Mapping Autosomal Recessive Intellectual Disability in Consanguineous Families

By

Muhammad Ilyas

Registration # 46-FBAS/PHDBT/F-15

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

2015-2020

Ph.D
616.042
MUM

Accession No TH23506

I Mental illness

II Reproduction

III Disabilities

IV Genes

Mapping Autosomal Recessive Intellectual Disability in Consanguineous Families

Researcher

Muhammad Ilyas

Reg. # 46-FBAS/PHDBT/F-15

Supervised By

Dr. Asif Mir
Associate Professor
Department of Biological Sciences,
FBAS, IIUI

Co-Supervised
Dr. Abdul Hammed
Chief Scientific Office
IB&GE, KRL, Islamabad

Department of Biological Sciences
Faculty of Basic and Applied Sciences
International Islamic University Islamabad, Pakistan
2015-2020


بسم الله الرحمن الرحيم

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

Date: 19-08-2020

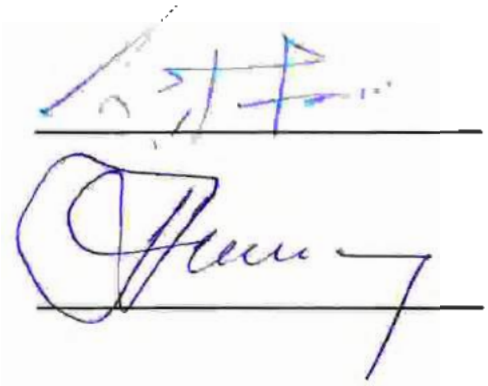
FINAL APPROVAL

It is certified that we have read the thesis submitted by Mr. Muhammad Ilyas and it is our judgment that this project is of sufficient standard to warrant its acceptance by the International Islamic University, Islamabad, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biotechnology.

COMMITTEE

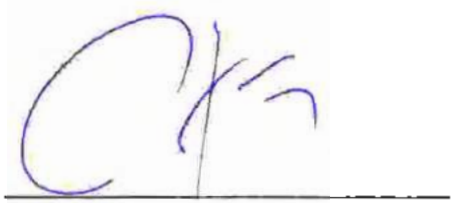
Supervisor

Dr. Asif Mir
Associate Professor
Department of Biological Sciences, IIUI



Co-Supervisor

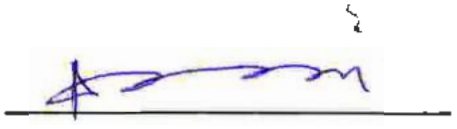
Dr. Abdul Hameed
Chief Scientific officer
IB&GE, KRL, Islamabad



External Examiner

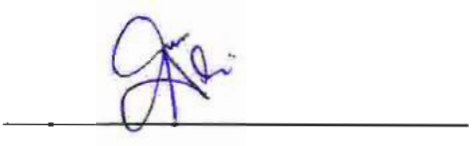
Dr. Muhammad Naeem
Professor of Biotechnology
Quaid-e-Azam University, "QAU" Islamabad

External Examiner



Dr. Bilal Haider Abbasi
Professor / Chairman
Quaid-e-Azam University, "QAU" Islamabad

Internal Examiner



Dr. Syed Ali Imran Bokhari
Assistant Professor
Department of Biological Sciences, IIUI

Department of Biological Sciences

Faculty of Basic and Applied Sciences

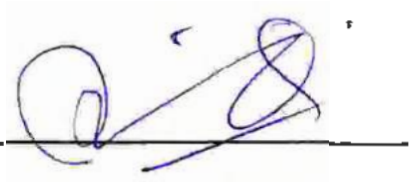
International Islamic University Islamabad, Pakistan

Date:_____

FINAL APPROVAL

It is certified that we have read the thesis submitted by Mr. Muhammad Ilyas and it is our judgment that this project is of sufficient standard to warrant its acceptance by the International Islamic University, Islamabad in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biotechnology.

Chairman:



Dr. Muhammad Arshad Malik
Associate Professor
Department of Biological Sciences
International Islamic University Islamabad, Pakistan

Professor Dr. Muhammad Sajid
Dean, Faculty of Basic and Applied Sciences
International Islamic University Islamabad, Pakistan

This dissertation is

Submitted to International Islamic University, Islamabad, Pakistan in partial fulfillment of the requirements of the doctorate degree in Biotechnology

DEDICATION

This work is dedicated to my sweet and beloved parents and my family whose constant support and guidance enabled me to achieve this milestone.

DECLARATION

It is certified that work done on this PhD Biotechnology research thesis is purely conducted by me. All the material is prepared by myself and has not been copied from anywhere; however some test and figures have been used which are properly referenced.

Date 17-08-2020

Muhammad Ilyas

CONTENTS LIST

LIST OF FIGURE

Figure No **Page No**

LIST OF TABLE

Table No **Page No**

LIST OF ABBREVIATION

AP	Fast-Alkaline phosphatase
ATP	Adenosine Triphosphate
BASAR	Board of Advanced Study and Research
BCL	Binary Base Call
CADD Depletion	Combined Annotation Dependent
CNS	Central Nervous System
CNVs	Copy Number Variation
CT Scan	Computed Tomography Scan
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
FISH	Fluorescence in situ hybridization
GABA	Gamma-amino butyric acid
GERP	Genomic Evolutionary Rate Profiling
ID	Intellectual Disability
IIU	International Islamic University
IQ	Intelligence Quotient
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
NCBI	National Centre for Biotechnology
NGS	Next Generation Sequencing
℃	Centigrade
OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase Chain Reaction

pH	Potential of Hydrogen
PK	Proteinase K
PKU	Phenylketonuria
RNA	Ribose nucleic Acid
SDS	Sodium Dodecyl Sulfate
SLO	Smith-Lemli-Opitz
STE buffer	Sodium Tris EDTA
STR	Short Tandem Repeats
TBE	Tris Boric Acid ETDA
TE	Tris EDTA
TEMD	Tetra methyl ethylene diamine
UCL	University College London
UCSC	University of California, Santa Cruz
UV	Ultra Violet Radiation
VCF	Variants Calling File
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WHO	World Health Organization

ACKNOWLEDGEMENT

In the Name of Allah, the Most Beneficent, the Most Merciful, Peace and Blessing be upon the lovely prophet Mohammed (PBUH). Alhamdulillah, all praises to Almighty Allah that has for the strengths and His blessing due to which I have been able to complete my research.

I would like to acknowledge my respectable and beloved parents, my siblings and family for their sincere care, support and prayers.

My deep gratitude, first And foremost, goes to my supervisor, Dr. Asif Mir, Ph.D. (Associate Professor Department of Biological Sciences, IIU, Islamabad). His assistance, guidance, and feedback have been instrumental throughout not only for this dissertation, but my entire time in the Department at in the Lab. My deep gratitude goes to my Co-supervisor Dr. Abdul Hameed, Ph.D. (Chief Scientific Officer, Institute of Biomedical and Genetic Engineering (IB & GE) KRL hospital Islamabad) for his support. I feel privileged to record my sincere thanks to all the faculty member of Department of Biological Science, International Islamic University Islamabad Pakistan.

I am also thankful to my foreign professor Prof. Henry Holden (University College London, UK) and his research Group (Stephanie Efthyiou, Vincenzo Salpietro, Roisin Sullivan) at Institute of neurology Queen Square UK for their help and support during HEC IRSIP program.

I want to acknowledge and pay my best regards to families for participating my research projects and their willingness to let monitor their lives to conduct this research. I sincerely hope that some of this work assists, at least in some small way, in the improvement of the lives of people with developmental and intellectual disabilities .Thank for your giving me the opportunity to meet and work with you. It was a genuine pleasure.

I would like to appreciate HEC IRSIP program and for finical support, for my research work abroad. Best wishes to all the fellows.

Muhammad Ilyas

ABSRTACT

Intellectual disability (ID) is a neurodevelopmental condition affecting 1–3% of the world's population. Genetic factors play a key role causing the congenital limits in intellectual functioning and adaptive behavior. The heterogeneity of ID makes it more problematic for genetic and clinical diagnosis. NGS facilitates the diagnosis of ID families.

To accomplish the research objectives of the study, 10 autosomal recessive ID families were analysed. Total numbers of samples were 60, out of those samples 19 were affected individuals. All ten families' probands were tested for FMR1 mutation, after exclusion of fragile X syndrome, WES was performed to identify, pathogenic variants, and small insertion/deletions on the affected individual of these families. Analysis of WES results in probands from MR-6 and MR-9 identified potential candidate gene causing ID. In family MR-6, *SLITRK3* gene was sharing homozygous truncated mutation (p.E606X) on chromosome 3, in all three affected individual reported to be involved in synapse and neurotransmitter. In family MR-9 *SDCBP2* gene also known as syntenin 2 gene sharing homozygosity among all affected probands. However, in both these families functional experiments are in progress to confirm these findings. In family MR-4 WES sequencing analysis identified novel missense mutation (p.P202L) in *VPS53*, a gene previously reported for ID, as the likely cause of ID in this family as well. The mutation was observed in both affected sibling and obligate carriers but not in healthy member of the family or in the population controls (n=100). In Family MR-5, novel missense homozygous mutation (p.R545C) in *NARS* gene at chromosome 18 involved in ID was observed in the affected sibling but not in healthy sibling. Family MR-7 exome analysis was performed a novel missense variant (p.H440Y) in *GLB1* gene on chromosome 3 was identified and the mutation impact was checked through in-silco analysis and mutation was further conformed by using segregation analysis of the family. All the affected siblings were homozygous and healthy sibling have normal allele. Family MR-8 WES analysis was performed after the data filtration and bioinformatics analysis a novel autosomal recessive missense mutation (p.T320K) in *MLC1* gene was identified on chromosome 22. WES analysis of two families' (chapter 10) resulted mutation (c.T2C; p.M1T) in *SEPSECS* gene at chromosome 4 disturbing start codon. *SEPSECS* gene mutations have been associated with

delayed development and ID. In families AS105 and AS110, novel nonsense variant (p.Q863X) was identified in *PIDD1* gene on chromosome 11 by using whole exome sequencing.

Chapter 1: Introduction

Chapter 1: Introduction

Brain is the control center of our whole body. Our brain consists of three different parts and each parts has specific role. Cerebrum composed of left and right hemispheres and it is largest part of the brain. Cerebrum performs higher function like vision and hearing, speech and control of movement. Cerebrum also play role in thinking planning and understanding the language. The second part of the brain is cerebellum. Cerebellum has important role in balance, motor control but also involved in emotional functions. The third part of the brain is brainstem. Brainstem consists of midbrain, pons and medulla. These parts has role in breathing maintain body temperature and also perform wake and sleep functions. Our brain is connected with our body through spinal cord. Midbrain, hindbrain and spinal cord make central nervous system by (Miller *et al.*, 2008). Nervous system is made up with specialized cells known as neurons and supporting cells known as glia cells. Neuron receive signal from sensory organs and transmit these signal to the targeted organs. Signaling pathways of neurons connected through nervous system. It consists of central nervous system (CNS) and peripheral nervous system (PNS). CNS comprised with brain and spinal cord, brain and CNS controls the body and regulates diverse function including organs in the body (Fitzgerald and folran, 2012). Genes controlling neuronal and brain functioning, variations in these genes leads to certain brain anomalies.

Intellectual disability is heterogamous group of genetic disorders described as significant impairments in intellectual and cognitive functioning, adaptive behaviour which occurs before <18 years of age (Palmer *et al.*, 2014). ID is also defined as an overall intelligence quotient IQ score lower than 70, associated with low functioning in adaptive behavior, such as social communication and daily life skills (Varcarolis, 2016). ID has emotional impact and affects 1 to 3% people in the world with higher prevalence in less developed nation as compared to developed countries (Maulik *et al.*, 2011; Musante and Ropers, 2014). The prevalence ratio of neurodevelopmental disorders slightly higher in less developed countries due to limited resources causing malnutrition in children and poor health facilities of the countries are the factors of ID. Consanguineous marriages is a contributing factor of the higher prevalence ratio of ID in Pakistan, Iran, Sadie Arabia, Nigeria and money other countries (Mokhtari and Bagga. 2003; Ropers and hamel, 2005). ID diagnosis is difficult because of its heterogeneity. ID diagnosis involved clinical examination of the patients, including medical history, birth defects, medical

history of the probands and IQ assessment. The child's with profound ID are identified in the first two years of age because they fail to achieve normal motor development. The most commonly used IQ test distinguishes two main categories of ID, patient with mild ID IQ range between 50 and 70, and severe ID with an IQ below 50 (Chelly *et al.*, 2006). Clinician also performs neurological examination CT scan or MRI to detect structural malformation of the brain (Moeschler and Shevell, 2014). The genetics etiology of ID is highly heterogeneous, large chromosomal abnormalities, submicroscopic copy number variants, and most severe forms occurs when mutations arises in single genes (Perou *et al.*, 2013; Musante and Ropers, 2014). The causes of majority cases of ID unknown and known factors of ID includes environmental and social factors, malnutrition, chemical and radiation exposures, multiple pregnancies, postnatal infections and insufficient medical facilities (Riazuddin *et al.*, 2017; Harripaul *et al.*, 2018). The remaining 50% ID patients are associated with inbred background of which 10–12% are X-linked ID, the remaining are thought to be caused by the autosomal dominant and autosomal recessive ID, as well as chromosomal abnormalities (Ropers and hamel, 2005; Hu *et al.*, 2016). Intellectual disability divided in syndromic and non syndromic types. Syndromic-ID is linked with multiple characteristics and neurological abnormalities while non-syndromic forms have no additional features. Non-syndromic x-linked ID extensively studied in the past few years (Chiurazzi *et al.*, 2008). X-linked ID is more prevalent in male as compared to female (Ropers and Hamel, 2005).

So far 140 x-linked genes have been reported causing x-linked ID in male (Schirwani *et al.*, 2018; Neri *et al.*, 2018). Recently 405 families with x-linked ID investigated through exom sequencing. Exome sequencing of these families reveals 7 novel x-linked ID genes in 19 families (Hu *et al.*, 2016). Genes with autosomal recessive traits appear to be more common for causing non syndromic ID in consanguineous population. On the basis omim and SysID database (http://sysid.cmbi.umcn.nl/) 399 genes causing autosomal recessive ID. Before 2006 only three genes show linkage with autosomal recessive ID these genes are *CC2D1A*, *CRBN* and *PRSS12* (Basel-Vanagaite *et al.*, 2006). Afterwards a large scale comprehensive sequencing approach was used to report novel pathogenic genes and variants in extended families. These studies further elevated through NGS methods. In the years of 2011, homozygosity mapping and WES sequencing was carried out on multiple inbred families with autosomal recessive ID. In this study through next generation sequencing twenty three reported and fifty novel candidate genes reported with autosomal recessive ID (Najmabadi *et al.*, 2011). Multiple Saudi Arabian

consanguineous families with different neurological disorders was studied and exome sequencing of resulted thirty three new candidate genes involved in different neurological disorders, (Alazami *et al.*, 2015). WES analysis of large group of kindred Pakistani families with ID shows 30 novel candidates' genes and 30 affecting reported ID gene in 60 families, (Riazuddin *et al.*, 2017).

2. PROBLEM STATEMENT

Intellectual disability is a disorder of mental health. People who suffer from intellectual disability have much more difficult than others in learning new things, and understand the concepts, face problem solving difficulties, concentration and memory. Thus, they need additional support to learn and achieve their full potential. People with ill mental health have the same rights and responsibilities as other members of the society.

The research is focused on the families based study at genetic level to identify the responsible genes involved in intellectual disability. Pathogenic variants affects genes and these affected genes leads to diseases, new born child's inherits these diseases from parents, errors occurs in genes due to unknown reasons. Examples of genetic diseases are Autosomal recessive intellectual disability in inbreeding families.

3. OBJECTIVE OF THE RESEARCH WORK

In this research study, probands selected from Pakistani families with intellectual disability underwent clinical diagnosis of ID. The research project covered probands disease inheritance pattern, phenotypes of the patients and genetic testing for the correct diagnosis.

Objectives of the work:

3.1 To identify the gene(s) involved in the Intellectual Disability

3.2 To determine the putative variables segregating with disease

3.3 Mutational analysis through bioinformatics determine the mutation impact

Chapter 2: Review of Literature

Chapter 2: Review of Literature

2.1 Classification of Neurological Disorders

WHO gives classification of all known human diseases including the genetic disorders of nervous system and central nerves system (CNS). All member state of WHO agree to use the international classification of diseases. Person cognitive functioning assessed through IQ test. Normal IQ level is above 70, mild IQ level range between (IQ: 50-70) and sever IQ level range between (IQ: 20-40) (WHO, 1993). The etiology of ID is genetic, non-genetic and environmental. Poverty malnutrition and use of alcohol also contribute to the disease. ID caused by environmental factors such as chemical/bio hazardous agents, exposures to chemicals at the time of pregnancy for example direct contacts with radiations that may cause ID. Another reason, such low neutrinos diets, cultural deprivation towards female child's, and disease during childhood including measles, meningitis, head injury and environmental toxins also cause nerve damages and stop growth of the nervous system, leading to ID (Winnepenninckx *et al.*, 2003). Mutation in single genes prenatal birth of child and abnormalities in chromosome also leads to intellectual disability. The etiological factor of ID is very broad and shown in Figure 2.1. Multiples Genetic factors involved in pathogenicity of disease such as epigenetic changes, a number of factors including obesity, high fasting glucose and many other metabolite, repeat of genetic codes and disease of mitochondria (Mclaren and Bryson, 1987). A core group of 40 non syndromic autosomal recessive ID gene(s) identified (Musante and Ropers, 2014), through the integration of new genes may be derived from path of these essential 40 known genes. Through this identification it will possible to understand the mechanisms involved in intellectual functioning of brain (Lee *et al.*, 2012). Environmental factor such as pollutants also play role in ID and excessive use of alcohol also play.

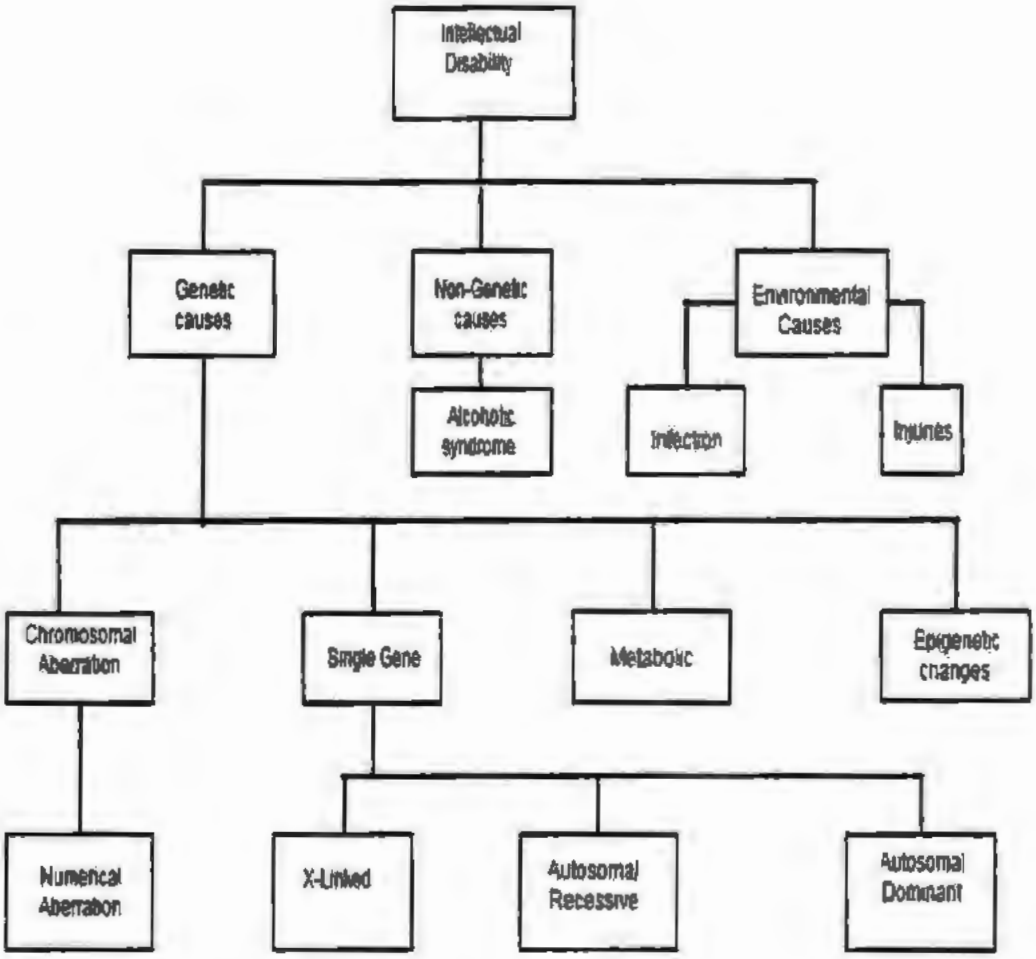


Figure 2.1 The diagrammatic representation and classification of ID. Multiple conditions causing ID includes genetic and environmental condition are the major causes of ID.

2.2 Diagnosis of Neurological Disorder

The diagnosis of Intellectual Disability is two steps process, in first step a detailed clinical evaluation is performed and the 2nd step is based on genetic testing e.g whole exome sequencing.

2.2.1 Clinical Finding of Intellectual Disability

Clinical findings of ID are very complex, multidisciplinary approach applied for the diagnoses. Diagnoses are based on family clinical records, birth history and complication during birth and family pedigree. Diagnosis also based on assessment of IQ level of patient. One third of diagnoses were conformed on the bases of clinical history in ID patient. Those Patients who are not diagnoses on the bases of clinical and physical examination but clinical and physical examination provide first step for further investigation to conform the ID in patients. Examinations of ID patients are based on MRI and CT scan tests detect gray matters and atrophic changes in certain part of the brain. Both the test result scans of the brain are taken to check the changes in brain and certain structural malformation also observed through MRI (van Karnebeek *et al.*, 2005). Automated pattern recognition system also used to identify neurological condition. Patient with neurological disorders also have gait disturbances (Stolze *et al.*, 2004). A neural network constructed to observe the changing patterns of certain part of brain when observed against wild type brain image (Brandt *et al.*, 2012).

2.2.2 Genetic Screening

Genetic testing of ID first step involves identification of defects at chromosomal level that cause several types of neurological anomalies. Linkage analysis, chromosome analysis, copy number variation (CNVs), duplications and translocation mapping are performed on the bases of clinical observation. Chromosome-banding method also knows as Cytogenetic technique performed for chromosomal screening, multiplex amplifiable probe hybridization and in situ labeling also used for chromosomal screening (Xu and Chen, 2003). Comparative genomic hybridization method applied to study the chromosomal copy number variation. It provides sufficient information about chromosomal changes including chromosomal losses and gains covering the entire genome (Weiss *et al.*, 1999). Telemetric rearrangement of chromosomes also associated with Intellectual

disability (ID). Telomere region deletion cannot be detected through conventional cytogenetic techniques. In situ labeling prints a new technique used to detect telemetric changes in ID patients (Bonifacio *et al.*, 2011). In recent years advanced techniques replace the conventional identification methods for the detection of genetic disorders. Microarrays used for the analysis of whole genome for copy no changes to find our deletion and duplication (Coughlin *et al.*, 2012). Exom sequencing used for large cohorts which outcomes help detection of new loci involved in neurological disorders (Riazuddin *et al.*, 2017).

2.3 Chromosomal Aberration

Genetic cause of ID was established for many years on the basis of identification of Down syndrome as a chromosomal abnormality. Intellectual research is to focus on to find out smaller chromosomal abnormalities that are linked with intellectual disability disease (Raymond and Tarpey, 2006). Chromosomal aberrations are responsible for ID cases. Changes at chromosomal level include numerical changes, partial changes and micro-deletions (Curry *et al.*, 1997). karyotyping analysis and FISH technique only find out up to 10% cases of intellectual disability (Stankiewicz *et al.*, 2007).

2.3.1 Numerical Aberration

A numerical chromosomal defect occurs due to monosomy or polyploidy in either cases deletion or addition of chromosomes from the normal set of 46. Chromosomal aberrations are detected by counting the chromosomes and arranging them in to set of normal 46 from blood cells under a microscope. Polyploidy are limited to Patus and Edwards syndromic condition known as trisomy 13 and trisomy 18. Down syndrome is major source of ID. Patau's syndrome and Edwards' syndrome have some common characteristic. Patau's and Edwards syndrome affected patient and children severely mentally retarded and died at some stage after birth. Chromosome 13. 18 and 21 contain less no of genes than any other chromosomes in genome, these three genes are gene poorest chromosomes in genome. Autosomal chromosome with monosomy is very fatal at the embryonic stage of life by (Winnepenninckx *et al.*, 2003). The main focus is to know the molecular mechanism and to make available better treatment for the patients and their families. Karyotyping analysis of the chromosomes gives clear picture of deletion or addition of 5 to 10

Mb genetic material from chromosomes. Gain or loss of short segment of DNA leads to developmental abnormality during embryonic stage (Barber, 2005).

2.3.2 Microdeletions

During the early 1990 microdeletions of the genome only detected with fluorescence in situ hybridization (FISH) and Microdeletions cannot be detected with visible light microscopy. Some common microdeletion syndrome also detected with in situ hybridization (FISH) and these syndromes are Williams and Prader Willi deletion syndrome. These syndromes are also associated with ID (Raymond and Tarpey, 2006). Cytogenetic analysis leads to identification of loss 7.50 Mb chromosomal locations $18q^2$, similarly deletion of copy no in the TCF4 gene causing ID in eight year girls with microcephaly, slow development of motor skills and stereotypical movements of hands. The deletion occurs during the developmental stages and clinical symptom disturbing many organs, founds in Pitt Hopkins syndrome (OMIM: 610954) (Stavropoulos *et al.*, 2010).

2.4 Abnormalities in Single Genes

Single gene disorders are classified on their heredity pattern. Different type of inheritance pattern causes mutation in single gene for example x-linked diseases. Genes that cause diseases identified on the bases of families study. Autosomal dominant pattern of disorder is more lethal and unlikely to appear in the population. New mutation arises from diseased genes and disease gene therefore dependent of the new mutation rate for autosomal dominant type of inheritance. Novel genes identified from the screening of individuals with ID. Some deletion and point mutation analysis in single gene that is ample to originate the disease phenotype for example Smith-Magenis syndrome (Slager *et al.*, 2003). Linkage analysis is use to identify causes of autosomal recessive intellectual disability in inbreed families. Significant number of autosomal recessive genes involved to cause Intellectual disability, the reporting of these new genes has been slow. Homozygosity mapping is method to report novel genes that causing autosomal recessive ID for the identification of the new genes families with autosomal recessive condition are required (Basel-Vanagaite *et al.*, 2006).

2.4.1 X-linked Intellectual Disability

X-linked genes accounts for 10-12% of the male with ID (Ropers and Hamel, 2005). X-linked ID is a common in male's population, clinically and genetically heterogeneous disorder. Genetic defect was reported in *NLGN4* (neuroligin 4, X linked; OMIM 300427) and *NLGN3* (neuroligin 3; OMIM 300366) in two brother that are linked with ID and autism, (Jamain *et al.*, 2003). After 2007 several novel genes linked with X-linked ID reported in literatures so for. *HUWE1* a novel micro duplication reported on chromosome Xp11.22 in six unrelated families with X-linked ID with (MIM 300697). This study shows that *HUWE1* point mutations are linked with X-linked ID (Froyen *et al.*, 2008). During the last decades more than 140 X-chromosome ID have be reported and identification of these genes suggests that further genes or loci are yet be identified (Hu *et al.*, 2016).

2.5 Autosomal Recessive ID

Autosomal recessive ID is common but causative genes, genotype–phenotype correlations are not fully understood. The heterogeneity of ARID and non-syndromic conditions marks it problematic to categories many families with similar genetic defect in order to achieve proper classification of ID (Marangi *et al.*, 2013). ARID arises in both syndromic and non syndromic condition. Syndromic type is linked with neurologic and skeletal disorder. Non-syndromic type causes learning disability, cognitive problems and intellectual dysfuctioning. Genes linked with non-syndromic types under studies to understand the normal variation in intelligence. When mutation arise in non-syndromic ID genes then these gene can create difference in learning ability without the existence of neural aberration. Distinction in IQ is linked with mutated genes causing variations in cognitive functioning. It is possible to compare mutated genes that causing ID with wild type genes to check the IQ of probands. More than 40 genes are associated with nonsyndromic ID (Musante and Ropers, 2014). On the basis of OMIM and SIS ID data bases (http://sysid.cmbi.umcn.nl/) so far 399 genes reported involved with autosomal recessive ID.

2.5.1 *PRSS12* Gene

PRSS12 OMIM no (MIM# 606709) is the first reported gene also known as Protease serine 12, which is involved in autosomal recessive ID identified in an Algerian family. *PRSS12* gene has

total (13) exon, and this gene is mapped on chromosome 4q26. *PRSS12* gene expresses serine protease neuro-trypsin, which function in neuronal growth. Mutation in this gene causes autosomal recessive non-Syndromic ID. In situ hybridization studies of neuro-trypsin was checked in human fetal brain and elevated level of gene expression was observed in period of 3 month of gestation. Few base-pair deletion in exon seven of *PRSS12* gene was observed in all affected proband with homozygous condition. All the patients in early two years of life were normal growth and development after two years neurological deterioration was observed in patients and their IQ was than below 50, and these children effected with ID (Molinari *et al.*, 2002).

2.5.2 *CRBN* gene

The (*CRBN*; MIM: 609262) gene was mapped in consanguineous German with non syndromic ID. Genes encodes ATP- dependant cereblon that is role in expression and assembly of Ca2+-activated K+ channels which control neurol excitability (Jo *et al.*, 2005).

2.5.3 *CC2D1A* gene

The (*CC2D1A;* MIM: 610055) gene was reported in consanguineous family that is responsible to cause NS-ARID. This gene contains 31, exon and mapped on chromosome 19p13 and express C2 domain protein, 1A protein. Mouse brain embryos encode *CC2D1A* protein. Frame shift mutation caused loss of 3589 base pairs starting from non-coding region intron 13 and stopping at intron 16 of *CC2D1A* gene leads to stop codon and truncated protein sequence was observed in all affected individual (G408fsX437) (Basel-Vanagaite *et al.*, 2006).

2.5.4 *MED23* Gene

The MED23 (MIM: 605042) gene is a mediator complex and regulate expression of protein coding gene. This gene contains 29 exon and located on chromosome 6q23.2. Missense mutation reported in this gene (p. R617Q) show co-segregation with non-syndromic autosomal recessive ID. A transcriptional defect of this gene was apparent when cell obtained from proband with neurological diseases (Hashimoto *et al.*, 2011).

2.5.5 *TRAPPC9* Gene

TRAPPC9 (MIM # 611966) gene is located on chromosome 8q24.3. Study of consanguineous Israeli Arab family, brain scan examination of probands displays malformations in axonal connectivity. Through homozygosity mapping a new genetic locus has been mapped for autosomal-recessive ID in this study loss of functional variation R475X in exon 7 at nucleotides position c.1422C>T was reported from three girls. This study show important roles of *TRAPPC9* in human brain development defects (Mochida *et al.*, 2009). Variant was reported in *TRAPPC9* gene in an extended consanguineous Pakistani family (Mir *et al.*, 2009). Novel pathogenic variants were identified in two Italian sisters causing gene loss of function (Marangi *et al.* 2013).

2.5.6 *TECR* Gene

TECR gene with OMIM no (MIM # 610057) is located on chromosome 19p13. A missense variant (p.P182L) was reported on exon 8 of *TECR* gene. Reported mutation change the amino acid leucine with proline at amino acid position 182 in *TECR* gene (Caliskan *et al.*, 2011).

2.5.7 *MAN1B1* Gene

MAN1B1 gene OMIM no (MIM # 604346) was mapped on chromosome 9q34.3 and reported from Iranian and Pakistani families. The *MAN1B1* gene encodes enzyme associated with development of glycan's in secretory pathways. In this gene missense mutation was reported at nucleotide position (c.G1189A) in 3 families from Pakistan (Rafiq *et al.*, 2011). Stop codon variant identified in two patients (c.G1418A) were borne in inbred Pakistani family (Rafiq et al., 2011).

2.5.8 *TUSC3* Gene

TUSC3 gene OMIM no (MIM# 601385) mapped on chromosomal position 8p22. This gene form protein that performs several function magnums uptake and function in embryonic development. Neuctotide variant in this genes are causing (ARID) autosomal recessive intellectual disability (Garshasbi *et al.*, 2008).

2.5.9 *ST3GAL3* Gene

ST3GAL3 (MIM: 606494) gene mapped on chromosomal position 1p34. *ST3GAL3* gene locus MRT4 previously identified to associate with autosomal recessive ID. It encodes protein beta-galactosidase-2, 3-sialyitransferase-II and this protein found in golgi membrane. This protein is responsible for attaining and maintaining higher cognitive functions. Two different variants reported in two families from Iran by using next generation sequencing. Iranian family with ID identified with a homozygous mutation in exon 14 c.1108G-T reported (Hu *et al.*, 2011).

2.5.10 *SEPSECS* Gene

SepSecS gene (MIM: 613811) located on chromosome no 4p15.2 play role in the production of SepSecS enzyme. SepSecS enzyme help in the formation of tRNA, it's required for the formation of protein selenocysteine (Sec). Selenocysteine protein plays important for the function of neurons and important for the normal brain development. Mutation in *SepSecS* gene causing profound mental retardation first time reported in two probands of Jewish Iraqi ancestry (Ben-Zeev et al., 2003). Mutation p.Y334C and compound heterozygous mutation p.A239T causing intellectual disability and pontocerebellar hypoplasia type reported in patients of Iraqi and Moroccan ancestry (Agamy et al., 2010). Sepsecs gene mutation p.N489V was reported in Jordan family causing Intellectual disability (Makrythanasis et al., 2014).

2.5.11 *GLB1* Gene

GM1 gangliosidosis (MIM# 230500) located on chromosome 3p22.3, is a lysosomal storage disorders of the *GLB1* gene (Suzuki, 1995; Caciotti *et al.*, 2005). The prevalence of GM gangliosidosis is 1 in 100,000 to 300,000 at birth (Regier and Tifft, 2013). GM gangliosidosis classified into four distinct types. Type 1 (MIM: 230500) GM1 gangliosidosis involved with cardiac involvement, early age of onset, hypotonia. Type 2 (MIM: 230600) disease phenotype characterized by slow progress of neurodegeration, seizures, muscle weakness. Type 3 (MIM: 230650) GM gangliosidosis less severe than type 1 and 2 (Callahan, 1999; Suzuki *et al.*, 2011). Morquio B (MIM: 253010) disease occurs due to the lack of β-galactosidase and it is an allelic from of GM1 gangliosidosis (Sohn *et al.*, 2012). Morquio B phenotypically characterized by cardiac involvement, manifest skeletal anomalies augmented urinary secretion of keratan sulfate,

no sign of CNS involvement is appeared (Suzuki *et al.*, 2011). GM1 gangliosidosis caused due to the storage of GM1 gangliosides and glycol-conjugates in the central nerves system (CNS). To date, total 185 variants were reported in *GLB1* gene (www.hgmd.cf.ac.uk).

2.5.12 *MLC1*

Megalencephalic leukoencephalopathy *(MLC1)* (MIM: 604004) located on chromosome 22q13.33, autosomal recessive disorder with disease occurrence ratio 1 out of 7,663. *MLC1* characterized by macrocephaly, slow start of neural indicators with ataxia, late decline of cognitive functioning and mild intellectual degeneration (Van der Knapp *et al.*, 1995; Leegwater *et al.*, 2001; Bonkowsky *et al.*, 2010). MLC1 probands MRI test shown sub-cortical cyst in the frontal temporal area of the brain (Van der Knapp *et al.*, 2012). *MLC1* occurs due to the defects in *MLC1* (Topcu *et al.*, 2000) and GlialCAM (Lopez-Hernandez *et al.*, 2011). It encodes amino acid and contains 8 Tran's membrane domains. *MLC1* gene regulates water and ions homeostasis to control the flow of fluids into cells to strength the cell adhesion in brain cells (Brignone *et al.*, 2015).

2.5.13 *VPS53*

VPS53 gene (MIM: 615851) located on chromosome 17p13.3, is a part of golgi associated retro grade protein (GARP) complex and function as tethering molecule between tans golgi network and endosome. Complex function to transport lysosomal from late ensomes to the TGN, retrograde sorting receptors degrade the retrograde vesicles and act as recycling machinery in TGN (Siniossoglou and Pelham, 2001; Conibear *et al.*, 2003; Pérez-Victoria *et al.*, 2008). Accumulations of lysosomal vesicles lead to the swelling with in the TGN due to dysfuctioning of tethering between lysosomal vesicles and GARP complex (Liewen *et al.*, 2005). The disease causing mechanism for *VPS53* mutation leading to the *PCH2* disorder is poorly understood. So for only two mutations of *VPS53* gene are reported in literature that causing *PCH2*.

2.5.14 *NARS*

NARS gene (OMIM: 108410) belongs to group of enzyme known as aminoacyl-tRNA synthetase. These enzymes involved in the formation of protein to provide energy to tRNSs with their associated amino acids. Aminoacyl-tRNA synthetases (ARSs) highly expressed to ensure the correct translation of genes into proteins (Rajendar *et al.*, 2018; Lee *et al.*, 2018). ARSs enzymes have very important function and it is highly expressed enzyme, but mutation in ARSs gene lined with various human diseases including this gene have been associated with combined oxidative phosphorylation deficiency (Meyer-Schuman, and Antonellis, 2017), like mutation in YARS can cause charcot marie tooth disease type c is a non-canonical function (Jordanova *et al.*, 2006). Variants in the tRNA synthetases gene have been linked with diseases including mitochondrial myopathy (*YARS2* [MIM 610957]), deafness (LARS2 [MIM 604544]), leukoencephalopathy (*MARS2* [MIM 609728]) (Oprescu *et al.*, 2017).

2.7 Autosomal Dominant ID

ADID is the most common due to genetic defects at the time of birth. Some CNV (copy-number variations) are also leads to intellectual disability. Genetic disorders such as tuberous sclerosis, neurofibromatosis and myotonic dystrophy were associated with autosomal dominant intellectual disability (Nelson, 2010), causes of non-syndromic autosomal dominant ID is not fully understood yet. Methyl binding domain 5 gene (*MBD5*; MIM: 611472) on chromosome 2q23.1 is associated with autosomal dominant ID. SNP microarray analysis of this gene shows a 200Kb deletion remove 6 exon from female, with sandal-toe and epilepsy (Wagenstaller *et al.*, 2007). AD intellectual disability genes on chromosome 6p21.3 contain (SynGAP1) that activate synapse (Hamdan *et al.*, 2009). *SYNGAP1* (MIM: 603384) encodes SynGAP, part of the NMDR complex was first checked because of heterozygous mutations in SynGAP were identified in mouse models and experimental data shows having trouble with talking and learning (Hamdan *et al.*, 2009). Variants in (SHANK2; MIM: 603290) have been reported in autosomal dominant ID. Patients with autosomal dominant ID also observed with autistic features. Variants in *SHANK2* on 11q13 reported through CNV analysis and gene sequencing method on large autism and autosomal dominant ID causing families (Berkel *et al.*, 2010).

2.8 Metabolic Causes of Intellectual Disability

The metabolic causes of ID are not fully understood. Normal brain functioning is also not fully understood yet and the defects of brain also not fully understood. In some disorders, the functioning and pathways are known for example phenylketonuria (PKU) but the brain disorders functioning are not known so for. The result for treated PKU, galactosemia, homocystinuria, and lysosomal disorders is not fully successful. Less familiar disorder with seizure including partial assembly of serine or creatine, harmful glucose carriage into the brain, defects causing malformations, includes Smith Lemli Opitz (SLO) syndrome and the congenital disorder of CDGs (glycosylation) play important role in delayed growth and developmental. Rare diseases with seizures or without seizures and disorders with deformities may primarily assume to be generic condition causing delayed development. Metabolic brain imaging magnetic resonance spectroscopy (MRS) also shows the map of biochemical abnormality within the brain (Kahler and Hahey, 2003). Many disorder originated due to abnormalities in the mitrochondrial genome for example Urea Cycle Disorders (Rohrbach and Clarke, 2007). Recently a study was performed on 213 patients with neurometabolic disorders and the patients were diagnosed percentage wise majority of the patients were male (54.3%). More than 71.4% of parent's patients with neurometabolic disorders were born in consanguineous marriages. More than 85% patients identified with developmental delay and 55% of patients suffering with different types of seizures. More than 200 patients with multiple neurometabolic disorders were diagnosed (Karimzadeh *et al.*, 2016).

Chapter 3: Materials and Methods

Chapter 3: Materials and Methods

3.1 Ethical Approval

The research work was carried out under the ethical standards and assent by the board of advanced studies and research (BASAR) International Islamic University Islamabad (IIUI), Pakistan by following the rules of Helsinki Principle.

3.2 Patient Cohorts

Families were enrolled on the basis of ID diagnosis. Written consent obtained from all the participants' including patient and parental guardian to collect detailed description of genetic relationships of the probands and healthy family members, Pedigrees of families was drawn on the basis of families' genetic history and no of probands present in families. Male and female individual was shown in square and circle correspondingly. Filled symbols of pedigree were representing affected individuals. 5 ml Blood samples were taken from all families members and stored in in BD vacationer EDTA (ethylene diamine tetra acetic acid) tubes.. Patient's clinical assessment was performed at local consultants.

3.3 DNA Isolation

Genomic DNA isolated from whole blood in the Human Molecular Genetic Lab department biological science at international Islamic University using a standard organic method (phenol chloroform method). In short, 5ml of human blood is added in 25ml of lysis buffer in a 50 ml falcon tube contains (NH_4cl 8.29 gram in 155mM concentration Khco3 1gram in 10mM concentration, EDTA 200ul in 0.5M concentration), incubated for 30 mints on ice, incubation step help cell lysis after cell lysis centrifugation was performed on 1200 rpm for 8 to 10 mints at 4°c. The pellet is left in the tube after discarding the supernatant from the tube. Pellet is suspended again with 5ml lysis buffer, incubated for 10 mints. In next step centrifugation were performed at 1200 rpm for 10 mints at 4 centigrade again and all supernatant is discorded leaving on the pellet. 4.75 ml of STE buffer contains (Sodium, Tris, EDTA) mixed in the tube containing the sample and vortexing is performed to mix the pellet, 250 ul of 10% of SDS was added in sample drop wise while slowly vortexing the sample. In the end 10ul proteinase K (PK) was added in tube and tube was placed on shaking water bath for overnight at 55°c. Next day sample

were taken out from shaker and equal volume of phenol (5ml) were added to the tubes and tube shaking is performed for 5 mints, after thoroughly mixing tube placed on ice container for 5 mints. Again centrifugation was performed at 3200 rpm at 4^0c for 25 mints, layer (upper layer inside tube) containing genomic DNA was transferred in new tube with cut tips without disturbing the layer. 5ml of chilled chloroform-isoamyl alcohol was added in samples, and shaking was performed for 10 mints and samples kept on ice 10 mints. In next step centrifugation of samples were performed at 3200 rpm for 10 mints at 4^0c. Upper portion of samples containing DNA was transferred in new tube. In the new tube equal volume of chilled absolute ethanol added with aqueous samples, 500ul of 5M ammonium acetate was mixed and DNA samples in the tube thoroughly inverted for many times. For DNA precipitations samples are placed at 15 mints at -70°c. After precipitations centrifugation is performed for 30 minutes at 3200 rpm. The upper layer of samples was removed from the tube. 3ml of 70% ethanol was mixed in the tube after vortexing the sample, centrifugation performed for 10 minutes at 3200 rpm. In final step supernatant was discarded and DNA left for air dried. The DNA was then dissolved in 300ul of 10mM tris-HCL (pH 8.0). Regent used in DNA extraction is mentioned in Table 3.1

Table 3.1: Solution for DNA Extraction

Solution A **Lysis buffer**	**Solution B** **STE buffer**	**Phenol**	**Chloroform** **Isoamyl Alcohol**	**TE** **Buffer**
Sucrose 27.36g	Nacl 4. 39 g	Buffer phenol with Tris Ph# 8	Chloroform (24ml)	1M Tris
Tris 0.30g	Tris_HCL ph 8			
$Mgcl_2$ 2.54 g Tritons (100x) 2.5ml	Na2EDTA 50ml 0.5 M		isoamyl alcohol(1ml)	0.5M EDTA
Ph # 8.00	Ph# 7.4			

3.4 DNA Quantification

NanoDrop 2000 Spectrophotometers measure concentrated DNA Samples with high clarity and reproducibility contains ND-1000 software. A NanoDrop absorbance measurement is measure DNA at a 260nm wavelength. The purity of DNA assess at the wavelength of 260/280 nm and

260/230nm. Absorbance wavelength of 260/280 considered equal to 1.8 and 260/230 equal to 2.0 represents clean DNA. The DNA concentration less than 1.7 specify the existence of protein contamination and ratios higher than 2 indicate the presence of phenol contamination. Maximum 2ul DNA is pipetted on to a measurement pedestal of the nanodrop. When the upper pedestal brought into contact with liquid sample that is on the lower pedestal, then flash lamp light cross through the liquid sample meanwhile a spectrometer analyze the light passing through the sample and the DNA concentration that is calculated and displayed in ng/µl.

3.5 Genotyping

Genetic defects identified with genotyping including defects causing intellectual disability in families for this linkage analysis were performed by using STR markers map the genes causing autosomal recessive ID and presence or absence of fragile X-linked syndrome involvement in the families was also checked using triplet repeat PCR test. Genotyping was done by using STR markers. Microsatellite (STR) markers were obtained from the genetic location database (LBD) department of psychiatry university hospital Zurich, Switzerland.

3.6 Polymerase Chain Reaction

Genomic DNA of 40ng concentration was used in PCR reaction, which was prepared from stock solution. PCR amplification of STR markers was performed using 10µL thermo scientific dream taq green master mix (2x) reaction volume of 20µl, containing 2ul of 40ng genomic DNA, 6µL DH_2O, 1µL of each primer and 10µL dream taq green master mix (2x) . PCR programming set for 35 cycles with following temperatures: 96°C for 4 mints for, 55°C for 40 sec, 72°c for 40sec, and in final step extension temperature run at 72°c for 7 mints. Another PCR was performed for selected variants after NGS data filtration using Roche Applied Sciences Faststart mastermix, total reaction mixture of 15µL contains 7.5µL Faststart mastermix, 5µL DH_2O and 0.75µL of each primer forward and reverse. PCR reaction mixture is mentioned in Table 3.2. The thermal cycling conditions are mention in the Table 3.3.

Table 3.2: Master mix for PCR.

S No	Solution	Quantity
1	DNA	2 µl
2	10x PCR buffer	1 µl
3	50mM MgCl2	0.3 µl
4	dNTP	0.8 µl
5	Taq	0.2 µl
6	Forward Primer	0.3 µl
7	Reverse Primer	0.3 µl
8	dH2O	5.1 µl

Table 3.3: PCR Conditions for linkage analysis.

Conditions	Temperatures	Time	Cycle
Initial denaturation	96^0C	5 mints	1x
Denaturation	94^0c	30 sec	35x
Annealing	55^0c	45 sec	
Extension	72^0c	30 sec	
Final Extension	72^0c	7 mints	1x

3.7 Polyacrylamide Electrophoresis

Amplified DNA sample were checked on 8% poly-acrylamide gel (Table 3.4). Poly acrylamide gel recipe was mixed in cylinder containing 50 ml of poly-acrylamide gel, 40 ul aps and 50 ul TEMD. The cylinder mixture solution was poured in electrophoresis glass plates. Comb for wells creation was placed between two glass plate and electrophoresis plate left for gel solidification at room temperature. The plate of PCR products was prepared and 6x loading dye, added in to the each wells of the plate containing DNA. The PCR product loaded into the wells and run starts for 180 minutes depending upon size of the STR markers at 80 watts. Once the electrophoresis processes completed then gel was stained with ethidium bromide (10 mg/mL) for visualization of

DNA bands on UV transilluminator and for photography by digital camera DC120 (Kodak, USA). 8% polyacrylamide gel composition was used in linkage analysis is mentioned in (table 3.4).

Table 3.4: 8% polyacrylamide gel composition.

Chemical	Composition
40% acryl amide Gel	62.5 ml
10x TBE	25 ml
TEMED	300 µl
25% APS	300 µl
dH_2O	162 ml

3.8 Primer Designing for Variants

For selected variants from families exome sequences gene specific primer generated with primer 3 software (http://primer3.ut.ee). Primer covering the mutation from forward and reverse side was selected see Table 3.5. Neuctotide sequences of target region for specific variants were obtained from ensemble genome browser data base.

Table 3.5: List of primer used for co-segregation analysis

Gene Name	Primer Forward	Primer Reverse
VPS53	TGTTTCTGGCTTTTCACCTGG	CTTGACAGCACTCATGTTTCAA
NARS	AGCTAACTCCTTTCATGGTGGA	TCTTTCCTCCACGCTTCTGG
PLCH2	GTGTCTCTCCCGTGTCCAG	ACTCTGATTTGTCCCCACCC
SLTRK3	TTCCTCTATCTTCCCGTGGC	AGCAGGCTGAGAATTAACACAG
GLB1	CGCTTTCATACATGTCTAGGGT	CTCATCCCCACCCTCACTG
CCT5	TTGTTTTGTGGTGTTTTCCTCA	CGAAGCTGTGAGATGCAATG
SEPSECS	TGGTTCCGTCGTGTGTCTG	TAGTCTCAAGCAGCGAAGAG
SDCBP2	GACACTGACTCTGGGGCC	CTGGGGTGAGACGGAGAG
MLC1	ACTCTGCTCACACCTCCTTC	CCCCACAGGCTTCTCACC
CCSAP	TATAGAGAGGCAGGGCTTGA	ACCTGTGTCCGTTTCTTTTGA
PGS1	TTTCCTGTCTGCACGTTTCC	GAAATCCCACTTGCTGCGAC

3.9 Agrose Gel Electrophoreses

Appropriate amount of agarose powder was taken, for 1.5% gel; 1.5g agrose weighted and transfer in 100 ml of TBE buffer in flask. The flask was covered with saran wrap, and was placed in microwave for 1-2 mints. The flask was visually checked to see agarose gel is properly melted. After melting the gel, gel was cool down for 5-10 mints and then 2ul of 10 mg/ml ethidium bromide was added in gel. After cooling gel was poured in gel caster; bubble was removed from the surface of the gel by using micropipette tip. The comb was removed from the gel after the solidification and 3ul of pcr product, 3 ul of 6X loading dye was loaded into eacg wells of gel. DNA ladder of 100bp was also loaded alongside the DNA samples.

3.10 PCR Product Purification

PCR product was cleaned with two different methods. In Pakistan pcr cleanup was performed with precipitation method, while in UCL institute of neurology pcr cleanup was performed with enzymatic method.

3.10.1 Precipitation Method

Pcr product was purified with ammonium acetate and absolute ethanol. For 15 ul of pcr product purification 3ul of 5mM ammonium acetate and 30ul of absolute was used. Tubes were placed on chill container for 20 minutes. After precipitation tubes were placed in centrifuge machine for centrifugation at 12000rpm for 8 minutes, after centrifugation waste was removed from the tube and DNA washing done with 70% ethanol (100 µl). Again centrifugation was performed at 13000 rpm for 10 mints, supernatants discarded. The genomic DNA was left for drying purposes for 1 hour. The dried samples were resuspened in 15 µl of deionized water. The purified product was used for sequencing PCR.

3.10.2 Enzymatic PCR Cleanup

Pcr product was also cleaned with enzymatic solution. The enzymatic mixture is made up with exo nuclease and alkaline phosphatase (Thermo scientific). The mixture removes unused dNTPs, single stranded DNA from pcr products. The pcr cleanup recipe is made in volume of 2ml in tube containing 100µl exonuclease I, 400µl Fast-Alkaline phosphatase (AP) and 1500µl of ultra-pure

water. The tube is placed at -20°C for the use of longer period of time. In order to clean the pcr product 2.8ul of the enzymatic solution was added in 7ul of pcr product in pcr plate. Pcr cleanup is performed on the PCR machine at 37°C for 30 mints followed by 80°C for 15 mints. Sequencing reaction was performed for specific targeted genes variants by using either forward or reverses primers. The sequencing reaction pcr product contains following chemicals in the ratio of big dye terminator v3.1 (0.5 µl) 5x sequencing buffer (Applied Biosystems) (2µl), 1µl of p. forward or p. reverse at 10 pmol/µl and clean 2µl DNA and 4.5µl of PCR grade water. The pcr plate centrifuged briefly before running on the thermal cycler. The Applied Biosystems optimized pcr program used to run the sequencing plate.

3.11 Purification Methods for Sequencing Reaction

Pcr product was purified with precipitation method in human molecular genetic lab international Islamic university Islamabad and Sephadex method was used to clean the sequencing reaction, in UCL institute of neurology queen square London.

3.11.1 Precipitation Method

For each 10 µl sequencing reaction product purification 1.5 µl of 250 mM EDTA and 60µl 99.9% ethanol was used. The sequencing reaction placed on ice for 20 mints. Centrifugation performed at 12000 rpm for 12 mints and supernatants were discarded from the tubes. 30µl of 70% ethanol was added in each tube for washing purposes. The samples were resuspened by using vortex. Centrifugation was performed again at 12000 rpm for 5 mints, supernatant were castoff, DNA tubes left for drying at room temperature. After drying, the DNA samples was mixed with in 10 µl of Hidi formamide.

3.11.2 Sephadex Purification Method

Alternately sequencing reaction was purified with Sephadex method. Sequencing plate prepared by consuming 2.9g Sephadex and 45ml of ultra-pure water in 50 ml falcon tube. The solution was thoroughly mixed and left at room temperature for at least 25-30 for complete hydration. Once the solution was hydrated properly, then 350 ul of solution was poured to 96 wells fiber filter plate size of 0.66mm. The plate was fixed on the empty collection for the purpose of centrifugation and plate was centrifuged for 3 mints at 700xg. The 96 well filter plates again was

place on new filter plate. The sequencing reactions were transfer onto Sephadex columns by using multi channels pippet. The plate ere centrifuged for 5 minutes at 910xg. Sequencing reaction filtered out from Sephadex columns and was collected in new sequencing plate. Before putting the plate in sequencer each well of the plate was filled with liquid. Sequencing plate was sealed and ready to go to the sequencer.

3.12 Sanger Sequencing For Co-Segregation Analysis

Sequencing plate map was designed on excel sheet and copied into the Sanger sequencing machine. Sequencing plate was placed on a 3730 genetic DNA analyzer manufactured by applied Biosystems, Foster City, CA, USA. Result was analyzed on Sequencher develop by gene codes corporation, MI, USA.

3.13 Next Generation Sequencing (NGS)

Next Generation sequencing has assisting research work and providing high through put data. NGS can read millions of sequence clusters base by base through addition of nucleotides, fluorescence reading and dye cleaving See figure 2.1.

3.14 WES Sequencing

DNA samples are sent to the Macrogen South Korea for exome sequencing at commercial level. The sequencing samples were made by following the preparation guide line of the Agilent SureSelect Target Enrichment Kit. The constructed libraries were sequenced with Illumina HiSeq4000 platform sequencer. The Illumina SureSelect target enrichment kit has 10X depth coverage of 93% of a DNA sample. All samples exome sequences were aligned to the mapping referenced genomeGRCH37.

3.14.1 WES Library Preparation

Genomic DNA samples were prepared by following the guide lines of SureSelect target enrichment system capture method. SureSelect biotinylated cRNA baits hybridized with genomic DNA. Region of interest selected through using magnetic streptavidin beads.

3.14.2 Clustering and Sequencing

Illumina sequencing platform uses a unique bridge amplification method to create clonal cluster. clonal cluster process occurs front side of the sequencing plate (flow cell).Sequencing plate cell containing of million unique clusters, adapter ligated fragments are also attached with flow cell surface. Adapter ligated fragment treated with regents for polymerase base extension. Polymerase based fragments extension start on the presence of all four nucleotides to achieve higher accuracy in clustering and sequencing.

3.15 Bioinformatics Pipeline for WES

The exome sequencing was performed on the selected families' probands only. The data generated from exome sequencing was filtered through with different bioinformatics software's for further analysis. Bioinformatics tools was used in data filtration was shown in flowing figure below.

TH23586

3.16.1 BCL File Format Data

The output file format of NGS data is BCL file. BCL file is binary file that contains base calls per cycle and the quality of the base call for every cycle sequenced by the sequencer.

3.16.2 Fastq Files

NGS raw sequencing data is available in the Fastq file format for bioinformatics analysis. Exome sequencing raw data were analysed by using FASTQC web based platform available online (http://bioinformatics.babraham.ac.uk/projects/fastqu/). Fastq data consist of four different lines for single read. First line of single reads start with "@" characters. The second line consist sequencing data, third lines contains only "+" symbol. Fourth lines contain quality score for each sequence read.

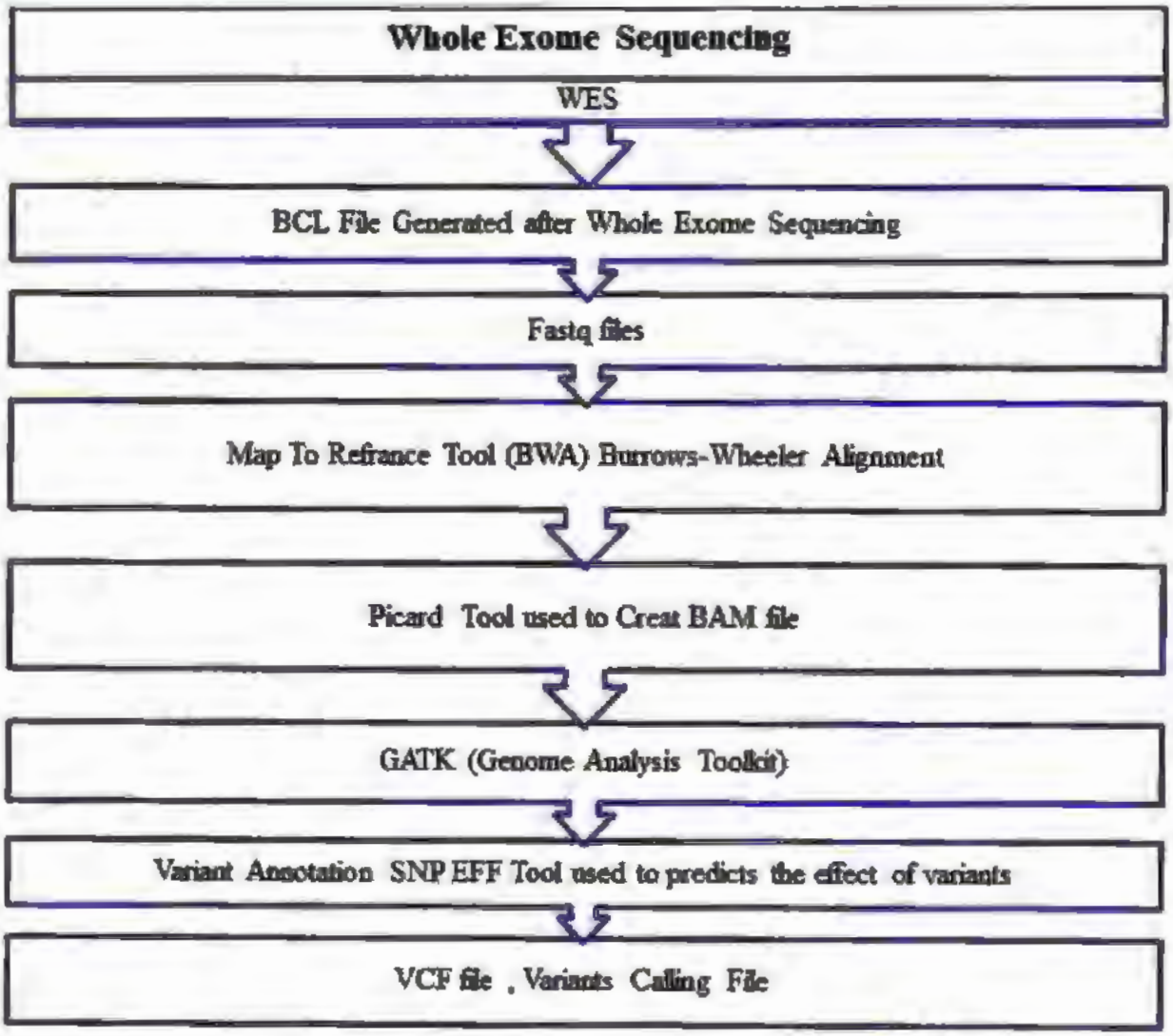


Figure 3.1: Whole exome sequencing method and data interpretation from Fastq file to VCF file.

3.17 Homozygosity Mapping

Online homozygosity mapper used on consanguineous families to identify autosomal recessive alleles shared by the affected probands. For Homozygosity mapping a web based Homozygosity mapper was used that allows to directly uploading the VCF file generated after bioinformatics pipeline of WES data. Homozygosity mapper detects homozygous region shared by probands of consanguineous families and display the region graphically. Shared region candidate genes were checked with Gene Distiller to find out the most causative gene that also must be present in exome file as well.

3.18 Tools for Predicting the Functional Effect of Genetic Variants

There are numerous tools for predicting pathogenicity of genetic variants. For these studies different tools have been used to calculate pathogenic score of filtered variants. Exome file contains know variants, synonymous variants, artifacts variants generated from sequencing errors and common polymorphism in general population. The exome excel file filtered and we have selected only rare missense, non-senses, and splice site variants. The frequency of <0.001% was applied for rare variants search at 1000gp see (figure 2.2). The GERP and CADD score was also accounted for pathogenicity of variants. The exac browser, mutation tester and other relevant data bases also applied on all the selected probands for rare variants. In this thesis filtering on a minor allele frequency of ID segregating according to autosomal recessive pattern was applied, remaining variants were evaluated (regardless of priority). After removing all non-informative variants around about 50 to 30 variants left with greater impact that might be have some kind of pathogenic effects. These variants further searched in data base like OMIM data bases, gnomad data base, Mutation Tester for further confirmation. The filtration criteria of variants was shown in the following figure no.

3.19 3D-Modeling of Mutations

Protein three dimensional modeling software and web based on line software provides plate forms to perform the Insilco modeling of mutated protein generated after WES sequencing and conformed to co-segregation analysis. In-silico analysis of mutated protein gives us the insight and impact of mutation on protein. By using Swiss modeling wild type and mutated protein structured was generated to check the impact of mutation in this study.

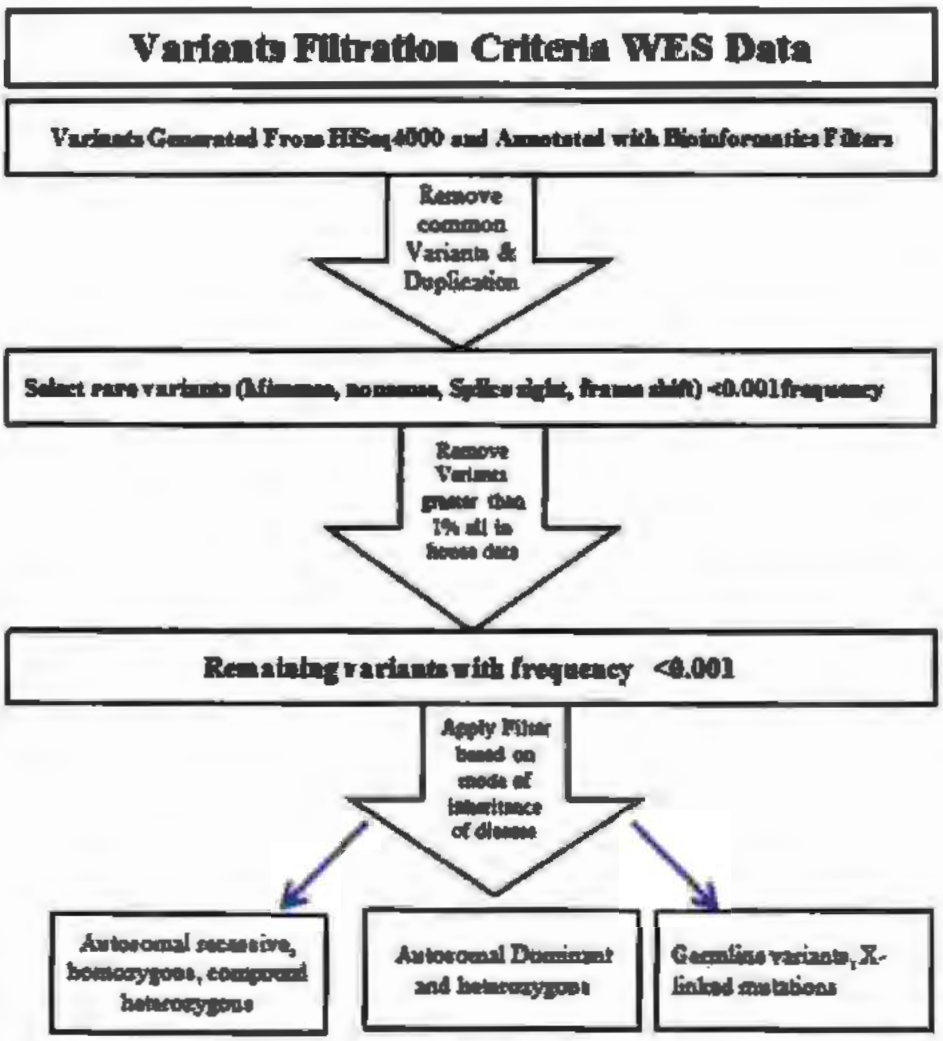


Figure 3.2: Whole exam data filtration method used to remove the artifacts, synonymous variants, intronic variants. For rare variants allelic frequency was set to <0.001 remove the common variants from the file.

Chapter 4: Results

Chapter 4: Results

Chapter 4.1: Novel *VPS53* Variant Causing PCCA2 in Consanguineous Pakistani Family

4.1.1 Introduction

PCH2 is a disorder that affects the development of the brain, development delay, lack of motor skills, microcephaly and intellectual disability manifested by the progressive cerebello cerebral atrophy (PCCA). There are two main type of ponotocerebellar hypoplasia (PCH). *PCH1* is caused by variant in *SepSecS* gene. *SepSecS* gene (OMIM # 613009) described in Moroccan families with progressive microcephaly, profound ID and generalized tonic clonic seizures (Ben-Zeev *et al.*, 2003; Agamy *et al.*, 2010). *PCH2* (MIM: 277470) is caused by variant in VPS53 gene (Feinstein *et al.*, 2014). The clinical symptoms of the *PCH2* are includes, microcephaly, developmental delay, intellectual disability and spasticity, some patients also suffer with seizures and lack of motor skill (Hady-Cohen *et al.*, 2018). *VPS53* gene (MIM: 615851) associated with golgi associated retrograde protein complex that act as tethering complex between endosomes and trans golgi network. It is required for the maintenance of the cycling of mannose-6-phosphate (M6P) receptors between the trans golgi network (TGN) and endosomes (Conibear and Stevens, 2000). GARP involved in the lysosomal transport of late ensomes to the TGN, lysosomal vesicles transported from late ensomes to TGN via GARP complex (Figure 1), retrograde sorting receptors localized in TGN degrade the retrograde vesicles and act as recycling machinery (Siniossoglou and Pelham, 2001; Conibear *et al.*, 2003; Pérez-Victoria *et al.*, 2008). Accumulations of lysosomal vesicles lead to the swelling with in the TGN due to dysfuctioning of tethering between lysosomal vesicles and GARP complex (Liewen *et al.*, 2005). GARP complex interact with Rab and SNARE proteins to permote retrograde transport to the TGN (Liewen *et al.*, 2005). One of the subunit of GARP complex *VPS51* is also act as facilitator between tethering and fusion machinery (Conibear *et al.*, 2003). The SNARE protein make interaction with 5 primer end of the vps53, vps54 subunits, its interaction with retrograde vesicles also requires 3 prime end of the small subunit vps53. Mutation in the 3 prime end of the VPS53 halts retrograde vesicles transportation (Vasan *et al.*, 2010). In fact GARP complex retrograde

vesicles transportation functions for TGN46 and Shiga toxin recycling proteins (Conibear and Stevens, 2000). The disease causing mechanism for vps53 mutation leading to the *PCH2* disorder is poorly understood so for. So for only two mutations of *VPS53* gene are reported in literature that causing *PCH2*. Further families need to be screen for detailed analysis of disease, in the population of Jews of Moroccan ancestry and families of Pakistani origin.

4.1.2 Summary of Family

The selected family (MR-4) was enrolled from the children hospital Multan, distt Multan Punjab Pakistan. Family pedigree was constructed based on the information provided by the parents and nearby blood relatives. Blood samples were collected from both affected and healthier family members. Written consent was obtained from all the members of the family and affected child informed consent was taken from parents and guardians. The study was approved from the Board of Advanced Study and Research (BASR) and ethical committee of International Islamic university Islamabad, Pakistan. Pedigree of the family of was shown in below Figure 4.1.1.

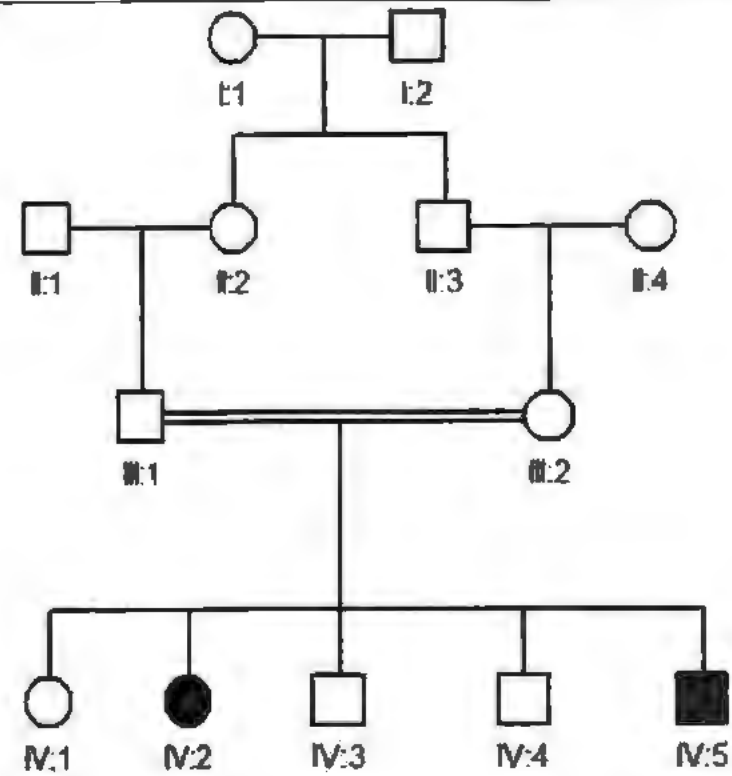


Figure 4.1.1: Family pedigree representing the two probands (IV: 2 and IV: 5). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals.

Figure 4.1.2: Proband (IV: 2) age was 12 years and proband (IV: 5) age was 9 years.

4.1.3 Clinical Details

The family (MR-4) with two probands (IV: 2 and IV: 5 see Figure 4.1.2) were the second and fifth children of healthy, consanguineous parents of Pakistani origin. Both start walking independently at the age of 14 to 15 months respectively. The proband (IV: 2) age was 12 years old at the time of medical examination. At 5 years of age, she developed maximum one seizure per day up to five minutes of duration. Type of seizure considered as generalized tonic-clonic seizure (GTC) and atonic seizure. Detailed clinical investigation suggests she has avoidance and fearful behavior, consciousness and alert responsive, able to talk. She has normal posture, normal fundoscopy, no speech problem, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Brain CT Scan demonstrated subtle hypodensity in gray matter more marked in left basal ganglia (Data not shown here). She is diagnosed with familial intellectual disability by the local neurologist. The time of clinical examination proband (IV: 5) was 9 year old and at 7 years of his age he developed multiple seizures per day. Type of seizure considered as focal motor. Detailed clinical investigation suggests he has aggressive behavior, no facial dysmorphic feature, head size is normal. He has normal posture, normal fundoscopy, no cerebellar signs, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Brain CT Scan demonstrated cerebellar atrophic changes, giant cisterna magna. Electroencephalogram test results demonstrated hypsarrhythmia pattern. He initially diagnosed with familial intellectual disability by the local neurologist.

4.4.4 Genetic Analysis

4.1.4.1 Linkage Analysis

DNA from all affected individuals as well as unaffected individual was analysed. Homozygous genetic variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.1.4.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 2 and IV: 5) come up with 137626 and 132172 variants, respectively with read depth around 60x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac data base (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 3.3. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All five variants putative score were checked in mutation taster (http://www.mutationtaster.org/), GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only five variants met these filtration criteria in both probands (IV: 2 and IV: 5). Out of these variants, three variants were non-synonymous and two were frame shift variants mentioned in table 4.1.1. Of these filtered variants only *VPS53* gene involved with profound intellectual disability, progressive cerebello-cerebral atrophy considered as neurodegenerative disorder. *VPS53* variant was not present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of *VPS53* gene variant considered as highly conserved. Mutation taster score for this variant is also considered as deleterious. *VPS53* variant selected for segregation analysis in family.

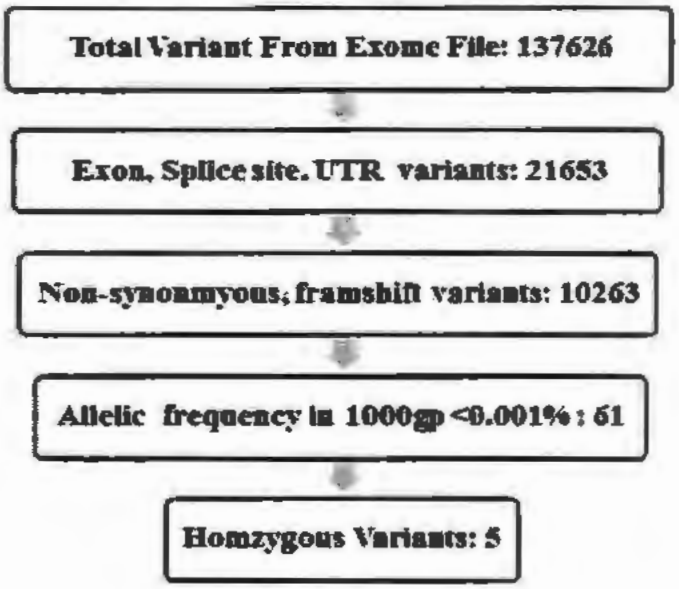


Figure 4.1.3: Method of filtering whole exome sequencing data of the patients (IV: 2 and IV: 5). Both the probands were sharing five homozygous variants only.

Table 4.1.1: Summery of variants found in proband (IV: 2 and IV: 5)

Gene	Ref. Seq. ID	cDNA Mutation	Protein consequences	dbSNP ID	GERP Score
NOP9	NM_001286367	c.483_484insGAGGAG	p.E169_D170insEE	rs113258190	NA
TRAK1	NM_001265609	c.1841_1842insGGAGGA	p.E624_G625insEE	rs10634555	NA
VPS53	NM_018289	c.C605T	p.P202L	NA	3.9
IP6K3	NM_054111	c.A14G	p.N5S	rs574159487	2.95
MLN	NM_001040109	c.T281C	p.L94P	NA	4

4.1.4.3 Co-segregation Analysis

Co-segregation analysis conform both affected siblings (IV: 2 and IV: 5) were homozygous, normal brother (IV: 3), heterozygous brother (IV: 4) and their unaffected parents (III: 1 and III: 2) were heterozygous for the *VPS53* variants, for sequencing result see figure (4.1.4). This c.C605T:p.P202L is novel variant not previously reported in databases, located in exon 8 of the *VPS53* gene, Ensemble transcript (ENST00000291074) is disease causing for this family.

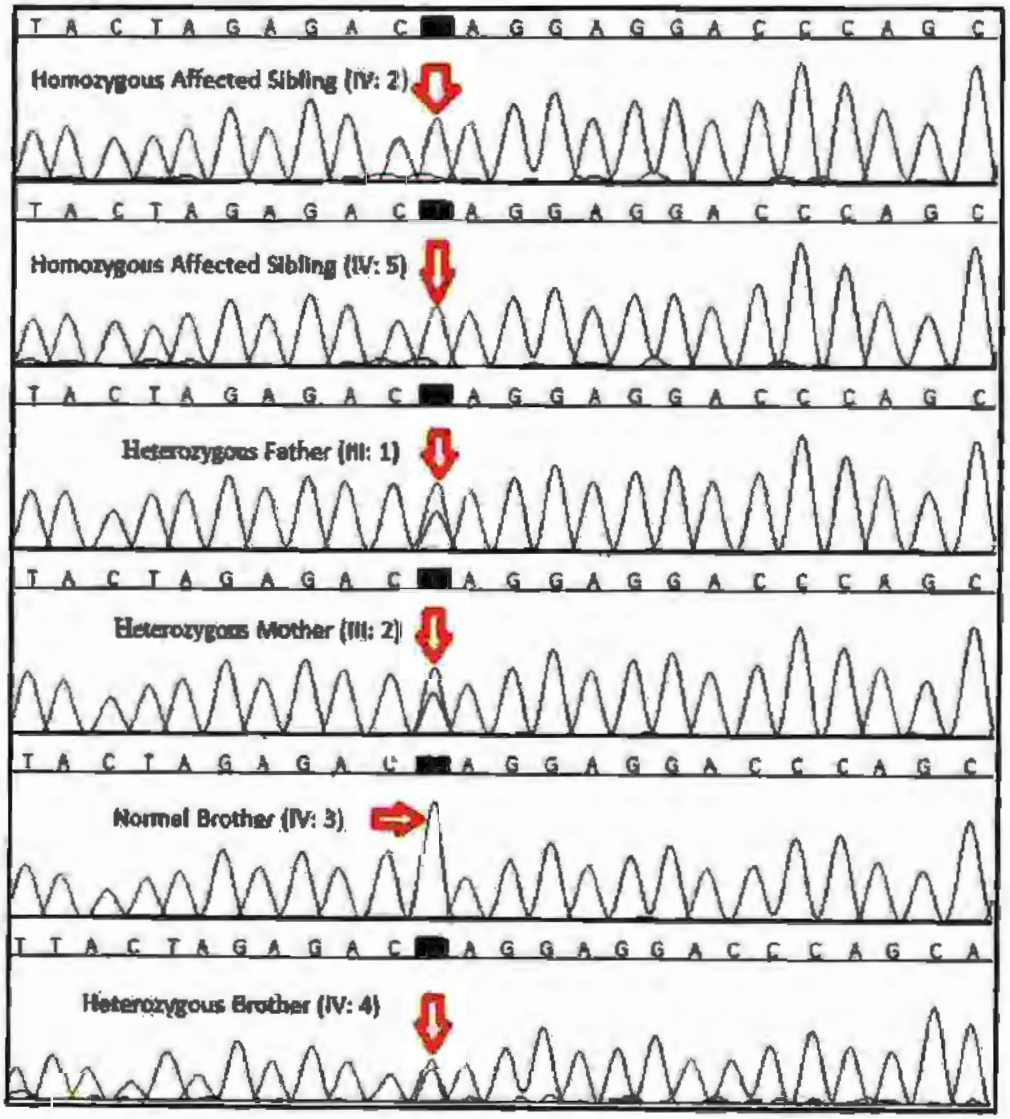


Figure 4.1.4: Chromatogram demonstrating segregation of mutation c.C605T in exon 8 of VPS53 in family A, Where both affected sibling were sharing homozygous region while parents were heterozygous. One brother (IV: 3) is normal and (IV: 4) is heterozygous.

4.1.5 In-silico Analysis

Insilco analysis of the missense mutation c.C605T; p.P202L by using Swiss modeling software (www.swiss-prot.org) discloses that the mutant amino acid effects protein structure of *VPS53* gene. As the mutation is nonsynonymous cause change in 3D structure of *VPS53* protein. The normal structure (A) and mutant structure (B) both visible in Figure 4.1.5.

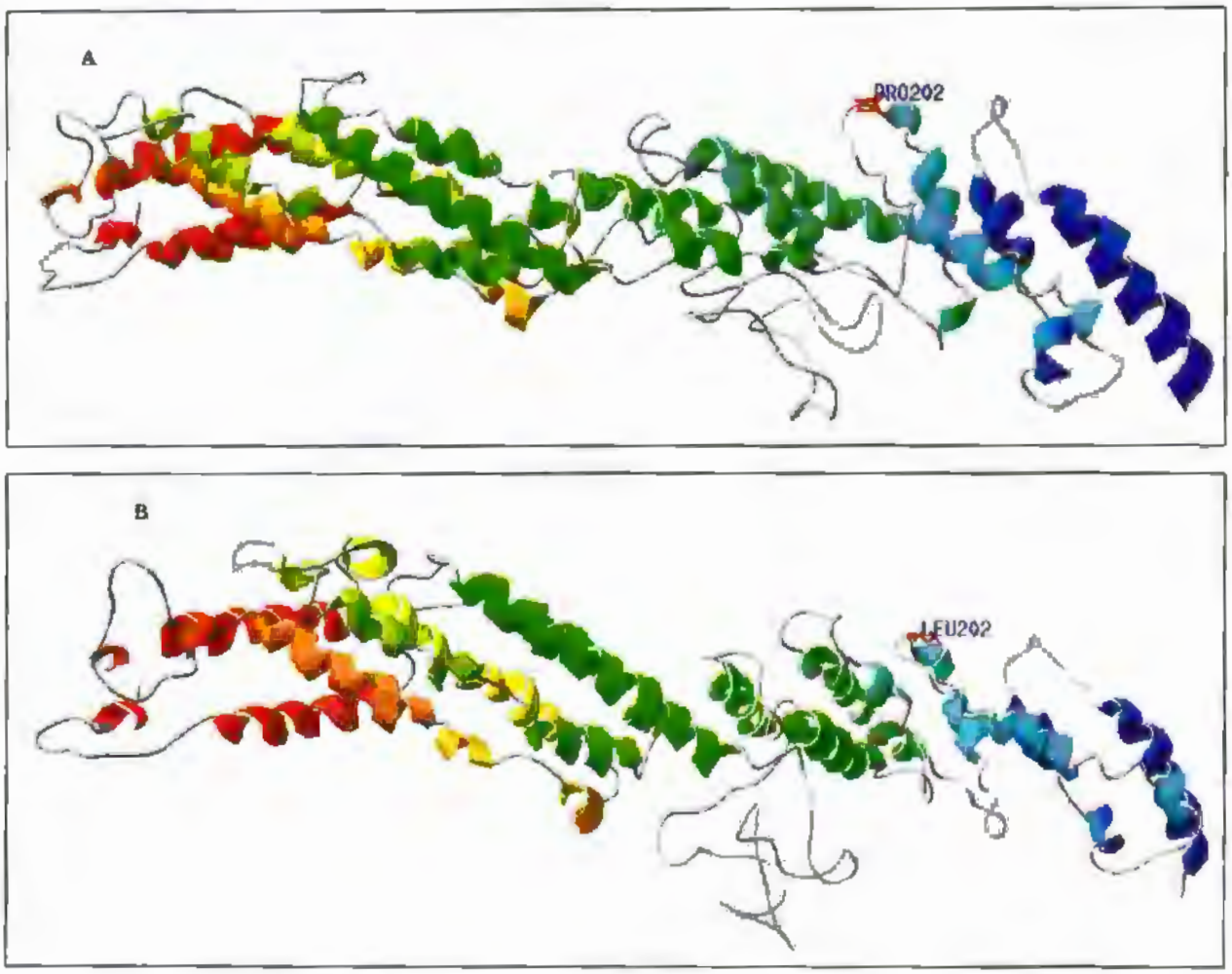


Figure 4.1.5: Swiss PDB viewer ribbon tertiary structure of *VPS53*. The diagram (A) is representing the wild type structure of *VPS53* protein. The diagram (B) is representing the mutated *VPS53* protein.

4.1.6 Discussion

Filtered rare variants (Table 4.1.1) selected for co-segregation analysis. Only *VPS53* variant was co-segregated with in family (MR-4) and a novel missense mutation at nucleotide position c.C605T in exon 8, substitutes the amino acid proline (CCA) with the leucine CTA (p.P202L). *VPS53* gene is located on chromosome 17 and consisting of 18 exons. Mutation was conformed to control sample of the population (n=100) and pathogenic score of mutation was checked by using bioinformatics tool like mutation taster and polyphen. The mutation was not present in public databases like exac, gnomad and the Z score of this mutation is 1.77 which is considered to be pathogenic mutation (https://gnomad.broadinstitute.org/). The genomic evolutionary rate profile (GERP) score for this mutation is 3.83 that is considered as conserved gene, and the gene does not tolerate mutation, GERP score above 2 to be considered as conserved for specific gene (Cooper *et al.*,2005). The Golgi associated retrograde protein complex (GARP) consists of four subunits, *VPS53*, *VPS52*, *VPS54* and ang2. The GARP complex plays an important role to direct retrograde vesicles from endosomes to trans Golgi network (TGN) (Conibear and Stevens, 2000; Liewen *et al.*, 2005). Retrograde protein function as degradation of lysosomal sorting receptor from endosomes to TGN. Dysfuctioning of tethering between retrograde vesicles and GARP complex results accumulation of lysosomal receptor molecules with in the TGN leads to swelling of lysosomes (Conibear and Stevens, 2000; Liewen *et al.*, 2005). Rab and SNARE fusion protein is localized in the TGN, GARP complex interact with these proteins, *VPS51* of GARP complex act as mediator between tethering and fusion machinery (Conibear *et al.*, 2003). The SNARE protein requires region specific interaction with GARP complex subunit. SNARE protein interacts with N-terminal regions in the *VPS53*, *VPS54* subunits and its interaction with retrograde vesicles also requires c-terminal region of the *VPS53* subunit of GARP complex. Mutation in the c-terminal region of the *VPS53* halts retrograde vesicles transportation (Vasan *et al.*, 2010). Previously reported mutation c.A2084; p. (gln695arg) substitutes the arginine with glycine at the conserved c-terminal region of *VPS53* gene. Splice site mutation c.1556+5 G>A result truncated protein with dysfunctioned *VPS53* and GARP complex. GARP complex act as tethering factor for retrograde, non-functional GARP complex leads to deposition and swelling of CD63 vesicles. CD63 is located in lysosomal membrane. Reported work suggests that mutation in *VPS53* subunit may leads to lysosomal storage disease or either non-lysosomal disorder like

PCCA disease (Feinstein *et al.*, 2014). Therefore given the role that in our family patient's phenotypes does not suggest lysosomal storage disorder of brain such as leukodystrophy. Patients with genetic history of intellectual disability may be affected with PCCA disease as previously reported that *VPS53* mutation causing disease in Jews of Moroccan ancestry peoples (Feinstein *et al.*, 2014). The mechanism of mutations yet needs to be study to fully understand the role of *VPS53* mutation role in PCCA disease.

Chapter 4.2: Novel Mutation in *NARS* gene Cause Neurodevelopment Delay, Seizures and Neuropathy Phenotype

4.2.1 Introduction

Aminoacyl-tRNA synthetases (ARSs) group of enzymes that are essential in the first step of gene translation into protein. The interfusion of tRNA to related amino acid required for the synthesis of gene into protein. This specific transformation is accomplished with multiple enzymes known as aminoacyl tRNA synthetases (ARSs), which are vastly expressed to ensure the correct translation of genes into proteins (Rajendar *et al.*, 2008; Lee *et al.*, 2018). The amino acylation reaction of protein was involved in two step reaction. In first step Amino acid, ARSs, and ATP reactant converts into aminoacyl adenylate and pyrophosphate products. In the second step adenylate conjugated amino acid residue placed on to 3 end of tRNA. The linkage formation between the amino acid and 3-OH (Hydroxyl group) in the ribosome ring establish the aminoacyl tRNA which act as a substrate in the protein translation during the reaction (Ibba and Soll, 2000). Despite the pre and post transfer proof reading ability of the ARSs enzymes, random mischarging of enzymes creates translational errors of protein during reaction. The coupling of tRNA to cognate amino acid is performed by group of ARSs enzymes. Each specific amino acid has designated (ARS) enzyme to catalyze a bond with a cognate tRNA. ARSs family consist of 37 genes, 17 genes functionally expressed in the cytoplasm, 17 genes expressed in the mitochondria and the remaining 3 genes encodes bi-functionally proteins that change tRNA (Antonellis and Green, 2008; Meyer-Schuman *et al.*, 2017). ARSs enzymes are divided into two classes. The class 1 enzymes are further divided into class a, b, and c and their classification based on the sequence conservation. The class 1 consists of monomers or dimers with catalytic domain that is built of parallel B-sheet. The class 2 enzymes consist of dimers or tetramers with conserved motifs and catalytic domain composed of anti-parallel b-sheet. The function of the class 1 and 2 enzymes is to attach the amino acid to the hydroxyl group for transesterification (Guo *et al.*, 2010). The classification of (ARSs) genes begins with one letter code of the amino acid that is designated for enzymes followed by (ARS), mitochondrial and cytoplasmic enzymes denoted with AARS and *AARS2* for mitochondria aminoacyl tRNA synthetase. The function of ARSs enzymes is vital and highly expressed but, mutations in ARS gene reported with multiple human diseases in both recessive and dominant manners. Molecular mechanisms of ARSs recessive

disease clearly express loss of function at genetic and functional level. Mutations in patient's leads to loss of ARSs gene function resulted impaired proteins. Molecular mechanisms of autosomal dominant disease causing mutations in ARSs gene are not clear. Autosomal dominant mutation in five genes (*YARS, GARS, AARS, HARS*, and *WARS*) encoding amino acyl tRNA synthetase cause dominant phenotypes (Meyer-Schuman, and Antonellis, 2017), like mutation in *YARS* causing Charcot-Marie-Tooth disease type C is a non-canonical function (Jordanova *et al.*, 2006). It is still important to find out if loss of non-conical function is the cause of ARSs related neuropathy in patients. Most of the mutation causing neuropathy shows functional impairment rather than gain of enzymatic function (Oprescu *et al.*, 2017). So it is supposed that gain of canonical function not causing ARSs related neuropathy (Meyer-Schuman and Antonellis, 2017). Impaired ARSs functions in dominant neuropathy studied with yeast complementation assays implicated those ARSs mutations causing neuropathy deficits with tRNA charging (Oprescu et al., 2017). ARSs missense mutations associated with neuropathy show damaging outcome the gene function, implicated that the deleterious effect is a part the disease pathogenesis in probands (Meyer-Schuman and Antonellis, 2017). Interestingly, disease phenotypes includes late onset neuropathy to more severe neuropathy and some progressive syndromes (Antonellis and Green, 2008; Meyer-Schuman *et al.*, 2017). NGS technology makes rapid the diagnosis of genetic disorders; however, many disorders with genetic background still required to be solved. The gene encodes ARSs have gain attention over the last few years due to the identification of mutations in these enzymes causing neurological disorders. Mutations in the tRNA synthetases are causing multiple disease e.g. mitochondrial myopathy (*YARS2*; MIM: 610957]), hearing loss (*LARS2*; MIM: 604544), leukoencephalopathy (*MARS2* [MIM: 609728]) (Oprescu *et al.*, 2017). By using WES sequencing we identified one family with two affected individuals that have autosomal recessive homozygous mutations in the asparaginyl-tRNA synthetase gene (*NARS*). Patients were negative for other genetic causes of neurodevelopmental disorders. The clinical phenotype in patients consisted of microcephaly, global developmental delay, intellectual disability, dysmorphism, neuropathy and seizures.

4.2.2 Family Background

The *NARS* family (MR-5) was enrolled from the children hospital Multan, distt Multan Punjab Pakistan. Family pedigree was constructed after interviewing the parents and nearby blood relatives. The blood collected from all the family members. The written research participant's forms collected from family members and affected child informed consent was also taken from their parents. Approval of the study was taken from the (BASR) and ethical committee of the univesity. Figure 4.2.1 representing the family pedigree.

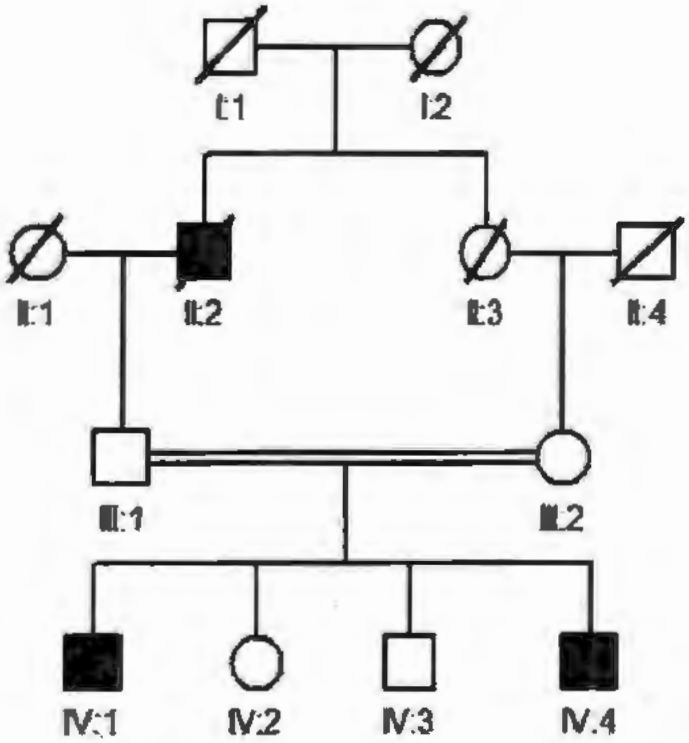


Figure 4.2.1: Family pedigree representing the two probands (IV: 1 and IV: 4). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals.

4.2.3 Clinical Details

NARS gene family (MR-5) with two probands (IV: 1 and IV: 4) were the first and fourth children of healthy, consanguineous Pakistani parents from district khanewal, Punjab. The proband (IV: 1) age was 21 months at the time of examination at that time she was not able to walk and was bedridden. She has developmental regression at the age of 9 months. She has normal posture, normal fundoscopy, able to speak few words only, no involuntary movement, no hepatosplenmegaly and cardiac involvement. The head size of the proband was 46 cm at the time of examination and considered as microcephaly. Brains MRI scan demonstrated cerebral emotional changes with reduced periventricular deep white matters. Clinically she was diagnosed with leukodystrophy disease. The proband (IV: 4) age was three years at the time of examination; he was not able to walk and was bedridden. He has developmental regression at early age when he was only 6 months old. He developed maximum ten seizures per day when he was only 1.5 years old . Type of seizure considered as generalized tonic-clonic seizure (GTC). He has decerebrate posture, normal fundoscopy, no speech, no involuntary movement, no hepatosplenmegaly and cardiac involvement. The head size of the proband was 45.5 cm at the time of examination and considered as microcephaly. Brains MRI scan demonstrated abnormal deep white matter signals involving subcortical prominent ventricular extra vent spaces. An electroencephalogram result shows diffuse slowing of background activity. Clinically he was diagnosed with leukodystrophy disease by the local neurologist.

4.2.4 Genetic Analysis

4.2.4.1 Linkage Analysis

DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.2.4.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 1 and IV: 4) come up with 137626 and 132172 variants respectively with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac browser (http://exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 4.2.2. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All eight variants putative score were checked in mutation taster (http://www.mutationtaster.org/), GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only eight variants met these filtration criteria in both probands (IV: 1 and IV: 4). Out of these variants six variants were non-synonymous and one was frame shift variant mentioned in Table 4.2.1. Of these filtered variants only NARS gene involved with neurodevelopment delay, seizures and neuropathy phenotype. NARS variant was not present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of NARS

gene variant considered as highly conserved. Mutation taster score for this variant is also considered as deleterious. NARS variant selected for validation with Sanger sequencing.

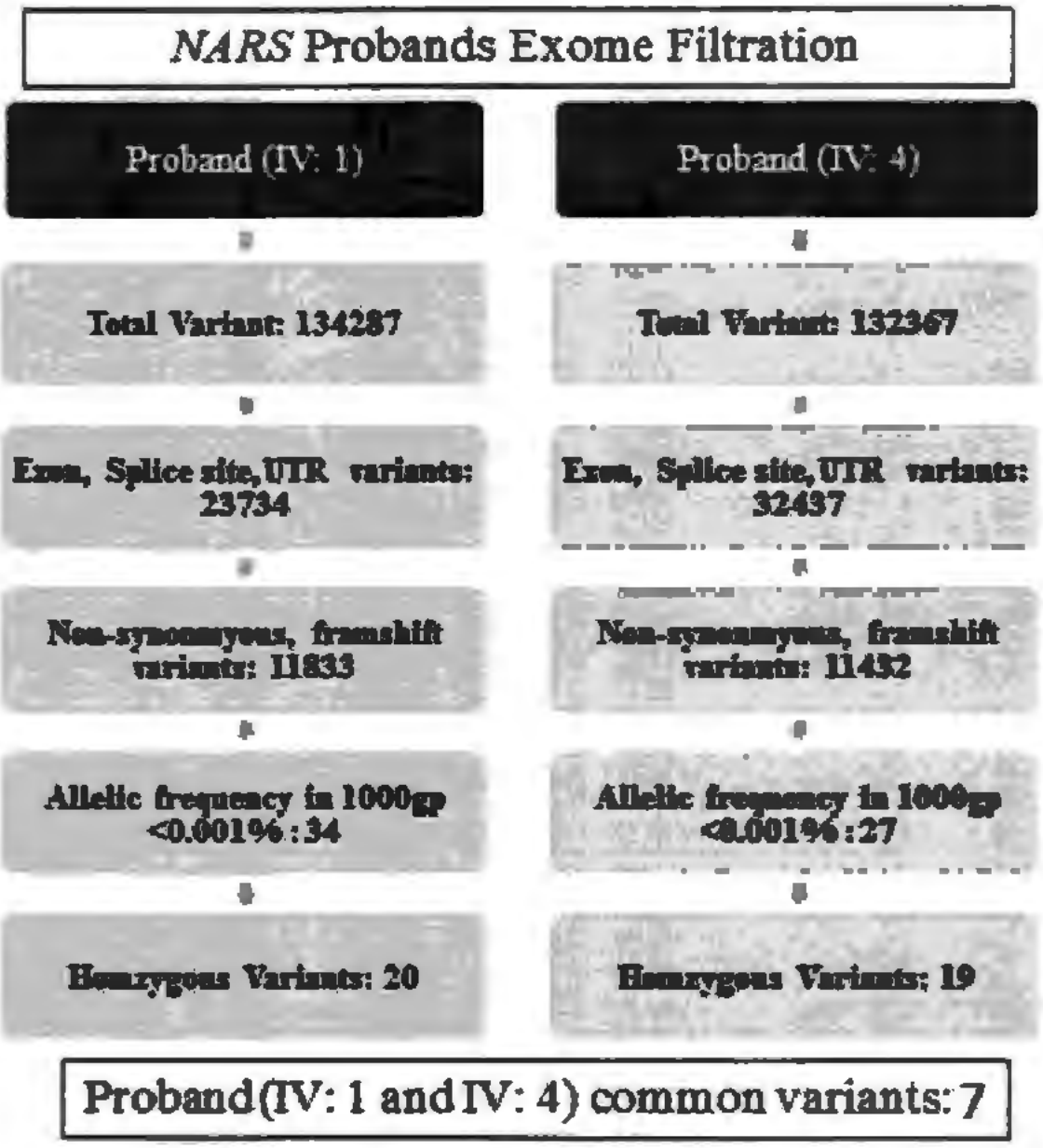


Figure 4.2.2: Exome sequencing filtration method and variants data of the individual (IV: 1 and IV: 4). The probands were sharing seven homozygous variants only.

Table 4.2.1: Summery of short listed variants in proband (IV: 1 and IV: 4)

Gene	Ref. Seq. ID	cDNA Mutation	Protein consequences	dbSNP ID	GERP Score
MORN1	NM_001301060	c.G491A	p.R164H	rs371494594	4.60
PLCH2	NM_001303012	c.G119A	p.R40H	rs528523454	4.60
DMBT1	NM_001320644	c.C655C	p.I219L	NA	-3.58
FAM24B	NM_001204364	c.G194C	p.C65S	rs5486425799487	-4.87
CTBP2	NM_022802	c.G994A	p.V332I	rs200267736	4.62
SLAIN1	NM_001242868	c.219_220insGG	p.L75Afs*125	rs201380414	NA
NARS	NM_004539	c.C1633T	p.R545C	NA	6

4.2.5 Co-Segregations Analysis

Sequencing analysis conform both affected siblings (IV: 1 and IV: 4) were homozygous, normal brother (IV: 2 and IV: 3) and their unaffected parents (III: 1) were heterozygous for the *NARS* variants, for sequencing result (see figure 4.2.3). This c.C1633T; p.R545C is novel variant not previously reported in databases, located in exon 18 of the *NARS* gene, Ensemble transcript (ENST00000256854) is disease causing for this family.

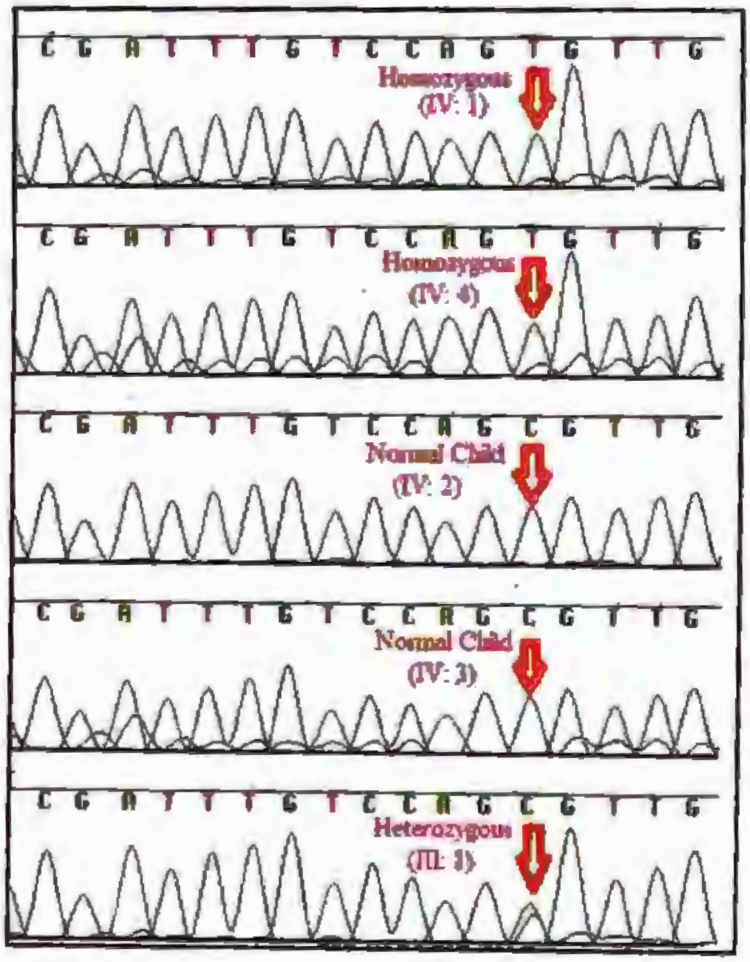


Figure 4.2.3: Co-segregation of mutation c.C1633T in exon 14 of NARS gene. Where both affected sibling were sharing homozygous region while father was heterozygous. Two other brother (IV: 2 and IV: 4) is normal.

4.2.6 In-silico Study

Insilco analysis of the mutation c.C1633T; p.R545C by using Swiss modeling software (www.swiss-prot.org), discloses that the mutant amino acid effects protein structure of *NARS* gene. The normal structure (A) and mutant structure (B) both visible in Figure 4.2.4

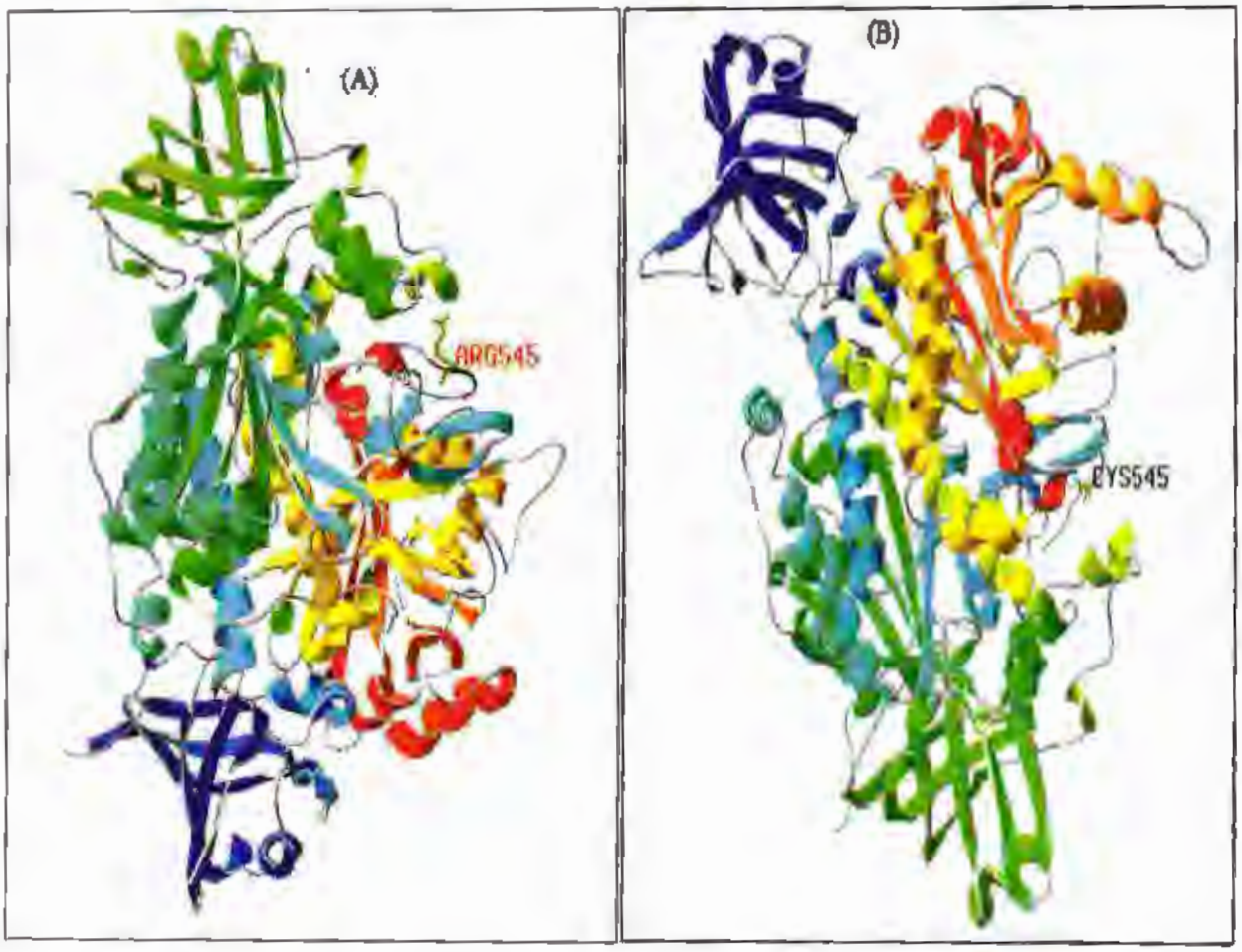


Figure 4.2.4: Swiss PDB viewer ribbon tertiary structure of *NARS* protein. The diagram (A) is representing the wild type structure of *NARS* protein. The diagram (B) is representing the mutated *NARS* protein. The amino acid cysteine replacing amino acid arginine at the position p.R454C; ensemble transcript for this mutation is (ENST00000256854).

4.2.7 Discussion

In this study missense variant reported in ancient tRNA enzyme asparaginyl-tRNA synthetase gene (*NARS*). There was a hotspot of homozygous recessive mutations located at p. R545C, within the last 40AA of the 548AA *NARS* gene with a severe clinical phenotype and death in childhood. The N- and C-terminus region mutation numbers, clusters and phenotypes suggests these are functionally important regions of the gene and disruption of ATP-binding/catalytic domain is critical in *NARS*. The mutation effect on *NARS* gene would seem to be widespread in the CNS and PNS as affected individuals had a broad neurodevelopmental phenotype characterized by motor and intellectual developmental delay, microcephaly, and seizures. The spectrum of diseased is severe in the p.R545C homozygous mutation. Mutation in the tRNA synthetizes functions exclusively in the cytoplasm, and many tRNA synthetase cofactors linked with multiple human diseases. Mutations in the tRNA synthetases connected with many human disease including the mid brain and hypo- myelination diseases. Aspartyl tRNA synthetase (*DARS*) gene with OMIM NO (MIM: 603084), non-syndromic deafness causing *KARS* gene, cytoplasmic AARS-encoding genes leads to neurodegeration of the PNS system e.g. tooth nerves pain, Glycyl-tRNA synthetase gene *AARS* (Omim: 601065), and *KARS* (Omim: 601421). ARSs have been associated with neurodevelopmental disorders and epilepsies e.g. Aspartyl tRNA synthetase gene (*DARS*), *AIMP* (OMIM: 260600), *AARS* (Omim: 601065), AARS2 (Omim: 615889), *DARS2* (MIM: 611105) (Oprescu *et al.*, 2017). *QARS* gene [MIM: 603727], were identified in two families with clinical phenol types such as progressive microcephaly, developmental delay, seizures, diffuse atrophy of the cerebral cortex (Zhang *et al.*, 2014). Mutations in 31 ARSs genes causing autosomal recessive disorders that show wide range of clinical phenotypes including early onset epileptic encephalopathy, myoclonic epilepsy, global developmental delay, microcephaly and seizures (Oprescu *et al.*, 2017). The pattern of inheritance of these diseases was either dominant or recessive. Research is needed to identify more neuropathy-associated ARSs mutations and loci and functional studied also needed to understand the full mechanism of dominant neuropathy mutation. Animal model study showing loss of functional affects will be important to define if additional studies needed to understand the pathogenicity of disease. It is important to decide if single enzyme or gene e.g. *MARS* is causing axonal neuropathy in dominant manners and to conform the disease animal model to neuropathy

will be very vital steps for the determination of disease pathogenicity. In future the functional work will help to understand the role of NARS mutations in disease but developing CRISPR/Cas9 animal models will be the next important step in understanding this protein function.

Chapter 4.3: Candidate Gene *SLITRK3* Mutation Causing Autosomal Recessive Intellectual Disability

4.3.1 Introduction

Synapse is a junction between two neurotransmitters, these neurotransmitters pass from presynaptic terminal to postsynaptic terminal via synapses. Different molecules involved in the multistep growth of synapse (Shen and Scheiffele, 2010; Siddiqui and Craig, 2011). Synaptic cell adhesion molecules (CAM), neurologins and *SLITRK3* molecules are neurotransmitters. CAMs are a membrane protein essential for the development of synapse connection. Neurologins and *SLITRK3* molecules bind with presynaptic protein tyrosine phosphates to mediate trans-synaptic proteins in CNS (Missler *et al.*, 2012; Lu *et al.*, 2017). *SLITRK3* a trans-membrane protein and has six different iso-forms with leucine-rich repeat domains (Aruga and Mikoshiba, 2003). Mutations in *SLITRK1* are causing Tourette's syndrome (Abelson *et al.*, 2005). *SLIRTK2* gene variants are associated with schizophrenia (Piton *et al.*, 2011). These study implicated that *SLITRK3* gene regulate synapse development and functions throughout the brain. *SLITRK3* act as inhibitory synapses development and tyrosine phosphates act as a presynaptic receptor for *SLITRK3*. Direct contact of *SLITRK3* with exons induced inhibitory presynaptic differentiation through trans binding to exons tyrosine phosphates in the brain. In-vitro and in-vivo study of *SLIRTK3* mice model implicated that *SLITRK3* proteins only act as inhibitory synapses (Takahashi *et al.*, 2012). *SLITRK3* membrane protein also requires for the maturation of GABAergic neurotransmitter (Jun li *et al.*, 2017). GABAergic neurotransmitter acts as inhibitory neurotransmitter in the central nervous system (CNS) in the mammalian brain. GABA neurotransmitter mediates inhibitory affects through distinct receptors, $GABA_A$, $GABA_B$ and $GABA_C$ (Bormann, 2000). Two type of inhibitory action take place in the CNS, phasic inhibition and tonic inhibition (McQuail *et al.*, 2015). GABA is synthesized from enzyme glutamate and transported to presynaptic terminal along the exon and captured into vesicles (Gonzalez-Burgos *et al.*, 2009). Depolarization of membrane cause the release of GABAergic neurotransmitters from presynaptic vesicles into the synaptic cleft. Release of neurotransmitters cause also activates receptor molecules on the postsynaptic membrane and the activation of postsynaptic membranes cause phasic inhibition in hippocampus of the mammalian brain (Farrant and Nusser, 2005). In the central nerves system balance between excitatory and inhibitory synapses required for the

brain functioning. Imbalance excitatory glutamate synapses and inhibitory GABAergic neurotransmitter was found to be pathogenic and cause neurological disorders (Kehrer *et al.*, 2008). GABAergic neurotransmitters are control neuronal excitation and act as inhibitory source in the brain. Dysfuctioning of GABAergic neurotransmitter cause neurological disorders (Ko *et al.*, 2015). Neurologins and *SLITRK3* are important for GABAergic synapses formation in the hippocampus (Chih *et al.*, 2005). Dysfuctioning of these two molecules increase seizures susceptibility in patients (Li *et al.*, 2017). In our studied family we identify a stop gain mutation that was found to disturbing the balance between excitatory and inhibitory synapse. Imbalance of excitatory and inhibitory synapses is the possible cause of the disease in our patients.

4.3.2 Family Background

The Slitrk3 family (MR-6) was enrolled from the children hospital Multan, distt Multan Punjab Pakistan. Family pedigree was constructed after interviewing the parents and nearby blood relatives. Blood samples were collected from the family members both affected and normal. The informed consent was obtained from all the members of the family and affected child informed consent was also taken from their parents. The study was approved from the (BASR) Board of Advanced Study and Research and ethical committee of International Islamic university. Pedigree of the family of was shown in below Figure 4.3.1.

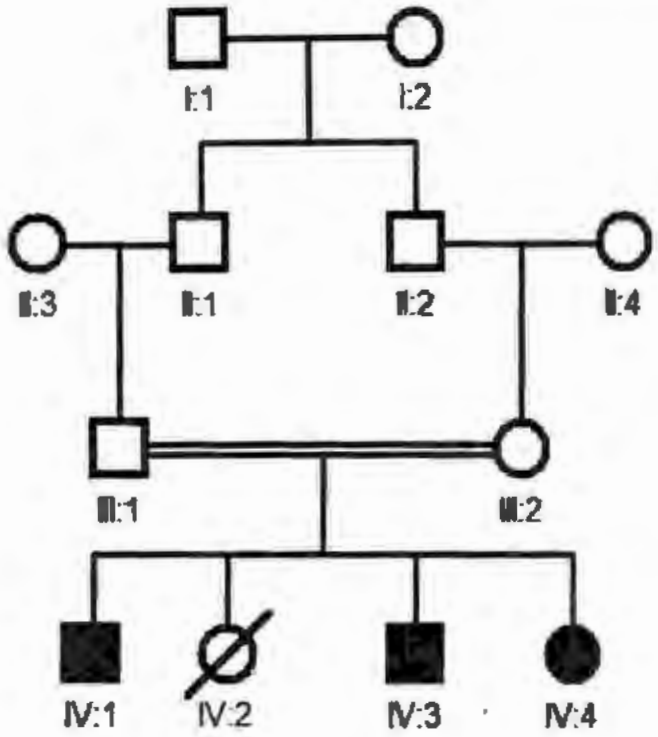


Figure 4.3.1: Family pedigree representing the three probands (IV: 1, IV:3 and IV: 4). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals.

4.3.3 Clinical Details

The family (MR-6) with three probands (IV: 1, IV: 3 and IV: 4 Figure 4.3.2, while IV: 4 pic was not included here) were the first, third and fourth children of healthy, consanguineous Pakistani parents from district Muzfar Ghar, Punjab. Proband (IV: 1) start walking independently at the age of 13 months. The proband (IV:1) age was 4 years when first came for clinical examination. He developed maximum five seizures per day up to 1-2 minutes of duration at the age of 2 years. Type of seizure considered as generalized tonic-clonic seizure (GT) and extensor spasm. Detailed clinical investigation suggests he has no behavior problems, but only able to talk few words. He has normal posture, normal fundoscopy, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Brain CT Scan demonstrated as unremarkable. He was diagnosed with familial intellectual disability, developmental delay with epilepsy by the local neurologist. The

proband (IV: 3) age was 3 years at the time of clinical examination. At the 18 months of age, he developed maximum 7-8 seizures per day up to 4-5 minutes of duration. Type of seizure considered as GTC. During the clinical investigation his behavior was normal, no facial dysmorphic feature, his head size is 44cm considered as microcephaly. He has normal posture, normal fundoscopy, no cerebellar signs, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Electroencephalogram test demonstrated encephalopathy. He was diagnosed with familial intellectual disability, global developmental delay by the local neurologist. The proband (IV: 4) age was 5 years when she went for clinical examination. She developed maximum 10 seizures per day up to 5 minutes of duration only when she was 6 months old. Type of seizure considered as GTC. She has no behavior problem, no facial dysmorphic feature, her head size was normal. She has normal posture, normal fundoscopy, no cerebellar signs, no involuntary movement, no hepatosplenmegaly and cardiac involvement. She was able to talk just few words with non-responsive cognition. Brain CT Scan read as unremarkable electroencephalogram test demonstrated focal encephalopathy activity from left parental area. She was diagnosed with familial intellectual disability, global developmental delay by the local neurologist.

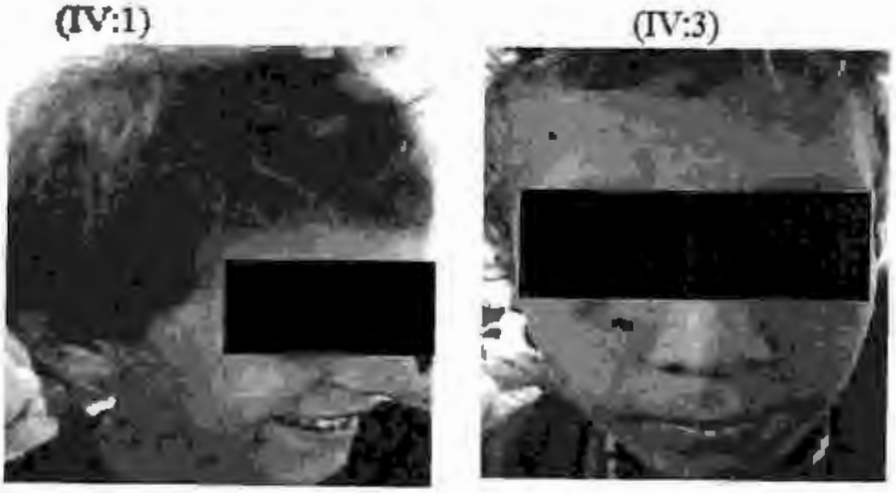


Figure 4.3.2: Phenotypic representation of the proband (IV: 1) and (IV: 3), during clinical investigation behavior of both the probands normal and they have no facial dysmorphic feature.

4.3.4 Genetic Analysis

4.3.4.1 Linkage Analysis

DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.3.4.2 Whole Exome Analysis

To identify underlying genetic causes the family (MR-6) probands selected for whole exome sequencing. WES results of the probands (IV: 1, IV: 3 and IV: 4) come up with 90382, 89732 and 90832, variants respectively with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed from the exome file. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac browser (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% shortlisted figure 4.3.3. For further conformation filtered variants checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All five variants pathogenic score checked in mutation taster data base (http://www.mutationtaster.org/), GERP and CADD phred pathogenic score checked by using ensembl online tool. Only four variants met these filtration criteria in probands (IV: 1, IV: 3 and IV: 4). Two variants were non-synonymous and 2 were frame shift variants mentioned in table 4.3.1. Of these filtered variants only *SLITRK3* gene involved with ID. *SLITRK3* variant was not present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of *SLITRK3* gene variant considered as highly conserved. Mutation taster score for this variant show deleterious. *SLITRK3* variant selected co-segregation analysis.

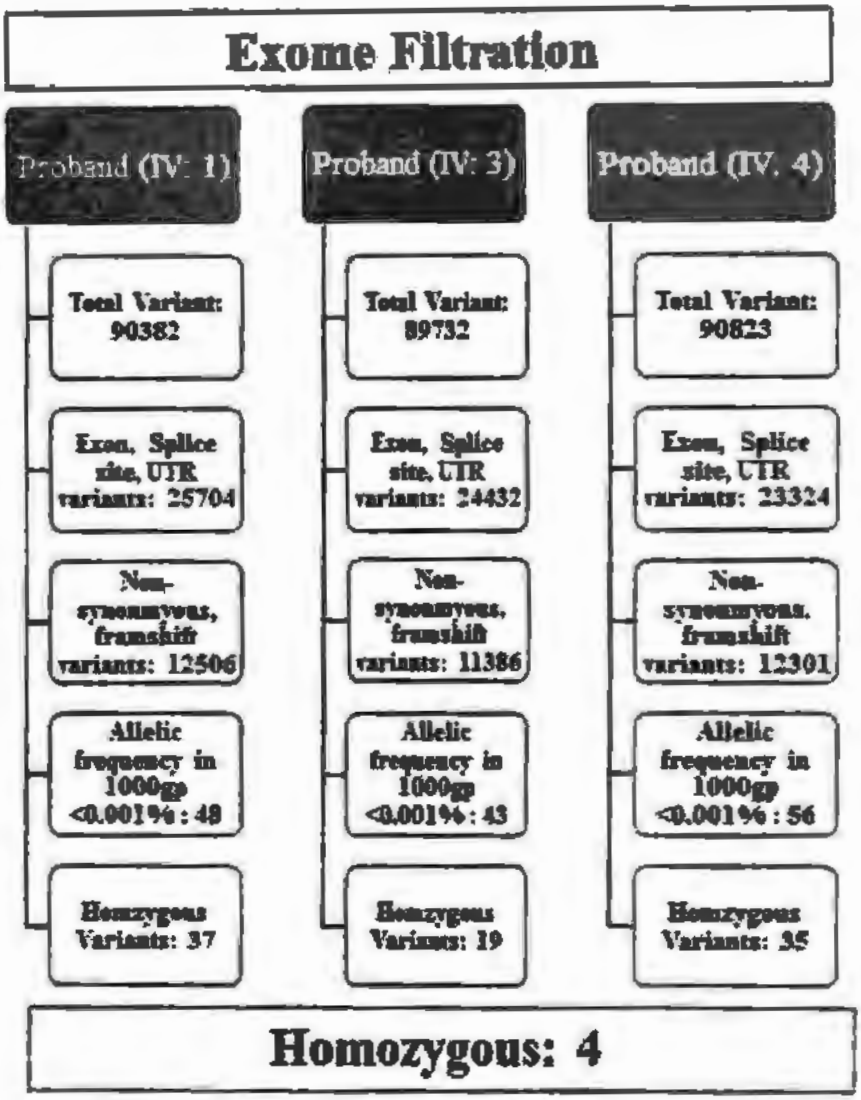


Figure 4.3.3: WES Data filtration steps of the patients (IV: 1, IV: 3 and IV: 4).

Table 4.3.1: Summery of variants found in proband as mention in pedigree of the family.

Gene	Ref. Seq. ID	cDNA Mutation	Protein	dbSNP ID	GERPScore
ITPK1	NM_001142593	c.G1138A	p.G380S	rs574548381	0.027
LNP1	NM_001085451	c.194_195insTCCTA	p.S80_H81insRL	rs71132521	NA
SLITRK3	NM_001318810	c.G1816T	p.E606X	NA	5.76
PHC3	NM_001308116	c.68_70del	p.T23Ifs*961	NA	NA

4.3.5 Co-segregation Analysis

Sequencing analysis conform all the affected children"s were homozygous. Both father (III: 1) and mother (III: 1) heterozygous for *SLITRK3* variants, for sequencing result see figure (4.3.4). This c.G1816T: p.E606X is novel variant not previously reported in databases, located in exon 2 of the *SLITRK3* gene, Ensemble transcript (ENST00000241274) is disease causing for this family.

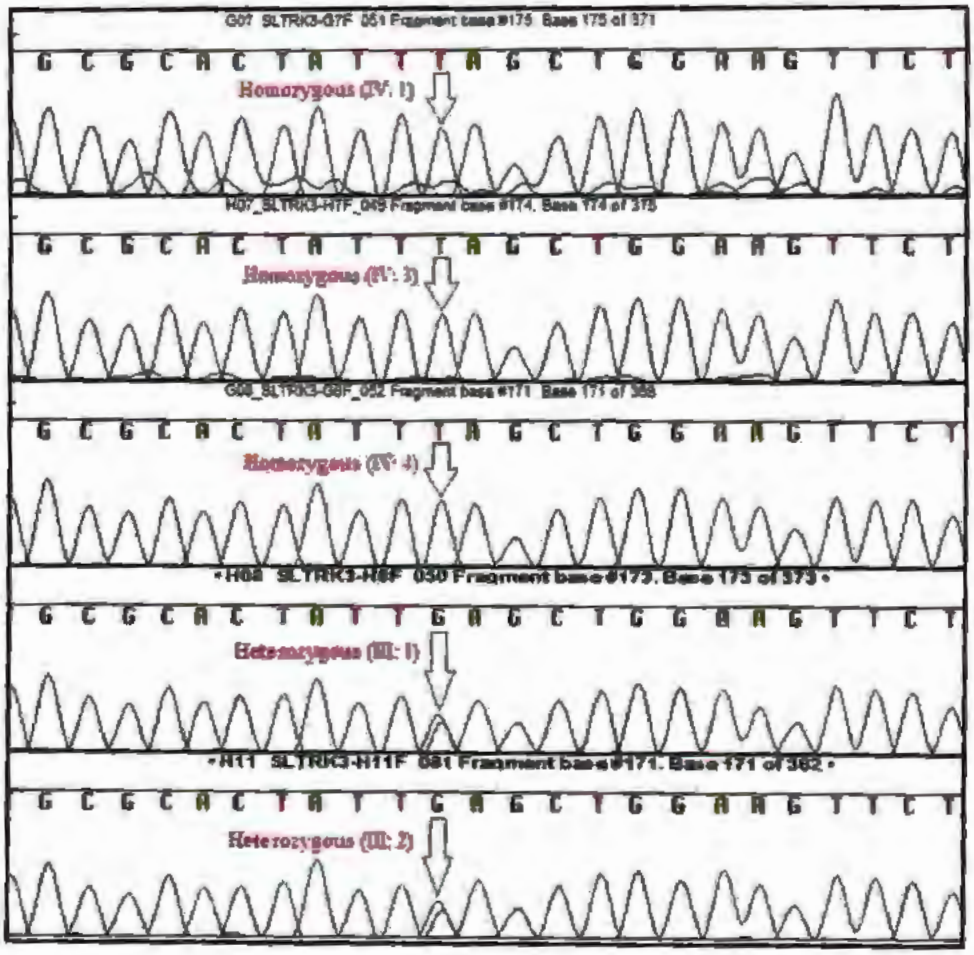


Figure 4.3.4: Chromatogram demonstrating segregation of mutation c.G1816T in exon 2 of *SLITRK3* gene. Where all three affected sibling were sharing homozygous region while parents were heterozygous.

4.3.6 Bioinformatics Analysis

Insilco analysis of the stop gain mutation c.G1816T: p.E606X by using I- TASSER software discloses that the mutant amino acid effects protein structure of *SLITRK3* gene. As the mutation is nonsynonymous cause change in 3D structure of *SLITRK3* protein. The normal structure (A) and mutant structure (B) both visible in Figure 4.3.5.

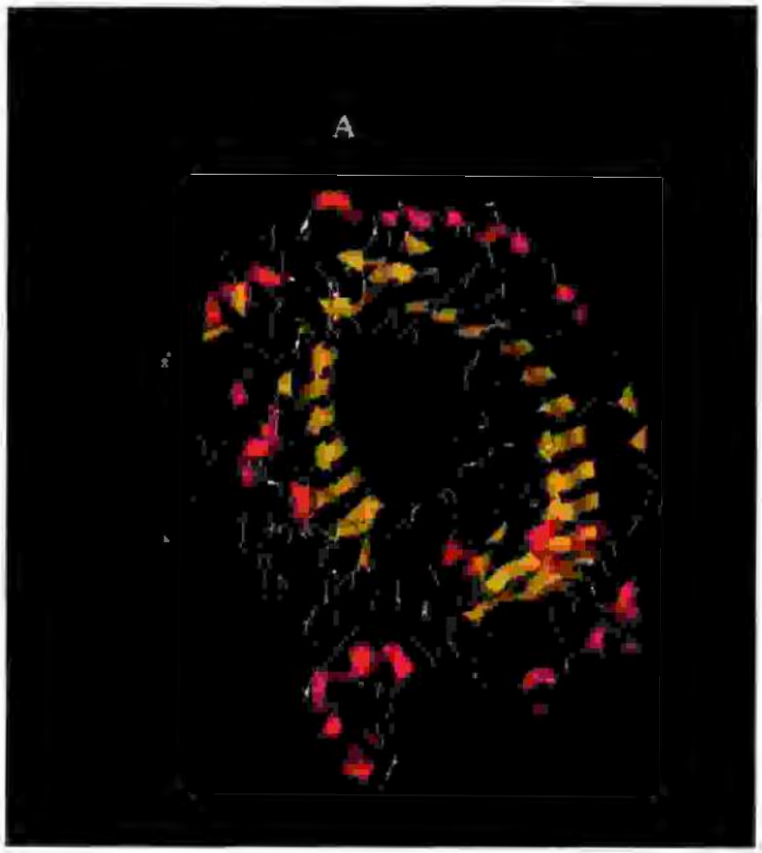


Figure 4.3.5: I-TASSER ribbon tertiary structure of *SLITRK3*. The diagram (A) is representing the wild type structure of *SLITRK3* protein.

4.3.7 Discussion

Normal development of excitatory and inhibitory synapses is essential for normal brain functioning and cognitive processing. Distorted growth of these synapses leads to the neurological disorders. Different type of interneurons collectively establish diverse inhibitory synapses onto cell with long exon known as principal neurons in the brain (Takahashi *et al.*, 2012). GABAergic interneurons transmit inhibitory synaptic signal to neurons in the CNC and keep balance between excitatory and inhibitory synapses (Huang *et al.*, 2007; Cline, 2005). Imbalance between excitatory and inhibitory synapses is the possible causes of neurological disorders (Rubenstein and Merzenich, 2003; Wassef *et al.*, 2003; Moehler, 2006). Maturation of GABAergic requires interneurons molecules NL2 and *SLITRK3* in the hippocampus (Chih *et al.*, 2004; Yim *et al.*, 2013). *SLITRK3* is integral membrane proteins with leucine rich domain function as connecting molecule when neuron cross from pre-synapse and inter into post-synapse area that control inhibitory synapse development with receptor tyrosine phosphatase (Takahashi *et al.*, 2012). *SLITRK3* variant causes imbalance between excitatory and inhibitory synapses that affect the brain functioning as described above. In this family WES analysis of three affected sibling was performed and four variants were selected for co-segregation (Table 4.3.1). Sequencing of the variants has identified a novel stopgain mutation (ENST00000475390: exon2: c.G1816T:p.E606X) in *SLITRK3* gene as the cause of disease in this family. Amino acid Glutamic acid (E) is replaced with stop codon (X) resulted truncated protein. Co-segregation analysis conformed all three affected sibling were homozygous while their parents were heterozygous (Figure 4.3.4). The mutation is not present in databases like HGMD, exac browser, Gnomad data base. In-silico analysis of mutation implicated strong evidence that mutation is disease causing for this family. Combined annotation dependent depletion (CADD) score for this variant is 38, considered as pathogenic (Rentzsch *et al.*, 2018). GERP score for this mutation is (5.76); suggest that mutation is present in highly conserved region in the gene and this gene is intolerant for the mutation. Mutation taster predicts that the variant is damaging as well for *SLITRK3* gene. Protein modeling of the *SLITRK3* gene reveals major structural change in the protein model due to the stopgain resulted truncated protein structure. The common symptoms in this family were intellectual disability, developmental delay and epilepsy. Head circumference in one of the patient was smaller than normal head size and patient considered as microcephaly.

Additionally seizures were present in all the patients. WES analysis is better option to diagnosed patients with neurological disorders.

Chapter 4.4: Mutation in *GLB1* Gene Causing Rare Autosomal Recessive GM1 Gangliosidosis Disease in Pakistani family

4.4.1 Introduction

Mutation in GLB1 gene cause autosomal recessive lysosomal storage disorders known as GM1 gangliosidosis (Suzuki, 1995; Caciotti *et al.*, 2005). The prevalence of Gangliosidosis vary from country to country, global prevalence ratio at birth about to be 1 in 100,000 to 300,000 (Regier, 2013). The incidence of GM1 gangliosidosis in Brazil has been 1 in 17,000 (Severini *et al.*, 1999) and in Maltese Islands 1 in 3700 live birth ratio (Lenicker *et al.*, 1997). GM1 gangliosidosis clinically can be classified into distinct phenotypes. Type 1 (MIM: 230500) GM1 gangliosidosis involved with cardiac involvement, early age of onset, hypotonia, severe degeneration of CNS and leads to death with in first 2 years of life. Type 2 (MIM: 230600) disease phenotype characterized by slow progress of neurodegeration, seizures, muscle weakness. The early diagnosis of type 2 GM1 gangliosidosis is challenging due to slow progress of disease. Type 3 (MIM: 230650) GM1 gangliosidosis is less pathogenic than type 1 and 2 (Callahan, 1999; Suzuki *et al.*, 2011). Morquio B (MIM: 253010) pathogenic condition occurs due to variation in *GLB1* gene due to the deficiency of β-galactosidase and it is an allelic from of GM1 gangliosidosis (Sohn *et al.*, 2012). Morquio B phenotypically characterized by cardiac involvement, manifest skeletal anomalies augmented urinary secretion of keratan sulfate, no sign of CNS involvement is appeared (Suzuki *et al.*, 2011). Morquio B patients' phenotypes similar with GM1 gangliosidosis as phenotypes described above. Another disorder of galactosialidosis (MIM# 256540) is a defect due to the combined deficiency of b-galactosidase and neuraminidase (Azzo *et al.*, 1992). The clinical phenotype of galactosialidosis in some extent shows similarities with gangliosidosis. Disease phenotypes GM1 gangliosidosis were including severe neuraminidase deficiency instead of fractional shortage of the beta galactosidase enzyme (Brunetti-Pierri and Scaglia, 2008). GM1 gangliosidosis caused due to the storage of GM1 gangliosides and glycol-conjugates in the central nerves system (CNS). The molecular mechanisms of the disease pathogenesis are not fully understood yet. Demyelination and apoptosis of neurons together with micro-gliosis, astro-gliosis are detected around the vacuolation in the brain (Tessitore *et al.*, 2004). To date, total 185 mutations have been reported in the *GLB1* gene (www.hgmd.cf.ac.uk). Novel mutation described in this section causing GM1 gangliosidosis in consanguineous family.

4.4.2 Background

The selected family (MR-7) was enrolled from the children hospital Multan, distt Multan Punjab Pakistan. Family pedigree was constructed based on the information provided by the parents and nearby blood relatives. Blood samples were collected from the family members both affected and healthier. The informed consent was obtained from all the members of the family and affected child informed consent was also taken from their parents. The study was approved from the (BASR) Board of Advanced Study and Research and ethical committee of international Islamic university Islamabad, Pakistan. Pedigree of the family of was shown in below Figure 4.4.1

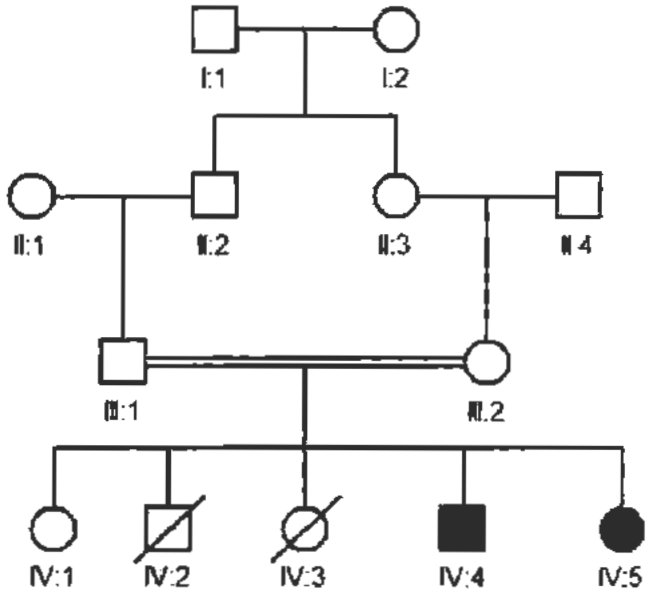


Figure 4.4.1: Family pedigree representing the two probands (IV: 4 and IV: 5). Pedigree squares symbolizing the male individuals and circle designates the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle represents deceased individuals

4.4.3 Clinical Details

The clinical manifestations of the family (MR-7) correlates with the identified mutation, in this family two probands (IV: 4 and IV: 5) were the fourth and fifth children of healthy, consanguineous parents from Pakistan. The proband (IV: 4) a 3 years old boy was brought for

medical examination in the neurology section of the children complex hospital Multan Pakistan due to the developmental regression and seizures. He developed seizures at the age of 1.5 years, maximum 10 seizures per day. Type of seizure considered as generalized tonic-clonic seizure (GTC). He gradually lost his motor skill and became bed ridden and his mental ability decreased with course of time. Detailed clinical investigation suggests his head size is (45.5 cm) considered as microcephaly. He has decerebrate posture, normal fundoscopy, not able to talk, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Brain MRI Scan demonstrated abnormal deep white mater signal in subcortical, prominent ventricular and extra vent spaces. EEG result shows diffuse slowing of background activity (See Figure 4.4.2). The proband (IV: 5) age was 21 months at the time of clinical examination. Detailed clinical investigation suggests she has developmental regression and the regression started at the age of 9 months, she has no behavior, no facial dysmorphic feature. She has normal posture, normal fundoscopy, no cerebellar signs, no involuntary movement, no hepatosplenmegaly and cardiac involvement. She is not able to walk (bed ridden), able to talk just few words, head circumferences is (46 cm) noted as microcephaly. Brain MRI Scan demonstrated cerebellar emotional changes with reduced periventricular deep white matter (See Figure 4.4.2, IV: 5).

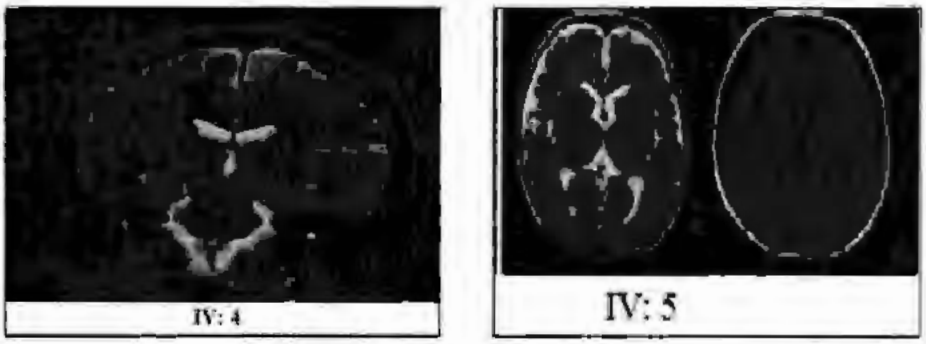


Figure 4.4.2: Brain MRI Scan of (IV: 4) demonstrated abnormal deep white mater signal and brain MRI of (IV: 5) demonstrated cerebellar emotional changes with reduced periventricular deep white matter.

4.4.4 Genetic Analysis

4.4.4.1 Linkage Analysis

DNA from all affected family (MR-7) members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.4.4.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 4 and IV: 5) come up with 78683 and 78321 variants respectively with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac browser (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 4.4.3. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All five variants putative score were checked in mutation taster (http://www.mutationtaster.org/). GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only five variants met these filtration criteria in both probands (IV: 4 and IV: 5). Out of these variants three variants were non-synonymous and two were frame shift variants mentioned in table 4.4.1. Of these filtered variants only GLB1 gene involved with accumulation of GM1 gangliosides in brain, abnormal deep white mater signal in subcortical area of the brain considered as neurodegenerative disorder. GLB1 variant was not present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of GLB1 gene variant considered as highly conserved. Mutation taster score for this variant is also considered as deleterious. Sequencing performed to check the co segregation of the family with GLB1 gene.

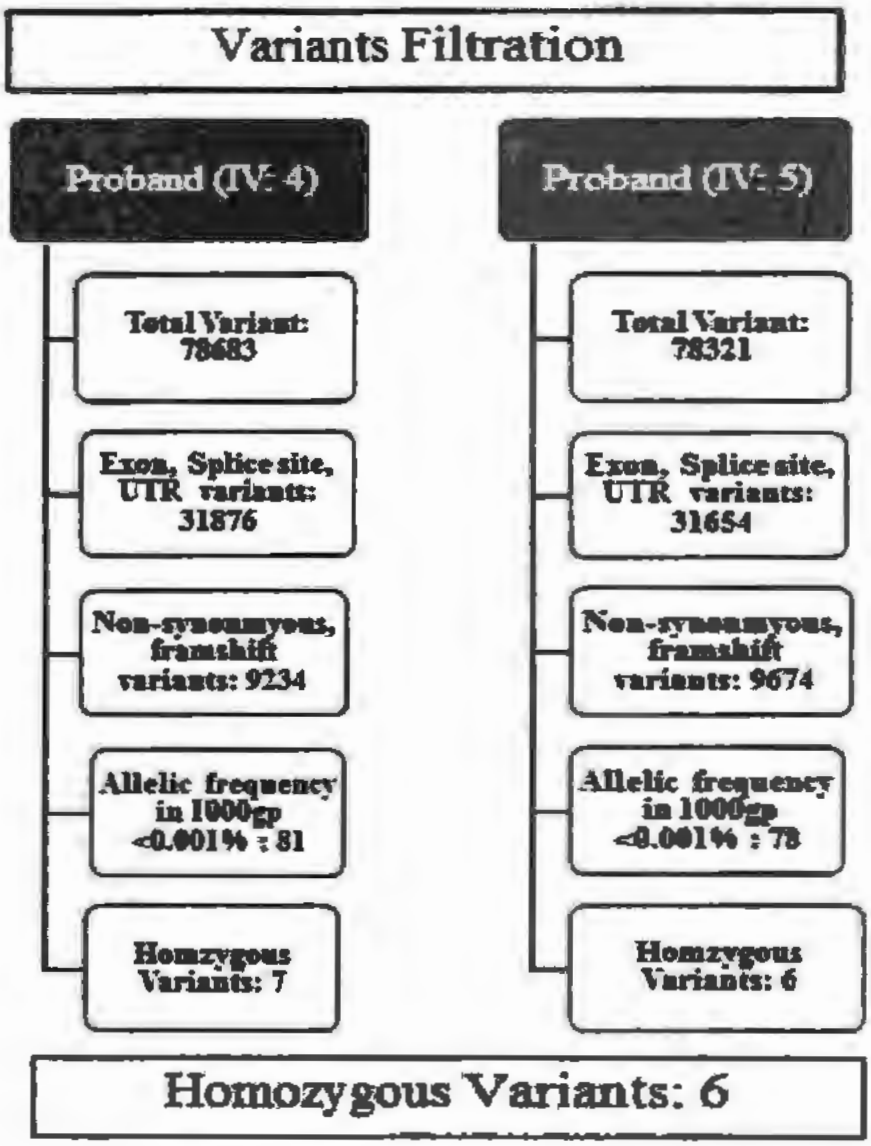


Figure 4.4.3: Whole exome sequencing (WES) data analysis steps of the individual (IV: 4 and IV: 5). Only 6 homozygous variants common in both patients.

Table 4.4.1: Summery of variants found in proband (IV: 4 and IV: 5)

Gene	Ref. Seq. ID	cDNA Mutation	Protein consequences	dbSNP ID	GERP Score
STAM	NM_001324286	c.G157A	p.A53T	NA	5.23
TRAK1	NM_001265609	c.1841_1842insGGA	:p.E624_G625insEE	rs10634555	NA
GLB1	NM_001135602	c.C1318T	p.H440Y	NA	5.53
LRIG1	NM_015541	c.G1210A	p.G404R	rs576361868	5.46
CCT5	NM_001306154	c.A556G	p.K186E	NA	5.50
FAM46A	NM_017633	exon2:c.131_132insCG	p.G45_G46insDP	NA	NA

4.4.5 Co-Segregation analysis

Sequencing analysis conform both affected siblings (IV: 4 and IV: 5) were homozygous, healthy sister (IV: 1), and their unaffected parents (III: 1 and III: 2) were heterozygous for the *GLB1* variants, for sequencing result see figure (4.4.4). This c.C1318T: p.H440Y is novel variant not reported in databases, located in exon 12 of the *GLB1* gene, Ensemble transcript (ENST00000307363) is disease causing for this family.

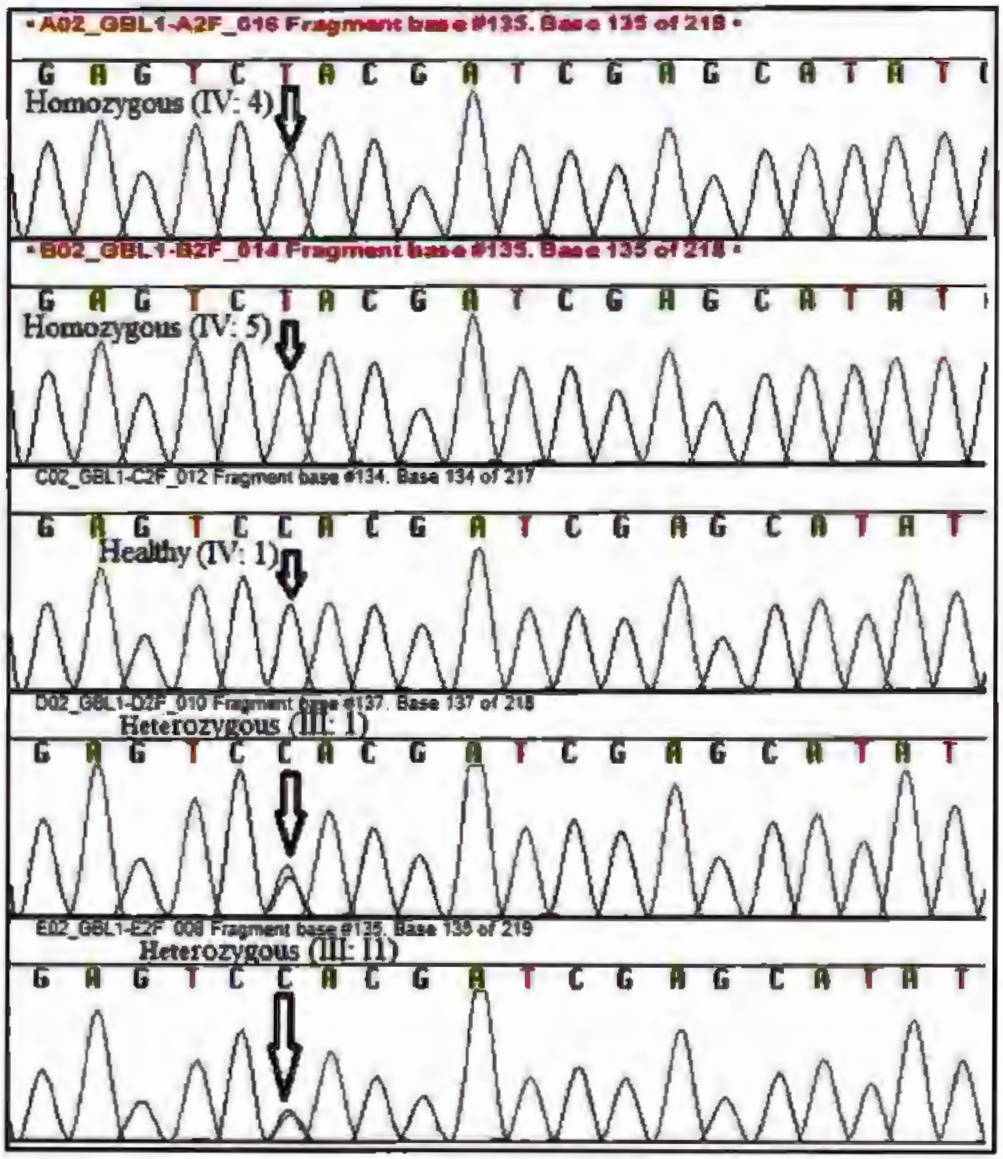


Figure 4.4.4: Chromatogram demonstrating segregation of mutation c.C1318T; p.H440Y in exon 13 of *GLB1*. Where both affected sibling were sharing homozygous region while parents were heterozygous and sister (IV: 1) normal.

4.4.7 In-silicon Study of Wild Type and Mutated Protein Structure of *GLB1*

Insilco analysis of the missense mutation c.C1318T; p.H440Y by using Swiss modeling software (www.swiss-prot.org), discloses that the mutant amino acid effects protein structure of *GLB1* gene. As the mutation is nonsynonymous cause change in 3D structure of *GLB1* protein. The normal structure (A) and mutant structure (B) both visible in Figure 4.4.5.

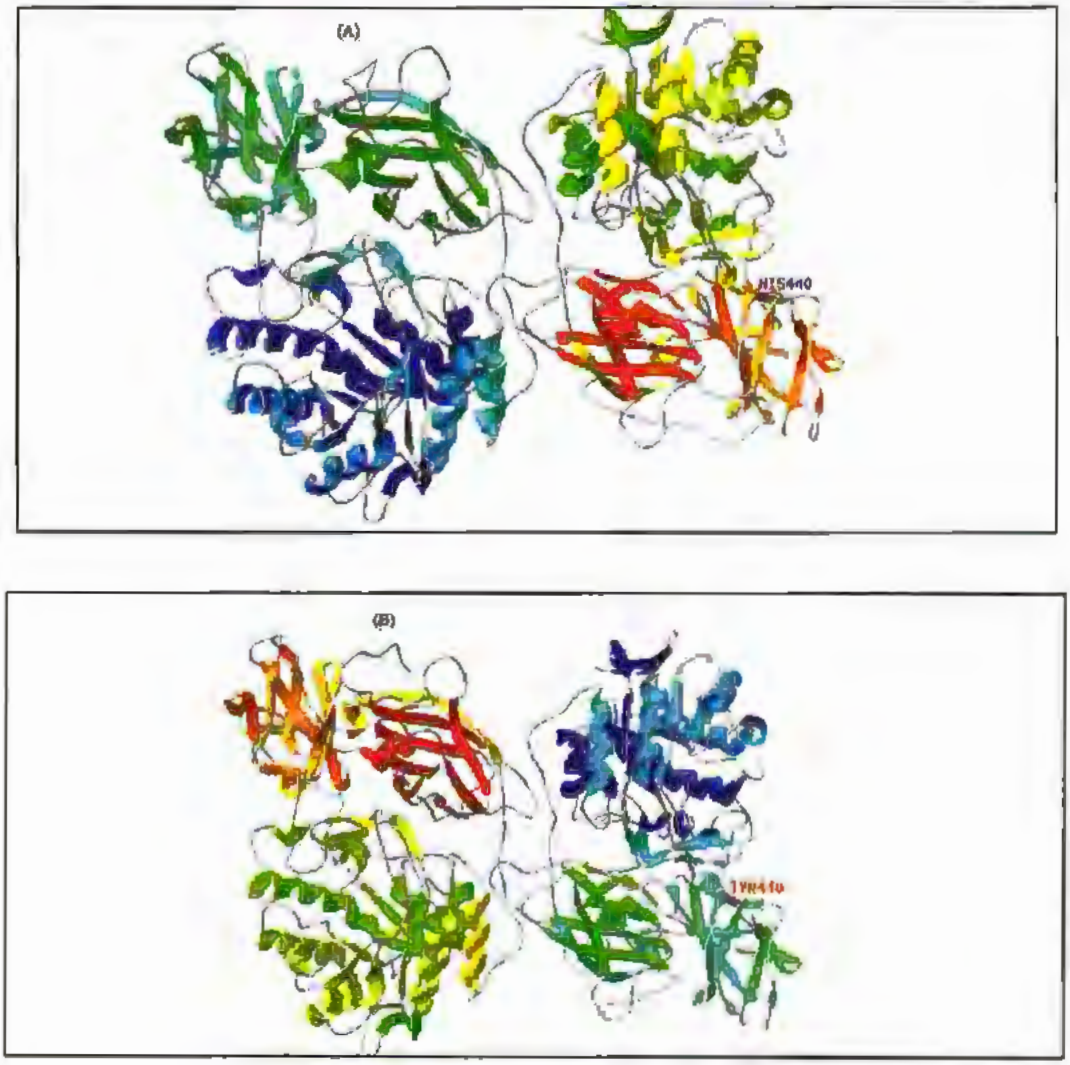


Figure 4.4.5: Swiss PDB viewer ribbon tertiary structure of *GLB1*. The diagram (A) is representing the wild type structure of *GLB1* protein. The diagram (B) is representing the mutated *GLB1* protein.

4.4.8 Discussion

GLB1 gene encodes an enzyme called beta galactosidase. B-galactosidase located in lysosomes and acts as a degrading component of few molecules such as keratan sulfate, glycolipids and GM1 gangliosides within lysosomes. The accumulation of GM1 gangliosides, glycolipids and keratan sulfate leads to the lysosomal storage disease due to the deficiency of dissoluteness of B-galactosidase enzyme (Kannebley *et al.*, 2015; Sandhoff and Harzer, 2013). The structure of B-galactosidase enzyme consists of three domains, TIM barrel domain, b-domain1 and b-domain 2. Triosephosphate isomerase (TIM) barrel domain consist of (49-359) amino acid residues, while the b-domain 1 is start from amino acid 397-514 and the b-domain 2 (545-647) amino acid residue (Ohto *et al.*, 2012). More than 185 mutations was reported in *GLB1* gene have been reported, nonsynonymous mutation, small insertion and deletion mutations from all around the world (Stenson *et al.*, 2017). Here we performed WES to identify pathogenic mutations in two sibling were born of consanguineous Pakistani family (MR-7) with early onset of developmental regression at the age of 6 month. WES analysis was performed, after data filtration we selected six variants for Sanger sequencing to conform the segregation. We identified novel nonsynonymous homozygous mutation in *GLB1* gene on (ENST00000307363:exon13:c.C1318T:p.H440Y), where amino acid histidine (CAC) was replaced with amino acid tyrosine (TAC) (figure 4.4.4). Our identified mutation was present in B-domain 1 of the b-galactosidase structure. B-domain consists of four beta sheets, mutations in the protein core region causing large structural changes in the protein structure of the b-galactosidase enzymes. These mutations may be affecting the degrading activity in lysosomes. Our reported variant is not present in public data bases like 1000gp, exac browser and gnomad data bases. In-silico study of this mutation generates high scores of deleteriousness, combined annotation dependent depletion (CADD) score (26) suggest that variant is pathogenic (Rentzsch *et al.*, 2018). GERP score for this mutation is (5.50); suggest that mutation is present in highly conserved region in the gene. Mutation Taster predicts that the variant is damaging, Polyphen2 predicts that the variant is pathogenic. A novel missense homozygous mutation in *GLB1* gene, co-segregating with the disease phenotypes was identified in two siblings see figure 4.4.5.

Chapter 4.5: Novel Variant in *MLC1* gene Causing Megalencephalic Leukoencephalopathy with Subcortical Cysts

4.5.1 Introduction

Megalencephalic leukoencephalopathy *MLC1* (MIM: 604004) is a autosomal recessive disorder with prevalence ratio 1 in 7663 live births, characterized by macrocephaly, with slow onset of progressive neurological signs including ataxia, slow onset of degradation of motor functioning and mild mental degeneration (Van der Knaap *et al.*,1995; Leegwater *et al.*, 2001; Bonkowsky *e tal.*, 2010). Affected individual with *MLC1* contains subcortical cysts and white matters in the anterior temporal region of the brain (Van der Knaap *et al.*, 2012). Brain MRI used for the diagnosis and presence of white matters in cerebral and subcortical cysts, specifically, in the area of temporal region of the brain, but often also in the frontal and parietal parts of the brain. Regardless of these severe neuro-radiological anomalies, *MLC1* presents a clinical evolution milder than other infantile forms of leukodystrophy with lysosomal defects. Brain biopsy analysis shows porous white matters associated in the presence of many fluid follicle that protecting medullary sheaths (Pascual-Castroviejo *et al.*, 2005). Variants in MLC1 and GlialCAM are leads to MLC disease (Topcu *et al.*, 2000; Lopez-Hernandez *et al.*, 2011). Approximately 75% *MLC* patients are genetically identified with homozygous or compound heterozygous variants in the *MLC1* gene. *MLC1* gene is localized in astrocyte glial cell and Bergmann glia cell area and express in brain (Ilja *et al.*, 2005; Teijido *et al.*, 2004). The GlialCAM gene synthesis cell adhesion proteins, it is highly expressed and functional in the brain especially in CNS (Lopez-Hernandez *et al.*, 2011). GlialCAM act as molecular chaperons and mediates chloride channels clc-2 (Jeworutzki *et al.*, 2011). *MLC1* gene encodes 377 amino acid and contains eight transmembrane domains. The function of the human *MLC1* gene is not fully elucidated, it connects neighboring astroglial cells at junctions, and the role of *MLC1* at cell junction is unknown. The brain alterations detected in *MLC1* patients, large amount of protein expression found in glial cells have led to hypothesize that this gene regulates ions and water homeostasis and control the flow of fluids into cells to strength the cell adhesion (Brignone *et al.*, 2015). Here, using WES we identified a novel homozygous mutation in Pakistani family causing autosomal recessive mutations in the *MLC1* gene. Patient was previously negative for other genetic causes of neurodevelopmental and neurological disorders.

4.5.2 Background

The selected family (MR-8) was enrolled from the children hospital Multan, distt Multan Punjab Pakistan. Family pedigree was constructed based on the information provided by the parents and nearby blood relatives. Blood samples were collected from the family members both affected and healthier. The informed consent was obtained from all the members of the family and affected child informed consent was also taken from their guardians. The study was approved from the (BASR) Board of Advanced Study and Research and ethical committee of International Islamic university Islamabad, Pakistan. Pedigree of the family of was shown in below Figure 4.5.1

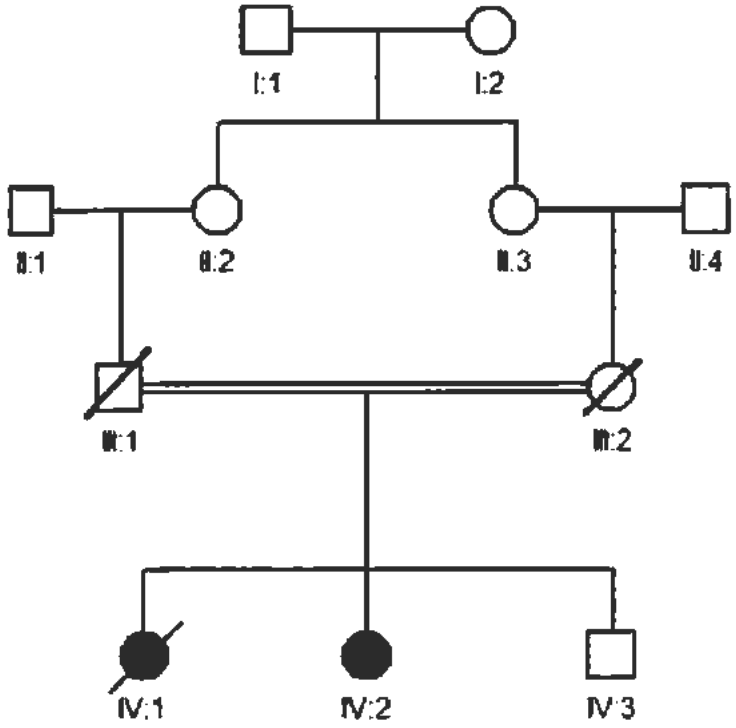
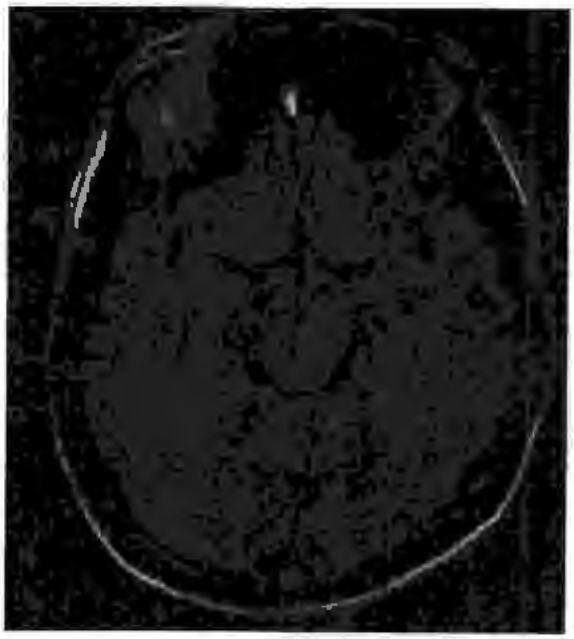


Figure 4.5.1: Family pedigree representing the proband (IV: 2) and healthy (IV: 3). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals

4.5.3 Clinical Details

A 7 year old patient (IV: 2) was the child of healthy consanguineous parents from district Mian Chanu Punjab. Her birth weight was 2.2 kg and head circumference 41cm. she developed

macrocephaly within the first few months of life, and thereafter showed slow motor deterioration, and cognitive decline. Her head size was 43 cm at the age of 4 months; the head circumferences indicated macrocephaly in child. She starts controlling her head at the age of 8 month. She start walking at the age of 15 month but with some difficulties, but with some difficulty. She was developed recurrent episodes of seizures at the 2.5 years of age, mental retardation and progressive motor dysfunction. Brain CT scans at 7 years old revealed extensive bilaterally symmetrical white matter changes and with subcortical cysts in the bilateral anterior temporal region of the brain (See Figure 4.5.2).



IV: 2

Figure 4.5.2: The brain MRI of patient (IV: 2) showing white matter changes in and with subcortical cysts in the bilateral anterior temporal region of the brain

4.5.4 Genetic Analysis

4.5.4.1 Linkage Analysis

Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.5.4.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 2) come up with 78963 variants with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac data base (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 4.5.3. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All 15 variants putative score were checked in mutation taster (http://www.mutationtaster.org/), GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only 15 variants met these filtration criteria in proband (IV: 2). All the variants were non-synonymous table 4.5.1. Of these filtered variants only *MLC1* gene involved with white matter changes and with subcortical cysts in the bilateral anterior temporal region of the brain. *MLC1* variant was not present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of

MLC1 gene variant considered as highly conserved. Mutation taster score for this variant is also considered as deleterious. *MLC1* variant selected for co-segregation analysis.

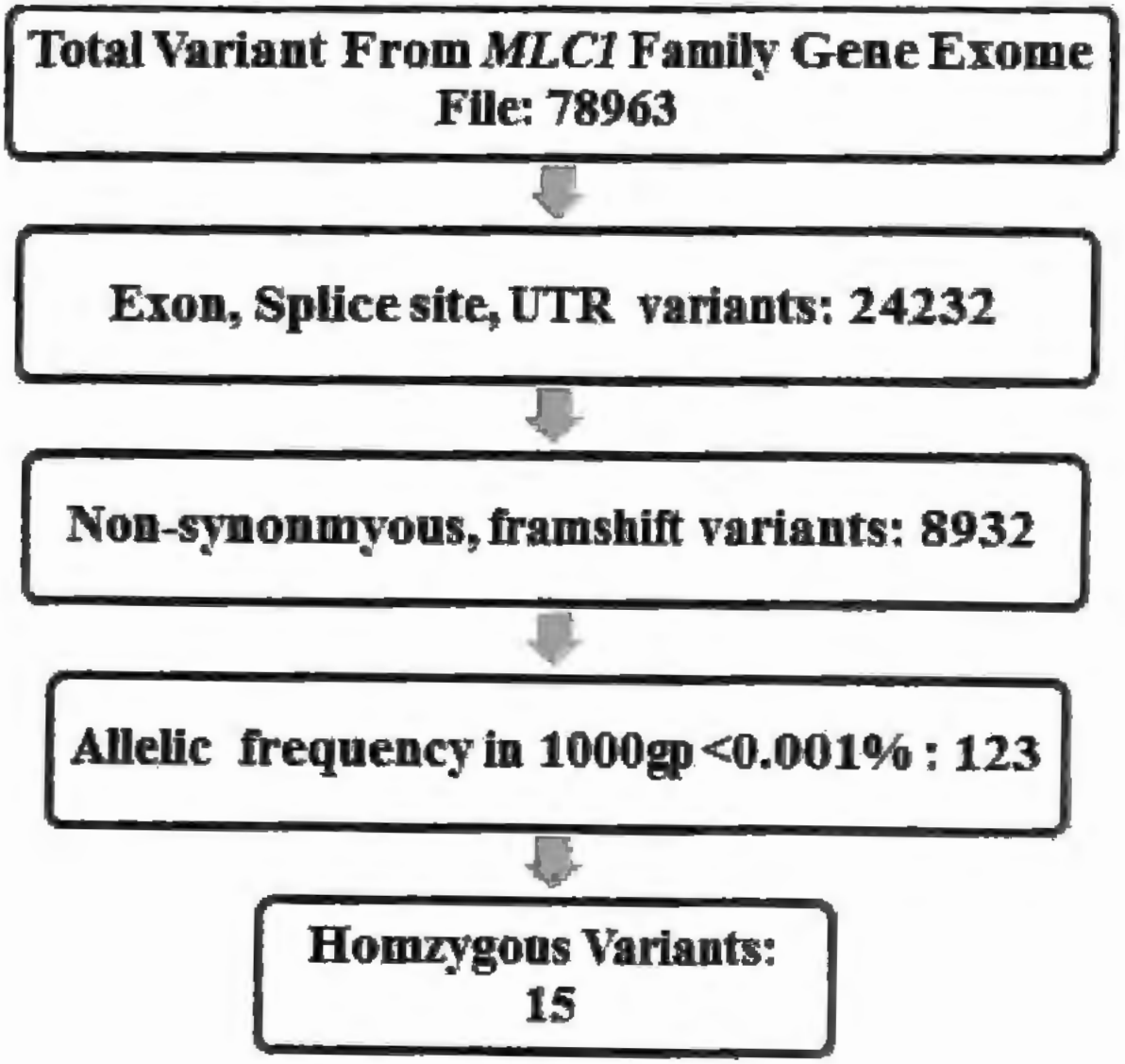


Figure 4.5.3: Method of filtering whole exome sequencing data of the individual (IV: 2). The Proband was filtered with 15 homozygous variants only.

Table 4.5.1: Summery of variants found in proband (IV: 2)

Gene	Ref. Seq. ID	cDNA Mutation	Protein sequences	dbSNP ID	GERP Score
RNF19B	NM_001127361	c.A185C	p.Q62P	rs113840389	0.30
ERICH2	NM_001002912	A2491C	p.I831L	rs537105429	-0.93
ERICH3	NM_001002912	c.G2490T	p.E830D	rs555839913	-1.01
ATN1	NM_001007026	c.C2782T	p.P928S	NA	3.79
MDM1	NM_001205028	c.C685T	p.R229C	rs150886543	2.64
EDNRB	NM_001122659	c.C214T	p.P72S	rs150750272	1.49
RNF219	NM_024546	c.A1777G	p.K593E	rs575630386	4.63
PLD2	NM_001243108	c.C91A	p.P31T	NA	2.49
PIRT	NM_001101387	c.G86C	p.S29T	rs561790761	5.46
PTPN4	NM_002830	c.G1094A	p.R365Q	rs532212706	4.22
SLC11A1	NM_000578	c.T806C	p.I269T	NA	3.95
MLC1	NM_015166	c. C959A	p. T320K	rs281875313	4
CCHCR1	NM_001105563	c.C361T	p.R121C	rs375819943	4.43
DNAH11	NM_001277115	c.G13500C	p.E4500D	NA	-0.81
PTCH1	NM_000264	c.G328A	p.A110T	rs199476091	3.12

4.5.5 Co-segregation Analysis

Sanger analysis conform one affected siblings (IV: 2) was homozygous and the other sibling was normal (IV: 3) for the *MLC1* variants, for sequencing result see Figure 4.5.4. This mutation c.C959A: p.T320K is novel variant not previously reported in databases, located in exon 11 of the *MLC1* gene, Ensemble transcript (ENST00000395876) is disease causing for this family.

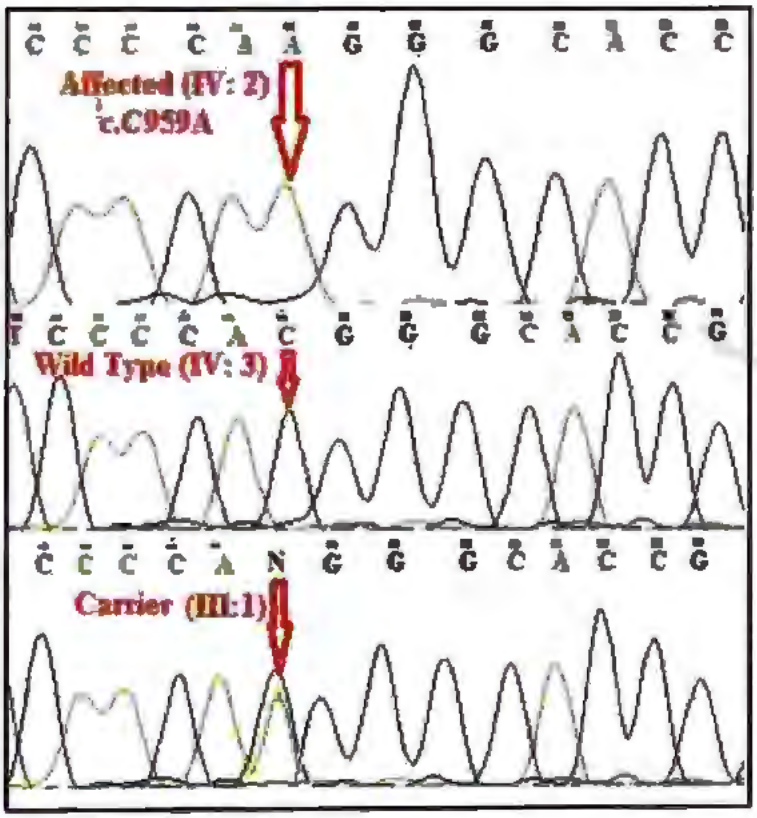


Figure 4.5.4: Chromatogram demonstrating segregation of mutation c. C959A in exon 11 of *MLC1* gene. Where one affected sibling (IV: 2) was sharing homozygous region while other sibling (IV: 3) is healthy.

4.5.6 In Silico Modeling of *MLC1*

Insilco analysis of the non-synonous mutation c. C959A; p. T320K by using Swiss modeling software (www.swiss-prot.org) discloses that the mutant amino acid effects protein structure of *MLC1* gene at the 320 amino acid locations. As the mutation is nonsynonymous cause change in 3D structure of *MLC1* protein. The normal structure (A) and mutant structure (B) both visible in Figure 4.5.5

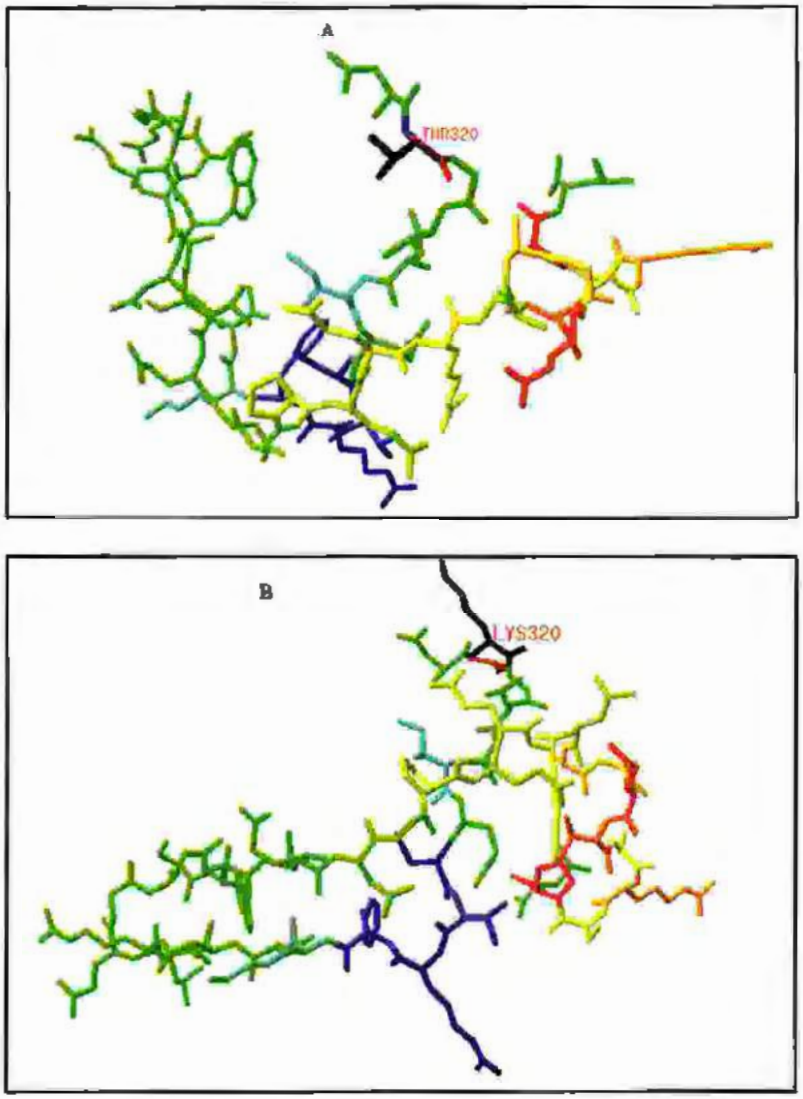


Figure 4.5.5: Swiss PDB viewer ribbon tertiary structure of *MLC1*. The diagram (A) is representing the wild type structure of *MLC1* protein. The diagram (B) is representing the mutated *MLC1* protein.

4.5.7 Discussion

Mutation in *MLC1* gene leads to the condition of megalencephalic leukoencephalopathy with subcortical cysts in brain (Van der Knaap *et al*., 1995). More than 80 mutations was reported in MLC1 gene so far. *MLC1* affect brain development and motor function in patients with movement and seizures problems (https://ghr.nlm.nih.gov/gene/MLC1#conditions). Mutations in *MLC1* are a known cause of megalencephalic leukoencephalopathy with subcortical cysts. In this study we analysed Pakistani consanguineous family with one affected patient. WES analysis was performed on the patient after data filtration we selected fifteen variants for di-deoxy sequencing to validate the co-segregation of the family to disease gene. We identified novel missense mutation in *MLC1* gene on (ENST00000311597:exon11:c.C959A:p.T320K), where amino acid threonine (ACG) was replaced with amino acid lysine (AAG) (figure 4.5.3). The c.C959A:p.T320K *MLC1* mutation has not been reported previously in public database like exac browser and gnomad. In-silico study of this mutation generates high scores of deleteriousness, combined annotation dependent depletion (CADD) score (30) suggest that variant is pathogenic (Rentzsch *et al*., 2018). GERP score for this mutation is (4); suggest that this gene is not tolerated mutation. Mutation Taster score conform that mutation is causing disease, SIFT and Polyphen2 score also conform mutation leading to disease. *MLC1* gene plays role in the synthesis of a protein that is highly expressing in brain. *MLC1* protein present in specialized type brain cells known as glial cells. These cells protect and maintain nerve cells in the brain. *MLC1* protein functions at junctions that connect glial cells. Glial cell adhesion molecules (GlialCAM) is a beta subunit of *MLC1*, interact with *MLC1* (Favre-Kontula *et al*., 2008; Lopez-Hernandez *et al*., 2011). GlialCAM act as escort molecule to bring *MLC1* molecule to cell-cell junction. Mutation in glialCAM cause trafficking defect of *MLC1*. it is not able to reach cell-cell junction without glialCAM (Lopez-Hernandez *et al*., 2011). Protein modeling of the wild type and mutated protein was performed (Figure 4.5.5). The risk of *MLC1* disease is increased due to the prevalence of consanguineous marriages. Molecular testing of *MLC1* gene is necessary to control the disease and prenatal diagnosis of the disease in kindred families is important.

Chapter 4.6: Homozygous Mutation in *SDCBP2* Gene Causing Intellectual Disability in large Consanguineous Pakistani Family.

4.6.1 Introduction

Syndecan binding protein (*SDCBP*) gene also known as syntenin is a multifunctional adapter protein involved in various range of functions, including trafficking of transmembrane proteins, cell adhesion, neural, and synaptic development, expressing in a different type of cells and involved in exosomes biogenesis (Dasgupta *et al.*, 2013; Friand *et al.*, 2015). *SDCBP* gene mapped on the chromosome 8p12 and contains 9 coding exons (Gross, 2014). This gene was first discovered when it binds to the cytoplasmic domains of the syndecans in experiment of yeast-2-hybrid screen for protein (Grootjans *et al.*, 1997). *SDCBP* protein is composed with 298 amino acid residues, this gene also contain two PDZ domains, PDZ1 domain start from the amino acid 114 to 193 and PDZ2 domain start from the amino acid 198 to 273 (Philley *et al.*, 2016). *SDCBP* has been localized in the cell to cell adhesion sites, microfilaments and in the nucleus. The elevated level of expressions of SDCBP/Syntenin was observed in fetal kidney, liver, and brain and lower level of expression were observed in all other tissues (Zimmermann et al., 2001). PDZ domain of *SDCBP* protein was act as an intracellular binding partner of glycine transporter 2. C-terminal amino acid residues of glycine transporter 2 and the PDZ2 domain of *SDCBP* interact to regulate the presynaptic localization of glycine transporter-2 to glycine releasing inhibitory interneurons known as glycine-rgic neurons (Ohno *et al.*, 2004). It is also involve in the formation and controlling of neuronal structure of membrane (Hirbec *et al.*, 2004). Trafficking of neuroglia cells along with exons is an important pre-requirement for myelination, but the mechanisms of oligodendrocyte trafficking is not fully understood yet. Proteoglycan NG2 is involved in oligodendrocyte migration. PDZ domain protein syntenin interacts with NG2 to facilitate the migration and that syntenin is necessary for regular transportation (Chatterjee *et al.*, 2008). Receptor Ephrin B belongs to enzyme tyrosine kinase family, play role in systematized postsynaptic development of the brain interacting to different receptor of glutamate. Receptors Ephrin B1 and B2 involved in presynaptic development via *SDCBP* (Mcclelland *et al.*, 2009). Neurological problems and disorders such as multiple sclerosis (MS) and stroke, blood brain dysfuctioning is very important event. The role of micro-RNA 155 was observed as negative regulator. Micro-RNA 155 modulated brain endothelial barrier function by targeting focal

adhesion components *SDCBP*/syntenin protein and it is also identified as a critical miRNA in neuroinflammation at the blood brain barrier (Lopez-Ramirez *et al.*, 2014). Here, we are reporting homozygous non-synonymous mutation causing neurological problems in consanguineous Pakistani family.

4.6.2 Background

The selected family (MR-9) was enrolled from the Distt hospital Dera Ghazi Khan, distt Dera Ghazi Khan Punjab Pakistan. Family pedigree was constructed based on the information provided by the parents. Research consent forms taken from family and affected children research participants forms obtained from guardians. Research work approved from ethical and (BASR) committee of International Islamic University (IIUI) Islamabad. Pedigree of the family of was shown in below Figure 4.6.1.

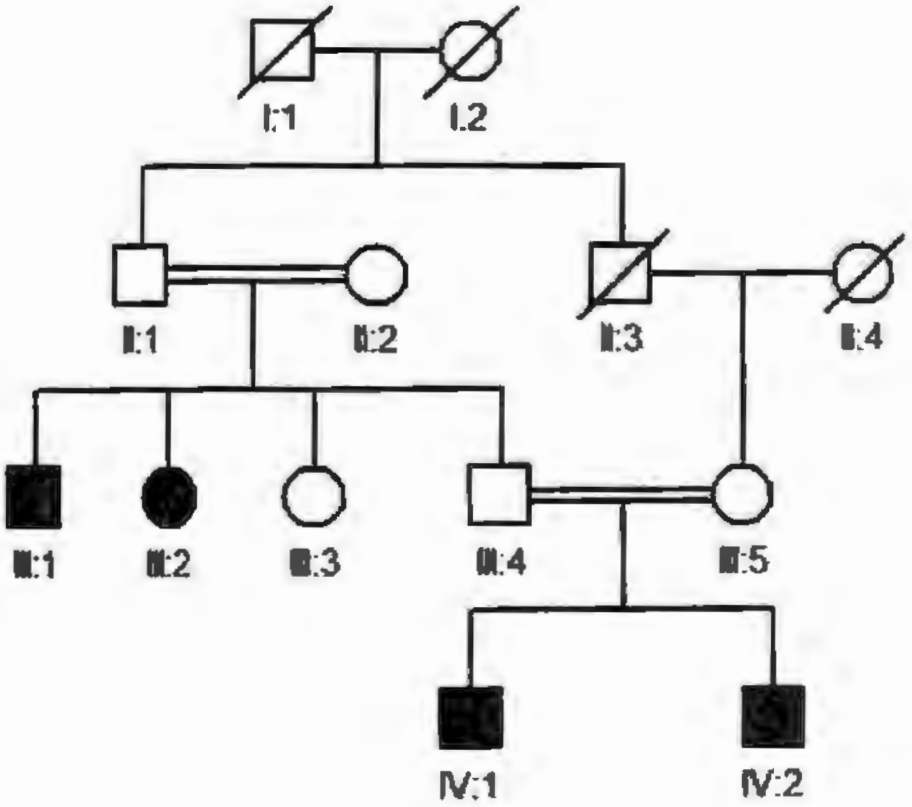


Figure 4.6.1: Family pedigree representing affected (III: 1, II: 2, IV: 1 and IV: 2). Pedigree squares mentioning the male individuals and circle indicates the female individuals in the family. All the filled square and circle designates the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals

4.6.3 Clinical Details

The nine member family (Figure 4.6.1) was sampled from rajan pur distt Punjab. In this family one female daughter (III: 4) was unaffected and she develop normal milestones. Four affected patients were present in this family (III: 1, III: 2, IV: 1 and IV: 2). Proband (III: 1) age was 30 years, with delayed milestones. In addition he can speak very late and had some facial dysmorphic features. Clinical analysis exhibited his IQ level was very low (ID: 25-35). He had no respiratory, genital and other problem. The proband (III: 2) was 20 years old female, with developmental delay, behaviors problems. She can speak only few words. Clinical examination was not performed on this patient. The proband (IV: 1) was 9 years old when he brought for medical examination. He developed 2-3 seizures per day at the age of 2 years up to 4-5 minutes of duration. Detailed clinical investigation suggests he has avoidance and aggressive behavior and not able to talk. He has normal posture, normal fundoscopy, no involuntary movement, no hepatosplenmegaly and cardiac involvement. Brain CT Scan demonstrated gray matter more marked in basal ganglia. He has also some teeth anomalies.

The proband (IV: 2) age was 7 years, when he was brought in the hospital for examination. He has developmental delay and delayed milestones. He has aggressive behavior, no facial dysmorphic feature, head size is normal. He has normal posture, normal fundoscopy, no cerebellar signs, no involuntary movement, no hepatosplenmegaly and cardiac involvement. He starts walking 14 months of age. His blood count test and biochemical tests were correct. Brain MRI demonstrated large ventricles than normal, no other symptoms appeared during the analysis.

4.6.4 Genetic Analysis

4.6.6.1 Linkage Analysis

DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.6.6.2 Whole Exome Analysis

To identify underlying genetic causes the family two (III: 1 and IV: 2) probands was selected for whole exome sequencing. WES results of the probands (III: 1 and IV: 2) yields 86544 and 86123 variants respectively with read depth around 100x covering all intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac data base (www.exac.broadinstitute). The variants with allelic frequencies <0.001% were shortlisted figure 4.6.3. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for high level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Filtered variants also searched in PubMed and OMIM (https://www.omim.org/), data base for any other disease association and phenotypes. All 6 variants putative score were checked in mutation taster (http://www.mutationtaster.org/), GERP data bases. CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only 6 variants met these filtration criteria in both probands (III: 1 and IV: 2). Out of these variants 4 variants were non-synonymous and one was frame shift variants mentioned in table 4.6.1. Off these filtered variants only SDCBP2 gene variant co-segregating with family probands, SDCBP2 variant was present in 1000 genome project data base but before this no homozygosity was reported with this specific variant. Genomic evolutionary rate profiling (GREPP) score of

SDCBP1 gene variant considered as highly conserved. Mutation taster score for this variant is also considered as deleterious.

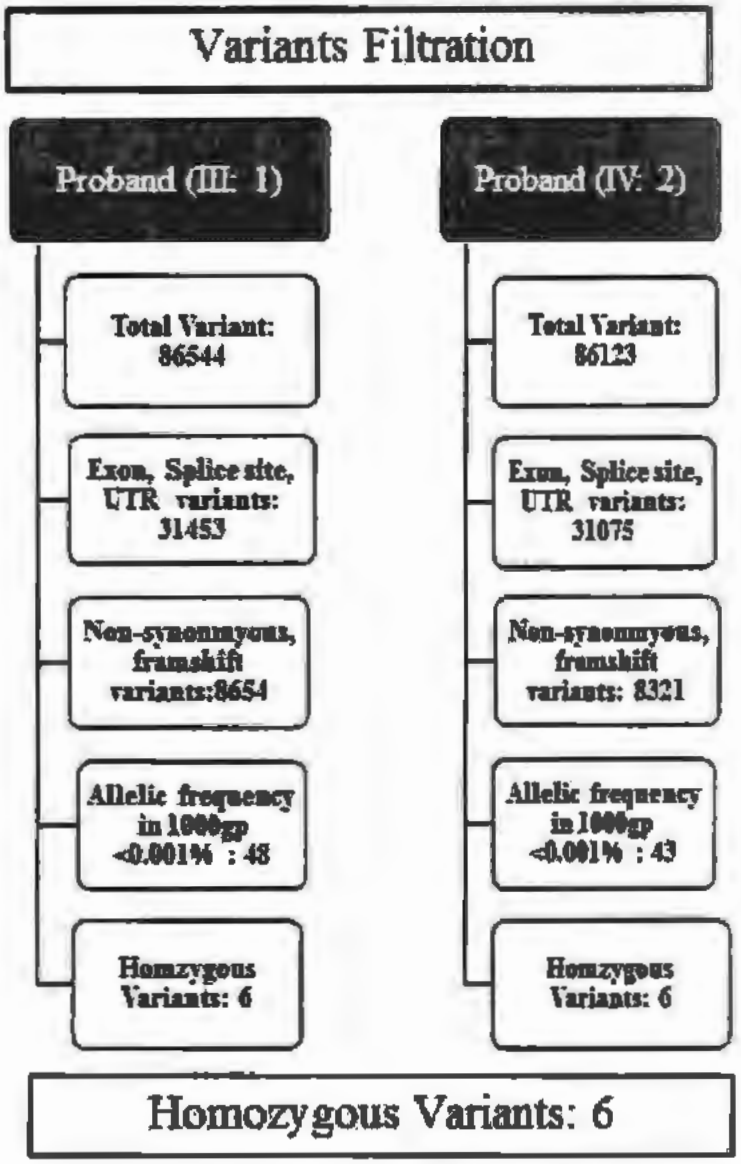


Figure 4.6.3: Probands (III: 1 and IV: 2), WES analysis results, both the probands were sharing 6 homozygous variants in this family.

Table 4.6.1: Summery of variants found in proband (III: 1 and IV: 2)

Gene	Ref. Seq. ID	cDNA Mutation	Protein	dbSNP ID	GERP Score
MLLT10	NM_001195626	c.G1587C	p.Q529H	NA	-6.61
DAOA	NM_172370	c.G178A	p.G60R	rs140575409	-0.598
SDCBP2	NM_015685	c.G556T	p.D186Y	rs201595244	4.51
IRX1	NM_024337	c.C92T	p.A31V	NA	2.41
SLC41A3	NM_001008486	c.A841G	p.I281V	NA	-9.07
LNP1	NM_001085451	c.194_195insTC	p.S80_H81insSD	rs71132521	NA

4.6.5 Co-Segregation analysis

Sanger analysis conform all affected probands (III: 1, III: 2, IV: 1 and IV: 2) were homozygous, female proband (III: 3) is normal for the *SDCBP2* variants, for sequencing result see figure (4.6.4). The variant c.G556T: p.D186Y is pathogenic variant on the bases of bioinformatics prediction scores and pathogenicity was further conformed through co-segregation analysis.

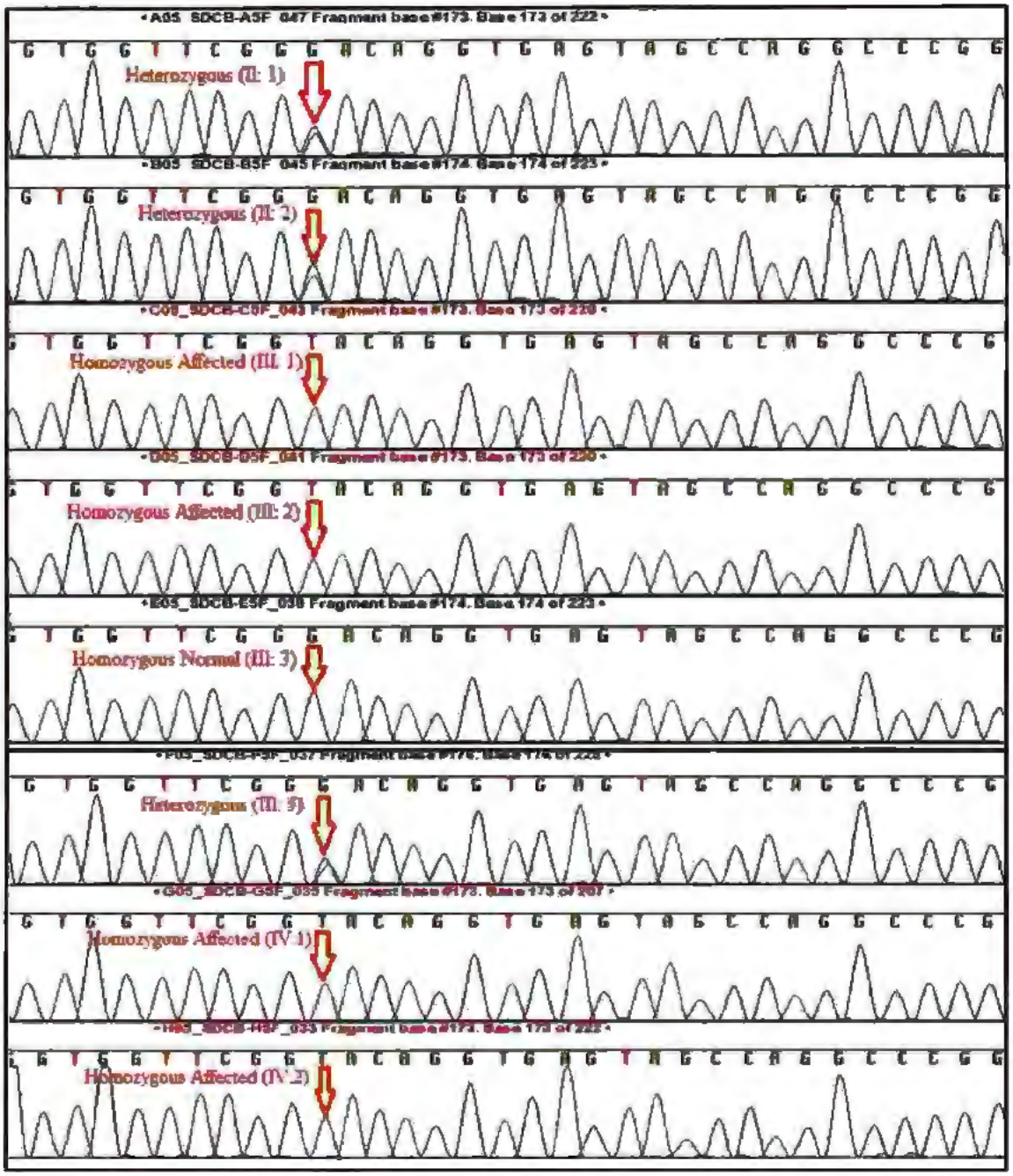


Figure 4.6.4: Chromatogram demonstrating segregation of mutation c.G556T; p.D186Y in exon 6 of *SDCBP2*. Affected patients were sharing homozygous region while parents were heterozygous. The female WT (III: 3) was normal.

4.6.6 In-Silico Analysis

Insilco analysis of the non-synomous mutation c.G556T: p.D186Y by using Swiss modeling software (www.swiss-prot.org), showed that the mutant amino acid effects protein structure of *SDCBP2* gene. The normal structure (A) and mutant structure (B) both shown in figure 4.6.5.

Figure 4.6.5: Swiss PDB viewer ribbon tertiary structure of *SDCBP2*. The diagram (A) is representing the wild type structure of *SDCBP2* protein. The diagram (B) is representing the mutated *SDCBP2* protein.

4.6.6.1 Protein-Protein Interaction

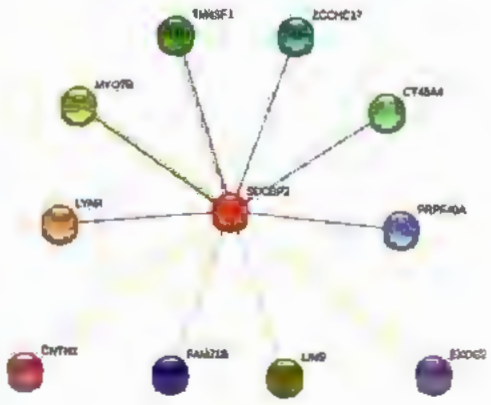


Figure 4.6.6: Syntenin-2 play a role in the cell division and cell survival and interact with contactin-associated protein 2 gene associated with Pitt-Hopkins like syndrome.

4.6.7 Discussion

Here, we are reporting a missense mutation in the syndecan binding protein gene (*SDCBP2*), in Pakistani consanguineous family. The clinical phenotypes of patients are development delay, speech problem, facial dysmorphic features and seizures. *SDCBP2* genes is belongs to the syndecan binding protein gene family. *SDCBP2* gene is multi-functional involved in cell- cell adhesions, exome biogenesis, neurotransmitter, neural, and synaptic development. Two patients were selected for exome sequencing, the data analysis of exome sequencing revealed six variants sharing homozygous variants (Table 4.6.1). All the variants score were checked by using bioinformatics mutation prediction tools. The mutation c.G556T: p.D186Y was present in exac browser, but no homozygosity was reported for this specific mutation in exac data base and genome aggregation database (gnomAD). That mutation is presented in highly conserved location on the gene according to the genomic evolutionary rate profiling (GERP) prediction score (4.5). In silico evaluation of the mutation prediction through, SIFT, Mutation Taster, Polyphen2 and Mutation Taster as 'deleterious' and 'disease causing', accordingly. After exome filtrations and bioinformatics analysis De-deoxy sequencing method used to conform co-segregation of this specific mutation to probands. Sanger sequencing conform the segregation all probands sharing homozygous alleles while parents are heterozygous and healthier child was showing normal allele (Figure 4.6.5). One hundred control samples also selected from this ethnic group to check whether the variant present in 100 control sample or not, control samples tested negative for this variant as well. More over *SDCBP2*/syntenin gene is involved in neuronal and pre synaptic development, mutation in this gene may leads to the dysfuctioning of neuronal and pre synaptic development that causing the disease in this family. *SDCBP2*/syntenin genes contain two PDZ domains, PDZ1 domain and PDZ2 domains. *SDCBP* PDZ-domain 1 and 2 interacts with multiple proteins including EphrinB1 and B2, glutamate receptor, neurofascin a multiple proteins and has several intracellular functions (McClelland *et al.*, 2009; Beekman and Coffer, 2008). Receptor ephrinB1 (EphB) is coordinate with glutamate receptor to directs the post synaptic development, control dendritic filopidial and mediate spinal formation. Ephrin receptor is play role in the formation of pre-synapse with SDCBP2 association (Mcclelland et al., 2009). By using web base bioinformatics tool STRING we check the protein-protein interaction of *SDCBP2* gene. The protein-protein interaction (Figure 4.6.6) of SDCBP2 revealed that this gene is interacting with *CNTNAP2* gene that contributes to the association of axonal domains at nodes

of ranvier by maintaining voltage gated potassium channels at the juxtaparanodal region. We believe this gene is disease causing for this family because of its diverse function and its interaction with *CNTNAP2* gene that is associated with pitt-hopkins like syndrome 1 and autism susceptibility 15. Functional work also will be performed for the conformation of this gene pathogenic nature in this family. Moreover, other probands and families with similar variants c.G556T: p.D186Y helps to clarify whether or not it is pathogenic. MLLT10 gene variant (NM_001195626: c.G1587C; p.Q529H) Sanger sequencing will be performed for co-segregation analysis.

Chapter 4.7: Start Loss Mutation Causing Progressive Cerebello-Cerebral Atrophy in Autosomal Recessive Intellectual disability Families

In this Chapter, two families (MR-10A and B) with congenital intellectual disability were selected for genetic analysis. The detail depiction of these families is explained below.

4.7.1 Introduction

Progressive cerebello-cerebral atrophy *PCCA* (MIM: 613009) is an autosomal recessive disorder of profound intellectual disability, spasticity, and microcephaly (Ben-Zeev *et al.*, 2003). *PCCA* type 1 is caused by the mutation in the phosphoserine-tRNA-selenocysteine (*SepSecS)* gene (Hady-Cohen *et al.*, 2018). *SepSecS* was the first gene reported to cause PCCA (Agamy *et al.*, 2010). *SepSecS* is a ubiquitously expressed enzyme involved in the formation of tRNA that enhances the formation of selenocysteine molecule in final reaction (Puppala *et al.*, 2011). The trace element selenium (Se) is found in proteins as selenocysteine (21st amino acid) function as an essential micronutrient in the diet. Selenocysteine is distinct from other amino acids and synthesized on the cognate tRNA in all domains of life (Palioura *et al.*, 2009). The selenocysteine biosynthesis mechanism in humans has been interpreted (Yuan *et al.*, 2006), *SepSecS* enzyme in complex with the tRNA molecule transfer selenocysteine to the $tRNA^{Sec}$ and $tRNA^{Sec}$ provided details insights into the reaction mechanism (Palioura *et al.*, 2009). Selenocysteine biosynthesis is a tree step biosynthesis process. In first step seryl-tRNA synthetase attach serine molecule to $tRNA^{Sec}$ in reaction. In second step Ser-$tRNA^{Sec}$ converted into o-phosphoseryl tRNA in the presence of co-factor pyridoxal phosphate. In the final step of reaction O-phosphoseryl tRNA converted into Sep-$tRNA^{Sec}$ in the presence of enzyme *SepSecS* that converted into selenocysteine protein (Figure: 4.7.1) (Agamy *et al.*, 2010).

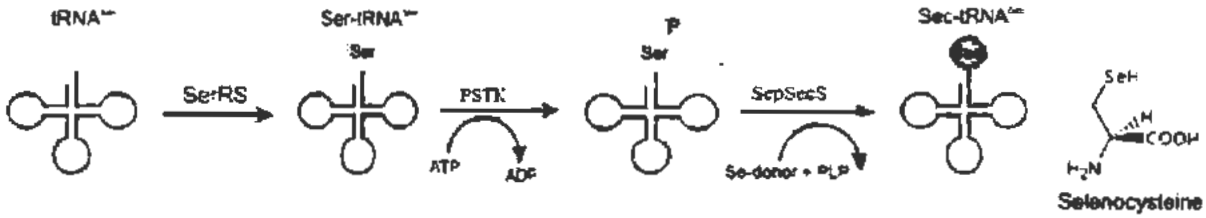


Figure 4.7.1: adopted from (Agamy *et al.*, 2010). Three step reaction biosynthesis of selenocysteine protein. Selenocysteine is distinctive amino acid and synthesized on the cognate tRNA.

Previously research groups reported that mutations in human sepsecs gene caused progressive cerebello-cerebral atrophy (*PCCA*) and term also associated with pontocerebellar hypoplasia type 2 (*PCH2*). But recently study reported that *PCH1* caused by the mutation in the sepsecs gene and *PCH2* is caused by the mutation in the *VPS53* gene (Hady-Cohen *et al.*, 2018). Total 6 mutations have been reported in the sepsecs gene, 4 were missense, one splicing and one deletion mutation (http://www.hgmd.cf.ac.uk/ac/gene.php?gene=SEPSECS). Whole exome sequencing was performed on two Pakistani families with similar phenotypes. WES sequencing revealed a start codon mutation c.T2C, p.M1T in the *SepSecS* gene in two Pakistani families.

4.7.2 Family A

The *SepSecS* family A was enrolled from the District hospital Dera Ghazi khan, Punjab Pakistan. Family pedigree was constructed on the bases of information provided by blood relatives. Blood samples were collected from the family members both affected and normal. Written consent taken from members and affected child informed consent was also taken from their parents. The study was approved from the (BASR) Board of Advanced Study and Research and ethical committee of International Islamic university Islamabad. Pedigree of the family of was shown in (Figure 4.7.2).

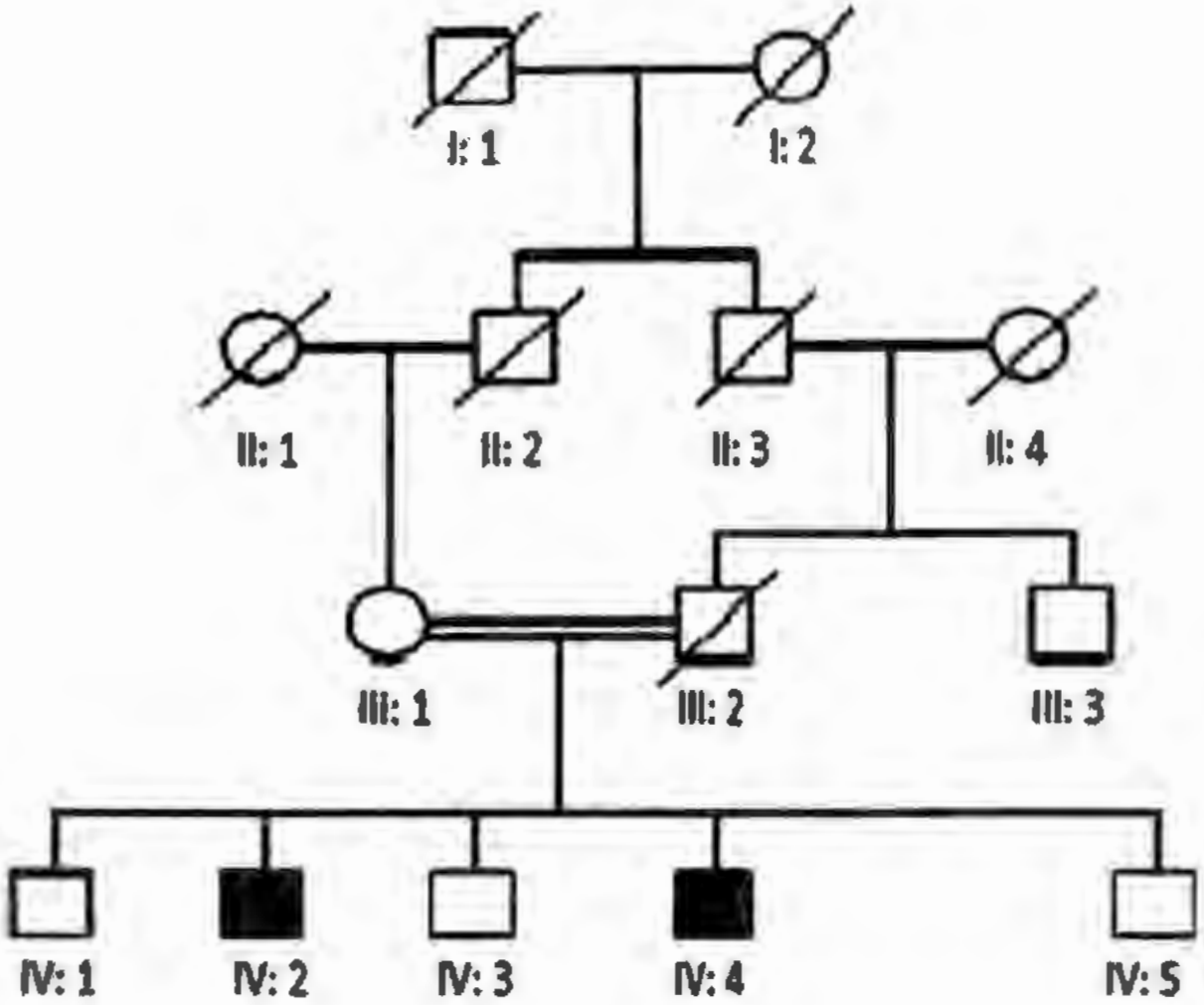


Figure 4.7.2: Family pedigree representing the two probands (IV: 2 and IV: 4). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals.

4.7.3 Clinical Details

The *SepSecS* family A having 2 affected male (IV-2 and IV-4) with cognitive impairment. Both patient IQ has very low. The proband (IV: 2) 26 years old when brought of clinical examination. Medical history of patient (IV: 2) revealed when he was 9 month old, he developed seizures. Type of seizure considered as generalized tonic seizure (GT). Detailed clinical investigation suggests he has feeding problem, not able to talk. He presented dysmorphic facial feature with broad nasal root and pointed nose, pointed chin, long narrow face and slight webbing of the neck {Figure 4.7.3 (IV: 2)}. He was diagnosed with familial intellectual disability, developmental delay by the local neurologist. The proband (IV: 4) age was 23 years at the time of clinical examination. At the 18 months of age, he developed maximum 3 seizures per day up to 6-5 minutes of duration. Type of seizure considered as GTC. During the clinical investigation his behavior was aggressive, he presented facial dysmorphic feature including, long face, pointed nose, and broad chin {Figure 4.7.3 (IV: 4)}. He has normal fundoscopy, no hepatosplenmegaly and cardiac involvement. He has no speech with non-responsive cognition. He was diagnosed with familial intellectual disability, developmental delay by the local neurologist. Electroencephalogram test demonstrated progressive slowing of the background in both the probands (IV: 2 and IV: 4). Brain CT Scan revealed progressive a trophy, cerebral white matter and decreased level of choline in the cerebellar hemispheres in both patients (Data not shown here).

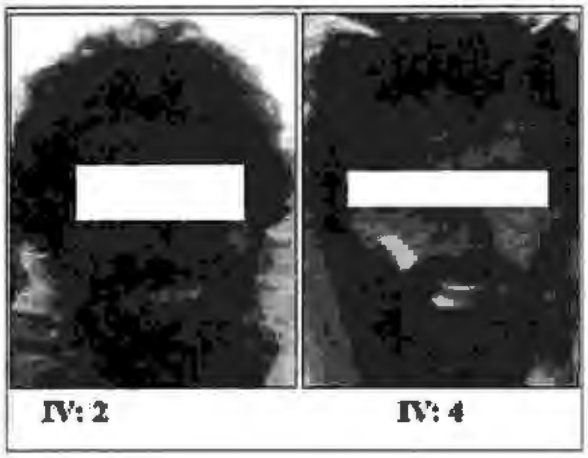


Figure 4.7.3: The probands (IV: 2) and (IV: 4) have aggressive behavior, developmental delay problem, facial dysmorphic feature including, long face, pointed nose, and broad chin.

4.7.4 Genetic Analysis

4.7.4.1 Linkage Analysis

DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.7.4.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 2 and IV: 4) come up with 133672 and 133952 variants respectively with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac data base (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 4.7.3. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for elevated level of expression of variants in brain both at RNA and protein level as compared to other organs or tissues. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All five variants putative score were checked in mutation taster (http://www.mutationtaster.org/). GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only five variants met these filtration criteria in both probands (IV: 2 and IV: 4). Out of these variants three variants were non-synonymous and two were frame shift variants mentioned in table 4.7.1. Of these filtered variants only *SepSecS* gene involved with profound intellectual disability, progressive cerebello-cerebral atrophy considered as neurodegenerative disorder. Two *SepSecS* variant was present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of *SepSecS* gene variant considered as

highly conserved. Mutation taster score for this variant is also considered as deleterious. *SepSec5* variant selected for Sanger sequencing.

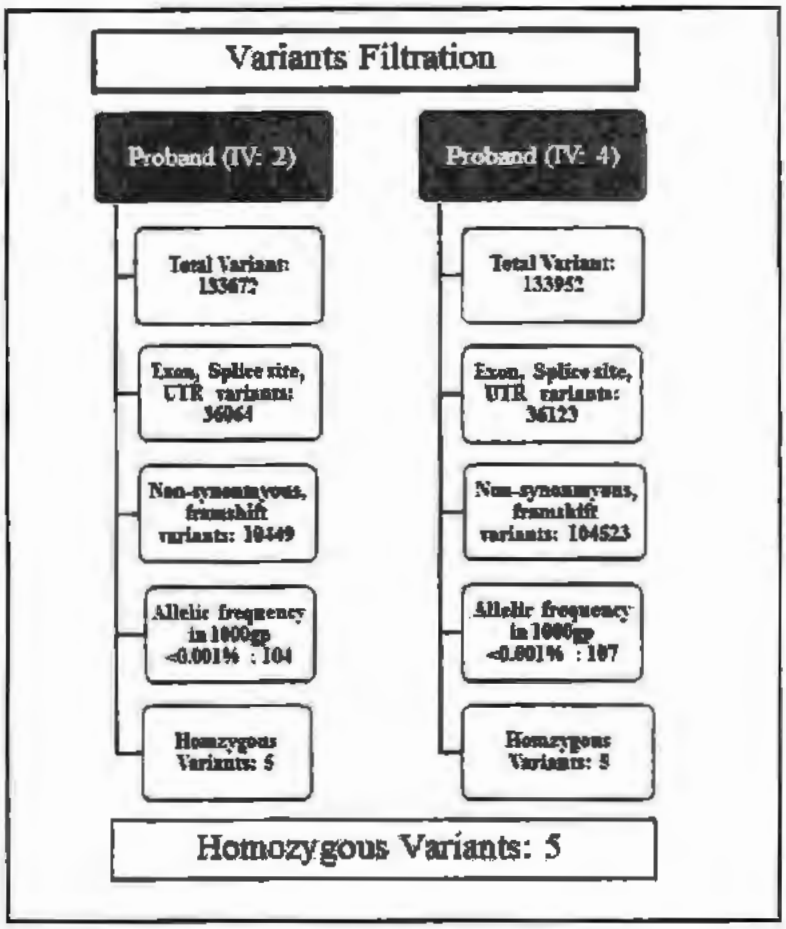


Figure 4.7.4: WES data filtration and analysis of the probands(IV: 2 and IV: 4), both the individual was sharing only five homozygous variants with allelic frequency less than 0.001%.

Table 4.7.1: Summery of short listed variants in patients (IV: 2 and IV: 4)

Gene	Ref. Seq. ID	cDNA Mutation	Protein	dbSNP ID	GERP Score
SEPSECS	NM_016955	c.T2C	p.M1T	rs532571185	NA
ARAP2	NM_015230	c.T4716G	p.N1572K	NA	5.63
UTP3	NM_020368	c.C182G	p.A61G	NA	4.45
PGAP2	NM_001346403	c.T615A	p.C205S	rs544741456	1.92
PKD2	NM_000297	c.A652C	p.M218L	rs2234917	3
CCSER1	NM_001145065	c.C301A	p.Q101K	rs543876065	4.88

4.7.5 Co-Segregation Analysis

Sequencing analysis conform both affected siblings (IV: 2 and IV: 4) were homozygous, three brother were heterozygous (IV: 1, IV: 3 and IV: 5) and their unaffected mother (III: 1) were heterozygous for the *SepSecS* variants, for sequencing result (Figure 4.7.5). The mutation c.T2C; p.M1T is variant that disturbing start codon, located in exon 1 of the *SepSecS* gene, Ensemble transcript (ENST00000302922) was a disease causing variant for this family.

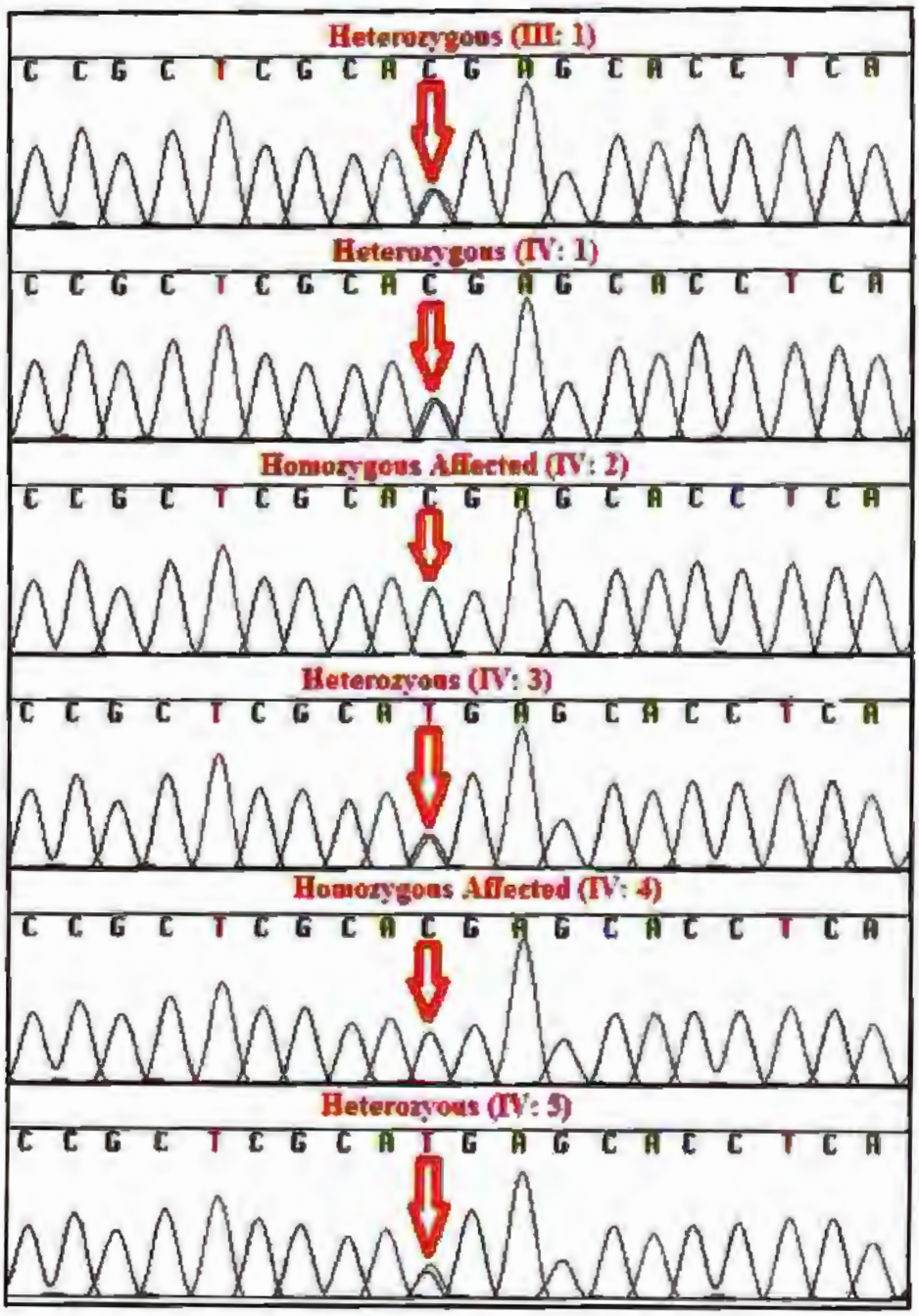


Figure 4.7.5: Chromatogram demonstrating segregation of mutation c.T2C; p.M1T in exon 1 of *SepSecS* gene. Where both affected sibling (IV: 2 and IV: 4) are sharing homozygous region while mother is heterozygous. Three other brothers (IV: 1, IV: 3 and IV: 5) are also heterozygous.

4.7.6 Bioinformatics Analysis

Insilco analysis was representing the 3D structure of the *SepSecS* gene by using Swiss modeling software (www.swiss-prot.org) Figure 4.7.6.

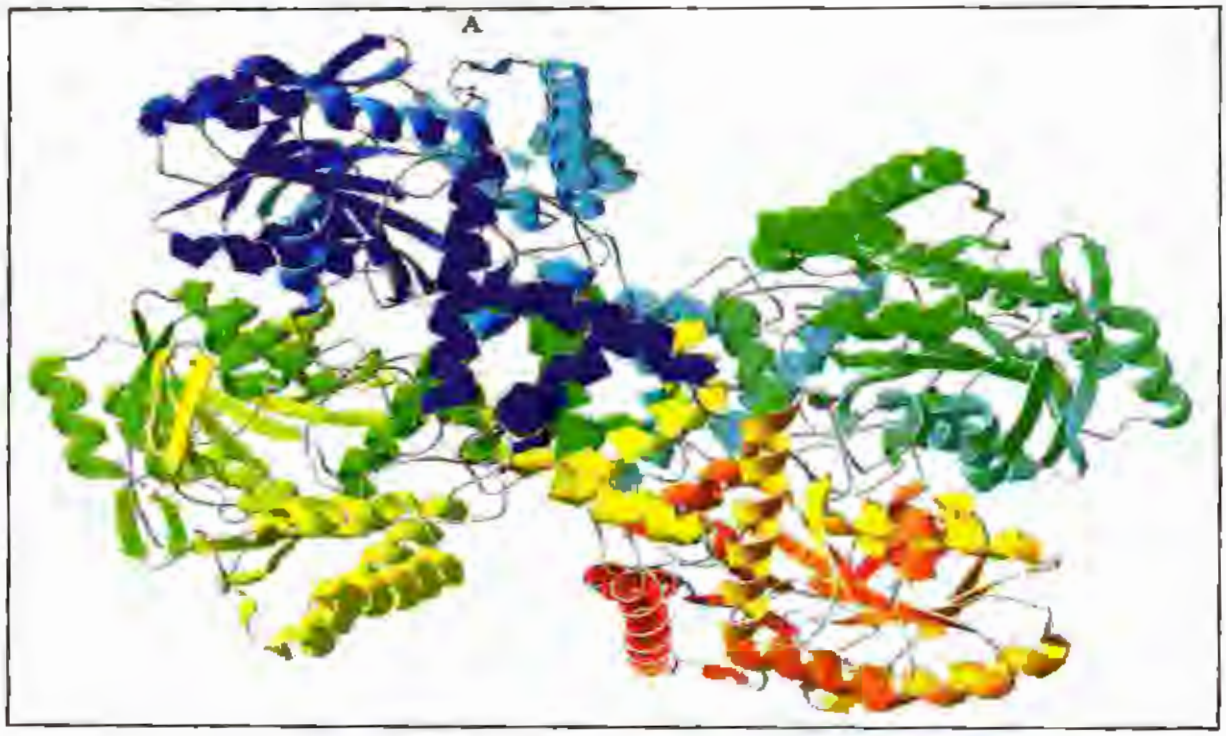


Figure 4.7.6: Swiss PDB viewer ribbon tertiary structure of *SepSecS*. The diagram (A) is representing the wild type structure of *SepSecS* protein.

4.7.7 Family B

The *SepSecS* family B was enrolled from the District hospital Dera Ghazi khan, Punjab Pakistan. Family pedigree was constructed on the bases of information provided by blood relatives. Blood samples were collected from the family members both affected and normal. The informed consent was obtained from all the members of the family and affected child informed consent was also taken from their parents. This research work approved from university ethical and BASR committee of IIUI.

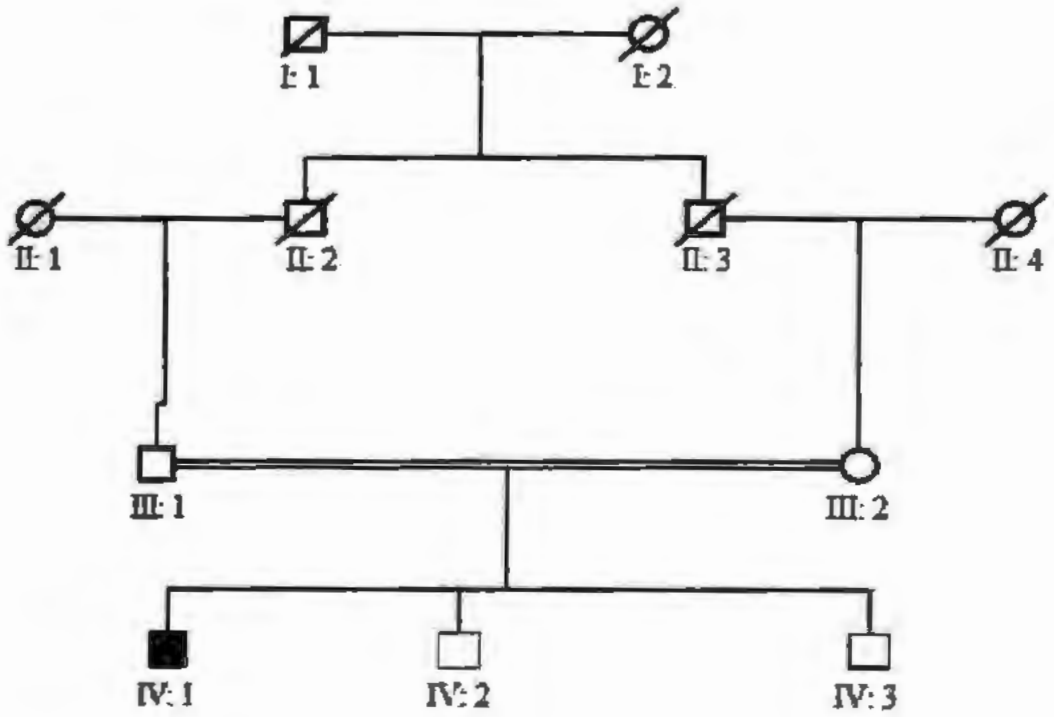


Figure 4.7.7: Family pedigree representing the one probands (IV: 1). Pedigree squares symbolizing the male individuals and circle denotes the female individuals in the family. All the filled square and circle denotes the probands and double line specifies consanguineous marriage. The crossed line on the square and circle denotes deceased individuals.

4.7.8 Clinical Details

The *SepSecS* family B has one affected child (IV-1) with cognitive impairment and macrocephaly. IQ level of the patient (IV: 1) was very low and considered as profound ID (rang: 20-25). The proband (IV: 1) was 9 year old when admitted in hospital for examination. At 9 month of age, he developed generalized tonic seizure (GT). Detailed clinical investigation suggests he has feeding problem, able to talk just few words. He presented dysmorphic facial feature with broad nasal root and pointed nose, pointed chin, large head size (Figure 4.7.8A). He was diagnosed with familial intellectual disability, developmental delay. He has normal fundoscopy, no hepatosplenmegaly and cardiac involvement. He has able to speak just few words with non-responsive cognition. Electroencephalogram test demonstrated progressive slowing of the background in the proband (IV: 1). Brain CT Scan revealed progressive a trophy, cerebral white matter in the cerebellar hemispheres in patients (IV: 1) (Figure 4.7.8B).

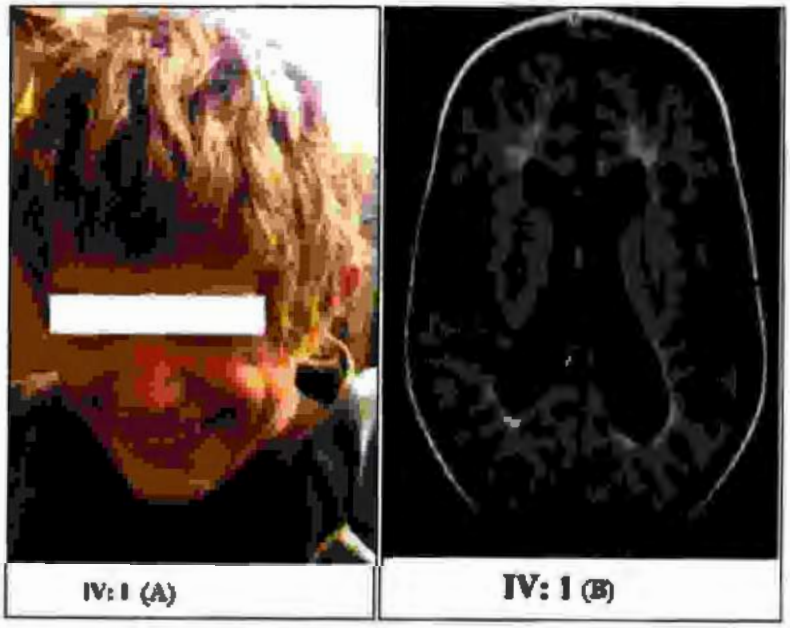


Figure 4.7.8: (A) The proband (IV: 1) phenotypes demonstrated dysmorphic facial feature with broad nose. (B). Brain CT Scan revealed progressive a trophy, cerebral white matter in the cerebellar hemispheres in patients

4.7.8 Genetic Analysis

4.7.8.1 Linkage Analysis

DNA from all affected family members as well as unaffected family members were analyzed. Homozygous genetics variants were hypothesized on the bases of pedigree. Intellectual disability knows as heterogeneous group of genetic disorder in which many genes are involved. Linkage analysis was performed using STR markers for most prevent genes causing intellectual disability in Pakistani population. Subjected family shows no linkage with targeted gene by using STR markers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.7.8.2 Whole Exome Analysis

To identify underlying genetic causes the family probands selected for whole exome sequencing. WES results of the probands (IV: 1) come up with 89561 variants with read depth around 100x including intronic and exonic variants. Initially all intronic variants were removed. Exome analysis targeted homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac data base (www.exac.broadinstitute.org). The variants with allelic frequencies <0.001% were shortlisted figure 4.7.9. For further conformation filtered variants were checked in the human protein atlas data base (https://www.proteinatlas.org/) for high level of expression of variants in brain both at RNA and protein level. Selected variants also searched in OMIM (https://www.omim.org/), data base for any other disease association. All eight variants putative score were checked in mutation taster (http://www.mutationtaster.org/). GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP). Only eight variants met these filtration criteria in proband (IV: 1). Out of these variants five variants were non-synonymous and two were frame shift variants mentioned in Table 4.7.2. Of these filtered variants only *SepSecS* gene involved with profound intellectual disability, progressive cerebello-cerebral atrophy considered as neurodegenerative disorder. Only two *SepSecS* variant was present in 1000 genome project data base. Genomic evolutionary rate profiling (GERP) score of *SepSecS* gene variant considered as

highly conserved. Mutation taster score for this variant is also considered as deleterious. *SepSec8* variant selected for Sanger sequencing.

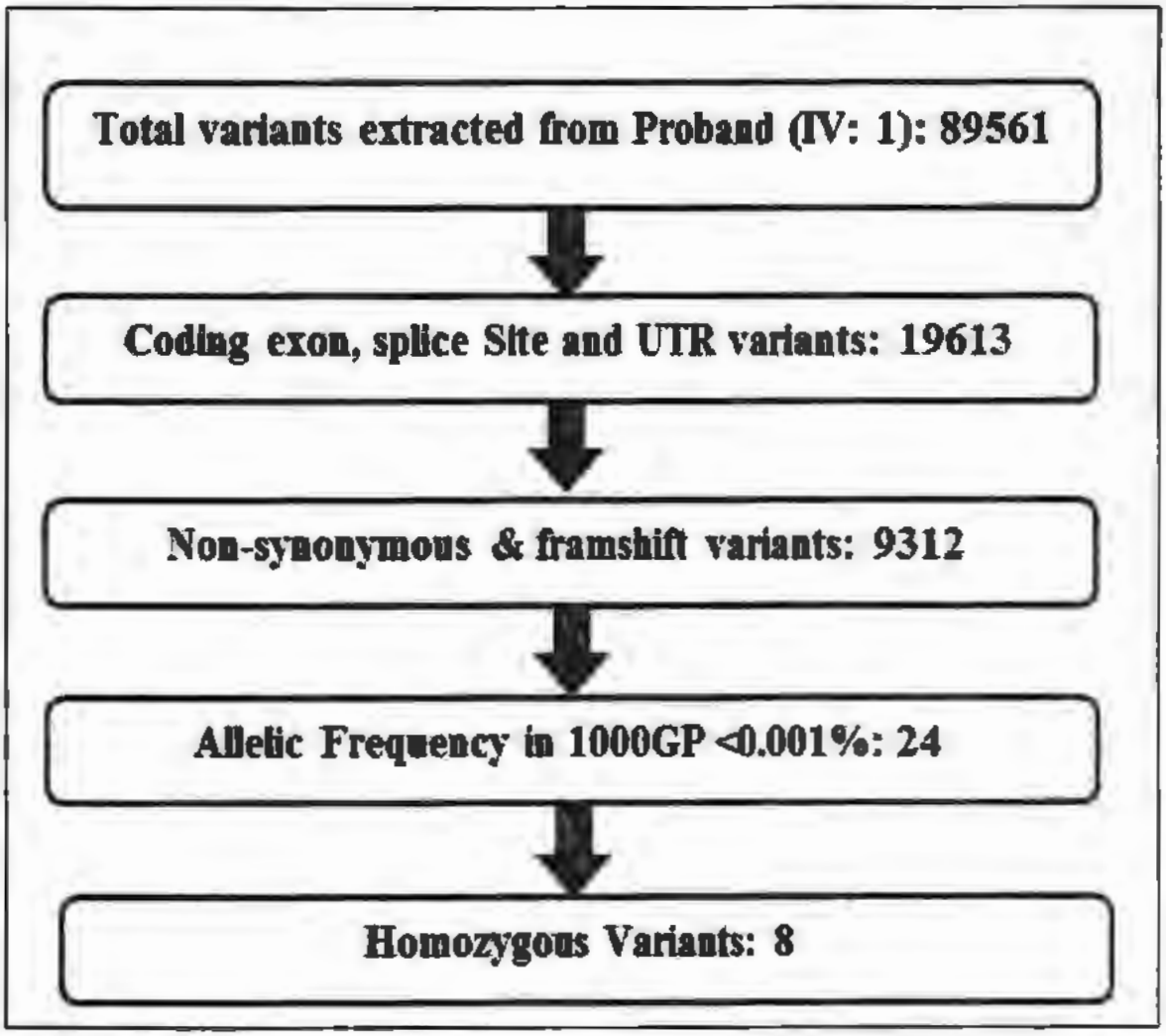


Figure 4.7.9: Method of filtering whole exome sequencing data of the individual (IV: 1). Both the individual was sharing only 8 homozygous variants with Allelic frequency less than 0.001%.

Table 4.7.2: Summery of short listed variants in family

Gene	Ref. Seq. ID	cDNA Mutation	Protein	dbSNP ID	GERP Score
ABCC2	NM_000392	c.T2508G	p.N836K	NA	1.92
PKD2L1	NM_001253837	c.G2246A	p.R749H	rs142439136	-6.61
NUDT2	NM_001244390	c.407_409del	p.E137Gfs*11	N/A	N/A
SEPSECS	NM_016955	c.T2C	p.M1T	rs532571185	NA
FADS6	NM_17812830	c.17_18insGA	p.P15_A16insTE	NA	NA
LNP1	NM_001085451	c.194_195insTC	p.S80_H81insSDRL	rs71132521	NA
TOP1MT	NM_000297	c.G1076C	p.R359P	rs150897789	2.31
EEF1D	NM_001130053	c.G883A	p.D295N	rs138059210	2.68
ZNF16	NM_006958	c.C917T	p.S306L	rs61732864	3.12

4.7.9 Co-Segregation Analysis

Sequencing analysis conform affected siblings (IV: 1) was homozygous, one brother was normal (IV: 2) and other brother (IV: 3) was heterozygous, father and mother both heterozygous, for sequencing result see figure (4.7.10). The mutation c.T2C: p.M1T is the mutation that disturbing start codon, located in exon 1 of the *SepSecS* gene, Ensemble transcript (ENST00000302922) was a disease causing variant for this family.

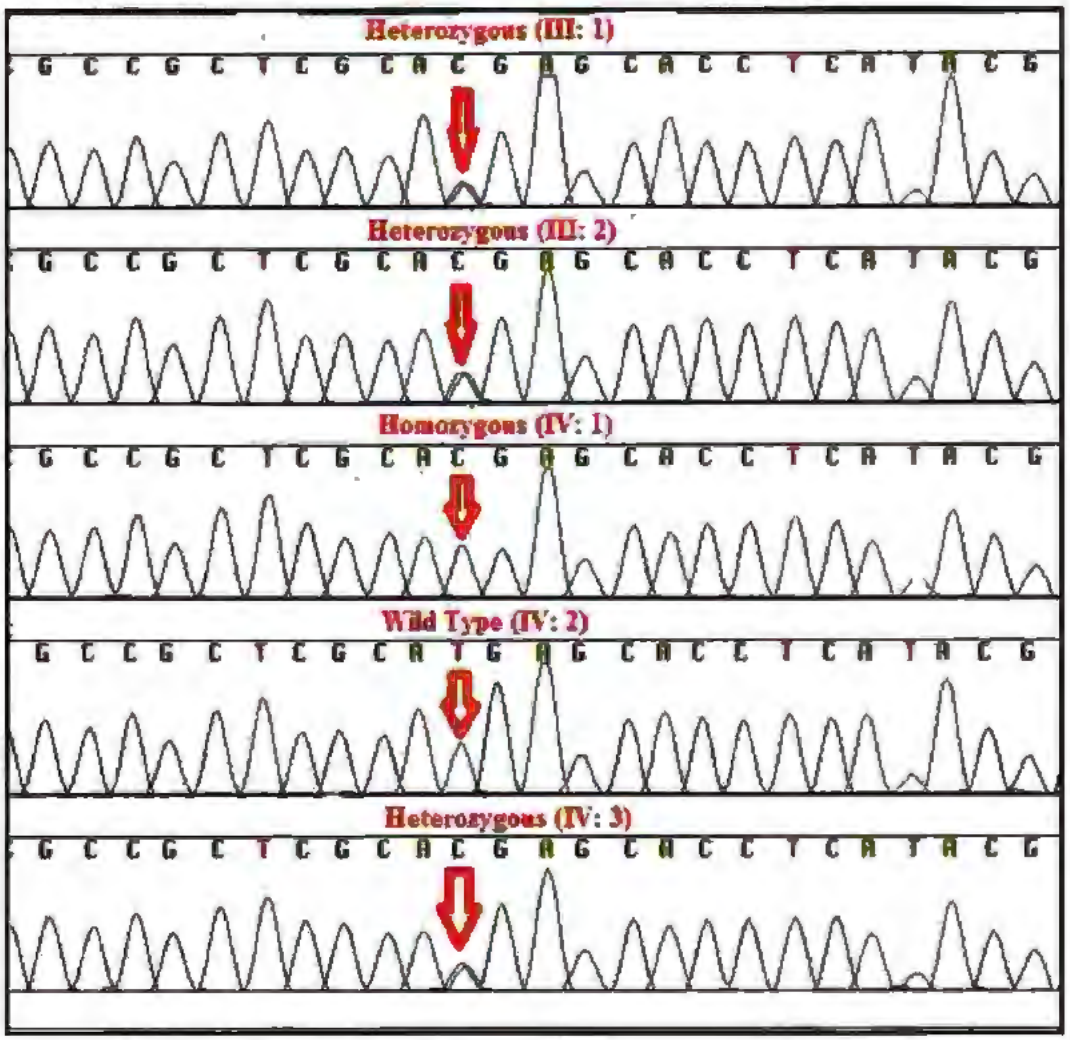


Figure 4.7.10: Chromatogram demonstrating segregation of mutation c.T2C: p.M1T in exon 1 of *SepSecS* gene. Where one affected sibling (IV: 1) was sharing homozygous region while one brother is normal (IV: 2) and the other is heterozygous. Parents were heterozygous (III: 1 and III: 2).

4.7.10 Discussion

Mutation in the *SepSecS* gene (OMIM: 163811) causing autosomal recessive ID and the mechanism of the *SepSecS* gene is well understood: sepsecs gene provides instruction for making sepsecs enzyme. The sepsecs enzyme is participates in the formation of tRNA molecule. This tRNA molecules play role in the formation of selenocysteine (sec) (Yuan *et al.*, 2006). Complex formation between *SepSecS* gene with tRNA molecules provide key insights into the biosynthesis pathway of selenocysteine protein (Palioura *et al.*, 2009).

In both the families WES analysis has leads to the identification start loss mutation in the *SepSecS* gene exon 1: c.T2C, p.MIT at the start codon in the following transcript (ENSG00000109618), amino acid methionine replaced with amino acid threonine. This variant was segregated with selected families, two affected individual (IV: 2 and IV: 4) from family A sharing homozygous region and while the 3 siblings was heterozygous. In Family B one affected patient was sharing homozygous allele while the parents were heterozygous for this variant.

The identified variant in *SepSecS* gene is disease causing variant for both the families. Both families' patient sharing similar phenotype and sample were collected from the same village. Sanger sequencing was performed on other variants (Table 4.7.1 and Table4.7.2) selected from patients exome data to check the segregation but none of the variants segregating with under studied families. Families' phenotypes were similar with our family phenotypes and carrying sepsecs gene mutations as reported in literature (Agamy *et al.*, 2010). So for there were four different mutations reported in literature that causing mutation in the sepsecs gene. A homozygous substitution mutation in the exon 8 of the *SepSecS* gene c. G1001A replacing amino acid Tyrosine (tyr334) with amino acid Cysteine (Cys334) in a conserved residue (Bee-Zeev *et al.*, 2003; Agamy *et al.*, 2010). Compound heterozygous mutation in which amino acid Alanine (ala239) was replaced with amino acid threonine (thr239) in the exon 6 of the sepsecs gene. The compound heterozygous mutation causes dysfuctioning of two alpha helices of the sepsecs gene (Agamy *et al.*, 2010). A homozygous mutation c.T1466A in exon 11 of the sepsecs gene was reported in 3 sibling of consanguineous Jordanian family. This reported mutation was replacing amino acid asparagine (asp489) with amino acid valine (val489). A nonsense mutation tyrosine (p.tyr429*) was reported in sepsecs gene causing cerebellocerebral a trophy (Makrythanasis *et al.*, 2014). However, if our reported variant is not disease causing variants for the family, then

there are number of consideration required first other families may need to be studied with start loss homozygous *SepSecS* mutations. We need to test further variant especially (PGAP2:NM_014489:exon5:c.TGC615AGC:p.C205S) in family, disease associated with PGAP2 gene is hyperphosphatasia with intellectual disability and in second family we need to test a frameshift deletion (NM_001244390:exon3:c.407_409del:p.E137Gfs*11) in NUDT2 gene involved in intellectual disability and developmental delay. Whole genome analysis (WGS) may be required to further explore the families.

Chapter 4.8: Candidate Intellectual Disability Gene (*PIDD*) Causing Loss of Function Mutation in Pakistani Families

4.8.1 Study Statement

Two Families AS105 and AS110 described in this chapter, both the families' clinical phenotypes related with intellectual disability, global developmental delay. Sequencing analysis of both the families resulted same loss of function mutation c.2587C>T; p.Q863X in p53-induced death domain protein (PIDD1/PIDD) gene. Both the families' results was published in (**Mol Psychiatry**) with our research collaborator (Harripaul *et al*., 2018).

4.8.2 Families Back Ground

4.8.2.1 Family AS105

Family AS105 figure (4.8.1A) belongs to (Khanabad), District Lower Dir, Khyber Pakhtunkhwa. Family AS105 was having intellectual disability phenotypes and global developmental delay. Family consist of 6 members from the second and third generation of the pedigree (II: 1, II: 2, III: 1, III: 2and III: 4), including two affected members (III: 2 and III: 3), both the affected probands share common phenotypes. The patients (III: 2) IQ is range severe (IQ: 35-40 to 50-55) and she was 6 years old at the time of medical examination. The proband (III: 3) age was 4 years at the time of medical examination. He has cognitive problems and facing difficulties to pay attention towards things.

4.8.2.2 Family AS110

Family AS110 (figure 11.1A) belongs to (village Bagh), District Lower Dir, Khyber Pakhtunkhwa. Family consists of 11 members, 8 were suffering with intellectual disability ranging from mild to severe. Most of affected individuals sharing similar phenotypes, with mild facial dysmorphic features. Blood sample of the available affected and normal individual collected from the families and consent was obtained from the individual as well as probands guardians. DNA extraction and molecular genetic analysis was performed in human molecular genetic lab.

4.8.3 Genetic Analysis

4.8.3.1 Linkage Analysis

DNA from all affected members as well as normal members of both the families were extracted and analyzed through standard protocol. Homozygous genetics variants were hypothesized on the bases of pedigree. Linkage analysis was performed in both the families with STR markers known for mapping ID gene in Pakistani population. Subjected families were shows no linkage with targeted gene by using STR makers. After excluding reported pathogenic genes the probands DNA subjected to WES analysis.

4.8.3.2 WES Analysis

WES analysis was performed in affected probands (II: 2) of family AS105 and proband (IV: 3) of family AS110 using sequencing facilities at Centre for Addiction and Mental Health (CAMH), Toronto Canada. WES data of both families affected probands was analyzed and all the benign variants including synonymous, intronic, and UTR variants were removed from the exome file. Exome data was filtered and rare variants only selected with allelic frequency <0.001%. Homozygous or compound heterozygous non-synonymous, framshift, splice site and coding indel variant with allelic frequencies of less than <0.001% in the 1000 genome project (1000genomes.org), exac browser (http://exac.broadinstitute.org) see (table 4.8.1) was selected for further analysis. Only c.2587C>T variant in PIDD gene was checked in the human protein atlas data base (https://www.proteinatlas.org/) for high level of expression in brain both at RNA and protein level. Selected variant also searched in OMIM (https://www.omim.org/), data base for any other disease association. PIDD gene variant pathogenic score were checked in mutation taster (http://www.mutationtaster.org/), GERP and CADD phred pathogenic score was also checked by using ensembl variant effect predictor tool (https://asia.ensembl.org/Tools/VEP) (Table 4.8.1). CADD score of this variant was (39), a score of higher than 20 or equal 20 indicates deleterious of mutation (Rentzsch *et al.*, 2018).

4.8.4 Results

After filtering for synonymous variations, total 153 genomic variants, including missense (N=95), frame shifts, splice site, stop gained, and stat loss (n=58), variants, were found to be present in both probands. Among these variants, only one nonsense variant e.2587C>T; p.Q863X (Table 4.8.1) was found common in two unrelated Pakistani families (AS105 and AS110; Figure 4.8.1). Co-segregation analysis conformed mutation segregate perfectly. PIDD gene encodes p53 induced death domain protein (PIDD); MIM 605247), and Q863X halts the death domain protein, through which PIDD gene networks with other death domain proteins particularly with RIP1 or CRADD. Non sense variants in CRADD have been previously been reported for NS-ARID (MRT34; MIM: 614499) and thus our finding support the involvement of PIDD related pathways in the etiology of ID.

Table 4.8.1: PIDD mutation in two Pakistani families

PIDD Gene Mutation							
Gene	**Nucleotide change**	**A.A change**	**SIFT**	**Polyphen2**	**CADD-Phred**	**GERP**	**Gnomad**
PIDD	c.2587C>T	p.Q863X	Deleterious	Damaging	39	3.78	Not Found

Figure 4.8.1: Families AS105 and AS110

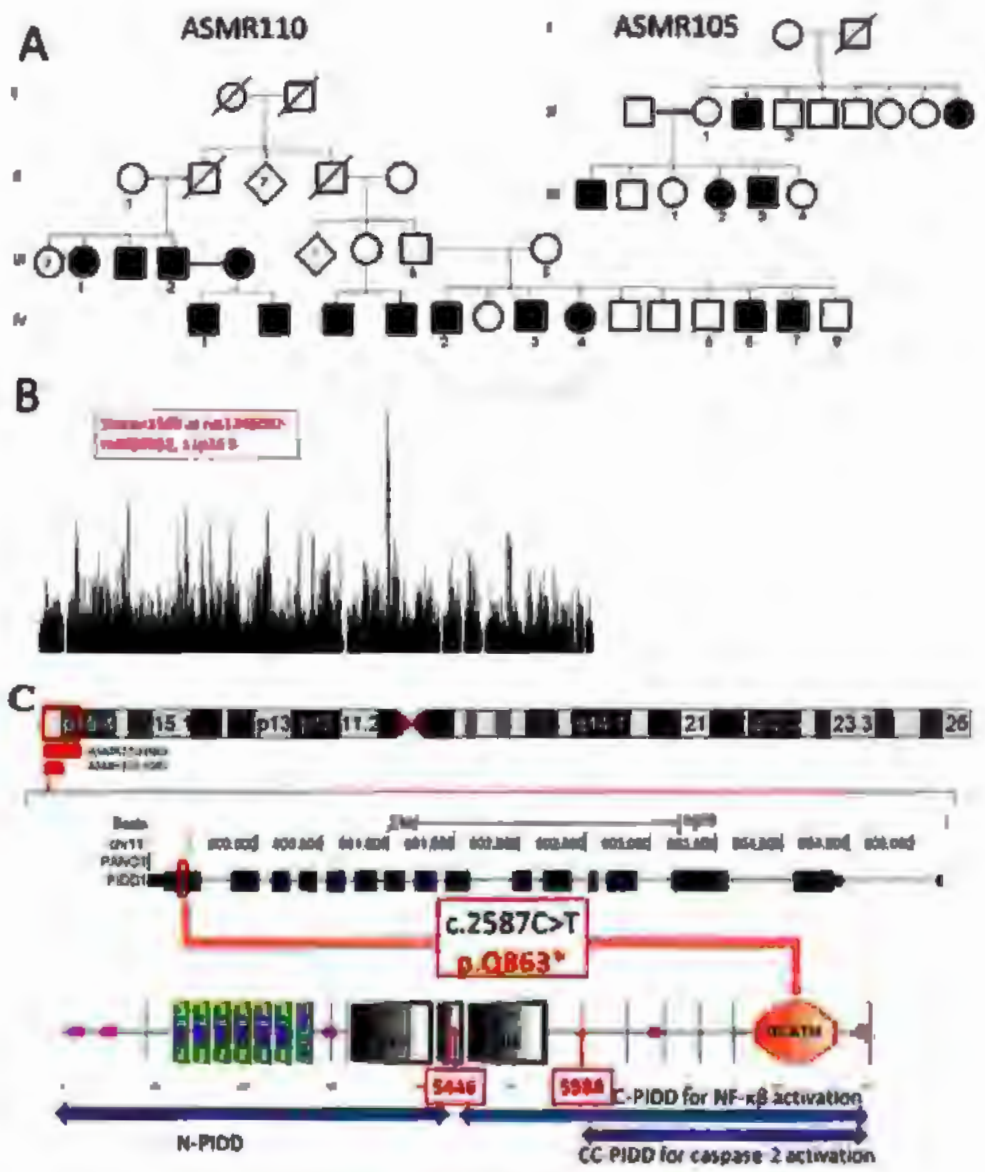


Figure 4.8.1: (A) Pedigrees and (B) Homozygosity Mapper output for both the families. The locus harboring the mutation is indicated with yellow shading. (C) Mutation representing the mutation location on chromosome 11 in DD domain.

4.8.5 Discussion

Exome data revealed a novel loss of function (nonsense) variant in new candidate ID gene (*PIDD*). The only loss function mutation identified in both families was located in the *PIDD* gene at chromosome 11. It consisted in a G to A transition at nucleotide 2587 (c.2587G<A), leading to the protein truncation p.Q863X. The variant was cross checked in online data basis for further confirmation. It was predicated to be pathogenic by different in-silico methods, was not previously reported and was not found in any data base including exac data base. The mutation was then examined in local population normal individuals and found to be absent in control samples (n=100), further supporting its pathogenicity.

PIDD1 gene mutations was never been reported before in intellectual disability families. This gene forms PIDDosome complex that activates *CASP2* and triggers apoptosis. This complex contains death domain (DD) protein and adaptor protein RAIDD, mutation in death domain stops the apoptosis process (Tinel and Tschopp, 2004). *PIDD* death domain also associated with other death domain proteins particularly with *CRADD* death domain. *CRADD* gene (MIM: 614499) have been associated with non syndromic autosomal recessive intellectual disability (Puffenberger *et al*., 2012) and this loss of function mutation support the involvement of PIDD-related pathways in the etiology of ID. *CRADD* gene loss of function and reduced caspase 2 play important role in neuronal apoptosis, reduced neuronal apoptosis may leads to some human brain neuronal storage disorders like megalencephaly (Mirzaa and Poduri, 2014). Human brain cortical malformations caused by reduced apoptosis, impact of reduced apoptosis in brain development previously described in different studies like in caspase 3 (Kuida and Zheng, 1996) and caspase 9 (Hakem and Hakem, 1998). *CRADD* signaling needed for the normal growth of brain neocortex and cognitive function. *CRADD* signaling also reduced level of caspase facilitated apoptosis during human brain development can result malformation in brain (Di Donato *et al*., 2016). Therefore, given the role PIDDosome complex associated with other death domains and play role in triggering CASP-2 and neuronal apoptosis. Mutation in *PIDD* gene may disrupt the functioning of apoptosis or casp-2 activates for normal brain functioning. In conclusion this study describes the role of *PIDD* gene associated with intellectual disability.

Chapter 5: Conclusion and Future Perspective

Chapter 5: Conclusion and Future Perspective

Current research focused on the identification of new candidate genes and novel mutation causing intellectual disability in consanguineous Pakistani families. The aim was to use linkage analysis and WES to identify the etiologies of ID families from Pakistani population. Through the work on these ID families, successfully identified the causal variants.

The key finding of the currents research includes the identification of three ID causing genes. *SLITRK3* gene involves in the pre synaptic and post synaptic neuronal activities and play role in neurotransmitters. *SDCBP* gene is the syndecan binding protein gene involves in the trafficking of transmembrane protein, neuronal and synaptic development. *PIDD1* gene encodes induced death domain protein; mutation in death domains halts the interaction of death domain proteins with other proteins especially *RAIDD* gene death domain protein associated with ID. In other families, five novel mutations identified including four missense mutation in four different genes and one start loss mutation in two families.

In future, genetic testing of ID will be significant for the diagnosis of ID patients. Early diagnosis will help the families regarding the treatment and management of the disease. Genetic testing of ID will also help the families in terms of genetic counseling but further knowledge needed to increase our understanding about brain functioning. Through genetic testing it will be possible to describe phenotypes on the basis of genotypes of the probands. On time diagnosis will be help to distinguish the psychiatric and genetic problems of the ID patients. Proper genetic counseling centers are required for the management of ID in Pakistani population. Drugs trials in animal model of ID give hope for the better treatment of ID which will be possible in near future.

In addition, many of the genes causing ID are likely to have founder effects that may be causing ID in Pakistani population, accelerating diagnosis of those genes and revealing carrier status to help family planning in future.

Gathering of correct phenotypes, mainly in the context of ID are very important and will help in correct diagnosis. Lastly the collaboration between Neurogenetic research groups will help to find out new genes and causal variants causing ID. Moreover ID should be treated in a comprehensive manner and patients suffering with ID should be provided social benefits that

today are not present for them. ID must be included in the political, social and economic agenda at government level to formulate the policies to prevent the disease.

REFERENCES

- Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM, Mathews CA, Pauls DL, Rašin MR, Gunel M, Davis NR. Sequence variants in SLITRK1 are associated with Tourette's syndrome. Science. 2005 Oct 14;310(5746):317-20.
- Agamy O, Zeev BB, Lev D, Marcus B, Fine D, Su D, Narkis G, Ofir R, Hoffmann C, Leshinsky-Silver E, Flusser H. Mutations disrupting selenocysteine formation cause progressive cerebello-cerebral atrophy. The American Journal of Human Genetics. 2010 Oct 8;87(4):538-44.
- Alazami AM, Patel N, Shamseldin HE, Anazi S, Al-Dosari MS, Alzahrani F, Hijazi H, Alshammari M, Aldahmesh MA, Salih MA and Faqeih E. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. Cell Reports. 2015, 10(2): 148-61.
- Antonellis A, Green ED. The role of aminoacyl-tRNA synthetases in genetic diseases. Annu. Rev. Genomics Hum. Genet.. 2008 Sep 22;9:87-107.
- Aruga J, Mikoshiba K. Identification and characterization of Slitrk, a novel neuronal transmembrane protein family controlling neurite outgrowth. Molecular and Cellular Neuroscience. 2003 Sep 1;24(1):117-29.
- Barber JC. Directly transmitted unbalanced chromosome abnormalities and euchromatic variants. Journal of Medical Genetics. 2005, 42(8): 609-29.
- Basel-Vanagaite L, Attia R, Yahav M, Ferland RJ, Anteki L, Walsh CA, Olender T, Straussberg R, Magal N, Taub E and Drasinover V. The CC2D1A, a member of a new gene family with C2 domains, is involved in autosomal recessive non-syndromic mental retardation. Journal of Medical Genetics. 2006, 43(3): 203-10.
- Beekman JM, Coffer PJ. The ins and outs of syntenin, a multifunctional intracellular adaptor protein. Journal of Cell science. 2008 May 1;121(9):1349-55.
- Ben-Zeev B, Hoffman C, Lev D, Watemberg N, Malinger G, Brand N, Lerman-Sagie T. Progressive cerebellocerebral atrophy: a new syndrome with microcephaly, mental retardation, and spastic quadriplegia. Journal of Medical Genetics. 2003 Aug 1;40(8):e96-

- Berkel S, Marshall CR, Weiss B, Howe J, Roeth R, Moog U, Endris V, Roberts W, Szatmari P, Pinto D and Bonin M. Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. Nature Genetics. 2010, 42(6): 489-91.
- Bonifacio S, Centrone C, Da Prato L, Scordo MR, Estienne M and Torricelli F. Use of primed in situ labeling (PRINS) for the detection of telomeric deletions associated with mental retardation. Cytogenetic and Genome Research. 2001, 93(1-2): 16-18.
- Bonkowsky JL, Nelson C, Kingston JL, Filloux FM, Mundorff MB, Srivastava R. The burden of inherited leukodystrophies in children. Neurology. 2010 Aug 24;75(8):718-25.
- Bormann J. The 'ABC' of GABA receptors. Trends in Pharmacological Sciences. 2000 Jan 1;21(1):16-9.
- Brandt T, Strupp M, Novozhilov S and Krafczyk S. Artificial neural network posturography detects the transition of vestibular neuritis to phobic postural vertigo. Journal of Neurology. 2012, 259(1): 182-84.
- Brignone MS, Lanciotti A, Camerini S, De Nuccio C, Petrucci TC, Visentin S, Ambrosini E. MLC1 protein: a likely link between leukodystrophies and brain channelopathies. Frontiers in Cellular Neuroscience. 2015 Apr 1;9:106.
- Caciotti A, Donati MA, Bardelli T, d'Azzo A, Massai G, Luciani L, Zammarchi E, Morrone A. Primary and secondary elastin-binding protein defect leads to impaired elastogenesis in fibroblasts from GM1-gangliosidosis patients. The American journal of pathology. 2005 Dec 1;167(6):1689-98.
- Çalışkan M, Chong JX, Uricchio L, Anderson R, Chen P, Sougnez C, Garimella K, Gabriel, SB, DePristo MA, Shakir K and Matern D. Exome sequencing reveals a novel mutation for autosomal recessive nonsyndromic Intellectual disability in the TECR gene on chromosome 19p13. Human Molecular Genetics. 2011, 189(4): 569-72.
- Callahan JW. Molecular basis of GM1 gangliosidosis and Morquio disease, type B. Structure–function studies of lysosomal β-galactosidase and the non-lysosomal β-galactosidase-like protein. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. 1999 Oct 8;1455(2-3):85-103.
- Carvill GL, Mefford HC. Next-generation sequencing in intellectual disability. Journal of Pediatric Genetics. 2015 Sep;4(03):128-35.

- Chatterjee N, Stegmüller J, Schätzle P, Karram K, Koroll M, Werner HB, Nave KA, Trotter J. Interaction of syntenin-1 and the NG2 proteoglycan in migratory oligodendrocyte precursor cells. Journal of Biological Chemistry. 2008 Mar 28;283(13):8310-7.
- Chelly J, Khelfaoui M, Francis F, Chérif B, Bienvenu T. Genetics and pathophysiology of mental retardation. European Journal of Human Genetics. 2006 Jun;14(6):701.
- Chih B, Engelman H, Scheiffele P. Control of excitatory and inhibitory synapse formation by neuroligins. Science. 2005 Feb 25;307(5713):1324-8.
- Chiurazzi P, Schwartz CE, Gecz J and Neri G. XLMR genes: update 2007. European Journal of Human Genetics. 2008, 16(4): 422-34.
- Cline H. Synaptogenesis: a balancing act between excitation and inhibition. Current Biology. 2005 Mar 29;15(6):R203-5.
- Conibear E, Cleck JN, Stevens TH. Vps51p mediates the association of the GARP (Vps52/53/54) complex with the late Golgi t-SNARE Tlg1p. Molecular biology of the cell. 2003 Apr 1;14(4):1610-23.
- Conibear E, Stevens TH. Vps52p, Vps53p, and Vps54p form a novel multisubunit complex required for protein sorting at the yeast late Golgi. Molecular biology of the cell. 2000 Jan 1;11(1):305-23.
- Conibear E, Stevens TH. Vps52p, Vps53p, and Vps54p form a novel multisubunit complex required for protein sorting at the yeast late Golgi. Molecular Biology of the cell. 2000 Jan 1;11(1):305-23.
- Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. Genome Research. 2005 Jul 1;15(7):901-13.
- Coughlin CR, Scharer GH, Shaikh TH. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. Genome Medicine. 2012 Oct;4(10):80.
- Curry CJ, Stevenson RE, Aughton D, Byrne J, Carey JC, Cassidy S, Cunniff C, Graham, JM, Jones MC, Kaback MM and Moeschler J. Evaluation of Intellectual disability Recommendations. American Journal of Medical Genetics. 1997, 72: 468-77.

- Dasgupta S, Menezes ME, Das SK, Emdad L, Janjic A, Bhatia S, Mukhopadhyay ND, Shao C, Sarkar D, Fisher PB. Novel role of MDA-9/syntenin in regulating urothelial cell proliferation by modulating EGFR signaling. Clinical Cancer Research. 2013 Sep 1;19(17):4621-33.
- Di Donato N, Jean YY, Maga AM, Krewson BD, Shupp AB, Avrutsky MI, Roy A, Collins S, Olds C, Willert RA, Czaja AM. Mutations in CRADD result in reduced caspase-2-mediated neuronal apoptosis and cause megalencephaly with a rare lissencephaly variant. The American Journal of Human Genetics. 2016 Nov 3;99(5):1117-29.
- Durkin MS, Hasan ZM, Hasan KZ. Prevalence and correlates of mental retardation among children in Karachi, Pakistan. American Journal of Epidemiology. 1998 Feb 1;147(3):281-8.
- Farrant M, Nusser Z. Variations on an inhibitory theme: phasic and tonic activation of GABA A receptors. Nature Reviews Neuroscience. 2005 Mar;6(3):215.
- Favre-Kontula L, Rolland A, Bernasconi L, Karmirantzou M, Power C, Antonsson B, Boschert U. GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system. Glia. 2008 Apr 15;56(6):633-45.
- Feinstein M, Flusser H, Lerman-Sagie T, Ben-Zeev B, Lev D, Agamy O, Cohen I, Kadir R, Sivan S, Leshinsky-Silver E, Markus B. VPS53 mutations cause progressive cerebello-cerebral atrophy type 2 (PCCA2). Journal of Medical Genetics. 2014 May 1;51(5):303-8.
- FitzGerald MJ and Folan-Curran J. Clinical Neuroanatomy and Related Neuroscience. WB Saunders Company. 2002.
- Friand V, David G, Zimmermann P. Syntenin and syndecan in the biogenesis of exosomes. Biology of the Cell. 2015 Oct;107(10):331-41.
- Frints SG, Froyen G, Marynen P, Fryns JP. X-linked mental retardation: vanishing boundaries between non-specific (MRX) and syndromic (MRXS) forms. Clinical genetics. 2002 Dec;62(6):423-32.
- Froyen G, Corbett M, Vandewalle J, Jarvela I, Lawrence O, Meldrum C, Bauters M, Govaerts K, Vandeleur L, Van Esch H and Chelly J. Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are

associated with Intellectual disability. The American Journal of Human Genetics. 2008, 82(2): 432-43.

- Garshasbi M, Hadavi V, Habibi H, Kahrizi K, Kariminejad R, Behjati F, Tzschach A, Najmabadi H, Ropers HH and Kuss AW. A defect in the TUSC3 gene is associated with autosomal recessive Intellectual disability. The American Journal of Human Genetics.2008, 82(5): 1158-64.
- Gonzalez-Burgos G, Rotaru DC, Zaitsev AV, Povysheva NV, Lewis DA. The GABA transporter GAT1 prevents spillover at proximal and distal GABA synapses onto primate prefrontal cortex neurons. Journal of Neurophysiology. 2009 Feb 1.
- Grootjans JJ, Zimmermann P, Reekmans G, Smets A, Degeest G, Dürr J, David G. Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. Proceedings of the National Academy of Sciences. 1997 Dec 9;94(25):13683-8.
- Gross, M. B. Personal Communication. Baltimore, Md. 8/29/2014.
- Hady-Cohen R, Ben-Pazi H, Adir V, Yosovich K, Blumkin L, Lerman-Sagie T, Lev D. Progressive cerebello-cerebral atrophy and progressive encephalopathy with edema, hypsarrhythmia and optic atrophy may be allelic syndromes. European Journal of Paediatric Neurology. 2018 Nov 1;22(6):1133-8.
- Hakem R, Hakem A, Duncan GS, Henderson JT, Woo M, Soengas MS, Elia A, de la Pompa JL, Kagi D, Khoo W, Potter J. Differential requirement for caspase 9 in apoptotic pathways in vivo. Cell. 1998 Aug 7;94(3):339-52.
- Hamamy H. Consanguineous marriages. Journal of Community Genetics. 2012 Jul 1;3(3):185-92.
- Hamamy HA, Masri AT, Al-Hadidy AM, Ajlouni KM. Consanguinity and genetic disorders. Profile from Jordan. Saudi Medical Journal. 2007;28(7):1015-7.
- Hamdan FF, Gauthier J, Spiegelman D, Noreau A, Yang Y, Pellerin S, Dobrzeniecka S, Côté M, Perreau-Linck E, Carmant L and D'Anjou G. Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. New England Journal of Medicine. 2009, 360(6): 599-605.
- Harripaul R, Vasli N, Mikhailov A, Rafiq MA, Mittal K, Windpassinger C, Sheikh TI, Noor A, Mahmood H, Downey S, Johnson M. Mapping autosomal recessive intellectual

disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. Molecular Psychiatry. 2018 Apr;23(4):973.

- Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly JM and Colleaux L. MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. Science. 2011, 333(6046): 1161-63.
- Hirbec H, Martin S, Henley JM. Syntenin is involved in the developmental regulation of neuronal membrane architecture. Molecular and Cellular Neuroscience. 2005 Apr 1;28(4):737-46.
- Hu H, Eggers K, Chen W, Garshasbi M, Motazacker MM, Wrogemann K, Kahrizi K, Tzschach A, Hosseini M, Bahman I and Hucho T. ST3GAL3 mutations impair the development of higher cognitive functions. The American Journal of Human Genetics. 2011, 89(3): 407-14.
- Hu H, Haas SA, Chelly J, Van Esch H, Raynaud M, De Brouwer AP, Weinert S, Froyen G, Frints SG, Laumonnier F, Zemojtel T. X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. Molecular psychiatry. 2016 Jan;21(1):133.
- Huang ZJ, Di Cristo G, Ango F. Development of GABA innervation in the cerebral and cerebellar cortices. Nature Reviews Neuroscience. 2007 Sep;8(9):673.
- Ibba M, Söll D. The renaissance of aminoacyl-tRNA synthesis. EMBO Reports. 2001 May 1;2(5):382-7.
- Ilja Boor PK, de Groot K, Waisfisz Q, Kamphorst W, Oudejans CB, Powers JM, Pronk JC, Scheper GC, van der Knaap MS. MLC1: a novel protein in distal astroglial processes. Journal of Neuropathology and Experimental Neurology. 2005 May 1;64(5):412-9.
- Jamain S, Quach H, Betancur C, Råstam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C and Bourgeron T. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. Nature Genetics. 2003, 34(1): 27-29.
- Jeworutzki E, Lagostena L, Elorza-Vidal X, López-Hernández T, Estévez R, Pusch M. GlialCAM, a ClC-2 Cl-channel subunit, activates the slow gate of ClC chloride channels. Biophysical Journal. 2014 Sep 2;107(5):1105-16.

- Jo S, Lee KH, Song S, Jung YK, Park CS. Identification and functional characterization of cereblon as a binding protein for large-conductance calcium-activated potassium channel in rat brain. Journal of Neurochemistry. 2005 Sep;94(5):1212-24.
- Jordanova A, Irobi J, Thomas FP, Van Dijck P, Meerschaert K, Dewil M, Dierick I, Jacobs A, De Vriendt E, Guergueltcheva V, Rao CV. Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. Nature genetics. 2006 Feb;38(2):197.
- Kahler SG, Fahey MC. Metabolic disorders and mental retardation. InAmerican Journal of Medical Genetics Part C: Seminars in Medical Genetics 2003 Feb 15 (Vol. 117, No. 1, pp. 31-41). New York: Wiley Subscription Services, Inc., A Wiley Company.
- Karimzadeh P, Jafari N, Biglari HN, Jabbehdari S, Zadeh SK, Abadi FA and Lotfi A. Neurometabolic Diagnosis in Children who referred as Neurodevelopmental Delay (A Practical Criteria, in Iranian Pediatric Patients). Iranian Journal of Child Neurology. 2016, 10(3): 73-81.
- Kehrer C, Maziashvili N, Dugladze T, Gloveli T. Altered excitatory-inhibitory balance in the NMDA-hypofunction model of schizophrenia. Frontiers in Molecular Neuroscience. 2008 Apr 8;1:6.
- Ko J, Choii G, Um JW. The balancing act of GABAergic synapse organizers. Trends in molecular medicine. 2015 Apr 1;21(4):256-68.
- Kuida K, Zheng TS, Na S, Kuan CY, Yang D, Karasuyama H, Rakic P, Flavell RA. Decreased apoptosis in the brain and premature lethality in CPP32-deficient mice. Nature. 1996 Nov;384(6607):368.
- Lee EY, Kim S, Kim MH. Aminoacyl-tRNA synthetases, therapeutic targets for infectious diseases. Biochemical pharmacology. 2018 Jun 8.
- Lee PH, Perlis RH, Jung JY, Byrne EM, Rueckert E, Siburian R, Haddad S, Mayerfeld CE, Heath AC, Pergadia ML and Madden PA. Multi-locus genome-wide association analysis supports the role of glutamatergic synaptic transmission in the etiology of major depressive disorder. Translational psychiatry. 2012, 2(10): 184.
- Leegwater PA, Yuan BQ, van der Steen J, Mulders J, Könst AA, Boor PI, Mejaski-Bosnjak V, van der Maarel SM, Frants RR, Oudejans CB, Schutgens RB. Mutations of MLC1 (KIAA0027), encoding a putative membrane protein, cause megalencephalic

leukoencephalopathy with subcortical cysts. The American Journal of Human Genetics. 2001 Apr 1;68(4):831-8.

- Leonard H, Wen X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. Mental Retardation and Developmental Disabilities Research Reviews. 2002;8(3):117-34.
- Li J, Han W, Pelkey KA, Duan J, Mao X, Wang YX, Craig MT, Dong L, Petralia RS, McBain CJ, Lu W. Molecular dissection of neuroligin 2 and Slitrk3 reveals an essential framework for GABAergic synapse development. Neuron. 2017 Nov 15;96(4):808-26.
- Liewen H, Meinhold-Heerlein I, Oliveira V, Schwarzenbacher R, Luo G, Wadle A, Jung M, Pfreundschuh M, Stenner-Liewen F. Characterization of the human GARP (Golgi associated retrograde protein) complex. Experimental Cell Research. 2005 May 15;306(1):24-34.
- López-Hernández T, Ridder MC, Montolio M, Capdevila-Nortes X, Polder E, Sirisi S, Duarri A, Schulte U, Fakler B, Nunes V, Scheper GC. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. The American Journal of Human Genetics. 2011 Apr 8;88(4):422-32.
- Lopez-Ramirez MA, Wu D, Pryce G, Simpson JE, Reijerkerk A, King-Robson J, Kay O, de Vries HE, Hirst MC, Sharrack B, Baker D. MicroRNA-155 negatively affects blood–brain barrier function during neuroinflammation. The FASEB Journal. 2014 Jun;28(6):2551-65.
- Lu, W., Bromley-Coolidge, S., and Li, J. (2017). Regulation of GABAergic synapse development by postsynaptic membrane proteins. Brain Res. Bull. 129, 30–42.
- Makrythanasis P, Nelis M, Santoni FA, Guipponi M, Vannier A, Béna F, Gimelli S, Stathaki E, Temtamy S, Mégarbané A, Masri A. Diagnostic exome sequencing to elucidate the genetic basis of likely recessive disorders in consanguineous families. Human mutation. 2014 Oct;35(10):1203-10.
- Marangi G, Leuzzi V, Manti F, Lattante S, Orteschi D, Pecile V, Neri G and Zollino M. TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. European Journal of Human Genetics. 2013, 21(2), 229-32.

- Maulik PK, Mascarenhas MN, Mathers CD, Dua T and Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. Research in Developmental Disabilities. 2011, 32(2): 419-36.
- McClelland AC, Sheffler-Collins SI, Kayser MS, Dalva MB. Ephrin-B1 and ephrin-B2 mediate EphB-dependent presynaptic development via syntenin-1. Proceedings of the National Academy of Sciences. 2009 Dec 1;106(48):20487-92.
- McLaren J, Bryson SE. Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. American Journal on Mental Retardation. 1987 Nov.
- McQuail JA, Frazier CJ, Bizon JL. Molecular aspects of age-related cognitive decline: the role of GABA signaling. Trends in molecular medicine. 2015 Jul 1;21(7):450-60.
- Meyer-Schuman R, Antonellis A. Emerging mechanisms of aminoacyl-tRNA synthetase mutations in recessive and dominant human disease. Human molecular genetics. 2017 Jun 15;26(R2):R114-27.
- Miller M, Kibiuk LV and Stuart D. Brain Facts: A Primer on the Brain and Nervous System. 2008,
- Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, Kahrizi K, Rafiq MA, Weksberg R, Nasr T and Naeem F. Identification of mutations in TRAPPC9, which encodes the NIK-and IKK-β-binding protein, in nonsyndromic autosomal-recessive Intellectual disability. The American Journal of Human Genetics. 2009, 85(6): 909-15.
- Mirzaa GM, Poduri A. Megalencephaly and hemimegalencephaly: breakthroughs in molecular etiology. InAmerican Journal of Medical Genetics Part C: Seminars in Medical Genetics 2014 Jun (Vol. 166, No. 2, pp. 156-172).
- Missler, M., S€udhof, T.C., and Biederer, T. (2012). Synaptic cell adhesion. Cold Spring Harb. Perspect. Biol. 4, a005694.
- Mochida GH, Mahajnah M, Hill AD, Basel-Vanagaite L, Gleason D, Hill RS, Bodell A, Crosier M, Straussberg R and Walsh CA. A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. The American Journal of Human Genetics. 2009, 85(6): 897-902.
- Moehler H. GABA A receptors in central nervous system disease: anxiety, epilepsy, and insomnia. Journal of Receptors and Signal Transduction. 2006 Jan 1;26(5-6):731-40.

- Moeschler JB, Shevell M. Comprehensive evaluation of the child with intellectual disability or global developmental delays. Pediatrics. 2014 Sep 1;134(3):e903-18.
- Mokhtari R, Bagga A. Consanguinity, genetic disorders and malformations in the Iranian population. Acta Biologica Szegediensis. 2003 Jan 1;47(1-4):47-50.
- Molinari F, Rio M, Meskenaite V, Encha-Razavi F, Augé J, Bacq D, Briault S, Vekemans, M, Munnich A, Attié-Bitach T and Sonderegger P. Truncating neurotrypsin mutation in autosomal recessive nonsyndromic mental retardation. Science. 2002, 298(5599): 1779-81.
- Musante L and Ropers HH. Genetics of recessive cognitive disorders. Trends in Genetics. 2014, 30(1): 32-39.
- Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P and Zecha A. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011, 478(7367): 57-63.
- Nelson DL. Mental retardation and intellectual disability. In Vogel and Motulsky's Human Genetics. 2009, 663-80.
- Neri G, Schwartz CE, Lubs HA, Stevenson RE. X-linked intellectual disability update 2017. American Journal of Medical Genetics Part A. 2018 Jun;176(6):1375-88.
- Ohno K, Koroll M, El Far O, Scholze P, Gomeza J, Betz H. The neuronal glycine transporter 2 interacts with the PDZ domain protein syntenin-1. Molecular and Cellular Neuroscience. 2004 Aug 1;26(4):518-29.
- Oprescu SN, Griffin LB, Beg AA, Antonellis A. Predicting the pathogenicity of aminoacyl-tRNA synthetase mutations. Methods. 2017 Jan 15;113:139-51.
- Palioura S, Sherrer RL, Steitz TA, Söll D, Simonović M. The human SepSecS-tRNASec complex reveals the mechanism of selenocysteine formation. Science. 2009 Jul 17;325(5938):321-5.
- Palmer E, Speirs H, Taylor PJ, Mullan G, Turner G, Einfeld S, Tonge B and Mowat D. Changing interpretation of chromosomal microarray over time in a community cohort with intellectual disability. American Journal of Medical Genetics Part A. 2014, 164(2): 377-85.

- Pascual-Castroviejo I, Pronk JC, Garcia-Segura JM, Gutiérrez-Molina M, Pascual-Pascual SI. Vacuolating megalencephalic leukoencephalopathy: 24 year follow-up of two siblings. Neurologia (Barcelona, Spain). 2005;20(1):33-40.
- Pérez-Victoria FJ, Mardones GA, Bonifacino JS. Requirement of the human GARP complex for mannose 6-phosphate-receptor-dependent sorting of cathepsin D to lysosomes. Molecular Biology of the Cell. 2008 Jun;19(6):2350-62.
- Perou, R., Bitsko, R.H., Blumberg, S.J., Pastor, P., Ghandour, R.M., Gfroerer, J.C., Hedden, S.L., Crosby, A.E., Visser, S.N., Schieve, L.A., Parks, S.E., Hall, J.E., Brody, D., Simile, C.M., Thompson, W.W., Baio, J., Avenevoli, S., Kogan, M.D., Huang, L.N., 2013. Mental health surveillance among children – United States, 2005–2011. Morbid. Mortal. Week. Rep. 62 (Suppl. 2), 1–35, Surveillance summaries (Washington, DC, 2002).
- Philley JV, Kannan A, Dasgupta S. MDA-9/Syntenin Control. Journal of Cellular Physiology. 2016 Mar;231(3):545-50.
- Piton A, Gauthier J, Hamdan FF, Lafreniere RG, Yang Y, Henrion E, Laurent S, Noreau A, Thibodeau P, Karemera L, Spiegelman D. Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. Molecular Psychiatry. 2011 Aug;16(8):867.
- Puffenberger EG, Jinks RN, Sougnez C, Cibulskis K, Willert RA, Achilly NP, Cassidy RP, Fiorentini CJ, Heiken KF, Lawrence JJ, Mahoney MH. Genetic mapping and exome sequencing identify variants associated with five novel diseases. PloS One. 2012 Jan 17;7(1):e28936.
- Puppala AK, French RL, Matthies D, Baxa U, Subramaniam S, Simonović M. Structural basis for early-onset neurological disorders caused by mutations in human selenocysteine synthase. Scientific Reports. 2016 Aug 31:6:32563.
- Rafiq MA, Kuss AW, Puettmann L, Noor A, Ramiah A, Ali G, Hu H, Kerio NA, Xiang Y, Garshasbi M and Khan MA. Mutations in the alpha 1, 2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. The American Journal of Human Genetics. 2016, 89(1): 176-82.

- Rajendran V, Kalita P, Shukla H, Kumar A, Tripathi T. Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. International journal of biological macromolecules. 2018 May 1;111:400-14.
- Raymond FL and Tarpey P. The genetics of mental retardation. Human Molecular Genetics. 2006, 15(2): 110-16.
- Regier D.S. and Tifft C.J. , GLB1-Related Disorders. In: Pagon R.A., Adam M.P., Ardinger H.H., Wallace S.E., Amemiya A., Bean L.J., et al., editors. GeneReviews(®) [Internet]. Seattle (WA) University of Washington, Seattle; 2013. Available from: http://www.ncbi.nlm.nih.gov/books/NBK164500/
- Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research. 2018 Oct 29;47(D1):D886-94.
- Riazuddin S, Hussain M, Razzaq A, Iqbal Z, Shahzad M, Polla DL, Song Y, van Beusekom E, Khan AA, Tomas-Roca L, Rashid M. Exome sequencing of Pakistani consanguineous families identifies 30 novel candidate genes for recessive intellectual disability. Molecular psychiatry. 2017 Nov;22(11):1604.
- Rohrbach M and Clarke JT. Treatment of Lysosomal Storage Disorders. Drugs. 2007, 67(18): 2697-16.
- Ropers HH and Hamel BC. X-linked mental retardation. Nature Reviews Genetics. 2005, 6(1): 46-57.
- Ropers HH. Genetics of early onset cognitive impairment. Annual review of genomics and human genetics. 2010 Sep 22:11:161-87.
- Rubenstein JL, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. Genes, Brain and Behavior. 2003 Oct;2(5):255-67.
- Saleha S, Sajid M, Zafar S, Pervaiz N. Genetic Defects Underlie the Non-syndromic Autosomal Recessive Intellectual Disability (NS-ARID). Open Life Sciences. 2017 Jan 1;12(1):167-77.
- Schirwani S, Wakeling E, Smith K, DDD Study, Balasubramanian M. Expanding the molecular basis and phenotypic spectrum of ZDHHC9-associated X-linked intellectual disability. American Journal of Medical Genetics Part A. 2018 May;176(5):1238-44.

- Shen K, Scheiffele P. Genetics and cell biology of building specific synaptic connectivity. Annual review of neuroscience. 2010 Jul 21;33:473-507.
- Siddiqui TJ, Craig AM. Synaptic organizing complexes. Current opinion in neurobiology. 2011 Feb 1;21(1):132-43.
- Siniossoglou S, Pelham HR. An effector of Ypt6p binds the SNARE Tlg1p and mediates selective fusion of vesicles with late Golgi membranes. The EMBO journal. 2001 Nov 1;20(21):5991-8.
- Slager RE, Newton TL, Vlangos CN, Finucane B and Elsea SH. Mutations in RAI1 associated with Smith–Magenis syndrome. Nature Genetics. 2003, 33(4): 466-68.
- Sohn YB, Park HD, Park SW, Kim SH, Cho SY, Ko AR, Ki CS, Yeau S, Jin DK. A Korean patient with Morquio B disease with a novel c. 13_14insA mutation in the GLB1 gene. Annals of Clinical & Laboratory Science. 2012 Dec 21;42(1):89-93.
- Stankiewicz P and Beaudet AL. Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic Intellectual disability .Current opinion in Genetics and Development. 2007, 17(3): 182-92.
- Stavropoulos DJ, MacGregor DL and Yoon G. Mosaic microdeletion 18q21 as a cause of mental retardation. European Journal of Medical Genetics. 2010, 53(6): 396-99.
- Stolze H, Klebe S, Zechlin C, Baecker C, Friege L and Deuschl G. Falls in frequent neurological diseases. Journal of Neurology. 2004, 251(1): 79-84.
- Suzuki Y, Nanba E, Matsuda I, Higaki K, Oshima A. p-Galactosidase Deficiency (P-Galactosidosis): GM1 Gangliosidosis and Morquio B Disease Part 16: Lysosomal disorders. Cap 151. Valle D, Beaudet A, Vogelstein B, Kinzler K, Antonarakis S, Ballabio A. The Online Metabolic and Molecular Bases of Inherited Disease. Mc Graw Hill. 2011.
- Suzuki Y. β-Galactosidase deficiency (-galactosidosis): GM1 gangliosidosis and Morquio B disease. The metabolic and molecular bases of inherited disease. 1995:2785-823.
- Takahashi H, Katayama KI, Sohya K, Miyamoto H, Prasad T, Matsumoto Y, Ota M, Yasuda H, Tsumoto T, Aruga J, Craig AM. Selective control of inhibitory synapse development by Slitrk3-PTPδ trans-synaptic interaction. Nature Neuroscience. 2012 Mar;15(3):389.

- Teijido O, Martínez A, Pusch M, Zorzano A, Soriano E, del Río JA, Palacín M, Estevez R. Localization and functional analyses of the MLC1 protein involved in megalencephalic leukoencephalopathy with subcortical cysts. Human Molecular Genetics. 2004 Sep 14;13(21):2581-94.
- Tinel A, Tschopp J. The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. Science. 2004 May 7;304(5672):843-6.
- Topçu M, Gartioux C, Ribierre F, Yalçinkaya C, Tokus E, Öztekin N, Beckmann JS, Ozguc M, Seboun E. Vacuoliting megalencephalic leukoencephalopathy with subcortical cysts, mapped to chromosome 22qtel. The American Journal of Human Genetics. 2000 Feb 1;66(2):733-9.
- Van der Knaap MS, Barth PG, Stroink HA, Van Nieuwenhuizen O, Arts WF, Hoogenraad F, Valk J. Leukoencephalopathy with swelling and a discrepantly mild clinical course in eight children. Annals of Neurology. 1995 Mar 1;37(3):324-34.
- Van der Knaap MS, Barth PG, Vrensen GF, Valk J. Histopathology of an infantile-onset spongiform leukoencephalopathy with a discrepantly mild clinical course. Acta neuropathologica. 1996 Jul 1;92(2):206-12.
- van der Knaap MS, Boor I, Estévez R. Megalencephalic leukoencephalopathy with subcortical cysts: chronic white matter oedema due to a defect in brain ion and water homoeostasis. The Lancet Neurology. 2012 Nov 1;11(11):973-85.
- Van Karnebeek CD, Scheper FY, Abeling NG, Alders M, Barth PG, Hoovers JM, Koevoets C, Wanders RJ and Hennekam RC. Etiology of mental retardation in children referred to a tertiary care center: a prospective study. American Journal on Mental Retardation. 2005, 110(4): 253-67.
- Varcarolis EM. Essentials of Psychiatric Mental Health Nursing-E-Book: A Communication Approach to Evidence-Based Care. Elsevier Health Sciences; 2016 Jul 22.
- Vasan N, Hutagalung A, Novick P, Reinisch KM. Structure of a C-terminal fragment of its Vps53 subunit suggests similarity of Golgi-associated retrograde protein (GARP) complex to a family of tethering complexes. Proceedings of the National Academy of Sciences. 2010 Aug 10;107(32):14176-81.

- Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Gläser D, Liebscher V, Meitinger T and Strom TM. Copy-Number Variations Measured by Single-Nucleotide–Polymorphism Oligonucleotide Arrays in Patients with Mental Retardation. The American Journal of Human Genetics. 2007. 81(4): 768-79.
- Wassef A, Baker J, Kochan LD. GABA and schizophrenia: a review of basic science and clinical studies. Journal of Clinical Psychopharmacology. 2003 Dec 1;23(6):601-40.
- Weiss MM, Hermsen MA, Meijer GA, Van Grieken NC, Baak JP, Kuipers EJ and Van Diest PJ. Comparative genomic hybridisation. Molecular Pathology. 1999, 52(5): 243.
- Winnepenninckx B, Rooms L, Kooy RF. Mental retardation: a review of the genetic causes. The British Journal of Development Disabilities. 2003, 49(96): 29-44.
- WHO (World Health Organization). The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research, 1993
- Xu J and Chen Z. Advances in molecular cytogenetics for the evaluation of mental retardation. In American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 2003. 117(1): 15-24.
- Yim YS, Kwon Y, Nam J, Yoon HI, Lee K, Kim DG, Kim E, Kim CH, Ko J. Slitrks control excitatory and inhibitory synapse formation with LAR receptor protein tyrosine phosphatases. Proceedings of the National Academy of Sciences. 2013 Mar 5;110(10):4057-62.
- Yuan J, Palioura S, Salazar JC, Su D, O'Donoghue P, Hohn MJ, Cardoso AM, Whitman WB, Söll D. RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. Proceedings of the National Academy of Sciences. 2006 Dec 12;103(50):18923-7
- Zhang, X.; Ling, J.; Barcia, G.; Jing, L.; Wu, J.; Barry, B. J.; Mochida, G. H.; Hill, R. S.; Weimer, J. M.; Stein, Q.; Poduri, A.; Partlow, J. N.; Ville, D.; Dulac, O.; Yu, T. W.; Lam, A. T.; Servattalab, S.; Rodriguez, J.; Boddaert, N.; Munnich, A.; Colleaux, L.; Zon, L. I.; Soll, D.; Walsh, C. A.; Nabbout, R., Mutations in QARS, encoding glutaminyl-tRNA synthetase, cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures. Am J Hum Genet 2014, 94 (4), 547-58.
- Zimmermann P, Tomatis D, Rosas M, Grootjans J, Leenaerts I, Degeest G, Reekmans G, Coomans C, David G. Characterization of syntenin, a syndecan-binding PDZ protein, as a

component of cell adhesion sites and microfilaments. Molecular biology of the cell. 2001 Feb 1;12(2):339-50.

Appendices

Appendix A-1

Research Participants Consent Form

Consent form Urdu

ریسرچ میں رضامندی کا فارم:

Whole Genome Analysis سے آٹوسومل ریسیسو انٹلیکچوئل ڈس ایبلٹی کے جینیٹک اسباب کی تحقیق

تحقیق کنندگان:

Dr. Asif Mir Phone No# +923019730

Muhammad Ilyas Cell No:- +923468928212

ریسرچ کا مقصد:

ریسرچ کی تفصیل:

1) آپ کے خون کے نمونے سے ڈی این اے نکالا جائے گا جس کی ترتیب کا تجزیہ کیا جائے گا۔

2) آپ کا انٹرویو لیا جائے گا جس میں آپ کے خاندان کی طبی تاریخ کے بارے میں سوالات پوچھے جائیں گے۔

ممکنہ خطرات:

ممکنہ فوائد:

موروثی خصوصیات کا تجزیہ آپ کی فراہم کردہ باتوں (Information) پر کیا جائے گا۔ اگر موروثی ٹیسٹوں کے نتائج آپ کی اپنے خاندان کے بارے میں بتائی گئی معلومات سے نہ ملیں تو اس کا مطلب ہے کہ یا تو شجرہ غلط ہے یا آپ نے جو اپنے خاندان کے بارے میں معلومات فراہم کی ہیں وہ غلط ہیں۔ مثال کے طور پر والدین نے یہ بتایا ہو کہ یہ بچہ ان کا اپنا ہے لیکن اصل میں انہوں نے اپنایا ہوا ہے اور اس کا والد کوئی اور ہے۔

ممکنہ فوائد:

ہو سکتا ہے کہ اس سٹڈی سے آپ کو فائدہ نہ ہو لیکن اس سٹڈی کے نتائج مستقبل میں دوسرے مریضوں اور ان کے خاندانوں کے لئے فائدہ مند ہو سکتے ہیں اور بیماری کے بارے میں جاری معلومات بڑھا سکتے ہیں۔ مزید یہ کہ اس سٹڈی کا ایک فائدہ یہ بھی ہو سکتا ہے کہ اس بیماری سے متاثرہ بچے کا اس کی پیدائش سے بہت پہلے پتہ چلانے کا ٹیسٹ دریافت ہو جائے۔

رازداری:

آپ کی شناخت صیغہ راز میں رکھی جائے گی اور کسی کو بھی اس بارے میں نہیں بتایا جائے گا۔ یہ معلومات نامے کا فارم آپ کی فائل میں لگا دیا جائے گا۔ آپ کی نشوونما کی جانچ پڑتال کے نتائج صرف آپ کے علاج اور مستقبل میں اس بیماری سے بچنے کے لئے کی جانے والی ریسرچ تک استعمال کئے جائیں گے۔ کوئی بھی بات آپ کے علاج میں ملوث عملے کو بتانے کے لئے آپ سے اجازت لی جائے گی۔ آپ کی موروثی معلومات کو مکمل طور پر صیغہ راز میں رکھا جائے گا۔ آپ کے خون کے نمونوں پر آپ کے نام کی بجائے کوڈ نمبر لگائے جائیں گے اور اس کوڈ نمبر کی فائل کو لاک میں رکھا جائے گا جس تک صرف تحقیق کرنے والے افراد کو ہی رسائی حاصل ہو گی۔ کوئی بھی معلومات جس سے آپ کی شناخت ظاہر ہو، آپ کی اجازت کے بغیر کسی کو نہیں بتائی جائے گی۔ تاہم اس بات کا ریکارڈ [illegible] گی۔ مزید یہ کہ اس سٹڈی کے نتائج سے متعلق کوئی بھی معلومات آپ کی علاج معالجے کی معلومات والی فائل کا حصہ بن جائے گی۔ آپ کے نمونوں سے حاصل خلوی ڈی این اے (DNA) اور سیلز (Cells) کو کوڈ نمبر لگا کر فیر مدت کے لئے محفوظ کر لیا جائے گا تاکہ اس سے آئندہ ہونے والی تحقیق کے لئے دستیاب ہو سکیں اور اس ریسرچ کو جاری رکھا جائے گا۔

رضاکارانہ شرکت:

اس ریسرچ میں شمولیت رضاکارانہ ہے۔ اگر آپ حصہ نہ لینا چاہیں تب بھی آپ کا علاج معالجہ پاکستان کے ہسپتال میں جاری رہے گا۔ اور اگر آپ کسی بھی وقت اپنے خون کا نمونہ اس سٹڈی سے نکلوانا چاہیں تو آپ کے خون کا نمونہ ضائع کر دیا جائے گا اور اس سے آپ کے اس ہسپتال میں علاج پر کوئی فرق نہیں پڑے گا۔ اس سٹڈی میں شمولیت رضاکارانہ ہے اور کسی کو بھی اس سٹڈی سے انفرادی طور پر فائدہ نہیں ہوگا۔

اجازت نامہ

میں اقرار کرتا / کرتی ہوں کہ مجھے اس ریسرچ کے بارے میں تفصیلاً بتایا گیا ہے اور میرے تمام سوالات کا تسلی بخش جواب دیا گیا ہے اور مجھے اس بات سے آگاہ کیا گیا ہے کہ اگر میں اس سٹڈی میں شمولیت نہ کروں تو اس سے میرے خاندان کے اس بیماری میں علاج معالجے پر کوئی فرق نہیں پڑے گا۔ اس کے علاوہ اس ریسرچ کے فوائد و نقصانات کی تفصیلات مجھے بتائی گئی ہیں۔ مجھے معلوم ہے کہ میں اس ریسرچ کے بارے میں کوئی بھی سوال اب یا مستقبل میں پوچھ سکتا ہوں۔ مجھے اس بات کا علم ہے کہ میں کسی وقت بھی اس ریسرچ کے بارے میں مشورہ لے سکتا ہوں۔ مجھے اس بات کی یقین دہانی کرائی گئی ہے کہ میری تمام معلومات صیغہ راز میں رکھی جائیں گی اور میری اجازت کے بغیر کسی بھی قسم کی کوئی ذاتی معلومات یا شناخت کسی کو بھی نہیں بتائی جائے گی جب تک یہ چیز قانوناً ضروری نہ ہو۔

میں اجازت دیتا / دیتی ہوں کہ:

1- میرے میڈیکل ریکارڈ کو اس ریسرچ سے متعلقہ ریسرچ میں استعمال کیا جا سکتا ہے۔

2- میرا خون یا تھوک کا نمونہ لیا جا سکتا ہے۔

3- ان نمونوں سے ڈی این اے (DNA) نکالا جا سکتا ہے اور اسے غیر معینہ مدت تک محفوظ کیا جا سکتا ہے اور اعصابی و نفسیاتی اور نشوونما کی ریسرچ میں استعمال کیا جا سکتا ہے۔

4- ڈی این اے (DNA) کو کوڈ لگا کر استعمال کیا جائے گا۔

5- اور اگر میرے بچے سے متعلقہ کوئی اہم معلومات حاصل ہوں تو مجھ سے رابطہ کیا جائے گا اور اگر میرے گھر کے ایڈریس یا فون نمبر میں کوئی تبدیلی ہوئی تو میں آپ کو مطلع کروں گا۔

میں اس ریسرچ میں شمولیت کی اجازت دیتا ہوں:

نام _____

دستخط _____

تاریخ _____

اجازت لینے والے کا نام _____

دستخط _____

اس ریسرچ سے متعلقہ سوالات کے لیے آپ رابطہ کر سکتے ہیں۔

Dr. Asif Mir

Phone no:- +929019736

Muhammad Ilyas

+923346928212

Consent form in English

Name of Participant:
DOB:
Sample ID:

Research Participant Informed Consent

Subject of study: - Mapping Autosomal Recessive Intellectual Disability in Consanguineous Families

Researchers: DR. Asif Mir Email: asif.mir@iiu.edu.pk

Muhammad Ilyas Cell No: +923346928212

Purpose of the Research:

We hope to use this study to help us find genes that may predispose to, intellectual disability and/or associated neurogenetic disorders. Intellectual disability (ID) causes are unknown in more than 50% of individuals or we know very little about ID disorder. Other 50% ID is caused by the genetic factors including aneuploidies, copy no vartions and mutation in specific genes. Genetic studies of the genes involved in these conditions or finding of the new genes involved in these condition are likely to further our understanding of the basis of Intellectual Disability. The outcome of this research covers the possible identification of new gene(s) involved in disease and help to understand the gene therapy and possible treatment of the disease.

Description of the Research:

1. If you are agreeing to participate our research then we may ask you to provide information to us.
    - Past Medical history
    - Behavior and development of the affected individual

2. If you agree to participate, a trained health care professional will take blood sample (20-30 cc or 5 Ml) from your or your child. This blood will be used to look at the pattern of changes in DNA

Potential Harms:

There may be a small amount of bleeding when blood is taken from a vein and there may be slight discomfort and bruising or redness that will usually disappear in a few days.

What are the risks of the study?

During the course of the study, if we identify any information that may have clinical significance to your child, one of the investigators will contact you and your child about these observations and arrangements will be made for counselling and assistance for you and your child in understanding the personal and family significance should you need it. This knowledge could cause psychological stress to you, your child, and your family. In rare cases, knowing about a presence of a genetic problem might possibly affect you and your child's health and/or life insurance coverage in the future. The interpretation of the genetic information will depend in part on the family information that you have provided. If the results of genetic tests do not fit with the information that you have given about your family, it may be that the test is faulty, or that the family information that you gave is wrong. For example, this might happen if the parents do not mention that their child was adopted, or that the biological father is different from the apparent father.

Potential Benefits:

Your child and your family may or may not benefit from participating in this study. It is possible that we will be able to identify the molecular or genomic cause of the intellectual disability disorder in your family, which may help with treatment options; this study may help other patients and their families, while we cannot know this for sure.

Confidentiality:

We will respect you & your child's privacy. No information about who you and your child is will be given to anyone. We will put a copy of this research consent form in your child's patient

health records. The data produced from this study will be stored in a secure, locked location. Only members of the research team will have access to the data. Following completion of the research study the data will be kept as long as required. Published study results will not reveal your child's identity. No information that reveals you or your child's identity will be released or published without your consent. In addition, information regarding the results of this research may become part of your child's health record if relevant to his or her medical care. The DNA isolated from you/your child's blood will be stored indefinitely with the number codes, so that as new genes are discovered, which are involved in Intellectual Disability, this DNA can be used to continue the research.

Participation in Research

If you choose to have you /your child take part in this study, you can withdraw from the study at any time. Your withdrawal from research will not affect the care and treatment of your child in Pakistan in any way.

Consent:

By signing this form, I agree that:

- You have explained this study to me. You have answered all my questions.
- You have explained the possible harms and benefits (if any) of this study.
- I understand that I have the right to refuse take part in the study and the right to stop at any time. My decision about taking part in the study will not affect my child's health care at Hospital.
- I am free now, and in the future, to ask questions about the study.
- I understand that no information about who my child is will be given to anyone or be published without first asking my permission.
- My child will provide a blood sample and both parents (when available) will provide Blood or saliva samples.
- Specific Code will be Assign to every DNA sample
- Blood samples used as a source of DNA, and stored indefinite time, for research into the genetics of Intellectual Disability.

I have read and understood pages 1 to 3 of this consent form. I agree, or consent, to take part in this study.

....

Printed Name of Parent/Legal Guardian Parent/Legal Guardian's signature Date

....

Pretended Name of Person Who Explain Consent Signature & Date

If you have any questions about this study, please call Or Email:

Dr Asif Mir

asif.mir@iiu.edu.pk

Muhammad Ilyas

Ilyas.Phdbt46@iiu.edu.pk

Appendix A-2 (Publications)

- **Ilyas M**, Efthymiou S, Salpietro V, Noureen N, Zafar F, Rauf S, **Mir A**, Houlden H. Novel variants underlying autosomal recessive intellectual disability in Pakistani consanguineous families. **BMC medical genetics**. 2020 Dec;21:1-8. **I.F (1.58)**

- Harripaul R, Vasli N, Mikhailov A, Rafiq MA, Mittal K, Windpassinger C, Sheikh TI, **Noor A**, Mahmood H, Downey S, Johnson M, Vleuten K, Bell L, **Ilyas M**, Khan F.S, Khan V, Moradi M, Ayaz M, Naeem F, Heidari A, Ahmed I, Ghadami S, Agha Z, Zeinali S, **Qamar** R, Mozhdehipanah H, John P, **Mir A**, Ansar M, French L, Ayub M, Vincent B.J. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. **Nature Molecular psychiatry**. 2018 Apr;23(4):973-84. **I.F: 12.38**

- **Ilyas M**, **Mir A**, Efthymiou S, Houlden H. The genetics of intellectual disability: advancing technology and gene editing. **F1000Research**. 2020;9. **I.F: 2.64**