Genetic Analysis of Families with Autosomal Recessive Non-Syndromic Mental Retardation and Primary Microcephaly Disorders

By

Jamshaid Mahmood Baig

Registration # 29-FBAS/PHDBT/F-14

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

2014-2020

Genetic Analysis of Families with Autosomal Recessive Non-Syndromic Mental Retardation and Primary Microcephaly Disorders

Researcher

Jamshaid Mahmood Baig

Reg. #29-FBAS/PHDBT/F-14

Supervised By

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

Co-Supervisor
Dr. Abdul Hameed
Chief Scientific Officer
IB&GE KRL

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

2014-2020

Date: _____

FINAL APPROVAL

It is certified that we have read the thesis submitted by Mr. Jamshaid Mahmood Baig 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, FBAS, IIUI

External Examiner

_____
Associate Professor /

External Examiner

_____

Internal Examiner

Dr. _____
Assistant Professor /
Department of Biological Sciences, FBAS, 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. Jamshaid Mahmood Baig 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. Arshad Malik _____

Associate Professor,
Department of Biological Sciences
International Islamic University Islamabad, Pakistan

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

This Thesis is

Submitted to International Islamic University, Islamabad, Pakistan in partial fulfillment of the requirements of the degree of Doctor of philosophy (Biotechnology)

DEDICATION

This work is dedicated to my beloved grandparent’s, father, mother and my family members 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 _____

Jamshaid Mahmood Baig

Table of Content

Contents

List of Figures

Figure No **Page No**

List of Tables

List of Abbreviations

ATP	Adenosine Triphosphate
BASAR	Board of Advanced Study and Research
CADD	Combined Annotation Dependent Depletion
CNS	Central Nervous System
CNVs	Copy Number Variations
CT Scan	Computed Tomography Scan
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
ID	Intellectual Disability
IIU	International Islamic University
IQ	Intelligence Quotient
MRI	Magnetic Resonance Imaging
NGS	Next Generation Sequencing
°C	Centigrade
PCR	Polymerase Chain Reaction
pH	Potential of Hydrogen
PK	Proteinase K
STE buffer	Sodium Tris EDTA
STR	Short Tandem Repeats
TBE	Tris Boric Acid ETDA
TE	Tris EDTA
TEMD	Tetra Methyl Ethylene Diamine
UCSC	University of California, Santa Cruz
UV	Ultra Violet Radiation
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Publications

Sajida Rasool , **Jamshaid Mahmood Baig**, Abubakar Mouwia, Ilyas Ahmad ,Maria Iqbal,Syeda Seema Waseem,Maria Arif,Uzma Abdullah, Ehtisham Ul Haq Makhdoom, Emrah Kaygusuz,Saif Ul Haq, Asif Mir, Muhammad Tariq, Stefan Honing, Shahid Mahmood Baig,Peter Numberg,Muhammad Sajid Hussain." An update of pathogenic variants in *ASPM, WDR62, CDK5RAP2,STIL ,CENPJ* and *CEP135* underlying autosomal recessive primary microcephaly in 32 consanguineous families from Pakistan***".*Mol Genet Genomic Med***./doi.org/10.1002/mgg3.1408,2020

Acknowledgement

In the name of **Allah**, the most gracious, the most merciful. All blessings and peace upon the **Prophet Muhammad** (s.a.w). Alhamdulillah, all praises to almighty **Allah** due to which I have been able to accomplish my PhD research work. My profound gratitude goes to my supervisor, **Dr**. **Asif Mir** (Associate Professor Department of Biological Sciences). His assistance, guidance, and feedback have been very important throughout for this dissertation.

I am greatly thankful to my Uncle, **Prof. Dr. Shahid Mahmood Baig (*SI*)** for his moral and academic support. His kind advice and guidelines made things happen in the perfect way.

I am also thankful to **Dr. Arshad Malik**, Chairman Dept. Of Biological Sciences in providing me the departmental facilitation.

I thanks to all the faculty member of Department of Biological Science, International Islamic University Islamabad Pakistan.

Special thanks to my friends, **Dr. Muhammad Abdul Rab Faisal Sultan**, and **Dr. Sajid Hussain**, in provding support and sincere guidance in data analysis and compilation of thesis.

My lab fellows, especially **Dr. Muhammad Ilyas**, **Mr. Abrar** and **Mr**. **Muzamil**, facilitated me in the lab.

I would like to express an eternal appreciation for my parents and family members who have always been there for me. The unconditional support and patience of my caring and loving father and mother.

I would also like to thanks my loving and caring wife **Zakia**, my cute son **Ibrahim** and little angel **Fatima** for being extremely supportive to me throughout this entire process and has made countless sacrifices to achieve this milestone. They have continually provided the requistie breaks from philosophy and the motivation to finsih my degree with expediency. Finally, special and profound thanks to my brothers **Shoaib**, **Junaid**, **Jawad** and sister **Zaib un Nisa**, who offered invaluabe support.

I am thankful to all the families who participated in this study. Their cooperation made the course of the study smoother and this study would have been impossible without their support.

Abstract

Genetic disorders are the leading cause of physical disabilities due to high rates of consanguinity, across the world. In Pakistan there is a tradition of consanguineous marriages, resulting in significant increase in the prevalence of genetic disorders i.e., mental retardation and primary microcephaly. Mental retardation, also known as intellectual disability (ID), is categorized by considerable limitations in both adaptive behaviors and cognitive function, developing before the age of 18. Whereas, dysequilibrium syndrome (DES) which is due to a mutations in the *VLDLR* gene, is a heterogeneous autosomal recessive genetic disease characterized by moderate to profound mental retardation. Primary hereditary microcephaly (MCPH) is a neurodevelopmental disorder characterized by reduction in brain volume. Additionally, microcephaly phenotype is linked with centrosomal proteins, pointing the importance of centrosome during embryonic brain development. This genetically heterogeneous disorder accompanied with mild to profound intellectual disability. Next generation sequencing (NGS) technologies have accelerated the process of identifying variants responsible for disease causation. In the current study, whole-exome sequencing (WES) was used as a method of choice because of the heterogeneous nature of disorders. Variants obtained from WES were prioritized based on their inheritance pattern, disease association, functional relevance and pathogenicity score predicted by different bioinformatics tools such as CADD, Mutation Taster and PolyPhen-2. All the variants were confirmed through Sanger sequencing to counter the errors. Total of 9 families, 5 segregating the phenotype of MCPH and 4 families manifesting ID, were selected from various regions of Pakistan and studied by using different genomic approaches. Out of 5 primary microcephaly families, pathogenic novel mutations of *ASPM* was identified in 4 families and remaining one family carries novel mutation in *STIL*. This gene encodes a centriole protein implicated in regulation of the mitotic spindle checkpoint. Genomic analyses of the four ID families revealed a novel missense mutation of *VLDLR.* Remaining three ID families showed three novel mutations in *GDI1, NSUN2* and *LINS1* which we found associated with mental retardation in Pakistani families. The present study provides further insights regarding genetic heterogeneity of mental retardation and primary microcephaly in Pakistan. Therefore this data will not only be helpful to understand the molecular mechanism involved but to develop improve strategies for identification of potential therapeutic targets.

Chapter 1. INTRODUCTION

1. Background

The field of genetics was revolutionized by the experiments of Gregor Mendel on pea plants, which defined the basic mechanism of inheritance patterns (Bateson and Mendel, 2013). After the completion of Human Genome Project in 2003 and very recently the 1000 Genomes Project, and the Exome Aggregation Consortium (ExAC) have boosted up research in the field of population genetics (Abecasis *et al*., 2010; Auton *et al*., 2015) and, the data obtained by these projects has not only offered information but also act as reference genome for polymorphisms and mutations, consequently enabled and facilitated researchers to indicate the genetic basis of inherited diseases in human.

1.1 Human Genome

Human genome comprises of double helical DNA molecules, composed of four different kinds of nucleotides; Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). These nucleotides are attached through a strong phosphodiester bond, whereas complimentary nitrogenous bases (A -T and C - G) are linked through hydrogen bond, in the DNA strand (Watson and Crick, 1953).

1.1.1 Genomic Variation and its Types

Genetic variations are the changes in the DNA base sequence of an individual. Rearrangements of large chromosomal fragments (inversions, translocations), copy number variations (CNV), single nucleotide variants (SNVs) and short Indels (insertions or deletions), are various types of genetic variations. SNVs denote single nucleotide differences among different individuals and if the variation frequency of minor allele in a particular populationis ≥1% then the phenomenon is called as single nucleotide polymorphism (SNP) (Fay *et al*., 2001). SNPs, in human genome, are the most common genetic variations. SNVs are declared as a mutant and 85% of known mutations occur in the coding region of the genome. Indels and Substitutions are two main types of mutations. Indels are insertion and deletion of small fragments (1-50 bp) in genome (Fay *et al*., 2001). The expression of these variations can be beneficial, harmful or neutral, causing a certain type of genetic disorder (Feuk *et al*., 2006). Frame shift mutations arise from indels producing non functional protein. Substitution of one nucleotide with another may result in aberrations in the amino acids sequence encoding protein.

To substitute a nucleotide with another can result in abnormal codon sequence that consequently code for another amino acid causing missense mutation or it may cause non-sense mutation by inducing a stop codon, causing premature termination of protein (Nakamura, 2009). Rearrangements of chromosomal region and CNVs are forms of structural variants (Haraksingh and Snyder, 2013; Feuk *et al*., 2006). Whereas, CNVs exhibit increase or decrease in copies/number of a particular chromosomal region. There are more than 200,000 entries in the human genome mutation database while 20 % of them comprises of indels.

1.2 Human Genetic Disorders

Human genome contains roughly 20,000-25,000 genes that are involved in the expression of our cell's proteins (International Human Genome Sequencing Consortium, 2004). Monogenic disorders exhibit familial inheritance and can manifest as autosomal dominant or recessive, and sex-linked inheritance.

Analysis of pedigrees with multiple affected individuals aids in determining the inheritance pattern of monogenic disorders. Monogenic diseases arise as a result of variations in a single gene having significant impact on the phenotype, in contrast with polygenic inheritance in which variations in multiple genes, each responsible for a small effect on the disease etiology and also altered by environmental factors (Lander and Schork, 1994). In single gene disorders, a specific variation may alter one or both copies of an autosomal gene producing recessive or dominant inheritance pattern respectively within families.

Variations in the x and Y chromosome exhibit as X-linked or Y-linked inheritance. (Chial, 2008). A particular pathogenic mutation may cause a specific phenotype in patient. So, there is no specific inheritance pattern in such kind of diseases. The completion of human genome prompted the researchers to shift their focus from single gene disorders to polygenic inheritance. The impetus behind moving towards polygenic inheritance are many.

1.3 Consanguinity and its relevance with Autosomal Recessive Disorders

Marriages among close cousins or relatives are called consanguineous marriages. Though such kind of marriages are restricted in some regions of the world, but still 20% of the global population practice it (Bittles, 1994), as introducing someone to the

family, out of their blood relatives, is regarded an unsecure act by interfering family's privacy and of structure.. The probability of occurrence of disease causing rare variants is more in offspring of consanguineous marriages (Modell and Darr, 2002). Autosomal recessive disorders are strictly related with lineage and almost 69% of the families with first cousin marriages have genetic disease with recessive fashion of inheritance (Hamamy *et al*., 2007). Western and US Societies practiced such marriages during the mid-nineteenth century (Bittles, 2003; Ottenheimer, 1996). Marriages between biological relatives are still largely practiced in various Muslims countries (El-Kheshen and Saadat, 2013; Khoury and Massad, 1992). Whereas, in Pakistan the proportion of consanguineous marriages is one of the highest in the world i.e. almost 76% with more than 80% among first cousins (Hussain and Bittles, 1998; Hashmi, 1997).Due to reasons such as high traditional preference for consanguineous marriages, low literacy rate, unawareness, socio-economic burden and large family sizes has greatly enhanced the incidence of genetic disorders and, thus increased the probability.

In Pakistan, a strong traditional preference for consanguineous marriages, large family sizes, unawareness and low literacy rate considerably increase incidence of genetic disorders and increase the likelihood for identification of novel disease genes and their causative variants.

1.4 Intellectual Disability

Intellectual disability (ID) is considered as significant limitations in both adaptive behaviors and cognitive function, developing before the age of 18 (Schalock *et al*., 2010). Intelligence quotient (IQ) is the clinical measure for limitations in cognitive function. A score of 70 or below on testing scale indicates intellectual impairment (Diagnostic and Statistical Manual of Mental Disorders, DSM-V).

Depending upon severity of the presenting phenotype, ID can be ordered into mild, moderate, severe, profound and unable to detect, considering the level of supplemented help and IQ scores. (Luckasson and Reeve, 2001). Currently, to categorize the severity of MR/ID, more attention has been given towards adaptive functioning like social, conceptual and practical skills) (Schalock and Luckasson, 2015).

1.5 Etiology of ID/MR

The etiology of ID/MR is complicated, comprising genetic and environment risk elements; the latter includes exposure to toxins and certain diseases, and perinatal complications (Srourand Shevell, 2014). Presently, various genetic risk factors of Syndromic-ID have been elucidated, such as single gene mutations in Rett and Coffin–Lowry and Fragile X syndromes, and chromosomal abnormality in Down syndrome (Delaunoy *et al*., 2001; Gecz *et al*., 1996; Gibbons and Higgs, 2000). On the contrary, understanding the etiology of NS-ID, is less advanced, due to a combination of complex environmental and genetic causative factors, difficulty in phenotyping of affected individuals and absence of sufficiently latest research technologies (Wieczorek, 2018).

1.6 Incidence of ID/MR

The incidence of ID/MR is about 1-2% of the world population with almost 60% of individuals with ID having a severe core activity restriction (Maulik *et al*., 2011). Previously, due to high incidence of ID in males compared to females, mostly S-ID cases were studied with a focus on X-chromosome (Stevenson and Schwartz, 2009). Whereas, due to reproductive disadvantage in affected individuals, Autosomal-dominant inheritance of NS-ID is comparatively rare (Kaufman *et al*., 2010).

1.7 Classification of MR

Mental Retardations is classified as syndromic and non-syndromic intellectual disability. Pattern of inheritance of intellectual disability can be autosomal (dominant/recessive) and X-linked. Autosomal recessive mode of inheritance is responsible for approximately 25% of the total genetic cases (Higgins *et al*., 2004). Micro-deletions and duplications in chromosome have also been documented as the cause of ID. Autosomal dominant/recessive, and X-linked ID are heterogeneous and further categorized into syndromic intellectual disability (S-ID) which is characterized by either malformations, neurological/dysmorphisms anomalies, whereas, non-syndromic intellectual disability (NS-ID) is only accompanied by the defect in cognitive abilities without any additional characteristics (Basel-Vanagaite, 2007).

However, classification of NS-ID is more difficult to diagnose or may be less apparent in some individuals with ID and continues to attract a debate as subtle physical

dysmorphisms, and neurological and neuropsychiatric disorders (Valnegri *et al*., 2012). Nevertheless, in children and adult age groups, about50% ID is considered non-syndromic with no visible physical signs (Gecz and Haan, 2012).

1.8 Genes Involved in ID/MR

According to SysID database [http://sysid.cmbi.umcn.nl/, 654 mutated genes causing 535 diseases related to intellectual disability, including 310 candidate genes. Autosomal recessive ID (ARID) appears to be more heterogeneous and it is expected that more than 3000 genes are involved in ID (Study *et al.*, 2017).

1.9 VLDLR and its relevance with Mental Retardation

Dysequilibrium syndrome (DES) is a heterogeneous autosomal recessive genetic cause characterized by moderate-to-profound mental retardation, cerebellar hypoplasia, non-progressive cerebellar ataxia, and delayed walk (Rasmussen, 1985). According to Boycott *et al*., 2005 DES is due to mutation in the *VLDLR* gene, present on chromosome 9. *VLDLR* is a member of low density lipoprotein receptor (LDLR OMIM 606945) gene family and also take part in brain reeling signaling pathway. VLDLR directs the migration of neuroblast during the development of cerebellum and cebral coxter. T (Trommsdorff *et al*., 1999; Tissir and Goffinet, 2003). VLDLR together with ApoER2 has a vital role in migrating radial neurons present in the brain. Due to the absence of these receptors, the cortical neuronal sheets are inverted, resulting in failure of Purkinje cell migration, consequently cerebellar growth is stopped (Trommsdorff *et al*., 1999).

1.10 Microcephaly

The term microcephaly consists of two Greek words, “micro”from “*mikros*”, mean small, and “cephaly”from “*Kephale*”, mean head. Microcephalic patients exhibit small head size with prominent phenotype (Woods, 2004).

1.11 Etiology of Microcephaly

The primary etiology could be: a failure or reduction in neurogenesis of neurons, chromosomal disorder, pathogenic events prior to birth (eg, prenatal formation of porencephalic cysts due to ischaemia/hypoxia, or rare mitochondrial mutations), and occasionally at very early onset of degenerative process like Aicardi-Goutieres syndrome (Woods and Parker, 2013). There are some environmental causes like;

intrauterine infections(first trimester cytomegalo virus (CMV) infection), fetal alcohol syndrome, malnutrition or hypothyroidism are some factors that must be excluded for diagnosis of autosomal recessive primary microcephaly, because these factors also cause microcephaly with mental retardation (Woods *et al*., 2005).Drugs taken during pregnancy; maternal phenylketonuria; birth asphyxia and prenatal radiation exposure forming the cause of the rare secondary microcephaly (Jackson *et al*., 1998). The autosomal recessive primary microcephaly has only a genetic cause without environmental factors.

Among different populations around the world, many of the causative genes (MCPH1-MCPH25) for the 25 loci mapped till now, code for kinetochore/spindle pole proteins or centriole/centrosome, are involved in cytokinesis, centrosome maturation, centriole biogenesis, kinetochore and centromere function. This represents that centrosome malfunction is one of the primary causes of MCPH (Jayaraman *et al*., 2018; Nano and Basto, 2017).

Additionally, microcephaly phenotypes are linked with centrosome proteins (POC1A, PCNT, CEP63, NIN) developing a strong genetic relation between centrosome malfunction during development and the aetiology of microcephaly (Saade *et al*., 2018).

1.12 Incidence of Microcephaly

MCPH is rare, with prevalence estimates between of 1:30,000 and 1:2,000,000 in European population, but much higher (One in 10,000) in some parts of Pakistan due to the high consanguinity rates (Woods *et al*., 2005).

MCPH loci are various, but they are individually rare. More than 86% of the registered families exhibit mutations in one of the three genes (MCPH1, ASPM, WDR62), and most of these variants are nonsense, with loss of function. However, no mutation was observed in known genes responsible for 20 to 30% of Pakistani and 40-60% European and American MCPH cases. Therefore, more genes and mechanisms associated with congenital microcephaly are supposed to exist and still to be identified (Zaqout *et al*., 2018).

1.13 Classification of Microcephaly

There are two types of microcephaly; primary and secondary microcephaly. Primary microcephaly is a static anomaly in which the growth of the brain is reduced during pregnancy (develops at prenatal stages), whereas in secondary microcephaly the brain size is normal at birth but subsequently fails to grow normally due to progressive neuro-degeneration. In contrast to primary microcephaly, the secondary microcephaly develops at postnatal stages (Woods, 2004).

1.14 Primary Hereditary (MCPH)

Primary hereditary microcephaly is a neurodevelopmental disorder characterized by reduction in brain volume, with an occipital-frontal circumference ≤ -2 standard deviations (SD) that of the age and sex-matched population at birth. This genetically heterogeneous disorder leads to mild to profound intellectual disability (Woods and Parker, 2013) with normal height (except in minor cases where the MCPH1 gene is involved), weight and neurological functions (Cox *et al*., 2006; Nicholas *et al*., 2009). MCPH is inherited in autosomal recessive form and heterozygous carriers show no symptoms of disease. Through regular ultrasonography, primary microcephaly is detectable before 36th weeks of gestation period (Zaqout *et al*., 2017). Magnetic Resonance Imaging (MRI) of affected individuals presents normal brain, but smaller and simplified cortical area (Pang *et al*., 2008).

1.15 Genes Involved in Microcephaly

To date, variations in 25 different genes are known to be causative variants for autosomal recessive primary microcephaly (MCPH; OMIM phenotypic series: PS251200) (Perez, *et al*., 2019).

1.16 Strategies for disease gene identification

Diverse strategies were applied to locate disease causing variants in the genome, responsible for various disorders.. The last decades of the century saw positional cloning as the predominant approach for finding disease causing variants. Homozygosity mapping by Linkage analysis was used to elucidate the candidate chromosomal region or gene of affected individuals of large families (Collins, 1995). Identification of new genes causing ID is difficult due to syndromic condition of the

diseases in many families (Kürekçi and Dinçer, 2013). A brief introduction of the strategies involved in gene identification is discussed below.

1.17 Homozygosity Mapping

To locate the recessive variants present within inbred families, homozygosity mapping is used as a standard technique (Sukumaran *et al*., 2017; Martin *et al*., 2014). For homogygosity mapping, the DNA makers (either in the form of a single base pair changes called single nucleotide polymorphism (SNP) or variable number of repeats of two or more base pairs called microsatellites) used are natural variants present in the genome of individual (Vink and Boomsma, 2002). The methodology consists of genotyping, linkage analysis and targeted sequencing of samples. Prior to the availability of reference human genome, the region of interest was usually cloned for the targeted sequencing. These techniques were named as positional cloning, where homozygosity mapping was a crucial part (Teare and Santibanez, 2014).

1.18 Next generation sequencing technologies (NGS)

To establish a novel genetic basis of disease, in a research setting, DNA sequencing with next-generation sequencing technologies has proven to be an effective alternative to locus-specific and gene-panel tests (Rauch *et al*., 2012).

To resolve the limitations of traditional gene identification methods, Next generation sequencing techniques (NGS) were introduced in the beginning of this century. These technologies have advantage over the Sanger sequencing method to sequence the whole human genome in considerably shorter time (Hedges *et al*., 2009).

Next generation sequencing (NGS) also known as deep sequencing, massively parallel DNA sequencing or high-throughput sequencing. The NGS technologies at hand has shared features with massively parallel sequencing of DNA involving clonal amplification or single molecules. Specific separation of DNA molecules is carried out in a flow cell through successive ligation of oligonucleotide or repetitive cycle of polymerase chain reaction (PCR) (Behjati and Tarpey, 2013; Voelkerding *et al*., 2009). During synthesis step, multiple sequencing techniques have diverse mechanism and protocol. Applications of NGS are various; consist of target panel sequencing whole-exome sequencing (WES) and whole-genome sequencing (WGS).

1.19 Whole-Exome Sequencing

Advancement in the technique of whole exome sequencing and analytical approaches enabled the researchers to detect the disease causing variants in common as well as rare genetic disorders (Majewski *et al*., 2011). In Patients with suspected genetic disorders, whole-exome sequencing is the method of choise for the identification of molecular defects (Yang *et al*., 2013). Whole exome sequencing (WES) is selectively used for the sequencing of coding regions (180,000 exon) (Hedges *et al*., 2009). since most of the single nucleotide rare variants responsible for inherited diseases are found in the exonic part of the genome (Botstein and Risch, 2003) and it has the more likelihood of identification of rare single nucleotide variants (Bamshad *et al*., 2011).

1.20 Research objectives

The prime objectives of current research is to uncover the underlying molecular genetic causes of non-syndromic autosomal recessive mental retardation (NS-ARMR) and primary microcephaly (MCPH).

Experiments were designed by using a blend of linkage analysis and whole-exome sequencing technology to search out novel pathogenic variants in affected cases of NS-ARMR and MCP from various regions of Pakistan.

The specific design of current research was:

1. Characterization of Pakistani families segregating non-syndromic autosomal recessive mental retardation (NS-ARMR) and primary microcephaly (MCPH) by using molecular approaches.
2. To investigate novel mutant loci and underlying causative gene variants in NS-ARMR and MCPH.
3. Establishing carrier screening tests for the families at risk.
4. To propose improved treatment strategies based on molecular knowledge generated in this study.

Chapter 2: MATERIAL AND METHODS

2. Material and Method

2.1 Collection of Data

Approval of genetic and clinical studies, carried out in current study, was taken from ethical department committee at International Islamic University Islamabad (IIUI).The ethical department committee of international Islamic university Islamabad provided the approval for collection of data for genetic and clinical studies for this research.Initially the researcher identified affected families in the field study of respective areas. The members of these families showed strange behavior, difficulty in learning, Reduced head circumference and decline in intelligence without apparent features of syndrome. All the families were contacted in their respective residence. Disease history of families were recorded by interviewing parents and pedigrees were drawn accordingly in accordance to published methods (Bennett *et al.*, 1995). With prior family consent, local hospital physician was consulted for concerning basic clinical information of proband and diagnosis confirmation. The blood samples were drawn for genomic DNA purification. Circles and squares were used to represent the females and males respectively, in the pedigrees. Whereas, shaded circles and squares described affected individuals, and unshaded showed healthy individuals. On the other hand slanting line over the circle and square represented the expired individuals. Double lines represented the consanguinity, while each generation was represented by Roman numbers. Pedigree helped us in understanding the inheritance pattern of disease for each family and affected individual.

2.2 Blood Sampled

Blood samples were taken in 10ml syringes from affected and healthy individuals, and shifted to 10ml K EDTA vacutainer. Collected samples will be labeled with different identification number and stored at 39 °C, until further proceeding.

2.3 Extraction of DNA

Extraction of DNA from blood samples of families were processed by following Organic method of DNA extraction to recover the genomic DNA.

2.3.1 Extraction of DNA by Phenol Chloroform method

For DNA extraction, standard chemical solutions (A, B, C and D) were used. The composition of following chemicals are mentioned in annexure.

2.3.2 Method

I. In 1.5 mL eppendorf tube half mL blood was taken with same volume of solution A. Placed solution mixture for 15 to 25 min at room temperature and kept vortexed regularly to get mix the solution A in blood. After that, sample mixtures centrifuged for 30 sec at 13,000 Rpm.

II. Supernatant discarded away after centrifugation. Carefully suspended and vortexed the pallet in 500 µL Solution A for 20min.

III. Again centrifuged the solution mixture for 30 min at 13,000 Rpm.

IV. After centrifugation carefully discarded the supernatant. Then 500 µL Solution B, 30 µL proteinase k, and 10 µL of SDS solution were added in eppendorf tubes and vortexed for 15 to 20 sec to mixed solutions in pallets thoroughly.

V. The mixture was incubated at 37 °C for 7 hours. Added 500µL solution D in solution mixture after the incubated time and centrifuged the solution mixture for 30 sec at 13,000 Rpm.

VI. Took new eppendorf tube and shifted about 700 µL supernatant volume. Later on, same volume of solution C was poured and mixed gently, and left the solution mixture for 4 to 5 min.

VII. Centrifuged the mixture for 30 sec at 13,000 Rpm.

VIII. Again took new eppendorf tube, added 900 µL supernatant and then poured 800 µL cooled Iso propyl alcohol. Left mixture for 8 to 10 min to precipitate the DNA.

IX. Centrifuged the mixture for 30 min at 13,000 Rpm, Added up 60 µL, 70% methyl alcohol after discarded the supernatant. Then added up 30µL absolute alcohol to further purify the DNA.

X. Finally stored the DNA at 4 degree centigrade by adding 200µL TE buffer.

2.4 DNA Analysis

2.4.1 Gel Electrophoresis

Genomic DNA were analyzed by preparing 1% agarose gel in 60 mL 1X TBE buffer. To envisaged the DNA 2% EtBr solution mixture was added up. After that gel was transferred in cast and waited 15 to 20 min till the gel got solidified. Then by preparing 2% agaorse gel, the amplified resultant PCR product was investigated on it. Before loading the DNA samples within the gel wells, prepared the samples of DNA for PCR by adding 2 to 3 µL bromophynol blue. Electrophoresis was carried out in 1X TBE buffer at around 120 electric potential for 20 min. After electrophoresis, the gel was putting on Dolphin DOC to envisaged the result of PCR.

2.4.2 Whole Exome Sequencing

Whole exome sequencing is a tool of next generation sequencing technology to validate the changes in coding and non coding regions of genes. In Mendelian inheritance disorders, this technology exhibits the coverage for over 95% coding region, consequently responsible for 85% genetic mutations (Rabbani *et al*., 2014). Whole Exom sequencing was performed in the selected families.

The human genome consists of 1% exome with approximately 30 million bp (Ng *et al*. 2009). The entire set of exome as a whole exome sequencing target is used by a probe hybridization method (Gnirke *et al*. 2009). In present research WES was performed in 5 MCPH and one MR family (MCP158, MCP159, 15 MCP160, MCP163 and MCP166) and a family with NS-ASRMR. DNA Samples (1 µg) was taken from every affected person from each family, cause of disorder was consider to acquire the sequence. Before the enrichment step, sequence library was generated by using standard procedures. Laterally, a paired-end2 100- bp procedure was used to conduct the sequencing on an Illumina HiSeq 2000. The WES data was analyzed by using the existing channel (www.varbank.ccg.uni-koeln.de). Variants of great priority were confirmed further through Sanger Sequencing analysis technique.

2.5 Sanger Sequencing

After characterization of significant possible variants, With in the pedigree, variants from WES were confirmed using this technology. For the amplification of desired

region of gene, to designed flanking primers primer-3-plus (www.primer-3-plus.com/cgi-bin/dev/primer.3.plus.cgi) tool was used.

2.6 DNA Polymerase Reaction

To amplify the target region, within the family, each individual template DNA was taken to carry out the PCR process.PCR was performed by adding DNA (50ng), 0.025microliter of platinumTm Primer (0.3 µM-reverse & forward), Buffer, MgCl2 (2mM), 1M Betaine, dNTPs (1Mm), polymeraze enzyme in 20µl reaction mixture.

After amplification in PCR machine, regions of DNA run on 1.9% agaroze gel. In order to check the specificity and size of amplified fragments, DNA Ladders were used.

2.7 Enzyme Purification

To enzymatically purify the PCR products, experiment was performed by adding 0.35µl exo-nuclease-I (19 U / µl from thermoscientific) & 0.9µl of thermosensitive basic phosphatase and 1µL of H2O along with 9µL PCR outcome, after amplification.Two successive cycles at 38 °C & 79 °C for 25 and 15 min, respectively, were observed to incubate the mixture product.

2.8 PCR Sequencing Reaction

For sequencing PCR reaction, purified outcome of PCR was homogenized with 2 µl reaction buffer (6X – 1 µl). Ready terminator reaction mixture, 5 µl H2O make 9 µl final volume. Reverse (1µl of 2mM) and forward (1µl of 2mM) primers reactions were performed separately.

Reaction conditions for PCR completed in 25 rounds. Every round comprised of given three phases:

I. Denatured the product for 10sec at 95 °C
II. Hybridization of primers took place for 5 sec at 52 °C
III. Amplification performed for 5 min at 58 °C

2.9 Ethyl-alchol Precipitation

The resultant product of PCR sequencing reaction performed by below mentioned procedure;

I. 29µL Ethyl-alcohol/sodium acetate added and mixed in each 9µL reaction mixture
II. Centrifuged the reaction mixture at 3200rpm for 30min at 5 °C
III. Sample plates poured and dried off by using tissue paper for 5min
IV. In every well, 45µL ethyl-alcohol (70%) was added up
V. Centrifugation was performed at 3200rpm for 9min at 4 °C
VI. Sample plates poured and dried up by using tissue paper for 5min
VII. Sample plates were centrifuged for 2min at 5000rpm
VIII. Re-suspended the product samples in 10µL H2O

AB13730xl DNA analyzer was used to sequence the prepared samples. Sequenced chromatographs were arranged and analysed by Sequence® V-6.0.

2.10 In Silico Analysis

To predict secondary and three dimensional protein models swiss model and I-Tasser (Roy *et al*., 2010; Zhang 2008) were used. Predicted models reliability was checked using RAMPAGE (Lovell *et al*., 2002). Models were visualized via UCSF Chimera (Pettersen *et al*., 2004).

Chapter 3: RESULTS

3. Results

3.1 Family I

The family was consanguineous with four individuals (III: 1 Father, III: 2 Mother, IV: 1 abnormal daughter and IV: 2 normal son) from Pakistan (Fig 3.1). Autosomal recessive mode of Mendalian inheritance was observed in phenotypes. After obtaining consent blood samples were drawn from all members of this family. .

3.1.1 Clinical features

Proband of this family is a 38-year-old female (IV: 1), physical examination of the patient (Fig 3.2) showed small head circumference of -7 SD, height 53.8 inches, weight 42 kg, facial distortion, normal vision, no motor and neurological delay. She was also found to be manifested by severe intellectual disability accompanied by speech impairment, low communication and social ability, or issues with daily life tasks like dressing or eating. No other anomalies were noted. Patients (IV:1) clinical variables are represented in the table 3.1.

3.1.2 Molecular study

Initially, Whole Exome sequencing analysis of *ASPM* gene in family I, identified frame shift mutation (c.1615_1616delGA) in exon 3 (Fig 3.3). At protein level this deletion leads to change of Glu (GAA) with Arg (AGA) at amino acid position 539 and results in frame shifts of reading frame with the introduction of premature stop codon (p.(Glu539Argfs*15). This frame shift mutation found to be located in the c-terminus of *ASPM* corresponding loci in abnormal members of the family. Parents of abnormal individuals (III.1 & III.2) were obligate carriers. Parallel to Sanger sequencing, Genome wide linkage-analysis was also performed by operating, Human Core Exome 12 v.1.1 array from Illumina (San Diego, CA) by genotyping available family members (patient along with her both parents). Approximately 40 linkage regions were obtained with highest LOD score 1.33. And one of them at chromosome 1 harbors *ASPM* (Fig 3.3).

Table 3.1: Clinical features of patients (IV: 1).

Parameters	**Normal values**	**Values noted in patient**
Weight (kg)	60	42
Height (cm)	14	12
BMI (kg/m^2)	66.4	60.2
OFC (cm)	55	46
OFC (SD)	NA	-7

Note: SD stands for standard deviation, NA means not available

,

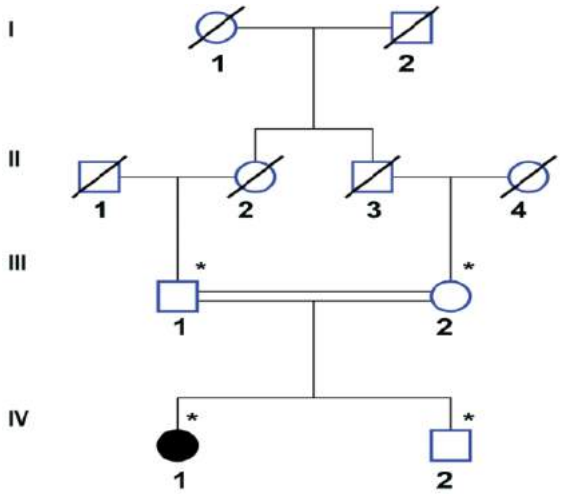

Figure 3.1: Family I pedigree .circle and square represents female and males respectively Square shows males whereas circles show females. Shaded square and circle show affected individuals. The cross lined square or circle shows deceased individual and double line represents the consanguineous marriage. Asterisks indicate individuals subjected for Sanger sequencing.

Figure 3.2: Photograph of affected individual (IV: 1) with sloping forehead and facial distortion

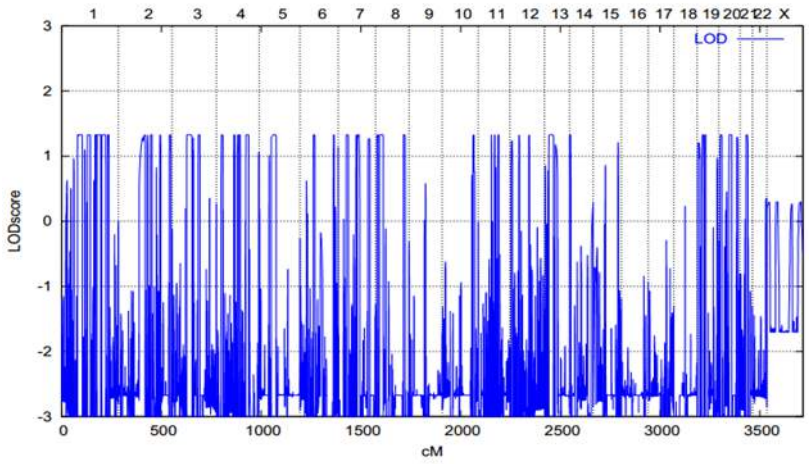

Figure 3.3: Genome-wide graphical view of the homozygous regions obtained by genotyping the available family members. LOD score calculated by ALEGRO was 1.33 at several regions of the genome, particularly at MCPH5 locus where *ASPM* resides.

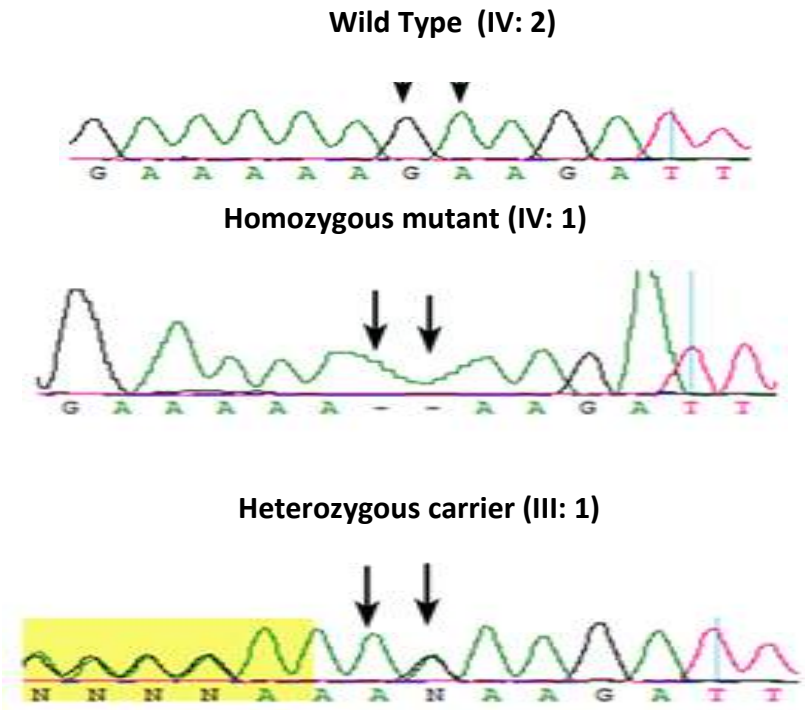

Figure 3.4: Sequence chromatogram of normal individual in upper panel, affected in middle panel and carrier in lower panel showing mutation (c.1615_1616delGA).

3.2 Family II

This consanguineous family consists of five individuals (Fig 3.5). After informed consent, sample of blood from these family members (III:1 Mother, III:2 Father, IV:1 affected son, IV:2 affected son and IV:3 unaffected son) were collected. Autosomal recessive mode of inheritance pattern showed by this family.

3.2.1 Clinical feature

Two abnormal male members of representing family (IV: 1 and IV: 2) manifesting microcephaly and severe intellectual disability (Fig 3.5 and 3.6). Ages of patient's were 31 (IV: 1) and 34 (IV: 2) years during the recording of data. Affected individual (IV: 1) showed head circumference (HC) of -6 SD, weight 58 kg, height 62.2 cm and showed hearing or speech problems, cognitive impairment, hyperactivity and difficulties in coordination and balancing but without any other ocular malformations. The second affected (IV: 2) person showed HC of -8 SD, height and weight was 61.3 cm, 43 kg, respectively. He was blind and depicted developmental delay, microcephaly, failure to thrive, spasticity and ventriculomegaly. Clinical analysis and BMI of the both patients are mentioned in Table 3.2.

3.2.2 Molecular study

To find the genetic determinant(s) of this family, linkage analysis was performed by genotyping 4 individuals (2 affected and both parents) on Illumina (San Diego, CA), Human Core Exome 12 v.1.1 array. Maximum possible LOD score of 2.53 was obtained with homozygous regions at chromosome 1, 5, 7, 8, 9, and 14 (Figure 3.7). One of the linkage regions at chromosome 1 was delimited by rs3008588 (195,987,908 bp, 207.99cM) and rs10900421 (204,646,295 bp, 218.38 cM). This 8.65 Mb, homozygous region contains several genes including *ASPM*. Therefore, sequence analysis of family 2 revealed a deletion of four nucleotide bases in coding region 18 of *ASPM* gene (c.8718_8721deITTTA) detected in either individuals of this family (IV:1 & IV:2 Fig 3.8). Either heterozygous parents shown mutation (III:1 and III:2) . This mutation alters CUU codon (Leu) with CGU codon (Arg) at position 2907 of ASPM, introducing early premature stop codon in subsequent 29 amino acids (p.Leu2907Argfs*30).

Table 3.2 Clinical features of patients (IV: 1) & (IV: 2).

Patient (IV: 1)	**Normal value**	**Values of the patients**
Weight (kg)	65	58
height (cm)	16	154
BMI (kg/m^2)	58.4	61.4
OFC (cm)	50	45
OFC (SD)	NA	-6
Pateint (IV: 2)		
Weight (kg)	58	43
height (cm)	165	152
BMI (kg/m^2)	64.3	66.7
OFC (cm)	50	46
OFC (SD)	NA	-8

Note: SD stands for standard deviation, NA means not available

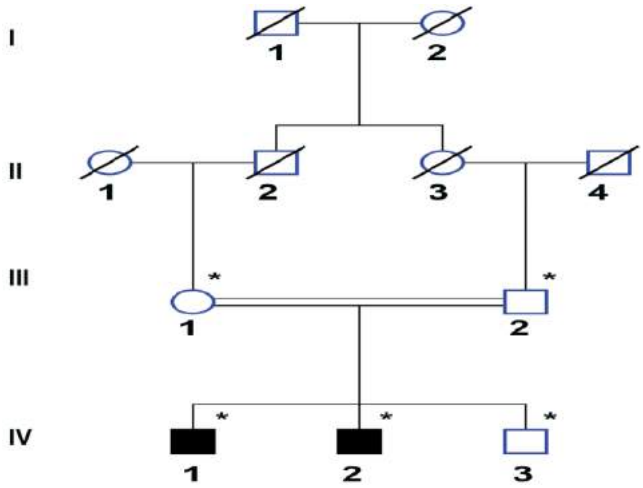

Figure 3. 5: Pedigree of four generations Pakistani family segregating primary microcephaly with two affected members. Asterisks indicate individuals subjected for genomic analyses.

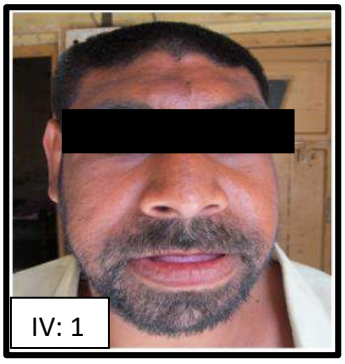

Figure 3.6: Photograph of affected individual (IV: 1) belong to Family II showing slopping forehead.

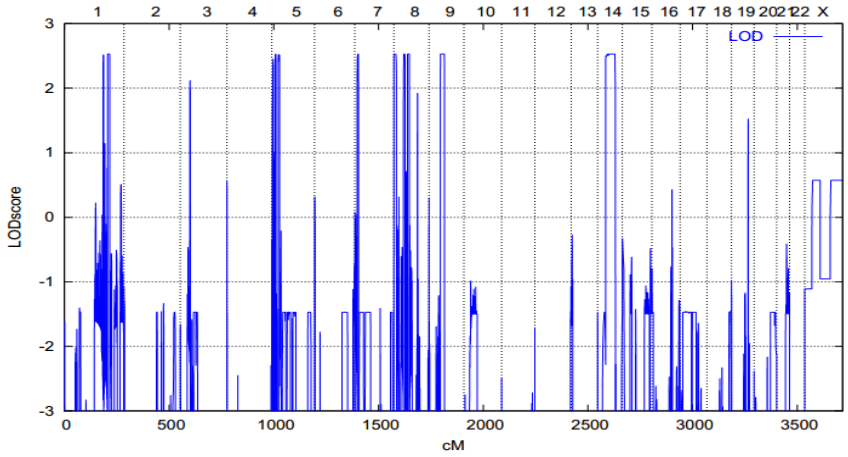

Figure 3.7: Genome wide graphical representation of family II, with a maximum attainable LOD score 2.53. *ASPM* is located at second linkage zone of chromosome 1 demarcated by rs3008588 (195,987,908 bp, 207.99 cM) and rs10900421 (204,646,295 bp, 218.38 cM) harbor *ASPM*. All the physical locations, are according to the NCBI genome build 37 (GRCh37; hg19).

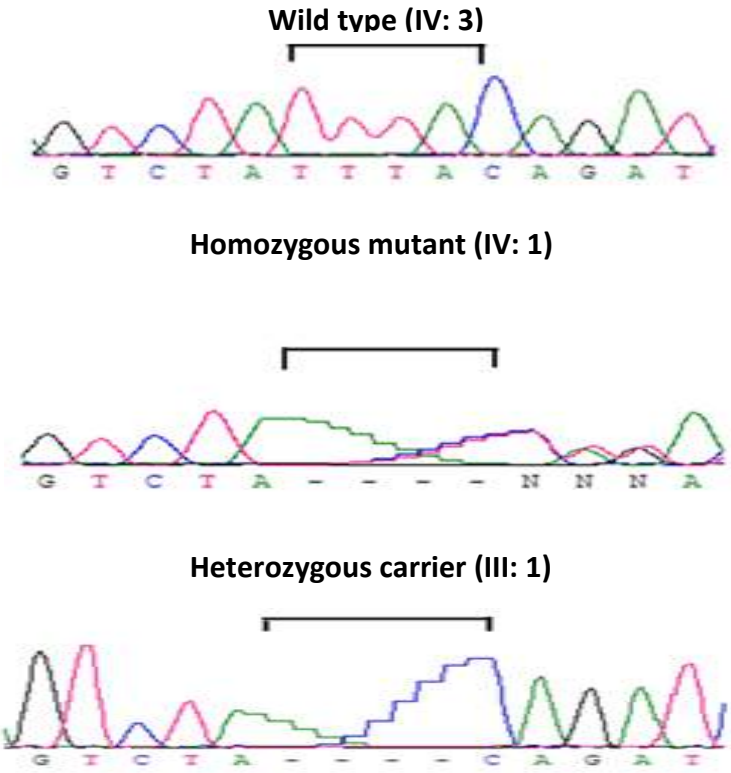

Figure 3.8 Sequence chromatogram of normal individual in upper panel, homozygous mutant in middle panel and carrier in lower panel showing mutation (c.8718_8721delTTTA).

3.3 Family III

Six generations consanguineous family was ascertained and consent was obtained from the concerned person of the family to collect blood samples and to conduct further study (Fig 3.9). Five random DNA samples, from 3 generations, were isolated (Fig 3.9). Mode of inheritance of the disease manifested autosomal recessive pattern.

3.3.1 Clinical feature

The extended pedigree of this family has two patients (V: 1) and (VI: 1) (Fig 3.9). Proband (V:1) is 25-year-old female with a head circumference of 17.6 inches, height 50.3 inches, weight 41 kg, sloping forehead, severe intellectual and social impairment as well as poor self-care skills. Another patient (VI:1) at the time of clinical examination was 9 year's old. Head circumference of this male affected member was 11.8 inches with normal facial features, slopping forehead, poor communication skills, learning problems and repetitive behavior (Fig 3.10).

3.3.2 Molecular study

Whole Exome sequencing, analysis of an affected individual, VI:1 of this family, revealed a nonsence mutation, c.3193C>T; p.(Gln1065*) of *ASPM*. This mutation revealed in exon 13 causes nonsence mutation of *ASPM*, replaces glutamine with stop codon (p.Gln1065*) at 1065 position (Figure 3.11).

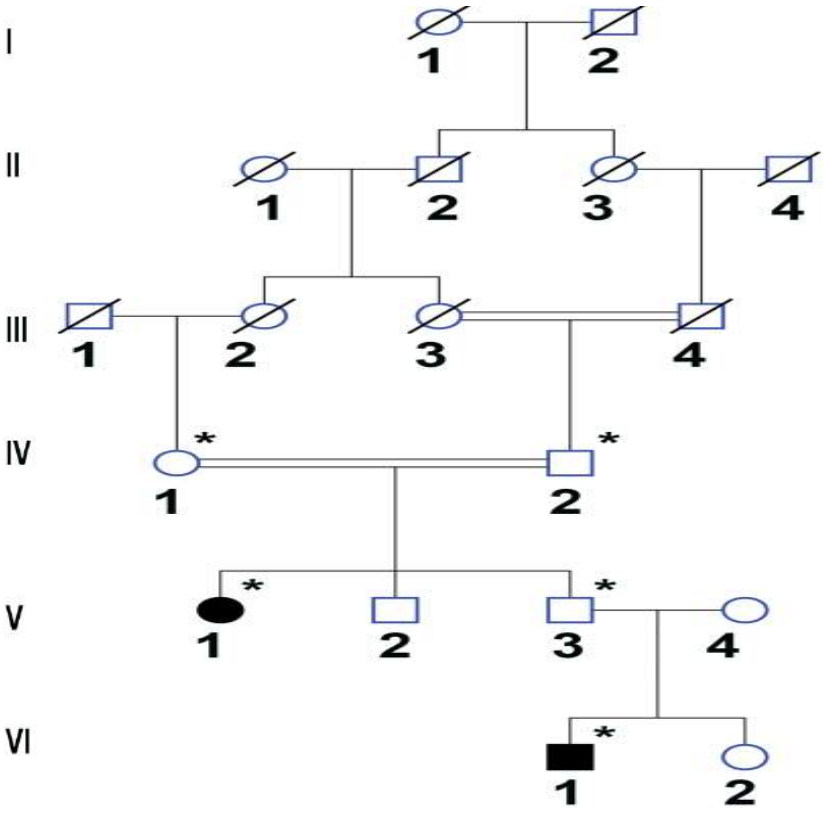

Figure 3.9: Six generations Pakistani pedigree with two affected members born in the fifth and sixth generations. Asterisks indicate individuals subjected for Sanger sequencing.

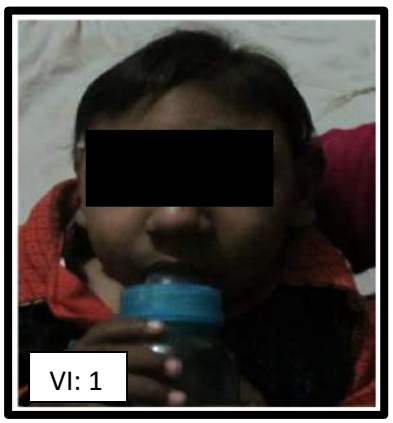

Figure 3.10: Photo of affected individual (VI: 1) showing microcephaly.

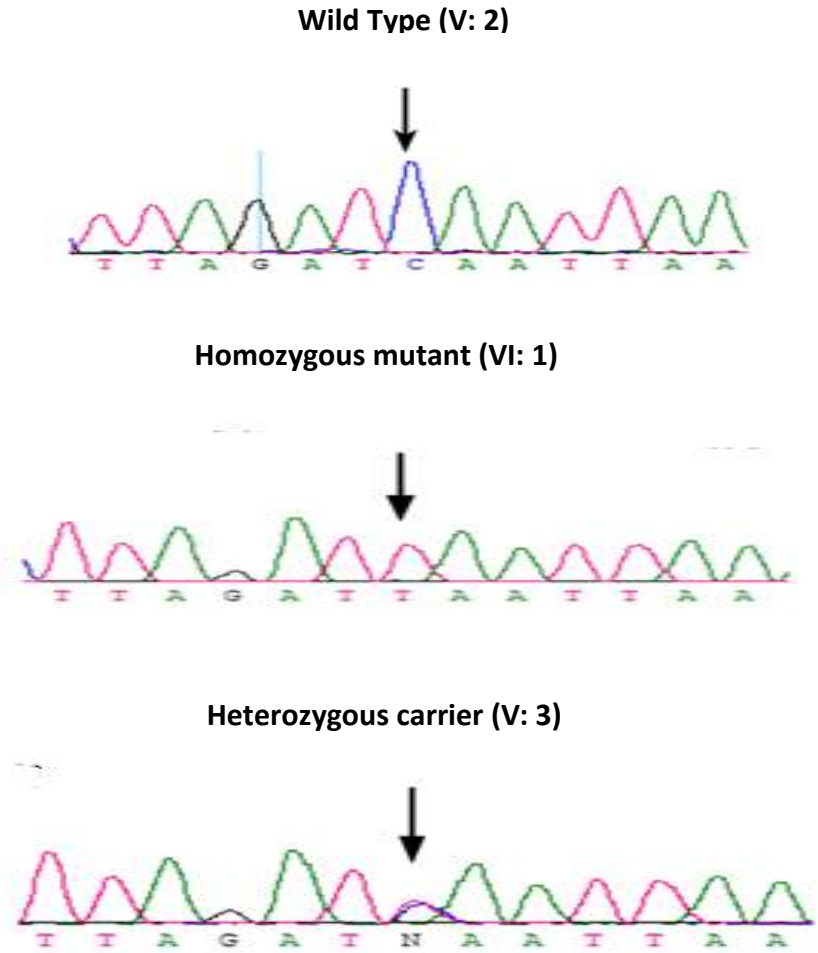

Figure 3.11 Sanger traces taken from part of exon 13 of *ASPM* belong to normal individual in upper panel, wild type (V: 2) in middle panel affected (V: 1) and carrier (VI: 3) in lower panel showing mutation (c.3193C>T).

3.4 Family IV

This family consists of five generations where two affected members were born in the fifth generation (Fig 3.12). A male and a female affected member were born from presumably unaffected parents which depicts autosomal recessive inheritance pattern (Fig 3.12).

3.4.1 Clinical features

Proband (V:1) is a 12-year-old male showing sloping forehead which head circumference of 16.9 inches, weight was 21 kg, height 45 inches, suffer from primary microcephaly and was able to perform everyday task such as eating, dressing, walking, but poor communication and cognitive skills were observed.

One affected female child (V: 3) of this consanguineous family also showed mild intellectual disability (Fig 3.13). She was facing difficulties to manage daily routine task but walked normally. Head was observed sloppy not consider to be normal size, because head circumference of the child was 13.9 inches. Other organs were observed normal and there was significant social interaction.

3.4.2 Molecular study

Initially, genome-wide linkage analysis was performed by genotyping the available individuals with the help of Illumina (San Diego, CA), Human Core Exome 12 v.1.1 array. The parametric curve analyzed by ALLEGRO showed 17 peaks with maximum 2.4, LOD score for such kind of pedigrees (Figure 3.14). Further analysis of the 17 homozygous segments showed that one of these regions at chromosome 1 limited by marker rs9662128 (40,150,156 bp, 68.22 cM) and rs4575030 (54,024,808 bp, 78.28 cM) contains *STIL* gene.

Laterally, individuals DNA was subjected to perform whole Exome sequencing.. Data analysis revealed frame shift mutation due to the duplication of nucleotide T in *STIL*. This duplication of one bp (c.3759dupT) was found to be located in exon 17 (Figure 3.15). At protein level this substitution leads to change of Pro (CCU) with Ser (AGU) At position 1254 (p. Pro1254Serfs*2) and causing frame shift which leads to premature stop codon. In both defective individuals, this mutation was existed in homozygous form/condition. (V: 1, V: 3), whereas, heterozygous state in parents (Fig 3.15).

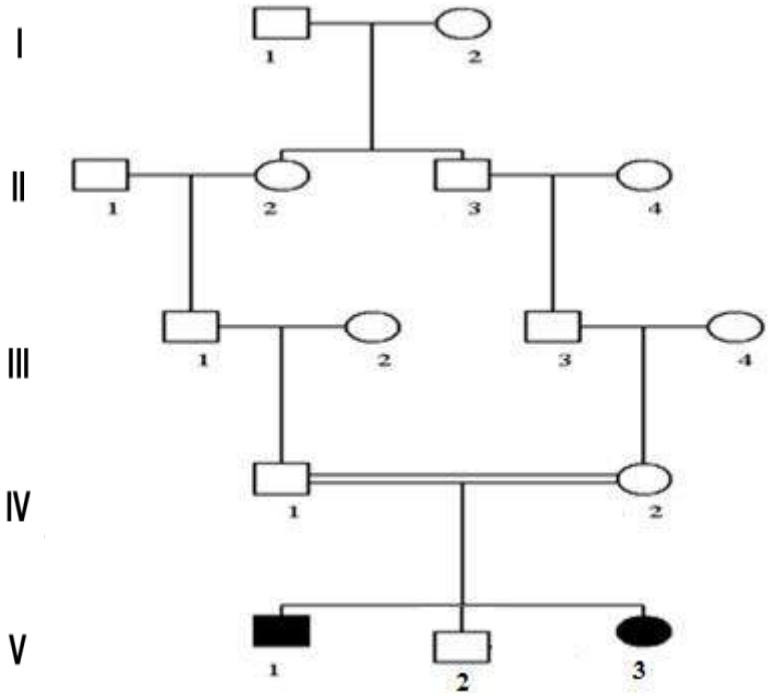

Figure 3.12: Pedigree of Family IV ascertained from remote region of Pakistan.

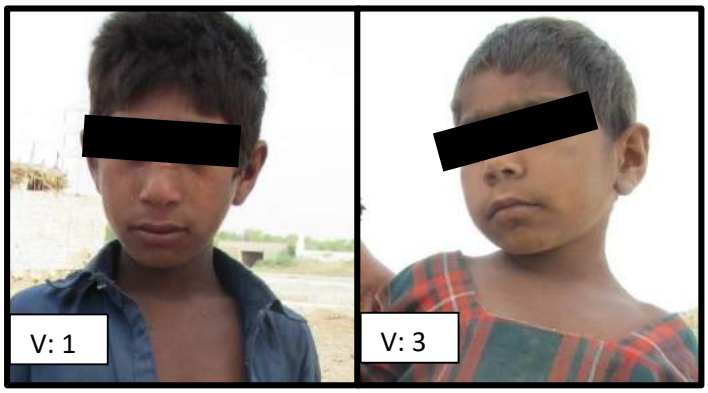

Figure 3.13 Images of abnormal individuals (V:1 boy and V:3 girl) showing slopping forehead which is a prominent feature of the primary microcephaly.

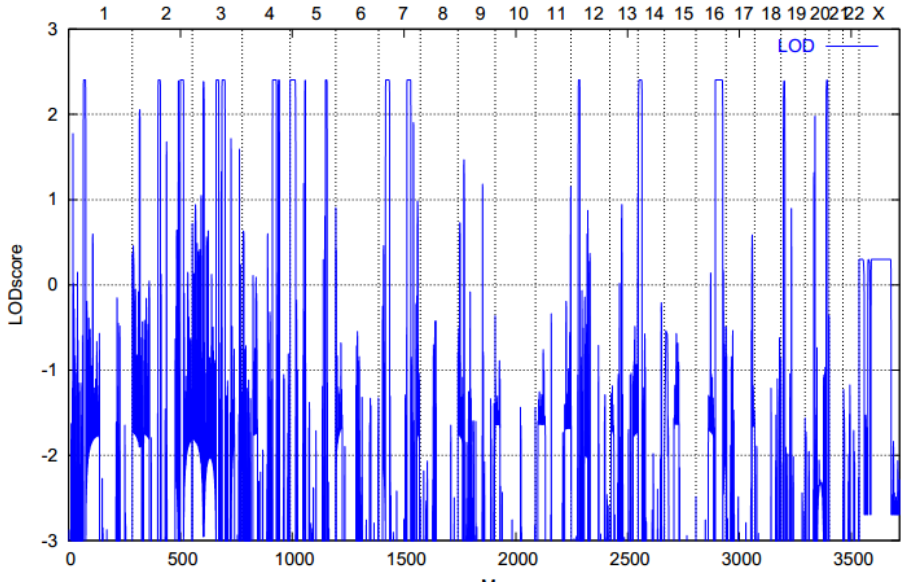

Figure 3.14: Genome wide graphical MCPH 7 locus representation, parametric linkage analyses results and showing LOD scores 2.4, on chromosomes 1, 2, 3, 4, 5, 7, 12, 14, 16, and 20. *STIL* gene resides in the linkage region of chromosome 1.

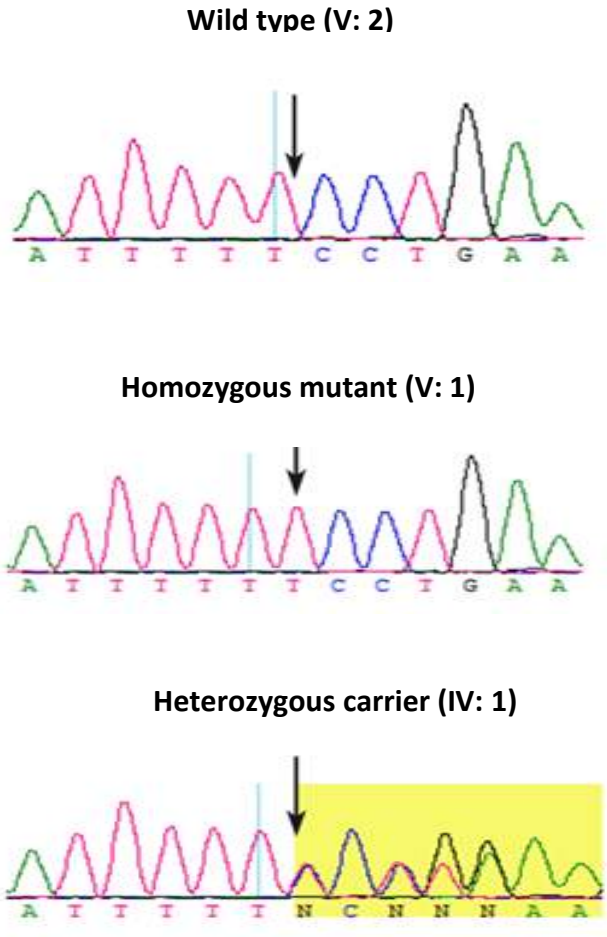

Figure 3.15: Sequence chromatogram of normal individual (V: 2) in upper panel, affected in middle panel (V:1) and carrier in lower panel (IV:1) showing mutation (c.3759dupT).

3.5 Family V

A Consanguineous family belongs to Pakistan where consanguineous marriages are common. This family has unaffected parents (IV:1 Father and IV:2 Mother) with their three abnormal children (V:1 affected son, V:2 affected daughter and V:3 affected son) (Fig 3.16). Blood samples were collected from these five selected individuals after obtaining consent for participation in this study.

3.5.1 Clinical features

At the time of clinical analysis, proband (V:1) of family V (Fig 3.17) was 25-year-old male, with body height 62 inches, weight 55 kg, and head circumference was 17.8 inches. No facial distortion was observed and no evidence for other abnormalities such as delayed development and epilepsy. His social interactions were not normal,. Based on clinical observation, the patient (V:1) was having primary microcephaly with profound intellectual disability.

Second affected individual (V:2) of family V is a 21-year-old female manifesting typical microcephaly along with facial dysmorphism. Further, we noted head circumference of 16.3 inches, height 48.5 inches and weight 46 kg. She was unable to perform daily tasks like eating, dressing, talking, walking. Social interaction was very weak.

Blood sample was drawn from the third male patient age 17 year (V:3) of family V (Fig 3.17). At that time his head circumference was 17.1 inches with sloppy forehead and enlarged face, height 58.9 inches, weight 50 kg, abnormal daily routine life and no interaction with public due to poor communication skills.

The three patients (V:1 affected son, V:2 affected daughter and V:3 affected son) who showed classical MCPH and severe mental retardation were selected for the study of genetic analyses.

3.5.2 Molecular study

Human Core Exome 12 v.1.1 array from illumina (San Diego, CA) was conducted to carry out homozygosity mapping. The calculation of parametric LOD score with ALLEGRO revealed 14 peaks with highest possible LOD score 1.2 (Figure 3.18). One of the homozygous segment at chromosome 1 carry *ASPM*. This linkage region is bounded by the makers rs2500097 (183,765,059 bp,199.19 cM) and rs10800790

(201,644,752 bp, 213.50 cM). Parallel to homozygoisty mapping, WES was also carried out on the DNA of one abnormal member. This analysis showed deletion of two homozygous nucleotides (c.8508_8509delAG) of *ASPM* in exon 18. All three (V:1, V:2, and V:3) abnormal individuals) were homozygous for this variant. This c.8508_8509delGA mutation changed the reading frame, consequently give rise to early premature stop codon, along the downstream 34 amino acids (p.(Lys2837Metfs*34)). This variation was existed in homozygous state in three abnormal individuals (V:1, V:2, and V:3), and heterozygous state in parents (Fig 3.19).

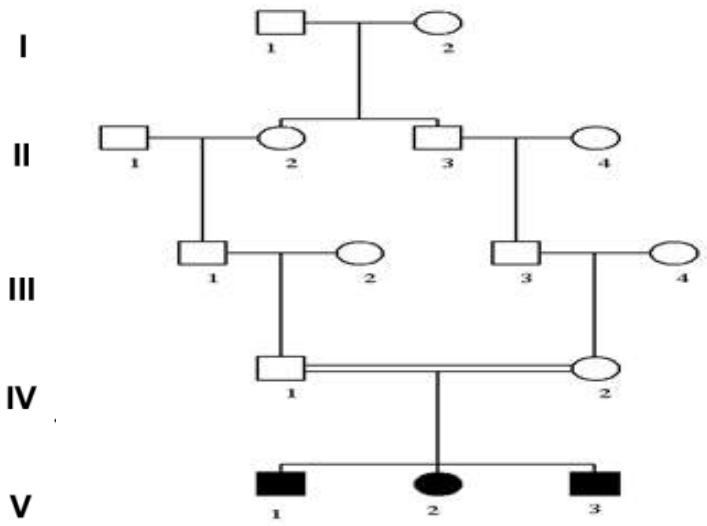

Figure 3.16: Family V pedigree representing autosomal recessive pattern of inheritance.

Figure 3.17: Photographs of abnormal individuals (V:1, V:3 male and V:2 female) showing reduced head circumference.

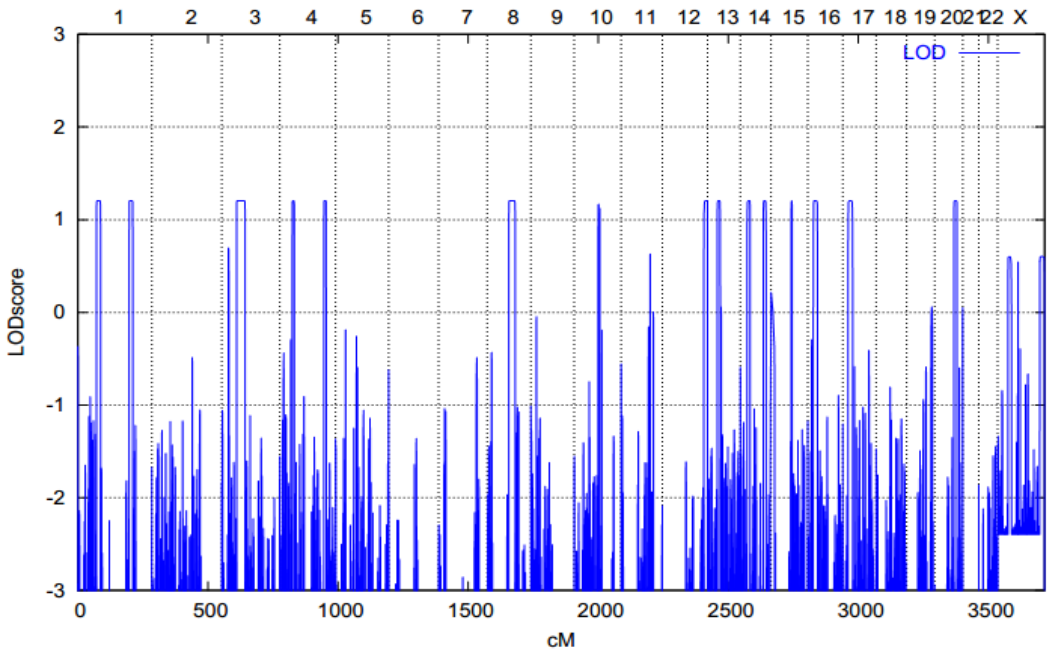

Figure 3.18: Genome-wide graphical representation of MCPH5 locus; parametric linkage analysis demonstration LOD scores of 1.2, detected at chromosome 1, 3, 4, 8, 10, 12, 13, 14, 15, 16, 17 and 20. *ASPM* is located within the second homozygous segment limited by rs2500097 (183,765,059 bp,199.19 cM) and rs10800790 (201,644,752 bp, 213.50 cM). All the physical locations, are according to the NCBI genome build 37 (GRCh37; hg19).

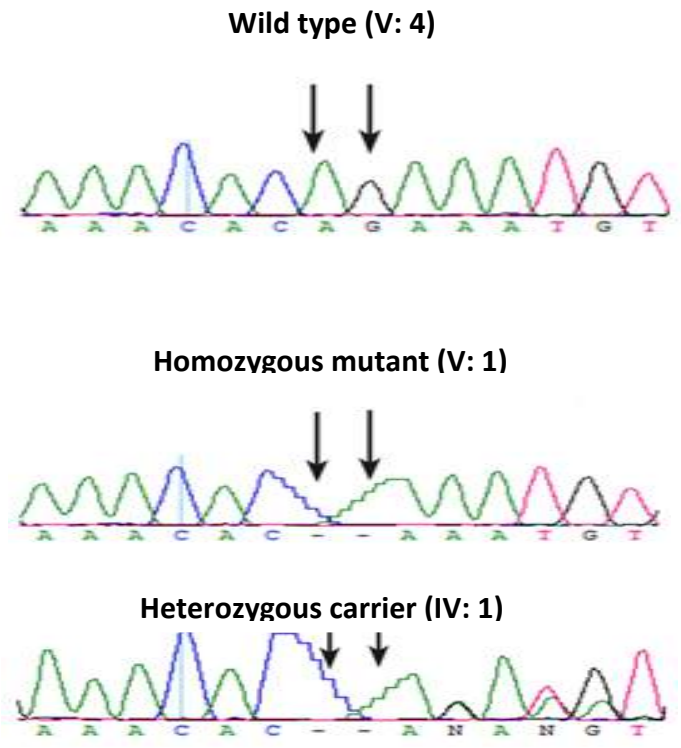

Figure 3.19: Sequence chromatogram of normal individual in upper panel, affected in middle panel and carrier in lower panel showing mutation (c.8508_8509delAGGA).

3.5.3 Discussion

Microcephalic patients exhibit small head size which is an obvious feature of this inherited disorder (Woods, 2004). The primary etiology could be: a failure of neurogenic mitosis or reduction in the number of neuronal cells due to aneuploidy or apoptosis and cell cycle delay (Woods and Parker, 2013). Some of the environmental causes; intra-uterine infection (first trimester cytomegalo virus (CMV) infection), alcohol consumption during pregnancy, malnutrition or hypothyroidism are some factors. These elements should be excluded out during the diagnosis of autosomal recessive microcephalic patients. As these elements also responsible for microcephaly with mental abnormality (Woods *et al*.,2005). MCPH is rare, with prevalence estimates between of 1:30,000 and 1:2,000,000 in European population, but much higher (1 in 10,000) in some parts of Pakistan due to high consanguinity rates (Woods *et al*., 2005).

ASPM gene carries the information for synthesizing a protein responsible for cell division. This particular protein is important for cell division in developing brain. Mutations in *ASPM* gene are linked with MCPH. ASPM protein involved in prior division of brain cells, consequently generate mature neuronal cells. This protein also assists in determining neurons number and brain size.

In family I, we identified a frame shift mutation (c.1615_1616delGA) in exon 3 of *ASPM* (Fig3.4). At protein level this substitution leads to premature stop codon (p.(Glu539Argfs*15)). This frame shift mutation is found to be located in the N-terminus of *ASPM* and was present in homozygous state in affected individuals of the family.

Sequence analysis of family II also revealed a frame shift mutation (p.(Leu2907Argfs*30)) due to deletion of four nucleotides (c.8718_8721deITTTA) in exon 18 (Figure 3.8) of *ASPM* detected in both affected individuals (IV:1 and IV:2) of this family.

Whole-exome sequencing of one affected member of family III revealed a nonsense mutation (p.(Gln1065*)) of *ASPM* in both affected individuals (V:1, VI:1).

In family IV whole-exome sequence was performed which shows nonsense mutation (p.(Pro1254Ser*fs2)) of *STIL* due to the duplication of one bp (c.3759dupT) in exon 17 (Fig 3.15).

Together with it, WES was also carried out for family V, on the DNA of one abnormal individual and analysis expressed deletion of two nucleotides (c.8508_8509delAG) of ASPM in exon 18. All three patients (V:1, V:2, and V:3) were homozygous for this variant. This c.8508_8509delAGGA (Figure 3.19) mutation changed the reading frame, consequently give rise to early premature stop codon, along the downstream 34 amino acids (p.(Lys2837Metfs*34)). This variation was existed in homozygous state in three abnormal individuals (V:1, V:2, and V:3), and heterozygous state in parents (Fig 3.19).

Five families were selected for microcephaly study. One patient from all families was subjected to exome sequencing. Exome sequence analysis was performed and variants selected for Sanger sequencing. All six variants were selected for segregation. *STIL* gene responsible for cytoplasmic protein coding, implicated in modulation of mitotic spindle check-point, whereas, mutation in this region leads to Microcephaly. *STIL* gene is associated with primary autosomal recessive microcephaly.

3.6 Family VI

This was a large family with multiple affected individuals (Fig 3.20). Eleven affected members, six females and 5 males, were noted in this family. Four of them died due to unknown medical reasons. Affected members were born in the 3rd, 4th and 5th generations of the family. Samples were collected from abnormal and healthy individuals of this family, with the consent of family head.

3.6.1 Clinical features

Multiple individuals of family VI (Fig 3.21) were affected and available members at the time of sampling were (IV:10, IV:12 affected male, IV:11 affected female) and with their parent's cousin marriage (III:5 mother, III:6 father). Individuals (IV:10 affected male, IV:11 affected female) showed mental retardation, aggressive behavior, no skills for communication, and were unable to feed and dress themselves. Both individuals (age-19 and 17 years) have sloppy forehead; IV:10 has head circumferences of 17.3 inches, height 58 inches, weight 55 kg and individual IV:11 has head circumference of 16.9 inches, weight 42 kg, height 49 inches. Another patient (IV:12) had reduced head circumference of 16 inches at the age of 13 years (Fig 3.21). Weight of patient was 21 kg and height of 46.6 inches. All abnormal individuals of this family also exhibited delayed language and aggressive behavior

3.6.2 Molecular study

Whole-exome sequence data analysis in family VI, identified mis-sense mutation (c.2104G>C; p.(Gly702Arg) of *VLDR*. At protein level this substitution leads to change of Gly (GGU) with Arg (AGA) at position 702 (p.(Gly702Arg)) (Figure 3.22). All the tested patients of this family were homozygous for this variant whereas both parents (III:5, III:6) were heterozygous (Fig 3.22). Among various species of vertebrates, Glycine is extremely conserved showing maximum functional significance (Fig 3.23).

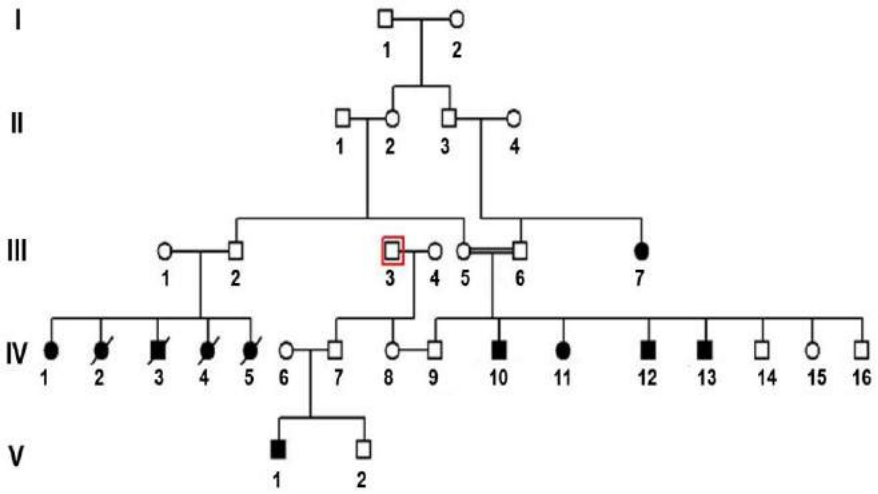

Figure 3.20: Family VI pedigree with eleven abnormal individuals.

Figure 3.21: Photos of the abnormal individuals (IV: 10 male, IV:11 female and V:11 male) showing reduced head circumference.

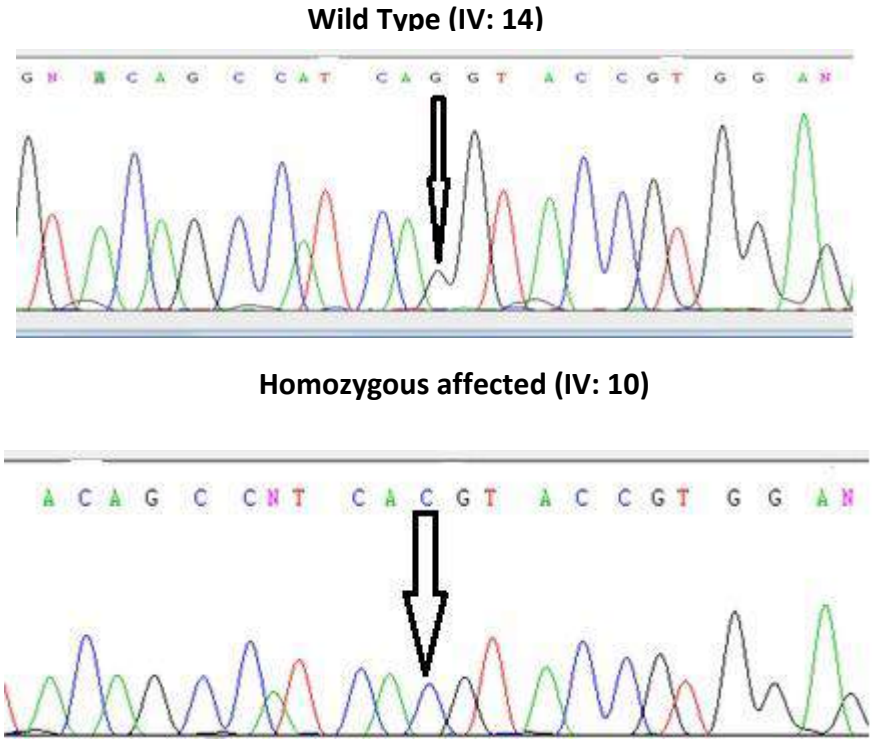

Figure 3.22: Sequence chromatogram of normal individual in upper panel, affected in lower panel showing mutation (c.2104G>C).

species	match	gene	aa alignment
Human			702 I I V Y H E L V Q P S G K N W C E E D M E N G G
mutated	not conserved		702 I I V Y H E L V Q P S R K N W C E E D M E N G
Ptroglodytes	all identical	ENSPTRG00000020732	702 I I V Y H E L V Q P S G K N W C E E D M E N G
Mmulatta	all identical	ENSMMUG00000015093	696 I I V Y H E L V Q P S G K N W C E E D M E N G
Fcatus	no homologue		
Mmusculus	all identical	ENSMUSG00000024924	702 I I V Y H E L V Q P S G K N W C E D D M E N G
Ggallus	all identical	ENSGALG00000010166	724 I I V Y H E L V Q P S G R N W C E E N M V N G
Trubripes	all identical	ENSTRUG00000004731	635 I I V Y H E L I Q L S G T N W C A E K G E N G
Drerio	all identical	ENSDARG00000006257	698 I I V Y H E L I Q L S G T N W C N E K M E N G
Dmelanogaster	no homologue		
Celegans	not conserved	T13C2.6	761 V R I Y H K Q A Q P L M Q N K C - - - - E N S
Xtropicalis	all identical	ENSXETG00000011709	702 V I V Y H E L V Q P L G R N W C N E Q I E N G

Figure 3.23: Conservation of glycine among different vertebrates.

A) **B)**

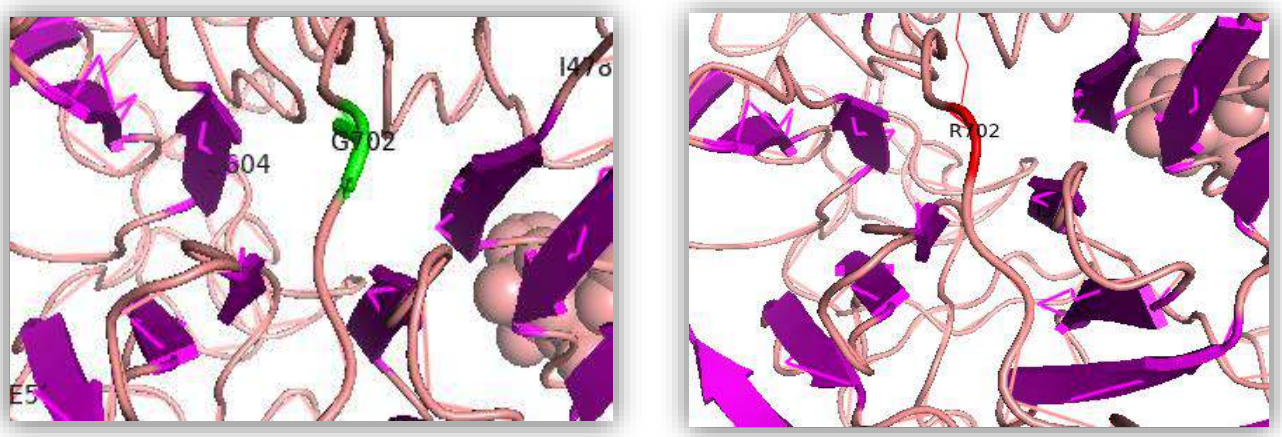

Figure 3.24**:** Zoomed 3D structure of ***VLDLR*** of human in ribbon presentation Sheets (shown in magenta color) and loops (shown in pink color). **A) wild-type *VLDLR*. Glycine** is present at position 702 shown in green color. **B) Mutant *VLDLR*. Glycine** is replaced by **Arginine** at position 702 shown in red color.

3.6.3 Discussion

Mutations in *VLDLR* are reported to cause *a*utosomal recessive congenital cerebellar ataxia and mental retardation (*Ozcelik et al., (2008*). Interestingly, this gene also showed association with Dementia (*Okuizumi et al., 1995*). Missense mutation identified in this study could be deleterious because the mutated nucleotide is the very last base of the exon 14 of *VLDLR*. Presumably, this change may result in the disruption of the normal splicing mechanism. Analysis of this variation in public databases of the genomic variations revealed this variant is not a polymorphism because it was absent not only in the dbSNP153 but also in gnomAD. Additionally, the mutant residue (p.(Gly702Arg)) is extremely conserved among different vertebrates (Figure 3.23) and this alteration is depicted to be deleterious and disease causing by number of *in silico* tools like Mutation Taster showed disease causing (score = 1.25), Polphen-2, probably damaging (score=1.000) and PROVEAN, deleterious (score= -5.783) and CADD score=35.

3.7 Family VII

This family belongs to coastal area of Baluchistan, Pakistan. Four abnormal individuals (V: 2, V: 4, V: 5 and V: 6) were born due to consanguineous marriage (Fig 3.24). Among them autosomal recessive mode of inheritance was observed. For genetic study, blood samples were drawn from all family members, with their consent.

3.7.1 Clinical features

A 9-year-old male proband (IV-2) is was born with consanguineous parents (Fig 3.25). The pregnancy and delivery were uneventful. He had mild developmental delay, speech impairment and at the age of 26 months begin to walk independently. When reached to eight months, patient showed poor vision and severe myopia (-12 diopters). No obvious dysmorphic features were found, except for prominent upper central incisors. Height and head circumference measurements were >-3 SD below mean for age and gender. The proband's younger sister (IV- 4) with 7-year-old, showed universal developmental delay, mild to moderate mental retardation, vision abnormality, short stature and microcephaly. Clinical features of the family were shown in table 3.3.

3.7.2 Molecular study

Variant filtration and prioritization in family VII revealed a new splice site mutation of *NSUN2*, a known gene of MRT5 locus. The mutation replaces guanine (G) with adenine (A) at the canonical splice donor site, c.1601+1G>A, of *NSUN2* situated in intron 14 of *NSUN2* (NM_017755.5) and mutation was confirmed by performing Sanger sequencing (Fig 3.26). This splice variant is identified in intron 14 of *NSUN2* (Fig 3.27). In parallel to WES, this family was also subjected for homozygosity mapping. Linkage analysis revealed three peaks approaching theoretical highest LOD score of 1.8 on chromosome 4, 5 & 14. The breakpoints of these three resultant homozygous segments are confined by the flanking SNP markers rs1032673 (134.98 cM; 129,465,489 bp) to rs2893315 (136.94 cM; 131,108,390 bp) at cytoband 4q28.2, rs10076494 (0.04 cM; 038,139 bp) to rs249264 (26.41 cM; 11,481,099 bp) at cytoband 5p15.33-p15.2 and rs11623181 (110.31 cM; 101,642,015) to rs2078693 (120.19 cM; 107,246,846 bp) at cytoband 14q32.32-q32.33. As expected, *NSUN2* resides in one of the identified homozygous segment of 11.44 Mb at chromosome 5p15.33-p15.2. We also screened all three identified homozygous regions for the possibility of any additional variation(s) and the data analysis did not show any interesting pathogenic

variation. Next, we analyzed the influence of c.1601+1G>A mutation on the splicing mechanism by using different algorithms. Human Splicing Finder (HSF), Neural Network**,** Mutation Taster and Splicing Regulation Online Graphical Engine (SROOGLE), all predicted that the splice donor site of *NSUN2* intron 14 is compromised. To check the mutation effect on gene, structure of protein was also predicted (Fig 3.28).

Table 3.3Clinical manifestations of the family carrying mutation in *NSUN2*.

Family VII		
Clinical features	**V.2**	**V.4**
Age (Year)	10	12
Gender	Male	Male
Head circumference (cm)	50	52
SD	-2	-1
Ht	126	140
ID severity	Mild	Mild
Motor development	Normal	Normal
Speaking skills	Normal	Normal

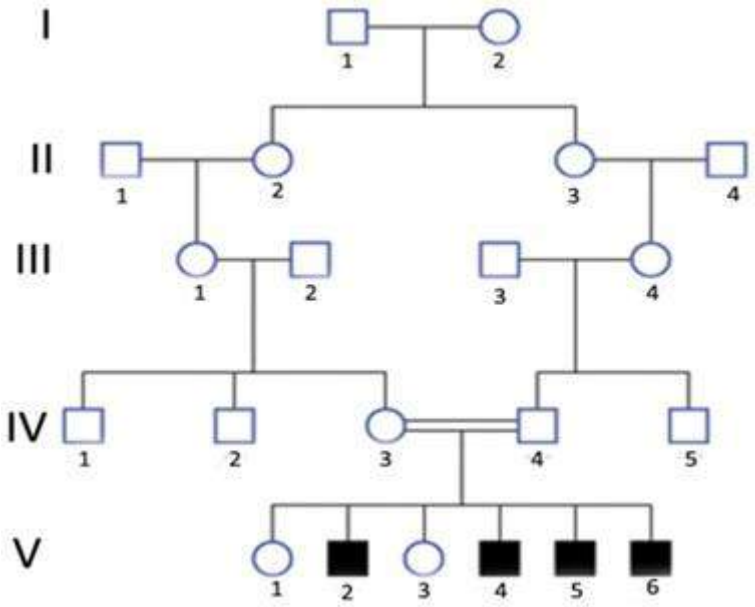

Figure 3.25: Family VII pedigree having mutation in *NSUN2*.

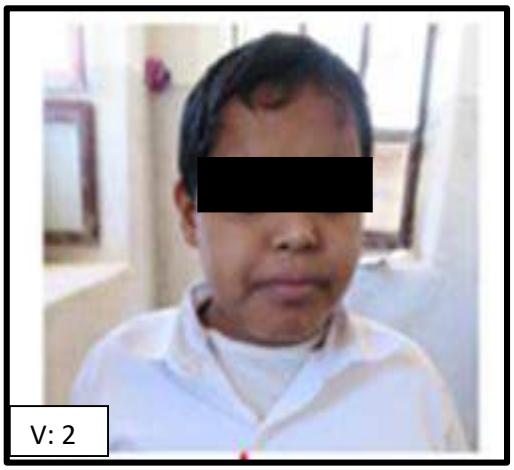

Figure 3.26: Photo of abnormal individual (V: 2) found mutated with *NSUN2*.

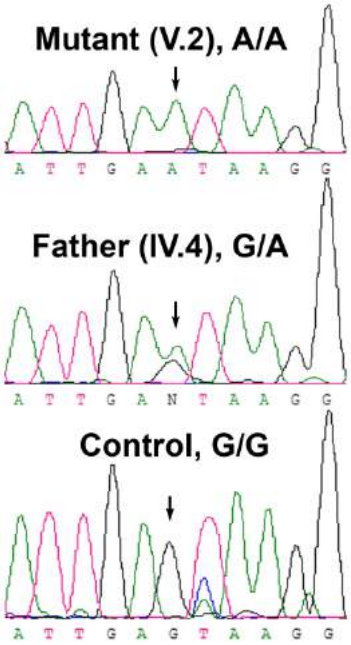

Figure 3.27: Segregation analysis of *NSUN2* mutation.

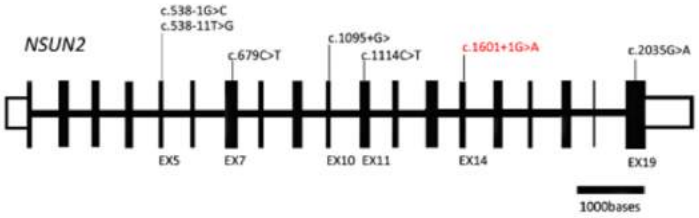

Figure 3.28: Genes structure of *NSUN2* with all known mutations. Identified mutation in this study is shown with red color. Vertical black lines show coding exons drawn with scale whereas horizontal shows intron which are not drawn with scale.

Figure 3.29: Protein structure *NSUN2* by using swiss model.

3.8 Family VIII

The family was identified from the, coastal area of Baluchistan, Pakistan. Pedigree of the family was framed using the data and history given by the parents (Fig 3.29). Blood samples from all family members were transferred in EDTA tube. The consent for research participation was obtained from all the members of the family. In case of children or intellectually disabled individuals, consent was sought from legal guardians or parents. The study was approved from the Board of Advanced Study and Research Committee of International Islamic University Islamabad Pakistan (BASAR).

3.8.1 Clinical features

The male proband (IV-1; fig 3.30) of 17 years old, belong to consanguineous parents delivered with uneventful pregnancy. Jaundice was noted at the age of one week that persisted for several months, however, was not severe to require admission to the hospital. He had psychomotor delay (walking acquired at 3 years and speech at 7 years). Clinical examination revealed mild hypotonia, joint laxity and jaundice. There were no dysmorphic features and no history of seizures. The elder sibling (IV-2; fig 3.30) is 17-year-old boy clinically observed with mild intellectual disability, speech impairment and without any history of neonatal jaundice. Clinical characteristics of the family are shown in Table 3.4.

3.8.2 Molecular study

Variant filtration and prioritization in family VIII revealed nonsense mutation (NM_001040616;c.2185A>T;p.(Lys729*)) of *LINS1,* a gene responsible for MRT27 locus of intellectual disability. The substituted nucleotide is located within exon 7 (Fig 3.32) of *LINS1* and resulted in stop codon, which subsequently produces truncated protein, if synthesized, with the loss of 29 amino acids at the C-terminus. The Mutation was confirmed by performing co-segregation analysis which followed the proposed inheritance pattern which is autosomal recessive (Fig 3.31). Normal protein structure of the *LINS1* gene was also predicted by using Swiss model (Fig 3.33)

Table 3.4: Clinical characteristics of the family representing mutation in *LINS1*

Family VIII		
Clinical features	**IV.1**	**IV.2**
Age	17 Year	14 Year
Gender	Male	Male
Head circumference	NA	NA
ID severity	Mild	Mild
Motor development	Delayed	Normal
Speaking skills	Delayed abnormal	Delayed abnormal

Note: NA means not available

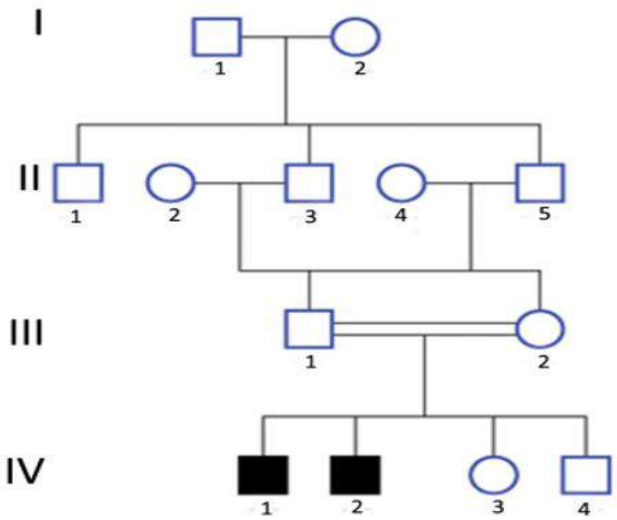

Figure 3.30: Family pedigree representing the probands (IV:1 & IV:2).

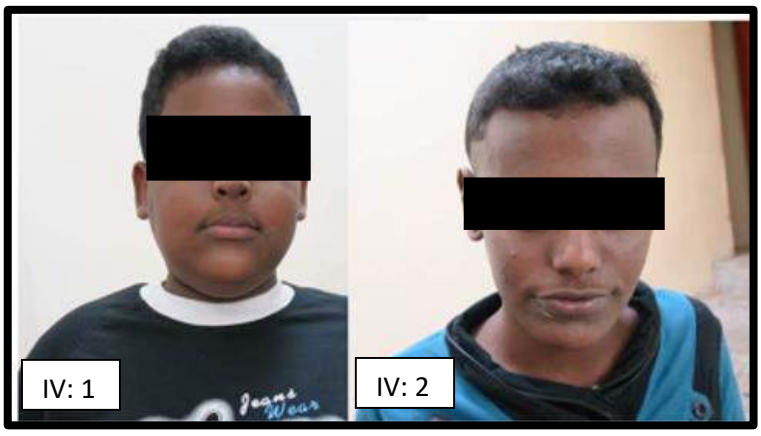

Figure 3.31: Photographs of patients (IV: 1 and IV: 2).

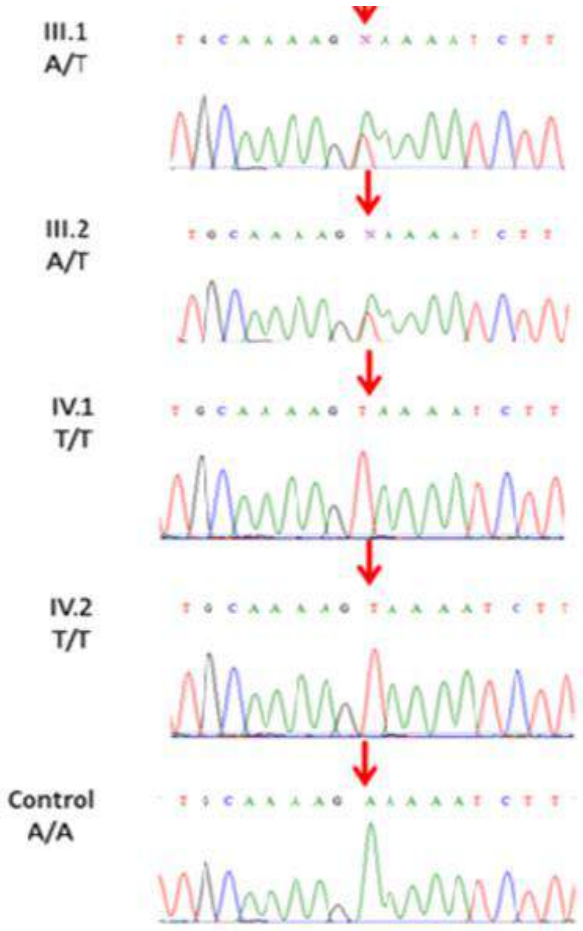

Figure 3.32: Co-segregation analyses in comparing with normal control. Both parents (III: 1 and III: 2) heterozygous, (IV: 1 and IV: 2) are homozygous affected and Lower panel showing normal control.

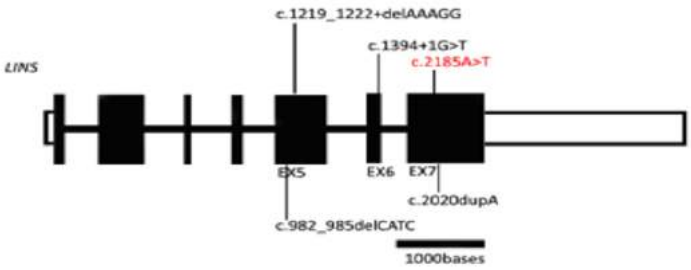

Figure 3.33: Genes structure of *LINS1* showing all the identified mutations. Mutation shown in red is identified in current study. Vertical black bars denote coding exons and white show non-coding. Horizontal line shows intron.

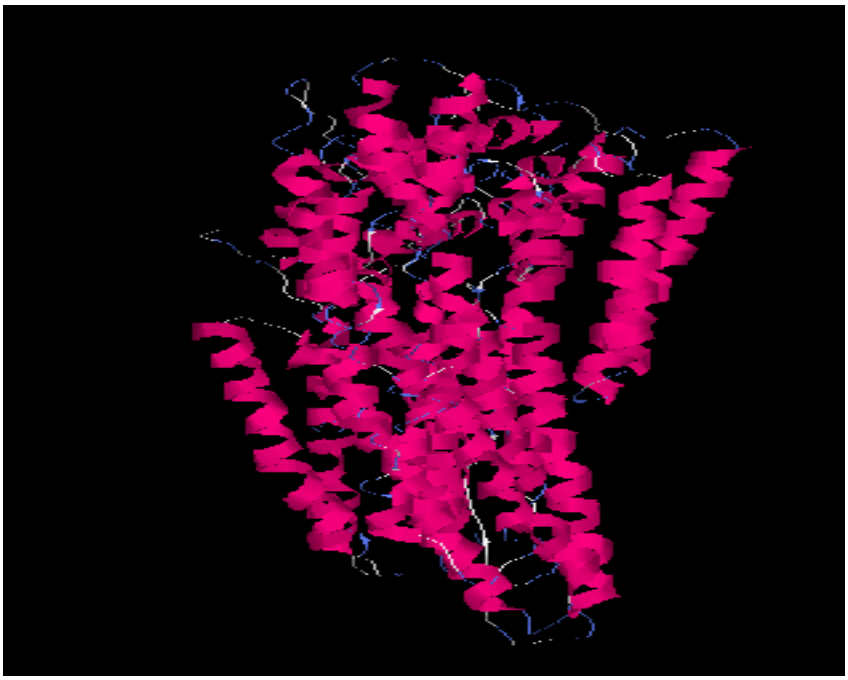

Figure 3.24 Wild-type protein structure of LINS1 was constructed by using Swiss modeling.

3.9 Family IX

This family was screened from the coastal area of Baluchistan province. Family pedigree drawn on the basis of information provided by the parents and nearby relevants (Fig 3.34). Clinical history of the patients was obtained from the parents and blood relatives. From abnormal and healthy individual of family, blood samples were transferred in EDTA tube. Research participation consent was obtained from all the members of the family.

3.9.1 Clinical features

The male proband (IV-2) of 10-yearsold born in consanguineous parents. The pregnancy and delivery were uneventful. In the first two years of his life he achieved normal motor milestones and showed no noticeable speech delay or defects. He displayed mild intellectual disability and was enrolled in special education school. On examination, no dysmorphic features were noted and the measurements of head circumference and height were normal. The other affected girl (IV: 3) showing similar phenotypes as described above (fig 3.35) had similar phenotypes. Clinical characteristics of the affected members are presented in Table 3.5.

3.9.2 Molecular study

Screening of whole-exome sequencing data of family IX for causative mutation, resulting in recognition of mis-sense mutation (c.412C>T;p.(Arg138Trp)) in *GDI1* (NM_001493.2), originally considered as causative gene of MRT41 locus. Mutant was validated by Sanger-sequencing of the family and mutation presented complete segregation with the family (Fig 3.36). The mutant nucleotide in the exon 5 of *GDI1* replaces a highly conserved arginine to tryptophan (Fig 3.37). PROVEAN, Polyphen-2 & Mutation Taster predicted this change as, deleterious probably damaging and disease causing, respectively. Intriguingly, Human Splicing Finder (HSF) predicts that the mutation may alter the exonic splice enhancer (ESE) site.

Table: 3.5. Clinical feature of the family, mutation in *GDI1*.

Family IX		
Clinical features	**IV.2**	**IV.3**
Age	9 Year	6 Year
Gender	Male	Female
Head circumference (cm)	48	47
SD	-3	-6
Height (cm)	110	98cm
ID severity	Mild-moderate	Mild-moderate
Motor development	Delayed	Normal
Speaking skills	NO	NO

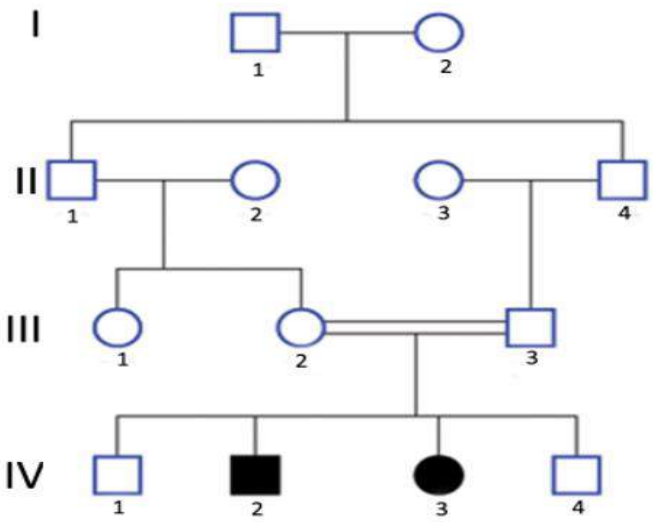

Figure 3.35: Pedigree of Family representing the probands (IV: 2 & IV: 3).

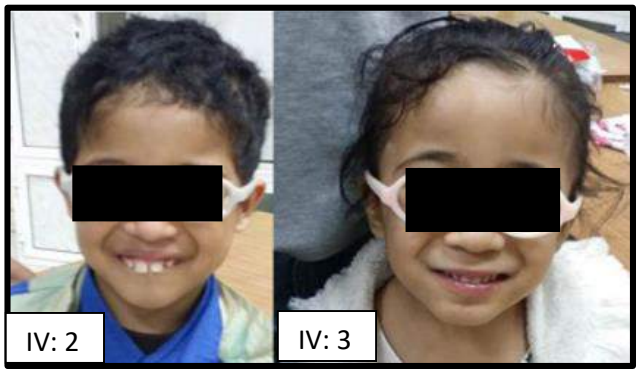

Figure 3.36: Phenotype of abnormal individuals (IV: 2 and IV: 3).

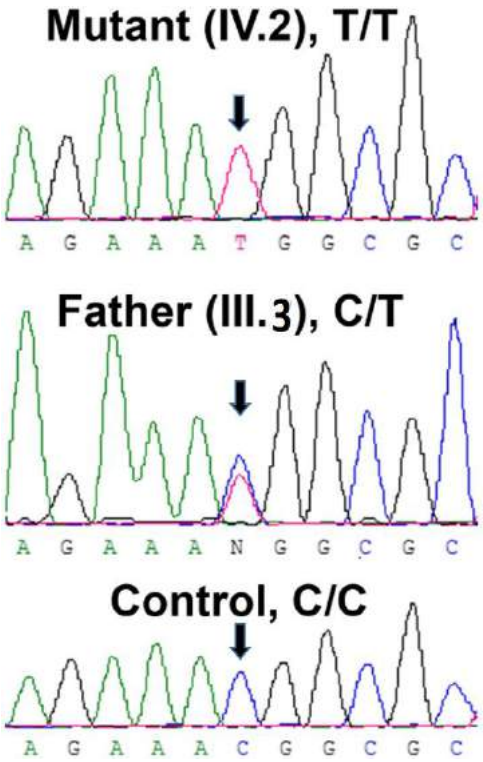

Figure 3.37 Sanger sequencing analyses of family showing co-segregation with mutation. Chromatogram showing wild type, homozygous affected (IV: 2) and heterozygous (III: 2) mutant.

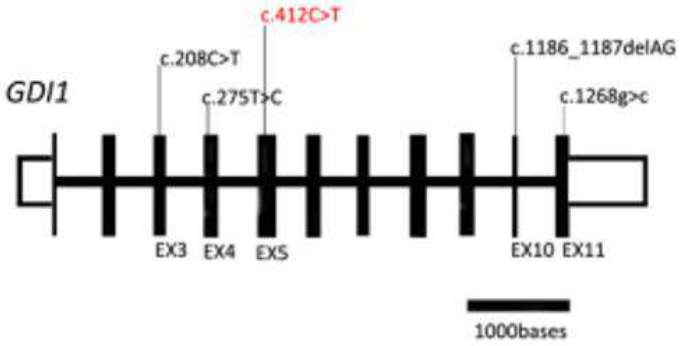

Figure 3.38: *GDI1* structure with vertical black bars showing coding exons and horizontal line shows intronic regions. Mutation recognized in current study is shown in red color.

A) **B)**

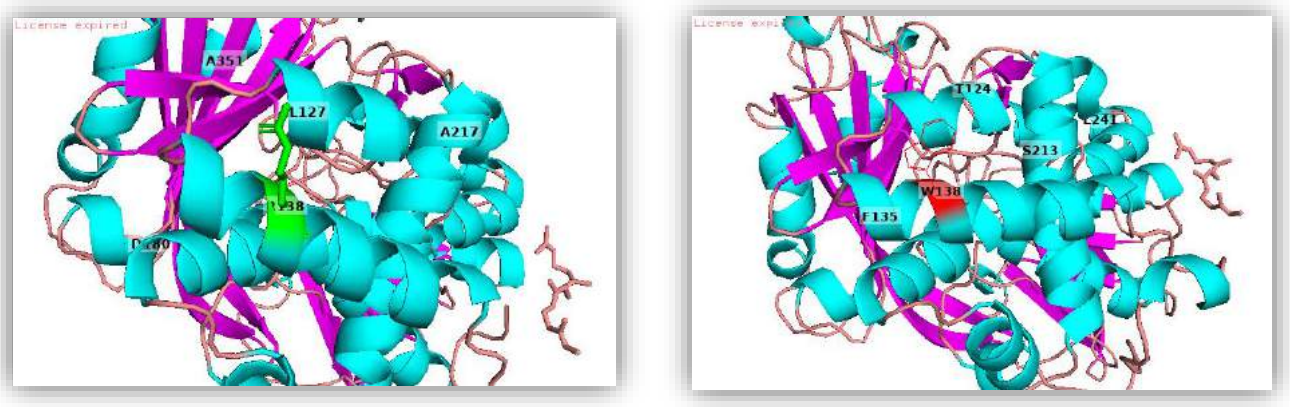

Figure 3.39**:** Zoomed 3D structure of **GDI1** of human in ribbon presentation. Helices (shown in cyan color), Sheets (shown in magenta color) and loops (shown in pink color). **A) wild-type *GDI1***. **Arginine** is present at position **138** shown in green color. **B) Mutant *GDI1*. Arginine** is replaced by **Tryptophan** at position **138** shown in red color.

Chapter 4: DISCUSSION

Discussion

In current study, we recruited 9 families manifesting microcephaly & intellectual disability. We employed a series of genomic analyses like homozygosity mapping, whole-exome and Sanger-sequencing to uncover the genetic determinants. In addition to this, several in silico tools were consulted to anticipate the pathogenicity of screened out variants. Polymorphic nature of these genetic culprits was ruled out by looking them in public databases of genomic variants. We identified and confirmed the association of three novel mutations of the GDI1, NSUN2 and LINS1 genes with ID in Pakistani families.

We have reported novel mutation in family IX. So far total five mutation reported on GDI1 gene (Fig 3.37) (Strobl-Wildemann G., 2011). *GDI1*is an X-chromosomal gene and expressed primarily in neural tissues which encodes brain enriched protein Rab GDP dissociation-inhibitor α (aGDI) protein that contribute to vesicle shuttling by regulating the activity of Rab GTPases (Grosshans *et al*., 2006). Several animal model studies reported that the *Gdi1* null mice showed damage associative learning process and attention, that plays a key role in cognitive impairment, noticed in XLID patients. As a matter of fact, in 2017 Lorenzo More and his colleagues provided strong evidence that the memory, interval-timing and learning and deficiencies noticed in *Gdi1* null mice are primarily concerned with the reduced glutamatergic release in different brain parts (hippocampal, striatal, cortical and amygdala) (Lorenzo M., 2017).

In family VIII we have reported novel variant causing mutation in *LINS1* gene. *LINS1* expressed in human brain, skeletal muscle, prostate, spleen and fetal kidney (Katoh, 2002). Independent of scientific studies, they predicted that variations in *LINS1* may cause the alteration of *ELAV1* (*HuR*) (embryonic lethal abnormal vision) expression during differentiation of neurons, resulting in intellectual disability. Through complicated mechanisms, *LINS1* has alterations effect of *ELAV1* (*HuR*) expression that act as a RNA binding protein, attaches to 3-UTR domain of mRNAs to enhance the stability. In humans this gene act as an ortholog of embryonic lethal abnormal vision in fruit-fly and highly expressed in nucleus specially in all post mitotic neuron in early stages of differentiation in embryo . According to high expression of HuR messenger RNA in the human brain, any other regulatory process of *ELAVI* (HuR) in brain cells or shuttling between the cytoplasm and the nucleus will affect the cell cytoplasm of neuron

origin, localization to axons, and neurons differentiation. This leads to minimize the growth of commissural axon, consequently affecting the function of central and peripheral neuron. Moreover, indirectly *LINS1* has effect on the foundation of anterior-posterior axis in the brain through cooperation with beta-catenin that facilitate the axon proliferation and neural progenitor , synaptogenesis and dendritogenesis through Wingless-Type MMTV Integration Site Family, Member 1 (WNT-1) signaling mechanism, at various phases of neuron tissue development and regulation. Hence, a variation in WNT-1 signaling mechanism is probably accountable for intellectual disability (Vargas *et al*., 2014; Sheth J., 2017). In our case as a consequence to nonsense mutation of *LINS1*, the truncated protein led to the inhibition of protein production. Therefore, this variation has a pathogenic effect on the function of protein. This mutation is considered as fifth mutation in *LINS1* that is associated with ID.

In family VII we also reported novel mutation in *NSUN2* gene. *NSUN2* gene has a potential role in ID with six known mutations identified in different ethnicities. Current gene responsible for coding of methyl-transferase that catalyzes the formation of 5 methylcytosine at C-34 of tRNA-leu (CAA) (Brzezicha B.,2006). During mitosis, methyl transferase enzyme has a critical molecular function in chromosomal segregation and in micro-tubular assembly (Blanco and Michaela Frye 2014). In short words, there are diverse mechanisms that play role of *NSUN2* in ID. One of them is through the function of *NSUN2* in regulation of protein translation that is required for synaptic flexibility, memory and learning. Therefore the absence of *NSUN2* expression in fetal brain during critical stages of brain development by proteomic shifts will cause phenotype of ID in affected individuals (Hussain *et al*., 2009; Chi *et al*., 2013). The other process may involve in defected methylation of hemi-methylated DNA, which is also one of targets of NSUN-2 activity (Hong C 2004). Knockout model of *NSUN2* ortholog in *Drosophila melanogaster* showed behavioral changes and STM (severe short-term memory) defects (Tweedie S. 2009). At the present time, in our study we reported seventh mutation of *NSUN2* that causes autosomal recessive intellectual disability as a consequence to impaired splice donor site.

References

Abecasis G. R., Altshuler D., Auton A., Brooks L. D., Durbin R. M., Gibbs R. A.,Hurles M. E., and McVean G. A. "A map of human genome variation from population-scale sequencing". Nature, 2010. 467: 1061-73

Auton A., Brooks L. D., Durbin R. M., Garrison E. P., Kang H. M., and Korbel J. O. "A global reference for human genetic variation". Nature, 2015. 526: 68-74

Bamshad M. J., Ng S. B., Bigham A. W., Tabor H. K., Emond M. J., Nickerson D. A. and Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. Nature Reviews Genetics, 2011. 12: 745

Basel-Vanagaite L. Genetics of autosomal recessive nonsyndromic mental retardation: recent advances. Clin Genet. 2007. 72: 167-74

Bateson W., and Mendel G. Mendel's principles of heredity. Courier Corporation, 2013.

Behjati S. and Tarpey P. S. "What is next generation sequencing?". Archives of disease in childhood-Education & practice edition, 2013. 98: 236-238

Bittles A. H. "The bases of western attitudes to consanguineous marriage," Dev Med Child Neurol, 2003. 45: 135-8

Bittles A. H. "The role and significance of consanguinity as a demographic variable" Population and development review, 1994. 61-584

Botstein D. and Risch N. "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease". Nat Genet, 2003. 33: 228-37

Boycott K. M., Flavelle S., Bureau A., Glass H. C., Fujiwara T. M., Wirrell E., Davey K., Chudley A. E., Scott J. N., McLeod D. R., Parboosingh J. S. Homozygous deletion of the very low density lipoprotein receptor gene causes autosomal recessive cerebellar hypoplasia with cerebral gyral simplification. Am J Hum Genet, 2005. 77: 477-483

Chial H. "Mendelian genetics: patterns of inheritance and single-gene disorders". Nature Education, 2008. 1: 63

Chiurazzi P., and Pirozzi F. Advances in understanding – genetic basis of intellectual disability. F1000Res. F1000 Fac Rev, 2016. 599

Collins F. S. "Positional cloning moves from perditional to traditional". Nat Genet, 1995. 9: 347-50

Cox J., Jackson A. P., Bond J., and Woods C. G. What primary microcephaly can tell us about brain growth. Trends Mol Med, 2006. 12: 358-366

Delauno J., Abidi F., Zeniou M., Jacquot S., Merienne K., Pannetier S., Hanauer A. Mutations in the X-linked RSK2 gene (RPS6KA3) in patients with Coffin-Lowry syndrome. Human Mutation, 2001. 17: 103–116

El-Kheshen G. and Saadat M. "Prevalence of consanguineous marriages among shi'a populations of Lebanon". J Biosoc Sci, 2013. 45: 675-82

Fay J. C., Wyckoff G. J., and Wu C. I. "Positive and negative selection on the human genome". Genetics, 2001. 158: 1227-34

Feuk L., Carson A. R., and Scherer S. W. "Structural variation in the human genome". Nat Rev Genet, 2006. 7: 85-97

Gecz J., and Haan E. New mutations and sporadic intellectual disability. Lancet, 2012. 380: 1630–1631

Gecz J., Gedeon A.K., Sutherland G.R., and Mulley J.C. Identification of the gene FMR2, associated with FRAXE mental retardation. Nature Genetics, 1996. 13: 105–108

Gibbons R.J., and Higgs D.R. Molecular-clinical spectrum of the ATR-X syndrome. American Journal of Medical Genetics, 2000. 97: 204–212

Hamamy H. A., Masri A. T., Al-Hadidy A. M., and Ajlouni K. M. "Consanguinity and genetic disorders. Profile from Jordan". Saudi Med J, 2007. 28: 1015-7

Haraksingh R. R., and Snyder M. P. "Impacts of variation in the human genome on gene regulation". J Mol Biol, 2013. 425: 3970-7

Hashmi M. A. "Frequency of consanguinity and its effect on congenital malformation--a hospital based study". J Pak Med Assoc, 1997. 47: 75-8

Hedges D. J., Burges D., Powell E., Almonte C., Huang J., Young S., Boese B., Schmidt M., Pericak-Vance M. A., Martin E., Zhang X., Harkins T.T., and Zuchner S. Exome Sequencing of a multi-generational human pedigree. PLoS One, 2009. 4: e8232

Higgins J. J., Pucilowska J., Lombardi R. Q., and Rooney J.P. Candidate genes for recessive non-syndromic mental retardation on chromosome 3p (MRT2A). Clin Genet, 2004. 65: 496-500

Hussain R. and Bittles A. H. "The prevalence and demographic characteristics of consanguineous marriages in Pakistan". J Biosoc Sci, 1998. 30: 261-75

Jackson A. P., McHale D. P., Campbell D. A., Jafri H., Rashid Y., Mannan J., Karbani G., Corry P., Levene M. I., Mueller R. F., Markham A. F., Lench N. J., and Woods C. G. Primary autosomal recessive microcephaly (MCPH1) maps to chromosome 8p22-pter. Am J Hum Genet, 1998. 63: 541-546

Jayaraman, D., Bae, B.-I. and Walsh, C. A. The genetics of primary microcephaly. Annu Rev Genomics Hum. Genet, 2018. 19, 177-200

Kaufman L., Ayub M., and Vincent J.B. The genetic basis of non-syndromic intellectual disability: A review. Journal of Neurodevelopmental Disorders, 2010. 2: 182–209

Khoury S. A. and Massad D. "Consanguineous marriage in Jordan". Am J Med Genet, 1992. 43: 769-75

Kochinke K., Zweier C., Nijhof B., Fenckova M., Cizek P., Honti F., Keerthikumar S., Oortveld M.A., Kleefstra T., Kramer J.M. and Webber C. Systematic phenomics analysis deconvolutes genes mutated in intellectual disability into biologically coherent modules. The American Journal of Human Genetics, 2016. 98: 149-164

Kürekçi G. K. and Dinçer P. "Exome Sequencing for The Identification of Mendelian Disease Genes". Correspondance, 2013.

Lander E. S. and Schork N. J. "Genetic dissection of complex traits". Science, 1994. 265: 2037-48

Lovell, S. C., Davis, I. W., Arendall III, W. B., De Bakker, P. I., Word, J. M., Prisant, M. G., ... & Richardson, D. C. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics. 2003. 50(3), 437-450.

Luckasson R., and Reeve A. Naming, defining, and classifying in mental retardation. Mental Retardation, 2001. 39: 47–52

Majewski J., Schwartzentruber J., Lalonde E., Montpetit A., and Jabado N. "What can exome sequencing do for you?". J Med Genet, 2011. 48: 580-9

Martin C.A., Ahmad I., Klingseisen A., Hussain M.S., Bicknell L.S., Leitch A., Nürnberg G., Toliat M.R., Murray J.E., Hunt D. and Khan F. Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. Nature genetics, 2014. 46: 1283

Maulik P.K., Mascarenhas M.N., Mathers C.D., Dua T., and Saxena S. Prevalence of intellectual disability: A meta-analysis of population-based studies. Research in Developmental Disabilities, 2011. 32: 419–436

Modell B., and Darr A. "Science and society: genetic counselling and customary consanguineous marriage". Nat Rev Genet, 2002. 3: 225-9

Nakamura Y. "DNA variations in human and medical genetics: 25 years of my experience." J Hum Genet, 2009. 54: 1-8

Nano M. and Basto R. Consequences of centrosome dysfunction during brain development. Adv Exp Med Biol, 2017. 1002: 19-45

Neri G., Schwartz C.E., Lubs H.A., and Stevenson R.E. X-linked intellectual disability update 2017. Am J Med Genet, 2018. 176A:1375–1388

Nicholas A. K., Swanson E. A., Cox J. J., Karbani G., Malik S., Springell K., Hampshire D., Ahmed M., Bond J., Di B. D., Fichera M., Romano C., Dobyns W. B., and Woods C. G. The molecular landscape of ASPM mutations in primary microcephaly. J Med Genet, 2009. 46: 249-253

Okuizumi, K., Onodera, O., Namba, Y., Ikeda, K., Yamamoto, T., Seki, K., Ueki, A., Nanko, S., Tanaka, H., Takahashi, H., Oyanagi, K., Mizusawa, H., Kanazawa, I., Tsuji, S. Genetic association of the very low density lipoprotein (VLDL) receptor gene with sporadic Alzheimer's disease. Nature Genet. 11: 207-209, 1995.

Ottenheimer M. Forbidden relatives: The American myth of cousin marriage: University of Illinois Press, 1996.

Ozcelik, T., Akarsu, N., Uz, E., Caglayan, S., Gulsuner, S., Onat, O. E., Tan, M., Tan, U. Mutations in the very low-density lipoprotein receptor VLDLR cause cerebellar hypoplasia and quadrupedal locomotion in humans.Proc. Nat. Acad. Sci. 105: 4232-4236, 2008.

Pang T., Atefy R., and Sheen V. "Malformations of cortical development". Neurologist, 2008. 14: 181-191

Perez Y., Bar-Yaacov R., Kadir R., Wormser O., Shelef I., Birk O. S., and Birnbaum R. Y. Mutations in the microtubule-associated protein MAP11 (C7orf43) cause microcephaly in humans and zebrafish. Brain, 2019.142: 574-585

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

Rasmussen F., Gustavson K-H., Sara V. R., Floderus Y. The dysequilibrium syndrome: a study of the etiology and pathogenesis. Clin Genet, 1985. 27: 191-195

Rauch A., Wieczorek D., Graf E., Wieland T., Endele S., Schwarzmayr T., Albrecht B., Bartholdi D., Beygo J., Di Donato N. and Dufke A. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. The Lancet, 2012. 380: 1674-1682

Ropers H.H. Genetics of early onset cognitive impairment. Annual review of genomics and human genetics, 2010. 11: 161-187

Roy A., Kucukural, A., & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols. 2010. 5(4), 7

Saade M., Blanco-Ameijeiras J., Gonzalez-Gobartt E., and Martí E. A centrosomal view of CNS growth. Development, 2018.145: 170613.

Schalock R. L., and Luckasson R. A systematic approach to subgroup classification in intellectual disability. Intellectual and Developmental Disabilities, 2015. 53: 358-366

Schalock R. L., Borthwick-Duffy S. A., Bradley V. J., Buntinx W. H., Coulter D. L., Craig E. M., and Shogren K. A. Intellectual disability: Definition, classification, and systems of supports. American Association on Intellectual and Developmental Disabilities. 444 North Capitol Street NW Suite 846, Washington, DC 20001, 2010.

Srour M., and Shevell M. Genetics and the investigation of developmental delay/intellectual disability. Archives of Disease in Childhood, 2014. 99: 386–389

Stevenson R.E., and Schwartz C.E. X-linked intellectual disability: Unique vulnerability of the male genome. Developmental Disabilities Research Reviews, 2009. 15: 361–368

Study D. D. D., McRae J.F., Clayton S., Fitzgerald T.W., Kaplanis J., Prigmore E., Rajan D., Sifrim A., Aitken S., Akawi N. and Alvi M. "Prevalence and architecture of de novo mutations in developmental disorders." Nature, 2017. 433-438

Sukumaran S.K., Stumpf M., Salamon S., Ahmad I., Bhattacharya K., Fischer S., Müller R., Altmüller J., Budde B., Thiele H. and Tariq M. CDK5RAP2 interaction with components of the Hippo signaling pathway may play a role in primary microcephaly. Molecular genetics and genomics, 2017. 292: 365-383

Teare M. D. and Santibanez Koref M. F. "Linkage analysis and the study of Mendelian disease in the era of whole exome and genome sequencing". Brief Funct Genomics, 2014. 13: 378-83

Tissir F., and Goffinet A. M. “Reelin and brain development”. Nav Rev Neurosci, 2003. 4: 496-505

Trommsdorff M., Gotthardt M., Hiesberger T., Shelton J., Stockinger W., Nimpf J., Hammer R. E., Richardson J. A., and Herz J. Reeler/Disabled-like disruption of

neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. Cell, 1999. 97: 689-701

Valnegri P., Sala C., and Passafaro M. Synaptic dysfunction and intellectual disability. In: M.R. Kreutz & C. Sala (Eds.), Synaptic plasticity: Dynamics, development and disease. New York, NY: Springer, 2012. 433–444.

Vink J. M. and Boomsma D. I. "Gene finding strategies". Biol Psychol, 2002 61: 53-71

Vissers L.E., Gilissen C. and Veltman J.A. Genetic studies in intellectual disability and related disorders. Nature Reviews Genetics, 2016. 17: 9

Voelkerding K. V., Dames S. A., and Durtschi J. D. "Next-generation sequencing: from basic research to diagnostics". Clin Chem, 2009. 55: 641-58

Watson J. D. and Crick F. H. "The structure of DNA," in Cold Spring Harbor symposia on quantitative biology, 1953. 123-131

Wieczorek D. Autosomal dominant intellectual disability. medizinische genetik, 2018. 30: 318-322

Woods C. G. and Parker A. Investigating microcephaly. Arch Dis Child, 2013. 98: 707-713

Woods C. G. Human microcephaly. Curr. Opin. Neurobiol, 2004. 14: 112-117

Woods C. G., Bond J., and Enard W. Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. Am J Hum Genet, 2005. 76: 717-728

Yang Y., Muzny D.M., Reid J.G., Bainbridge M.N., Willis A., Ward P.A., Braxton A., Beuten J., Xia F., Niu Z. and Hardison M. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. New England Journal of Medicine, 2013. 369: 1502-1511

Zaqout S., Morris-Rosendahl D., and Kaindl A. M. "Autosomal Recessive Primary Microcephaly (MCPH): An Update". Neuropediatrics, 2017.

Appendix

Composition of solution's used in DNA extraction

1. **Solution A:** 0.32 M sucrose, 10 mM Tris pH 7.5, 5 mM MgCl2, 1% Triton X-100.
2. **Solution B:** 10 mMTris pH 7.5, 400 mM NaCl, 2 mM EDTA pH 8.0.
3. **Solution C:** Buffered phenol.
4. **Solution D:** Chloroform (24 ml), Iso-amyl Alcohol (1ml).