: Clinical, Electroencephalographic and Neuroradiological Description of Four Genetically Confirmed Cases," *Neuropediatrics*, vol. 49, no. S 01, pp. S1–S12, Apr. 2018.

[73] E. Ebrahimi, S. Ghazala, … G. P.-J. of cellular, and undefined 2018, "Structural analysis of missense mutations in galactokinase 1 (GALK1) leading to galactosemia type-2.," *europepmc.org*.

[74] D. Li, J. Yang, H. Ye, Y. Tang, Y. Luo, and Y. Lu, "Target region sequencing in sporadic congenital cataracts reveals a new genotype-phenotype relation for the GALK1 gene," *Acta Ophthalmol.*, vol. 93, p. n/a-n/a, Oct. 2015.

[75] F. Radmanesh *et al.*, "Mutations in LAMB1 cause cobblestone brain malformation without muscular or ocular abnormalities.," *Am. J. Hum. Genet.*, vol. 92, no. 3, pp. 468–74, Mar. 2013.

[76] D. Tonduti, I. Dorboz, F. Renaldo, J. M.-P.- Neurology, and undefined 2015, "Cystic leukoencephalopathy with cortical dysplasia related to LAMB1 mutations," *AAN Enterp.*

[77] T. Okazaki, Y. Saito, T. Hayashida, … S. A.-C., and undefined 2018, "Bilateral cerebellar cysts and cerebral white matter lesions with cortical dysgenesis: Expanding the phenotype of LAMB1 gene mutations," *Wiley Online Libr.*

[78] S. Weitzer, T. Hanada, J. M. Penninger, and J. Martinez, "CLP1 as a novel player in linking tRNA splicing to neurodegenerative disorders," *Wiley Interdiscip. Rev. RNA*, vol. 6, no. 1, pp. 47–63, Jan. 2015.

[79] A. E. Schaffer *et al.*, "CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration," *Cell*, vol. 157, no. 3, pp. 651–663, 2014.

[80] E. M. Jenkinson *et al.*, "Comprehensive molecular screening strategy of *OCLN* in band-like calcification with simplified gyration and polymicrogyria," *Clin. Genet.*, vol. 93, no. 2, pp. 228–234, Feb. 2018.

[81] M. Abdel-Hamid, G. Abdel-Salam, … M. I.-J. of human, and undefined 2017, "Band-like calcification with simplified gyration and polymicrogyria: report of 10 new families and identification of five novel OCLN mutations," *nature.com*.

[82] Y. Du *et al.*, "Increased cerebral expressions of MMPs, CLDN5, OCLN, ZO1 and AQPs are associated with brain edema following fatal heat stroke," *nature.com*.

[83] P. Shepard, B. Barshop, … M. B.-G. in, and undefined 2015, "Consanguinity and rare mutations outside of MCCC genes underlie nonspecific phenotypes of MCCD," *nature.com*.

[84] Y. Kim, G. Seo, G. Kim, … H. Y.-J. of G., and undefined 2017, "Novel heterozygous MCCC1 mutations identified in a patient with 3-methylcrotonyl-coenzyme A carboxylase deficiency," *e-sciencecentral.org*.

[85] P. Ordonez, S. Kunzelmann, H. Groom, M. Y.-S. reports, and undefined 2017, "SAMHD1 enhances nucleoside-analogue efficacy against HIV-1 in myeloid cells," *nature.com*.

[86] J. Antonucci, C. Gelais, S. de Silva, J. Y.-N. medicine, and undefined 2016, "SAMHD1-mediated HIV-1 restriction in cells does not involve ribonuclease activity," *nature.com*.

[87] M. Cabello-Lobato, S. Wang, C. S.-T. in Genetics, and undefined 2017, "SAMHD1 Sheds Moonlight on DNA Double-Strand Break Repair," *Elsevier*.

[88] Y. J. Crow *et al.*, "Characterization of human disease phenotypes associated with mutations in *TREX1* , *RNASEH2A* , *RNASEH2B* , *RNASEH2C* , *SAMHD1* , *ADAR* , and *IFIH1*," *Am. J. Med. Genet. Part A*, vol. 167, no. 2, pp. 296–312, Feb. 2015.

[89] G. Yigit *et al.*, "A syndrome of microcephaly, short stature, polysyndactyly, and dental anomalies caused by a homozygous *KATNB1* mutation," *Am. J. Med. Genet. Part A*, vol. 170, no. 3, pp. 728–733, Mar. 2016.

[90] E. Faqeih, L. Bastaki, … R. R.-A. J. of, and undefined 2015, "Novel STAMBP mutation and additional findings in an Arabic family," *Wiley Online Libr.*

[91] I. Hori, F. Miya, Y. Negishi, A. Hattori, … N. A.-J. of human, and undefined 2018, "A novel homozygous missense mutation in the SH3-binding motif of STAMBP causing microcephaly-capillary malformation syndrome," *nature.com*.

[92] W. Goh, G. Ow, … V. K.-A. journal of, and undefined 2015, "DLAT subunit of the pyruvate dehydrogenase complex is upregulated in gastric cancer-implications in cancer therapy," *ncbi.nlm.nih.gov*.

[93] R. Sinkey, S. Ho, … S. S.-, and undefined 2016, "A Tissue-Based Study of VEGFR2 in Human Placentas: Associations with Pyruvate Dehydrogenase Component E2 (DLAT) and Sirtuin-4 (SIRT4).," *SAGE Publ. INC 2455 …*.

[94] E. Plaisier and P. Ronco, "COL4A1-related disorders," 2016.

[95] K. Rannikmäe, G. Davies, P. Thomson, S. B.- Neurology, and undefined 2015, "Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease," *AAN Enterp.*

[96] M. Mao, M. Kiss, Y. Ou, & D. G.-D. models, and undefined 2017, "Genetic dissection of anterior segment dysgenesis caused by a Col4a1 mutation," *dmm.biologists.org*.

[97] A. Vereczkei, O. Abdul-Rahman, … Z. H.-P. in N., and undefined 2019, "Association of purinergic receptor P2RX7 gene polymorphisms with depression symptoms," *Elsevier*.

[98] A. Saghazadeh, S. Tonekaboni, … H. N.-A. M., and undefined 2018, "A Novel Homozygous ATP8A2 Variant in a Patient With Phenotypic Features of Dysequilibrium Syndrome," *acta.tums.ac.ir*.

[99] S. Quintas, O. Moldovan, … T. dos S.-E. J. of, and undefined 2017, "New syndrome associated with ATP8A2 gene mutations: Encephalopathy, intellectual disability, severe hypotonia, chorea and optic atrophy Whole-exome," *ejpn-journal.com*.

[100] S. Mikkelsen, L. Mogensen, R. Gantzel, and B. Vilsen, "N905 of TM6 is crucial to the function of the mammalian flippase ATP8A2: Poster at the 15th International Conference on Na, K-ATPase and Related Transport," 2017.

[101] L. Mogensen, S. Mikkelsen, R. Gantzel, and J. Coleman, "Mutational Analysis of a Peripheral Pathway for Phospholipid Transport in ATP8A2: Poster at the 15th International Conference on Na, K-ATPase and Related," 2017.

8. Appendix

APPENDIX A: Ethical Committee Report
APPENDIX B: Consent Form Example

APPENDIX A: Ethical Committee Report

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Türkiye'de Akraba Evliliklerine Bağlı Nörogenetik Hastalık Yükünün Araştırılmasına Yeni Genomik Yaklaşımlar
VARSA ARAŞTIRMANIN PROTOKOL KODU	-
ETİK KURUL PROTOKOL NUMARASI	302-SBKAEK

ETİK KURUL BİLGİLERİ	ETİK KURULUN ADI	Dokuz Eylül Üniversitesi Klinik Araştırmalar Etik Kurulu
	AÇIK ADRESİ:	Dokuz Eylül Üniversitesi Sağlık Yerleşkesi Dekanlık Binası Kat:2 İnciraltı 35340 İZMİR-TÜRKİYE
	TELEFON	0 232 4122254 - 0 232 4122258
	FAKS	0232 4122243
	E-POSTA	etikkurul@deu.edu.tr

BAŞVURU BİLGİLERİ	KOORDİNATÖR/SORUMLU ARAŞTIRMACI UNVANI/ADI/SOYADI	Prof.Dr.Semra HIZ			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	Nöroloji			
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	Dokuz Eylül Üniversitesi Çocuk Nöroloji A.D			
	VARSA IDARI SORUMLU UNVANI/ADI/SOYADI	-			
	DESTEKLEYİCİ				
	PROJE YÜRÜTÜCÜSÜ UNVANI/ADI/SOYADI (TÜBİTAK vb. gibi kaynaklardan destek alanlar için)				
	DESTEKLEYİCİNİN YASAL TEMSİLCİSİ				
	ARAŞTIRMANIN FAZI VE TÜRÜ	FAZ 1	☐		
		FAZ 2	☐		
		FAZ 3	☐		
		FAZ 4	☐		
		Gözlemsel ilaç çalışması	☐		
		Tıbbi cihaz klinik araştırması	☐		
		İn vitro tıbbi tanı cihazları ile yapılan performans değerlendirme çalışmaları	☐		
		İlaç dışı klinik araştırma	☒		
		Diğer ise belirtiniz			
	ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ ☒	ÇOK MERKEZLİ ☐	ULUSAL ☒	ULUSLARARASI ☐

	Belge Adı	Tarihi	Versiyon Numarası	Dili
DEĞERLENDİRİLEN BELGELER	ARAŞTIRMA PROTOKOLÜ	mevcut		Türkçe ☐ İngilizce ☐ Diğer ☐
	Bilgilendirilmiş Gönüllü Olur Formu	mevcut		Türkçe ☐ İngilizce ☐ Diğer ☐
	OLGU RAPOR FORMU	mevcut		Türkçe ☐ İngilizce ☐ Diğer ☐
	ARAŞTIRMA BROŞÜRÜ			Türkçe ☐ İngilizce ☐ Diğer ☐
DEĞERLENDİRİLEN DİĞER BELGELER	**Belge Adı**			**Açıklama**
	SİGORTA	☐		
	ARAŞTIRMA BÜTÇESİ	☐	Mevcut 28.01.2016	
	BİYOLOJİK MATERYEL TRANSFER FORMU	☐		
	İLAN	☐		
	YILLIK BİLDİRİM	☐		
	SONUÇ RAPORU	☐		
	GÜVENLİLİK BİLDİRİMLERİ	☐		
	DİĞER:	☐	-Araştırma akış şeması -Biyolojik Materyal Transfer Formu 29.01.2016 tarihli imzalı -Araştırmacı özgeçmişleri	

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof.Dr.Ayşegül Yıldız
İmza:

Mukaddes AKKEÇE
Enstitü Sekreteri

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Türkiye'de Akraba Evliliklerine Bağlı Nörogenetik Hastalık Yükünün Araştırılmasına Yeni Genomik Yaklaşımlar
VARSA ARAŞTIRMANIN PROTOKOL KODU	-
ETİK KURUL PROTOKOL NUMARASI	302-SBKAEK

KARAR BİLGİLERİ	Karar No:2016/03-01	Tarih:11.02.2016
	Yukarıda bilgileri verilen başvuru dosyası ile ilgili belgeler araştırmanın/çalışmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup araştırmanın/çalışmanın başvuru dosyasında belirtilen merkezlerde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplantıya katılan etik kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir.	

KLİNİK ARAŞTIRMALAR ETİK KURULU	
ETİK KURULUN ÇALIŞMA ESASI	İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik, İyi Klinik Uygulamaları Kılavuzu
BAŞKANIN UNVANI / ADI / SOYADI:	Prof.Dr.Ayşegül Yıldız

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet		Araştırma ile ilişki		Katılım *		İmza
			E	K	E	H	E	H	
Prof.Dr.Ayşegül YILDIZ	Psikiyatri	DEU Tıp Fakültesi Psikiyatri Anabilim Dalı	☐	☒	☐	☒	☐	☐	
Prof.Dr.Hülya ELLİDOKUZ	Halk Sağlığı	DEU Onkoloji Enstitüsü Prevantif Onkoloji A.D.	☐	☒	☐	☒	☐	☐	
Prof.Dr.Nuray DUMAN	Çocuk Sağlığı ve Hastalıkları (Yeni Doğan)	DEU Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı	☐	☒	☐	☒	☐	☐	
Prof.Dr.Hale ÖREN	Çocuk Sağlığı ve Hastalıkları (Çocuk Hematoloji)	DEU Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı	☐	☒	☐	☒	☐	☐	
Prof.Dr.A.Necati GÖKMEN	Anesteziyoloji ve Reanimasyon	Anesteziyoloji ve Reanimasyon A.D	☒	☐	☐	☒	☐	☐	ANG
Prof.Dr.Taner DAĞCI	Fizyoloji	Ege Üniversitesi Tıp Fakültesi	☒	☐	☐	☒	☐	☐	katılacağ
Doç.Dr.Pembe KESKİNOĞLU	Biyoistatistik	DEU Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim A.D	☐	☒	☐	☒	☐	☐	
Doç.Dr.Erdem YAKA	Nöroloji	DEU Tıp Fakültesi Nöroloji A.D	☒	☐	☐	☒	☐	☐	
Doç.Dr.Uğur Önsel TÜRK	Kardiyoloji	Ege Üniversitesi İlaç ve Farmakokinetik Arş-Uyg.Merk	☒	☐	☐	☒	☐	☐	
Doç.Dr.Yasemin BAŞKIN	Temel Onkoloji	DEU Onkoloji Enstitüsü Temel Onkoloji A.D	☐	☒	☐	☒	☐	☐	
Yard.Doç.Dr.Yasemin ERAÇ	Farmakoloji	Ege Üniversitesi Eczacılık Fakültesi Farmakoloji Anabilim Dalı	☐	☒	☐	☒	☐	☐	
Av.Semra MARMARA	Hukuk	DEÜ Rektörlüğü	☐	☒	☐	☒	☐	☐	katılacağ
Av.Nazan PEDÜKCOŞKUN	Hukuk	Alsancak Nevvar Salih İşgören Hastanesi	☐	☒	☐	☒	☐	☐	
Hayat ALBAYRAK	Emekli	Sağlık Mesleği Mensubu Olmayan Üye	☐	☒	☐	☒	☐	☐	

*:Toplantıda Bulunma

Etik Kurul Başkanının
Unvanı/Adı/Soyadı: Prof.Dr.Ayşegül Yıldız
İmza:

Mukaddes AKKEÇELİ
Enstitü Sekreteri

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

APPENDIX B: Consent Form Example

ASGARİ BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU

(<12 yaş ve/veya zihinsel geriliği olan çocuklar için)

Araştırmanın adı: Türkiyede'ki akraba evliliklerine bağlı nörogenetik hastalık yükünün araştırılmasında yeni genomik yaklaşımlar

Değerli Veli,

Türkiye'de akraba evlilikleri oranı yüksektir. Anne-baba arasında akrabalık olması bu çiftin doğacak olan çocuklarında bazı kalıtsal hastalıkların ortaya çıkışını kolaylaştırmaktadır. Bu hastalıkların önemli bir bölümü beyin ve sinir sistemini etkilemekte ve değişik derecelerde özürlülüğe neden olabilmektedir. Bu durum aile ve çocuk için yıpratıcı olup üzücü sonuçlarla sonlanabilmekte, ayrıca hastaların tanı, tedavi ve bakımları hem aileler, hem de ülke ekonomisi açısından ciddi yük oluşturabilmektedir. Biz, Türkiye'deki akraba evliliklerine bağlı ortaya çıkan nörolojik kalıtsal hastalıkların genetik nedenlerinin araştırılması amacı ile bir araştırma yapıyoruz. Araştırmanın adı, "Türkiyede'ki akraba evliliklerine bağlı nörogenetik hastalık yükünün araştırılmasında yeni genomik yaklaşımlar"dır. Bu araştırmanın sonucunda elde edilen bilgiler ile ailelere sonraki gebeliklerden sağlıklı çocuklarının doğması amacına yönelik olarak genetik danışma ve erken teşhis olanaklarının sunulması ve akraba evliliğine bağlı genetik hastalık tesbit edilen hastaların tedavilerine yönelik gelecekteki bilimsel araştırmalara temel oluşturma yönünde faydalar elde edilecektir. Araştırmaya katılım gönüllülük esasına dayalıdır. Sizi bu araştırmaya katılmaya davet ediyoruz. Aşağıda bu araştırmanın nasıl yapılacağı konusunda bilgiler yer almaktadır. Çalışmaya katılmadan once yeterli zaman ayırarak bu bilgileri okumanızı rica ederiz. Bu çalışmaya katılmak isterseniz imzalı onayınızı almamız gerekmektedir. Eğer başka bir klinik çalışmada yer alıyorsanız bu çalışmaya katılamayacağınızı bildirmek isteriz.

Araştırma nasıl yapılacak:

Bu araştırma kan örneğinden yapılacaktır. Bu araştırmanın yapılabilmesi için çocuğunuzdan 2-3 çay kaşığı kadar küçük bir kan örneği alınacaktır. Kan alım işlemi dışında çocuğunuzdan başka herhangi bir şey istenmeyecektir. Çocuğunuzun takip ve tedavileri aynen devam edecektir. Kan alınırken ufak bir ağrı ve kızarıklık olabilir. Bu durumlar geçici olup önemli bir risk ve yan etki oluşturmaz. Bu araştırma 250 çocukta ve anne ile babalarında yapılacak ve 2,5 yıl sürecektir. Alınan kan örneğinden çocuğunuzda mevcut olduğunu düşündüğümüz kalıtsal hastalığın genetik nedenlerini araştıracağız. Bu araştırmalar yurtiçindeki ve yurtdışındaki laboratuvarlarda yapılacaktır. Bu kan tahlilleri için sizden herhangi bir ücret talep edilmeyecek, araştırmaya ait hiçbir gider Sosyal Güvenlik Kurumu'na veya sağlık sigortanıza yansıtılmayacaktır. Size de herhangi bir ücret ödenmeyecektir. Çocuğunuz herhangi bir risk veya rahatsızlığa maruz kalmayacak ve başka herhangi bir sorumluluğu da olmayacaktır. Bu araştırmadan elde edilecek sonuçlar çocuğunuza doğrudan bir yarar sağlamayacaktır. Ancak bundan sonra doğacak çocuklarınız için faydalı olabilir.

Çalışmadan çekilme:

Bu çalışmaya katılmanız tamamiyle sizin isteğinize bağlıdır. Hiçbir baskıya tabi olmadan, tamamen gönüllü olarak katılmalısınız. Bu çalışmaya katılmak zorunda olmadığınız gibi baştan kabul etseniz de sonradan çalışmadan ayrılmayı isteyebilirsiniz. Bunun için kimse size kızmaz ve küsmez, önceden olduğundan farklı bir muamele görmezsiniz. Çocuğunuza tanınan tüm haklar yürürlükte kalır, göreceği izlem ve tedavi hiçbir şekilde etkilenmez.

Veri gizliliği ve tıbbi kayıtlara erişim:

Bu formu imzalayarak çalışmanın verilerini toplama ve kullanma iznini bize vermektesiniz. Bu iznin bir son tarihi bulunmamaktadır, ancak siz herhangi bir zamanda çalışma doktorunuzu haberdar etmek kaydıyla olurunuzu çekebilirsiniz. Olurunuzu çektiğiniz takdirde artık çalışma verileri kullanılmayacak ve paylaşılmayacaktır. Bu çalışmada çocuğunuzun kimliğini ortaya çıkarabilecek tüm kayıtlar gizli tutulacaktır. Araştırma verilerini diğer doktorlarla paylaşabiliriz, ancak bu durumda çocuğunuzun kimlik bilgileri gizli tutulacaktır. Genom verileriniz ve sağlık bilgileriniz sıkı güvenlikli bir veri tabanında saklanacaktır. Bu verileri sadece kullanım izni almak için başvuran ve belli bir araştırma projesinde kullanımına onay verilen araştırmacılar kullanabilecektir. Genom verileriniz ve sağlık bilgilerinizde, isminiz ya da sizi belirlemede kullanılabilecek hiçbir bilgi yer almayacaktır. Veritabanındaki bilgileri kullanmalarına izin verilen araştırmacılar, sizin kimliğinizi belirleme girişiminde bulunmama taahhüdünde bulunacaklardır.

Veri paylaşımı:

Örnekleriniz, genom verileriniz ve sağlık bilgileriniz saklanacak ve diğer araştırmacılarla paylaşılabilecektir. Bu örnek ve bilgiler farklı araştırmalarda, örneğin bazı hastalıklara (kalp hastalığı, kanser, ya da psikiyatrik bozukluklar gibi) nelerin yol açtığını anlamada, yeni bilimsel yöntemler geliştirilmesinde, ya da farklı insan gruplarının nerelerden geldiklerini anlamaya çalışan araştırmalarda kullanılabilecektir.

Eğer çalışmaya katılmayı kabul ederseniz çalışmanın yararları ve olası riskleri konusunda bilgilendiğinizi gösteren bu formun aşağıdaki ilgili bölümünü imzalamanız istenecektir. Ayrıca bu çalışmanın genetik bir araştırma olmasından dolayı alınan kan örneğinin bu amaçla kullanılmasına izin verip vermediğinizi yine aşağıdaki ilgili bölümde belirtmeniz istenecektir. Sizin olurunuz doğrultusunda kandan çalışılan genetik analiz sonrası genetik örnek saklanacak veya siz saklanmasını istemezseniz imha edilecektir. İmzaladıktan sonra bu formun bir kopyası size verilecektir. Aklınıza şimdi gelen veya daha sonra gelebilecek olan soruları istediğiniz zaman aşağıdaki iletişim numaralarından ulaşarak bize sorabilirsiniz.

Prof. Dr. Ayşe Semra Hız Kurul

Uzm. Dr. Ayşe İpek Polat

Tel no: 0 232 4126215 – 0 232 4126226 – 0 506 2922590

Adres: Dokuz Eylül Üniversitesi Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı Çocuk Nörolojisi Bilim Dalı Balçova İzmir

Gönüllünün Beyanı:

Bu "Bilgilendirilmiş Gönüllü Olur Formu"ndaki tüm açıklamaları okudum. Bana yukarıda konusu ve amacı belirtilen araştırma hakkında aşağıda ismi belirtilen hekim tarafından yazılı ve sözlü açıklama yapıldı. Bu araştırmaya katılmak üzere davet edildim. Eğer bu araştırmaya katılırsam hekimim ile aramızda kalması gereken bilgilerin gizliliğine büyük özen ve saygı gösterileceğine inanıyorum. Araştırma sonuçlarının eğitim ve bilimsel amaçlarla kullanımı sırasında kişisel bilgilerimin ihtimamla korunacağı konusunda bana yeterli güven verildi. Araştırmanın yürütülmesi esnasında herhangi bir sebep göstermeden çalışmadan çekilebilirim. Ancak araştırmacıları zor durumda bırakmamak için çalışmadan çekileceğimi önceden bildirmemin uygun olacağının bilincindeyim. Çalışma için yapılacak harcamalarla ilgili ben veya ailem herhangi bir parasal sorumluluk altına girmiyoruz. Bana veya aileme bir ödeme de yapılmayacaktır. Eğer katılmayı reddedersem, bu durumda çocuğumun tıbbi bakımına ve hekimleri ile olan ilişkisine herhangi bir zarar gelmeyeceğini biliyorum.

Kendi başıma belli bir düşünme süresi sonunda bana yapılan tüm açıklamaları ayrıntılarıyla anlamış bulunmaktayım. Bu koşullarda söz konusu araştırmaya kendi rızamla, hiçbir baskı ve zorlama olmaksızın katılmayı kabul ediyorum.

"Türkiyede'ki akraba evliliklerine bağlı nörogenetik hastalık yükünün araştırılmasında yeni genomik yaklaşımlar" araştırması kapsamında çocuğumdan alınan kan örneklerinin,

☐ sadece yukarıda bahse geçen araştırmada kullanılmasına izin veriyorum.

☐ ileride yapılması planlanan tüm araştırmalarda kullanılmasına izin veriyorum.

☐ hiçbir koşulda kullanılmasına izin vermiyorum.

Gönüllü;	**Hastanın velisi;**
Adı-Soyadı:	Adı-Soyadı:
Adresi:	Adresi:
Tel:	Tel:
Tarih:	Tarih;
İmzası	İmzası

Olur alma işlemine tanıklık eden görevlinin;	**Gönüllüye bilgi veren araştırmacının;**
Adı-Soyadı:	Adı-Soyadı:
Adresi:	Adresi:
Tel:	Tel:
Tarih:	Tarih:
İmzası	İmzası

CURRICULUM VITAE

Name Surname	Elmasnur YILMAZ
Address:	Izmir Biomedicine and Genome Center Dokuz Eylül University Health Campus Balcova 35340 Izmir/TURKEY
Mobile:	+90 545 460 07 91
e-mail:	elmasnur.yilmaz@msfr.ibg.edu.tr

EDUCATION

Country Code	Unviersity	Department	Degree	Year of Graduation
TR	Dokuz Eylül University	Molecular Biology and Genetics	MSc.	2019
TR	Istanbul Technical University	Molecular Biology and Genetics	BSc.	2016

RESEARCH INTEREST

Since I was a child, I have always known the brain is a greatest mystery of our life. I have always been driven to understand how our brain can operate bewildering functions. Following my experiences working with neuropediatric patients with severe neurodevelopmental disorders, I have become motivated even more to unveil the brain's mysteries to create novel platforms for understanding and treating neurological disorders. Now, as a graduate student, I am exceedingly eager to unite my longstanding interest in neuroscience with genomics to explore the brain. I am especially interested in the disease-causing variant and gene discovery studies related to neurogenetic diseases.