Elucidating the Molecular Genetic Basis of Intellectual Disability Families

By

Iftikhar Ahmed

Registration # 60-FBAS/PHDBT/S-16

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

2022

Human genetics
Molecular Genetic
Genetics Diseases

Elucidating the Molecular Genetic Basis of Intellectual Disability Families

Submitted By

Iftikhar Ahmed

Reg. # 60-FBAS/PHDBT/S-16

Supervised By

Dr. Asif Mir

Associate Professor

Department of Biological Science,

FBAS, IIUI

Department of Biological Sciences

Faculty of Basic and Applied Sciences

International Islamic University Islamabad, Pakistan

2022

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

Date: 23/2/2022

FINAL APPROVAL

It is certified that we have read the thesis entitled "Elucidating the Molecular Genetic Basis of Intellectual Disability Families" submitted by Mr. Iftikhar Ahmed (Reg# 60-FBAS/PHDBT/S16) 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
Prof. Dr. Mushtaq Ahmed
Chairman, Department of Plant Sciences
Faculty of Biological Sciences
Qauid-i-Azam University, Islamabad

External Examiner
Dr. Muhammad Jawad Hassan
Chairman/Associate Professor, Department of Biological Sciences
NUMS, Islamabad

Internal Examiner
Dr. Asma Gul
Associate Professor
Department of Biological Sciences, FBAS, IIUI

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

Date: 23/2/2022

FINAL APPROVAL

It is certified that we have read the thesis entitled "Elucidating the Molecular Genetic Basis of Intellectual Disability Families" submitted by Mr. Iftikhar Ahmed (Reg# 60-FBAS/PHDBT/S16) 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

Chairman

Prof: Dr. Muhammad Irfan Khan
Department of Biological Sciences
Faculty of Basic and Applied Sciences
International Islamic University Islamabad, Pakistan

Dean

Prof: Dr. Muhammad Irfan Khan
Faculty of Basic and Applied Sciences
International Islamic University Islamabad, Pakistan

This dissertation is

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

DEDICATION

This work is dedicated to my sweet and beloved parents, to my family, to my wife, my son Rehan ullah and my daughter Areeba Iftikhar whose constant support and guidance enabled me to achieve this milestone.

DECLARATION

It is certified that work done on this PhD Biotechnology research thesis entitled "Elucidating the Molecular Genetic Basis of Intellectual Disability Families" 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 28/02/2022.

Iftikhar Ahmed

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AAMR	American Association of Mental Retardation
ARMR	Autosomal Recessive Mental Retardation
ARID	Autosomal Recessive Intellectual Disability
AR	Autosomal Recessive
Bp	Base pair
^0C	Degree Centigrade
Cm	Centimeter
CNS	Central nervous system
DNA	Deoxy-ribonucleic Acid
dNTP	di-nucleotide Tri-phosphate
EDTA	Ethylenediamine Tetra Acetic Acid
EXOSAP	Exonuclease 1 shrimp alkaline phosphatase
IQ	Intelligence quotient
ID	Intellectual Disability
Kb	Kilo base pair
µL	Micro liter
µM	Micro meter
Mg	Milligram
Ml	Milliliter
mM	Milli molar
NCBI	National Center for Biotechnology Information
Ng	Nano gram
NSID	Non-syndromic Intellectual disability
OMIM	Online Mendalian Inheritance in men
PCR	Polymerase Chain reaction
PAGE	Polyacraylamide gel electrophoresis
PNS	Peripheral nervous system
RPM	Round per mints

SNP	Single nucleotide polymorphism
STR	Microsatellite markers
TE	Tris-EDTA
TEMED	Tetra methylethylene diamine
WES	Whole exome sequencing

ACKNOWLEDGEMENT

First of all I pay my submissive gratitude in the domain of Almighty Allah who is the supreme authority over the universe and nothing goes unaccounted under his domain. I am thankful to Almighty Allah that has bestowed me the strength and qualities due to which I have been able to complete my research.

I express my profound gratitude to my respectable and beloved parents, wife, children Rehan ullah and Areeba Iftikhar and my family for their sincere care, support and prayers that encouraged me a lot, in getting through the research work.

Especially I am grateful to my supervisor Dr. Asif Mir (Associate Professor Department of Biological Science, IIU, Islamabad) for his ideas, valuable suggestions. He has always guided whenever I stuck and encouraged me during the whole period. I am also thankful to my foreign professor Dr. Andrew Henry Crosby (University of Exeter, UK) for his valuable suggestions and good ideas that help me to move in right direction during HEC IRSIP program.

I feel privileged to record my sincere thanks to all the faculty member of Department of Biological Science, International Islamic University Islamabad Pakistan, Including, Dr. Zafar Mahmood Khalid, Dr. Jabbar Zaman Khan Khattak, Dr. Bashir Ahmad, Dr. Ali Imran Bukhari and Dr. Arshad Malik for their critical advice and guidance without which this research work would not have been possible.

I would like to appreciate IRSIP HEC program and also those who have supported, guided and advised me in my research. Many thanks to all the fellows, Abrar Ahmad, Muhammad Ilyas, Jahangir Khan.

Iftikhar Ahmed

ABSTRACT

Cognition is due to cellular, biological and several molecular events in CNS and a slight change in any of these process finally leads to cognitive impairment or ID. A great challenge was posed for clinicians and scientist for their research due to clinical and heterogeneity of Intellectual Disability.

To accomplish the research objectives of this study, fifteen families diagnosed as ID were analyzed in which six families proband (MR200, 205, 206, 208, 212 and MR295) were exomed, while family (MR200) designated with Intellectual disability, whole exome sequencing and sequence analysis showed a novel missense mutation c.3418G>T in the exon 17 of AFF2 gene (located on Chr: X) and co-segregated with diseased phenotype among family members. Bioinformatics analysis identified change in the secondary structure and amino acids confirmation in protein due this missense mutation, which affects the overall function of protein. Whole exome sequencing analysis revealed a novel (RPTN) gene mutation in family-MR212. Co-segregation confirmed c.592G>T (p.Glu198*) mutation in RPTN gene (located on Chr:1). Normal structure of RPTN was observed with a total number of pockets as 126 respectively. A huge decrease in pockets numbers has been observed as the length of mutated protein due to stop gain codon only. 24 pockets remain as compared to normal protein with 126 pockets. This change showed implication on the overall confirmation and function of protein. Genetics analysis in family-MR295 with Intellectual Disability found to be caused due to missense mutations in the ROR2 gene leading to diseased phenotype. The number of active site remained same in the case of mutation but in some active sites got disrupted. This change showed implication of mutation on overall conformation of protein which eventually altered interaction and function of protein causing pathogenicity. In family MR205, selected filtered variants were found co-segregated among the selected gene variants. More extensive analysis of data is required in the future to find the causative variants. Family designated as MR206 and MR208, also not provided any co-segregation pathway among the selected variants during analysis. Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR201). Proband (V: 6) share common homozygous region (chr16:59,003,011-74,060,285) total of 15,057,275 bp, comprising total 187 genes. Out of 187 genes, 46 genes based on their expression in brain were selected for screening

in this family to find the pattern of inheritance of disease gene with phenotype. Microarray genotyping was performed. Affected individuals (V: 1, V: 2, V: 3 and V: 4) share common homozygous region (chr15:38,657,924-54,623,246) total of 15,965,323 bp, contained total 246 genes. Based on their expression in brain 8 genes were selected for screening in this family to find the exact pattern of mutation. DNA of the affected individual from family MR202 (IV: 3 and IV: 4), MR203 (IV: 2 and IV: 3), MR207 (V: 2 and V: 5), MR209 (III: 3 and IV: 1), MR213 (IV: 1, IV: 2 and IV: 3), MR214 (IV: 1 and IV: 2) were subjected to Sanger sequencing for AFF2 and RPTN mutation due to belonging to same localities, But no pathogenic variant was identified in these two genes. Thus, the molecular genetic analysis revealed here molecular basis of Intellectual disability in these families and define role of genes for normal neuronal development. More detailed study of the disease pattern underlying these conditions may in turn providing treatment therapies, pre-natal diagnosis for affected families in near future.

Introduction

Human nervous system is a complex meshwork of nerve and non-nerve cells that are constantly involved in communication and coordination among the different body parts and with external environment. Morphologically and physiologically, nervous system consist of two major division; a central nervous system (CNS) that is composed of brain and spinal cord, and a peripheral nervous system (PNS) that consist of nerve cells (neurons). Brain performs multiple functions among which "memory and learning" are the utmost functions that maintain the integrity of life and enable a person to live in the society. Memory or learning is of two kinds; declarative (consist of conscious through/feeling) and no-declarative (associated with procedural knowledge), and involve structural and functional adaptation in synapses. Neurologists that declarative memory stored and operated through hippocampus, while amygdale, striatum and nucleus are associated with non-declarative form of memory. The advancement in learning or cognition is the result of a phenomenon termed as synaptic plasticity in which new synapses are formed and the old synopsis get strengthen (Flint, 1999).

Intra familial matting's, also termed as consanguineous marriage is preferred in many countries of North and sub Saharan-Africa, Middle East, Central and South Asia (Bittels and Black, 2010). The large population size of the India sub-continent, which consists of Pakistan, India and Bangladesh, provide a significant chance for studies of genetic disorders (Gadgil et al., 1998). Additionally, the Pakistan population (over 140 millions), with over 60% consanguinity (>80 among first cousin) and large families size is very valuable for genetic studies. Moreover, Pakistan is inhabited by people of diverse ethnicities (Pukhtoon, Sindhi, Punjabi, Balochi, Saraiki, Kashmiri, Gilgati etc) having unique familial and social features (Mehdi et al., 1999). There are three main reasons for consanguineous are (religion, economics, and cultural). Religion put a great effect on consanguinity (Maulik et al., 2011). Inbreeding increases the likelihood of developing early-onset (congenital) autosomal recessive genetic disorders, including neuro-developmental disorders, which seriously impact the socio-economic status and quality of life. Autosomal recessive intellectual disability accounts for around 1/4 of the total genetic ID cases (Higgins et al., 2004).

1.1 Intellectual Disability

Cognition is due to cellular, biological and several molecular events in CNS and a slight change in any of these process finally leads to cognitive impairment or ID (Chelly et al., 2006). In addition, developmental delay or intellectual disability is also believed to be the result of a defect in synapse formation and plasticity.; therefore it is also termed as neuro-developmental disorders (Fedulov et al., 2007).

1.2 Prevalence of Intellectual disability

The prevalence of ID in the population is 2-3% (Uyguner et al., 2007). The clinical and heterogeneity of ID poses a major research challenge for clinicians and researchers. (Jennifer and Linda, 2004); (Basel, 2008). Intellectual disability is linked with about 900 genetic disorders (Opitz, 1996); (Castellvi and Mila., 2001).

1.3 Clinical Features

Clinically, intellectual disability occurs in syndromic or nonsyndromic form, malformation of eye, otorynolaryngologic, skin, bone or others are present in association with cognitive dysfunction in syndromic form, while in nonsyndromic Intellectual disability, only cognitive or learning impairment is present. There are several types of ID, such as Joubert's syndrome, Kahrizi's syndrome, and most of them have limited knowledge. To date, non-syndromic intellectual disability (NSID) is the least studies, and therefore only 47 loci and 26 genes have been identified. Comparing X-linked genes with autosomal chromosomes (chromosomes other than sex chromosomes), there are many genes that control the brain, and their functions are not yet known. Even in non-syndromic intellectual disability, 21 loci remain unidentified and the causative gene needs to be identified. Given the high rate of consanguineous marriages, 21 potentially uncharacterized loci (many more still under the tip of ice berg) and increased incidence of genetic ID conditions in Pakistan.

Intellectual disability, previously termed as mental retardation, is a manifestation of brain dysfunction, which affects the ability to learn of individual and cause cognitive dysfunction. Lower IQ (<70), (Table 1.1) impairment of adaptive skills and disease onset under 18 years of age level are main significant features of intellectual disability. Prevalence rate of Intellectual disability among general population is 1-3%, but its

percentage is high in the developing countries especially in regions where consanguineous marriages ratio is high such as in Pakistan and in the other Arabs countries. In the developing countries the main causes of Intellectual disability are parental consanguinity, under nutrition and cultural deprivation. Clinically Intellectual disability has been divided into two types, nonsyndromic Intellectual disability and syndromic form. On the basis on IQ level Intellectual disability may be, mild if IQ level is less than 70 (<70), moderate if IQ level: > 35-40 <50-55), sever if (IQ level: >20-25 < 35-40) and profound if (IQ level: <20-25) (American Psychiatric Association, fourth edition and ICD -10 classification (WHO, 1992)).

1.4 Diagnosis

Proper analysis of ID is an essential requirement to beginning a suitable assessment related to service referral, suitable ongoing management directed at likely co-morbidities and treatment that meets the needs of families (Shevell, 2006; Shevell, 2009). The diagnosis is usually expressed primarily on the basis of clinical outcome. The soundness of such analysis is associated with the degree of direct involvement with these patients by the clinician. Validity is augmented by clinical remarks, several efforts from reliable third party informers (e.g. mentors), frequent observations over time and contribution either parallel or subsequently from an interdisciplinary expert team proposing corresponding skill sets (e.g. occupational, physicians, therapists, physiotherapists, psychologists, and speech language pathologists). Associated understanding to social, cultural and linguistic variety is also particularly relevant. Certainly, such variable environments may prevent the convenience for management of uniform assessments that are the trademark of an objective verifying diagnosis of intellectual disability (Shevell, 2010). A selection of extensively used standardized measure for cognitive function exist (Majnemer, 2012).

1.5 Genetics of Intellectual Disability

The cause of ID as suitable for their description as 'symptom complexes' is fairly heterogeneous. Both genetics and environmental factors contribute equally in the development of disease. The environmental factors may be undernourishment, social deprivation, toxic materials exposure, brain injuries, complications in pregnancy, and sickness in early childhood etc. Currently, About a quarter to half of known causes appear

to be of genetic origin. There are many genetic causes, including chromosomal abnormalities (such as aneuploidy), small deletions/ duplicate / rearrangements (such as changes in copy number variation), and single-gene disorders. (Van, 2011). To date almost 560 genes have been associated with intellectual disability (Hamosh, omim.org.). All modes of Mendelian inheritance (i.e. autosomal dominant, autosomal recessive and X-linked) have been associated with the intellectual disability. Although the number of specific genes associated with intellectual disability has improved significantly over the past decade, an accurate molecular diagnosis of many suspected genetic causes is currently lacking. Presently, only a few well described single gene associations with a highly documented clinical phenotype (i.e. Fragile X-FMR1 and Rett syndrome-MECP2) are regularly verified at a molecular level through diagnostic assessment (Table 1.2) (Sherr et al., 2013), although this number is set to increase with the imminent introduction of next generation sequencing-based clinical diagnostic panels. In addition to the nuclear genome, defects in the mitochondrial genome may lead to syndromic intellectual disability, presenting a maternal inheritance pattern. (Srour and Shevell 2014).). Despite the fact that there may be sign for an incomprehensible range of genes involved, there appears to be a union of gene action into a number of isolated effective process network leading to failure of ID phenotype integration. These elementary procedures consist of cellular signaling cascades, inter-neuronal connectivity, neuronal proliferation, neuronal migration, and the extensive guideline of genetic/epigenetic transcription and translation (Srour and Shevell 2014). Improved identification of genes responsible for global developmental delay and mental retardation advances the understanding of the molecular basis of learning and memory, which is important for understanding human cognition and intelligence nervous perspective viewpoint (Kaufman et al., 2010; Van, 2011).). An increasing awareness of molecular pathways will permit the ultimate sensible selection of candidate gene and pharmacological methodologies.

1.6 Classification of ID:

Single gene defects leads to the development of monogenic Intellectual disability. On the basis of mode of inheritance (autosomal recessive or autosomal dominant) and position of mutated gene on chromosome (autosomal or X linked) monogenic Intellectual disability is classified into two groups.

1.6.1 X-linked Intellectual Disability

This intellectual disability group accounts for 10-20% of all cases of intellectual disability, with 1 in 600 men affected more than women. (Winnepenninckx et al., 2003). X Linked Intellectual disability is either inherited in dominant manner (defect in single allele copy) or in recessive pattern (both copies of allele are defective). Currently 217 X linked Intellectual disability loci have been mapped and eighty three genes causing Intellectual disability has been identified. Linkage and mapping studies has determined more than 217 X-linked loci in which 98 are mapped in syndromic families (with 33 mapped and 37 cloned genes) 51 loci in neuromuscular disorders (with 13 mapped and 27 cloned genes) and 68 loci in non-syndromic disorders (with 49 mapped and 19 cloned genes) (Liski et al., 2008). Though amongst the genes identified 11 are involved in both syndromic as well as in non-syndromic XLID. In case of mutations in OPJN1 are linked with nonsyndromic X-linked cerebellar hypoplasia and Mental retardation (Portes et al., 2004). In the same way PQBP1 gene mutation is accountable for renpenning syndrome (including Intellectual disability, Microcephaly and Microorchidism) and nonsyndromic X linked Mental retardation (Lenski et al., 2004).

1.6.2 Autosomal Dominant Intellectual Disability

Intellectual disability of this category is a rare condition mostly because the Intellectual disable individual does not reproduce mainly and therefore it is extremely difficult to get the autosomal-dominant pedigrees. The autosomal dominant genes are frequently mapped during chromosomal breakpoints assays in patients with balanced translocations duplications or inversion (Basel, 2007). The online-Mendelian-inheritance-in-man (OMIM) database has catalogued seven gene/loci involved in autosomal dominant intellectual disability. Genes/loci include "MRD2 on 9p24 contains cytokinesis 8 gene

dedicator" (DOCK8) (Griggs et al., 2008), MRD1locus on chromosome 2q23.1 contains the gene for methyl binding protein CpG 5 (MBD8) (Wagenstaller et al., 2007), MRD3 on 16q24.3 harbor cadherin 15 gene (CDH15) (Bhalla et al., 2008), MRD5 on 6p21.3 include synaptic RAS GTPase activating protein 1 (SYNGAP1) (Hamdan et al., 2009), MRD4 on 11q24.2 encounter KIN of IRRE like gene 3 (KIRREL3) (Bhalla et al., 2008), glutamate receptor ionotropic, N-methyl-D-aspartate, subunit 2B gene (GRIN2B) in MRD6 on 12p12 (Endele et al., 2010) and double specificity tyrosine phosphorylation regulated Kinase 1A (DYRK1A) in MRD7 on 21q22.1 (Van et al., 2011)".

1.6.3 Autosomal Recessive Intellectual Disability

ARID is responsible between 1/4th to 1/2 of the total inherited disorders of Intellectual disability (Higgins et al., 2004). Clinically, autosomal recessive intellectual disability has been divided into nonsyndromic or syndromic. Additional characteristic like malformations of the skin, eyes are associated with cognitive disability in syndromic forms of Intellectual disability, whereas in nonsyndromic form of Intellectual disability only learning or cognitive impairment is present. The syndromic cases have reported e.g. bardet biedl and Joubort syndrome, kahrizi syndrome, and the majority of these are less diagnosed. Therefore, we have incomplete information about the molecular basis of nonsyndromic Intellectual disability and similarly only forty seven loci and twenty six genes have been known. Comparing the ratio of autosomal gene (22 chromosomes), and X-linked genes (on a single X chromosome), Many (hidden) genes involved in the regulation of physiology in the human brain are still unknown.. In the case of nonsyndromic Intellectual disability, twenty one loci are still unknown and mutation in these genes responsible for diseased phenotype is yet to be identified.

Numerous autosomal recessive ID genes have been identified through exome sequencing which includes, methyltransferase like 23 (METTL23), DDHD domain containing 2 (DDHD2), calpain 10 (CAPN10), solute carrier family 6 (neutral amino acid transporter) member 17 (SLC6A17) (51-53) and serine/threonine/tyrosine interacting-like 1 (STYXL1). Among the ARID genes identified so far, 32 genes have been identified by using whole exome sequencing technology (Hill et al., 2016). Based on the recent studies it is anticipated that approximately 300 genes are associated with ARID, but majority of

these genes are associated with syndromic forms of ID (Musante and Ropers, 2014; Vissers et al., 2016). Similar studies carried out in Pakistan result in the identification of novel ID genes like MAN1B1 (Rafiq et al., 2011), TRAPPC9 (Mir et al., 2009), NSUN2 (Khan et al., 2012),FBXO31 (Mir et al., 2014), METTL23 (Bernkopf et al., 2014), FMN2 (Law et al., 2014), KMT2B, ZNF589, HHAT (Agha et al., 2014), SLC6A17 (Iqbal et al., 2015), DCPS, EDC3 (Ahmed et al., 2015), LMAN2L (Rafiullah et al., 2016).

.**Table 1. 1:** Intellectual disability Classification on the basis of IQ Level.

Class	IQ (%)	Population (AAMR 2005)
Severe ID	>19-<35	3-4%
Mild ID	>49-<70	85%

Table 1. 2: Some causes of developmental delay (Rauch et al., 2006)

List of common developmental delay	Causes
Trisomy 21 (Down syndrome)	Extra copy of chromosome 21
Micro deletion 22q11.2 (common 3 Mb)	3 Mb Deletion on chromosome 22q11.2
Williams–Beuren syndrome	Deletion of genes CLIP2, ELN, GTF2I, GTF2IRD1, and LIMK1
Fragile X syndrome	Mutations in the FMR1 gene
Cohen syndrome	Mutations in the VPS13B gene
Monosomy 1p36.3	Deletion of chromosome 1p36.3
Angelman syndrome	Mutations in the UBE3A gene
Coffin–Lowry syndrome	Mutations in the RPS6KA3 gene
Neurofibromatosis 1	Mutations in the NF1 gene
Prader–Willi syndrome	Deletion of chromosome 15q11-13
Atypical distal micro deletions 22q11.2	Deletion of chromosome 22q11.2
Coffin–Siris syndrome	Mutations in the ARID1A, ARID1B, SMARCA4, SMARCB1, or SMARCE1 genes
Kabuki syndrome	Mutations in the KMT2D or KDM6A gene
Mucopolysaccharidosis I	Mutations in the IDUA gene
Pallister–Killian syndrome	Presence of isochromosome 12p
Smith–Lemli–Opitz syndrome	Mutations in the DHCR7 gene
Sotos syndrome	Mutations in the NSD1 gene
Smith–Magenis syndrome	Deletion of chromosome 17p11.2
Trisomy 13 (Patau syndrome)	Extra copy of chromosome 13

Table 1.3: Function of reported genes.

Gene	Function
PRSS12	This gene encodes for Neurotrypsin consisting of 875 amino acids is a neuronal serine protease (Molinari *et al.*, 2002). Mutation in this gene cause severe intellectual disability (Rejf *et al.*, 2007).
CRBN	Mutation in CRBN gene cause NS-ARID. It consists of 11 exons and encodes a protein of 442 aa (3p26.2). Its expression is high in the human hippocampus (a chief region that is involved in long-term potentiating and learning) (Higgins *et al.*, 2004).
CC2D1A	The cytogenetic location of *CC2D1A* gene was 19p13.12 and contains 31 exons. This gene encodes for protein composed of 950 amino acids, comprise of 2 conserved regions i.e., a DM14 domain (unique, function unknown) and a C2 domain (function in calcium-dependent phospholipid binding). *CC2D1A* has highly expressed in the hippocampus and cerebral cortex in human (Basel *et al.*, 2005).
GRIK2	*GRIK2* encodes for protein consisting of 908 aa which is highly expressed in cerebral cortex and cerebellum in human brain (Bahn *et al.*, 1994).
TUSC3	*TUSC3* has a wide range of expression in human tissues, including brain. (Molinari *et al.*, 2008). Oligosaccharide transfer action of *TUSC3* is significant for normal

	cognitive- development.
TRAPPC9	First identified by (Mir *et at.*, 2009) in 6 consanguineous families in Pakistan. It encodes for protein (NIBP) which is highly expressed in cerebral cortex neurons. Mild to severe ID with whimsical microcephaly results from mutation in *TRAPCC9* gene. It has been reported that there is a great decrease grey matter in the people with defective *TRAPCC9*.
TECR	(Topper *et al.*, 2011) was the first to report TECR. It contains 13 exons and positioned on chromosome 19 at 19p13.12 locus. It has been reported that individuals with mutation in *TECR* gene have speech and developmental delays.
ST3GAL3	*ST3GAL3* was first reported in 2 Iranian families. It is composed of 12 exons which encodes for protein (beta-galactoside) of Golgi membrane.
MED23	First reported by (Hashimoto *et al.*, 2011) in Algerian family. It has 29 exons. Homozygous recessive missense mutation is involved in this gene causing Intellectual disability, result from the substitution of Arganine with Glutamine.
MAN1B1	*MAN1B1* was first reported by (Rafiq *et al.*, 2011) in one Iranian and 4 Pakistani families. Missense and nonsense mutation in

	this gene causes severe intellectual disability from mild to severe type.
FBXO31	*FBXO31* gene was first reported by (Mir *et al.*, 2014) in a Pakistani family having consanguineous marriage. Main function of *FBXO31* gene is tumor suppression by forming SCF-FBXO31 complexes that is important for the normal cell cycle. It has been reported that framshift mutation in *FBXO31* gene causes the premature termination codon.
CCDC82	*CCDC82* has high expression in brain and was first reported by (Harripaul *et al.*, 2017) in consanguineous Pakistani families.
MAPK8	MAPKs are a member of serine-threonine protein kinases family that participates in a main signaling system by means of which cells converts extracellular stimuli into intracellular response (Derijard *et al.*, 1994).
C12ORF4	Homozygous recessive mutation in the *C12ORF4* gene on chromosome 12p13 cause MRT66 (autosomal recessive mental retardation) (Philips *et al.*, 2017).
MPDZ	(Al-Dosari *et al.*, 2013) reported a homozygous c.628C-T change in exon 6 of the *MPDZ* gene, resulting in Hydrocephalus, Congenital, 2, with brain abnormalities.

PIDD1	(Tinel and Tschopp ; 2004) reported that apoptosis initiated by genotoxins has been regulated by the complex formed by PIDD and caspase-2. It has been reported as a candidate gene in Intellectual disability by (Harripaul *et al.*, 2017).

Objective of the Study:

The objectives of this research project was to find the genetic factors related with neurologic disorders through whole exome sequencing and homozygosity mapping and their functional analysis through cloning and localization.

The following are the objectives of the proposed molecular and functional genetics studies:;

1- The overall objective of the investigative study is to identify familial causative disease genes and underlying pathogenic mutations that characterize intellectual disability through positional cloning approaches (homozygous mapping) and whole exome sequencing.

2- To elucidate the functional consequences of identified genetic defect through in vitro-mutagenesis, cloning and cell based assays.

3- The planned study was focused on determining the genotype-phenotype correlation in order to help genetic counseling and molecular diagnostics.

4-To determine the genetic heterogeneity of individual neurodevelopmental diseases in the population of Pakistan by recruiting more patients which help to determine the recurrence rate of the mutation and its founding effect.

Materials and Methods

Materials and Methods

2.1. Families Ascertainment and Informed Consent Form Filling

Fifteen families with Intellectual disability were recruited in the present study. Prior to study, informed consents were signed by the patients take parts in the research study and the study was approved from Board of Advance Studies and Research and from ethical Committee Department of Biological Science, International Islamic University Islamabad. Research was performed in the HMG Lab, Department of Biological Science IIUI and in collaboration with University of Exeter Medical School, Medical Research (Level 4), RILD Wellcome Wolfson centre, Royal Devon & Exeter NHS Foundation Trust, Exeter UK.

2.2. Clinical Evaluation

Identified families were visited at their place of residence. In order to obtain maximum information about the disease, normal healthy members of the recruited families were interviewed. Local Psychiatrist confirmed the diseased phenotype.

2.3. Drawing Pedigrees

To depict the inheritance pattern of the disease with this family as per (Bennett *et al.*, 1995) pedigrees were drawn. In the pedigree, squares represent the male while circles represent the female. Affected individuals are represented by filled circles and filled squares. Cross line on the circles and squares represent the deceased individuals in the family. Consanguineous marriages are represented by double line in the pedigree.

2.4. Blood Collection and DNA Extraction

Peripheral blood samples from both the affected and normal members were collected and has been transferred to vacationer (ACD) tubes (standard potassium (EDTA 10 ml BD)). DNA from collected blood samples were extracted by organic method.

2.4.1. DNA Extraction by Organic Method

Phenol-chloroform method (Sambrook and Russell, 2001) was used to extract DNA from whole blood samples.

1. 300 µl blood was taken in 1.5 ml eppendorf tube and mixed with 500 µl of lysis buffer. Tubes were kept at room temperature for 20 minutes. In order to mix the blood and lysis buffer in the tube containing lysate, shake the tubes at regular interval. The tubes were centrifuged at 10000 rpm for 2 minutes in a micro centrifuge (model 1-14, Sigma, Germany 2010-2011).
2. Supernatant was discarded and pallets were re-suspended in 400 µl of lysis buffer and vortexed the tubes for 20 minutes in order to mix lysis buffer with pellets. The tubes containing pellets and lysis buffer were centrifuge again at 13,000 rpm for 1 minute.
3. Supernatant was discarded with care. Add 400 µl of solution "B", 12 µl of 20% "SDS" solution and 25 µl of proteanase-K (25 mg/ml). Eppendorf tubes were vortexed for 15 seconds to mix the solution and pellets. Tubes containing solutions and pellets were incubated at 37° C for overnight.
4. Next day after incubation 500 µl of solution "D" was added and mixed them by shaking vigorously. Centrifuge the tubes at 13,000 rpm for 1 minute.
5. Supernatant was transferred to new 1.5 ml eppendorf tube. Equal volume of solution "C" was added and tubes were left for 3-4 minutes after mixing the solution in the tube by shaking. Tubes are centrifuged again at 13,000 rpm for 1 minute.
6. Supernatant was collected in new eppendorf tube and 700 µl of chilled iso-propanol was added. Tubes were centrifuged again at 13,000 rpm for 1 minute.
7. Supernatant was discarded and 20 µl of 70% ethanol was added in order to remove the impurities from DNA.
8. 100 µl of TE buffer was added and stored the tubes at 4°C.

Table 2. 1: Solutions composition for DNA Extraction.

Lysis buffer	0.32 M sucrose, 10 mM Tris pH 7.5, 5 mM $MgCl_2$, 1% Triton X-100
Solution B	10 mM Tris pH 7.5, 400 mM NaCl, 2mM EDTA pH 8.0
Solution C	Buffered phenol
Solution D	Chloroform (24 ml), Iso-amyl-Alcohol (1 ml)

2.5. Targeted Genotyping

An initial association search was performed using a set of available microsatellite markers for known genes associated with intellectual disability to find genetic defects in families with intellectual disability in Human Molecular Genetic Lab IIUI. Targeted genotyping was done as listed in Table 2.8.

2.6. Polyacrylamide Gel Electrophoresis

Primer designed for each marker that run for affected with their normal parents by PCR. Result of amplified PCR products were determine by 1.8% agarose gel follow by electrophoresis. Microsatellite DNA can be measured by the size of an individual and compared to its parents. The gel mixture was crystallized by adding 50 ml of an 8% PAGE gel mixture containing 70 µl of TEMED. 70 µL of APS (25%) was added to the previously prepared mixture and TEMED and 25 mL mixture was poured, the comb was placed in the plate and kept for about 1 hour leaving the remaining gel in the beaker. To see if it has been polymerised.

2.6.1. Protocol for Electrophoresis

1 liter of 1X TBE was heated in the microwave for 6 minutes. The gel comb is carefully removed. (The buffer was poured to cover the strands at the bottom of the electrophoresis tank). The gel glass was carefully placed in the buffer and it was ensured that there were no air bubbles at the bottom of the gel glass. The buffer were poured into the chamber until they covered both sides under the gel glass. The gel runs at 100 mA. 2 µl of PCR sample and 2 µl of loading dye were mixed "denaturing loading dye consisted of , 70 µl xylene

cyanol, 70 µl bromophenol blue, 10 ml formamide". Mixed product was loaded in each well. Glass was removed from gel and on the left side mark was made.

2.7. Micro-array Analysis

Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 and 2.1 arrays in families which did not show linkage to known genes.

2.8. Whole Exome Sequence Analysis

WES was carried out at University of Exeter Medical School, "Medical Research (Level 4), RILD Wellcome Wolfson Centre, Royal Devon & Exeter NHS Foundation Trust, Barrack Road, Exeter, EX2 5DW, UK". WES capture of a single family member of each families was executed by using the "Agilent SureSelect Human All Exon V4 exome enrichment kit". "Sequence reads were aligned to the reference human genome with the Novoalign software package" (Novocraft Technologies Sdn Bhd). Duplicates were excluded from the final evaluation because of "PCR cloning or optical replication" and variants mapping to multiple locations from downstream assessment. Sequence coverage was measured with customized scripts and the BedTools package276. SNP and small insertion deletions were identified and quality filtered within the SamTools software package277 and in-house software tools. Annotation was done by using Annovar tool 278.

2.9. Primer design

Ensemble Genome Browser (GRCh37 assembly, Feb 2009) was used to get gene sequence (http://www.ensembl.org/Homo_sapiens/Info/Index?db=core). Primer3 software version 0.4.0 (http://frodo.wi.mit.edu/primer3/) was used for primers designing with the following criteria:

1. 20 to 24 bases of primers size were selected for particular temperature that depend on the size of the primers.
2. In order to maintained efficiency of PCR (55-620C) for both primers (Forward and Reverse) annealing temperature were kept constant.
3. Guanine – Cytosine (GC) ratio was kept between 40-60%.

UCSC Genome Bioinformatics website (website http://genome.ucsc.edu/) was used for Blast searches and for In silico PCR analysis

2.10. Polymerase Chain Reaction (PCR)

For in vitro cloning of DNA, that allow exponential production of particular lengths of DNA from only small volume of starting material, PCR is used. Primers for candidate genes were amplified in PCR. Primers were prepared from "dry oligonucleotides".

2.10.1. Preparation of Master Mix

Master mixture was prepared by mixing forward and reverse primers, Taq DNA polymerase, distelled water and Taq Green buffer for each PCR reaction. In each Eppendorf tubes (0.2 ml) 9.2 µl of master mix was put in each individual tubes, 0.8 micro liter of DNA was poured into negative and positive control tubes (Table 2.2), 0.8 µl of dH2O was used in each tubes.

To confirm absence of contamination in each PCR reaction, 1 negative control and 1 control reaction was setup for each PCR.

Table 2. *2:* Standard reaction mixture for one tube.

Forward primer (5 pmol/µl)	0.4
Reverse primer (5 pmol/µl)	0.4
dNTP (5mM)	0.4
Dream Taq Green buffer	1.0
Dream *Taq polymerase*	0.1
Double distilled water (ddH_2O)	6.9
DNA (control or sample) or ddH_2O	0.8
Total	10

2.10.2. Cycling Conditions for PCR

denaturation, primer annealing, and DNA synthesis (extension) are three cycles during PCR reaction. For touchdown PCR, programmes were set. For each subsequent program annealing temperatures were lowered by 20C. For all PCR reactions, the denaturation and extension temperatures were set to 950 ° C and 720 ° C, respectively..

The final annealing temperature was set 50°C lower than the lower Tm of the primer pair. To avoid non-specific PCR products, PCR was optimized by performing sequential PCR or by increasing the temperature in 10°C increments. Reactions were split into three programs for each PCR. The first two programs consisted of two cycles of denaturation, annealing and elongation, and the three programs consisted of 45 cycles. Each step (i.e. segment) of the cycle was performed within 30 seconds (Table 2.3).

Table 2. 3: Cycling conditions for PCR.

1	Denaturation	95	2	30
	Annealing	TD+4		
	Extension	72		
2	Denaturation	95	2	30
	Annealing	TD+2		
	Extension	72		
3	Denaturation	95	45	30
	Annealing	TD		
	Extension	72		

2.11. Agarose Gel Electrophoresis

To get desired region of DNA that was amplified during PCR reaction, Agrose gel electrophoresis was used, separating PCR products by size. In order to make 1.8% agarose gel 1.8 g of agarose powder was added into 100 ml of 1X TBE buffer in beaker. In order to get clear solution, mixture was heated in a microwave oven at high temperature. Running tap water was used to cool down the heated mixture. 5.0 µl of ethidium bromide solution (0.1 µg/ml final concentration) was added to the mixture before it was poured out to set. To form the loading wells comb was placed across the gel and was left for 30 minutes. 5 µl of PCR product was mixed with 3.0 µl of gel loading dye, and was then loaded into the gel. A 1 kilobase plus (kb+) ladder was also loaded with the PCR product to estimate the size of the PCR product. The loaded gel was run in 1-XTBE buffer solution and a voltage of 100 V was applied for 30 min. PCR products were visualized and photographed using ultraviolet light..

2.12. EXOSAP PCR product purification protocol

Purification was done using EXOSAP, a product that consist Exonuclease-I (EXO) with Shrimp Alkaline Phosphatase (SAP) (Table 2.4). Single stranded DNA in a 3' to 5' direction was degraded by Exonuclease-1, which release deoxyribonucleoside 5-monophosphates in orderly and dNTPs were released from the independent primers. Phosphate groups from 5'- and 3'was released by Shrimp Alkaline Phosphatase (SAP) from extra nucleotides. In 5 µl of each PCR product, 2 µl of EXOSAP was added. The sample was then incubated at 37 ° C for 30 minutes. The sample was exposed to 95 ° C for 5 minutes to inactivate EXOSAP (enzyme). (Table 2.5).

Table 2. 4: Composition of EXOSAP solution.

Exonuclease-I	2.5
Shrimp Alkaline Phosphatase	25
dH_2O	972.5

Table 2. 5: Conditions for the EXOSAP programme.

Temperature	Time (Minutes)
37^0C	30
95^0C	5

2.13. Sanger Sequencing

Sequencing was performed analyze the co-segregation of candidate genes in affected individuals. EXOSAP protocol were used to purify the PCR products. After that, PCR Purified products were then put into the reaction mixture for sequencing (Table 2.6). This was performed in a 96-well plate. For forward and reverse primer amplification, The separate PCR reaction was set up such that the combination of fluorescent deoxynucleotides (ddNTPs) occurred in only one double-stranded DNA strand..

The following conditions were used for PCR sequencing:

950C for 30 seconds, 500C for 15 seconds, and 600C for four minutes. This was programmed for 25 cycles (Table 2.7).

BIGDYE® X TerminatorTM Purification Kit (ABI, Applied Bio-systems) were used to purify amplified products according to the instructions manufacturer's. Then the Samples were placed in the automated sequencer. CLC sequence viewer (http://www.clcbio.com/products/clc-sequence-viewer/) software programme was used for data analysis.

Table 2. 6: Components for sequencing amplification.

Forward or Reverse primer (5 pmol/µl)	1.0
Purified PCR product	2.0
Big dye	1.0
Big dye buffer	2.0
ddH_2O	4.0
Total	10

Table 2. 7: Conditions for sequencing reaction.

Denaturation	95	30 seconds	
Annealing	50	15 seconds	25
Extension	60	4 minutes	

2.14. Sequence Analysis

Sequencing reactions were analysed using 3130*xL* Genetic Analyzer (Applied Biosystems®). Sequencing files obtained from the 3130*xL* Genetic Analyzer (Applied Biosystems®) were analysed using CLC Sequence Viewer (CLC bio).

2.15. Bioinformatics Analysis

Psipred (McGuffin et al., 2000) and I-Tasser (Roy et al., 2010; Zhang 2008) were used for the prediction of secondary and three dimensional protein models. RAMPAGE (Lovell et al., 2002) was used to check the reliability of predicted models reliability. UCSF Chimera (Pettersen et al., 2004) were used for visualization of models. Pockets on three-dimensional structures of proteins were identified using CASTp server (Binkowski et al.,

2003; Dundas et al., 2006). Meta SNP (Capriotti et al., 2013), I-Mutant2.0 (Emidio et al., 2005), and PREDICT SNP (Bendl et al., 2014) were used to the effect of mutation on stability of protein and to determine whether the mutation has an impact on normal function of protein or not.

2.16. Predicting Pathogenicity of the Mutations

Online tools including mutation Taster, SIFT, Polyphene were used to identify the pathogenic effects of sequence variants.

Table 2. 1: List of microsatellite markers for linkage analysis

Gene	Markers	forward + reverse	Annealing Temp.
TRAPPC9	D8S256	GTTCAAGGGCTCAGGGTTCT	55°C
		CTTCCACCTTTAGCCAAGGA	
	D8S1204	CCACTCTCATCTACCACCAC	
		TACCCCTTCTCCCAACTTAC	
	D8S1743	TGAATACTTAAAGTTTGCTAAAAGG	
		TGAGGTCTGCTGCTGG	
	D8S1837	AATGAAAGGCTGACCTCC	
		ACCCAGATTGCTTATGCTC	
FBXO31	D16S305	CCTCCCAGGTTCAGGCAATTCTTCT	55°C
		TAGGCGACAGAGTGGGACTCCTTA	
	D16S402	TTTTGTAACCATGTACCCCC	
		ATTTATAGGGCCATGACCAG	
	D16S520	GCTTAGTCATACGAGCGG	
		TCCACAGCCATGTAAACC	
	D16S2621	GTCATATGGGCCAATTCCC	
		TACCGCGTAGTGAGACTGTG	
TUSC3	D8S261	TGCCACTGTCTTGAAAATCC	55°C
		TATGGCCCAGCAATGTGTAT	
	D8S549	TCTGATTAGCCAACCTCTAACA	
		CTTACAAACACCACTATGCATG	
	D8S1731	CCAAGCAAATCATGGAAATC	
		AGCAAACTTATCCCACAAGG	
	D8S1827	GACAGAATCATGTGGCCTTT	
		TTTTGTAAAATGTAAAATTGGCTTT	

PRSS12	D4S191	GATCATAGCTTCCTCTCTTTGAATA	55°C
		AAAATGAGCCCACTAGAGATGGTA	
	D4S2392	CAGCCTTCCAAATAGCTTGA	
		TAGTAAGCAGCTCATTTTGGC	
	D4S3024	CTGGAAGCCAGGTAGGA	
		AACACTTAGAACTTGCAGCC	
NSUN2	D5S406	AAACATACCTCTTCCCACACA	55°C
		GATGCTAACTGCTGACTATGC	
	D5S580	TAGTCTCTTCATGACTTGGTA	
		CTGCATTCTAGCCTGGGC	
	D5S630	CTCTTCGCTCTTCTTTCTCC	
		TCACTGCTTTACCTTTCAGTG	
	D5S1981	CCTGTACCAATCCATGC	
		GAGCCATGTGAGTGTCC	
	D5S2505	TGTTGGAAGACTTCTCAGCC	
		CACACATGCTGTGTCTCTCA	

Table 2. 2: Primers used for ID genes sequence

Gene	Primers (forward + reverse)	Product size (bp)	Annealing Temperature
AFF2	CAGGCAATCAGAATACCAGTGA	245bp	59°C
	TGATCTGGGGTAAGCAGCTT		
CTSA	CCTCCCGTAGATGATCCGAG	204bp	58°C
	CAGTAGTGGAGGTGCTTGGA		
FUT2	ACCAGTGCTAGCCTCAACAT	247bp	57°C
	GTAGTTCTGCCAGGGGATCC		
ANKIB1	ACCAGACCCTCAGTATAAATTGG	431bp	58°C
	TCTACACAGGCTTTCATCCCT		
HRNR	GCCAGATCCATGCTGAGTGC	289bp	64°C
	CCAGGAGGGATCTAGCACAGGC		
MSRA	TGTCTCTCTAGCATGGGAGC	272bp	57°C
	TGCCCAATTTGTTGAAAGCA		
SSC5D	TCCGAATTCCACTGTGACCT	498bp	58°C
	AGCAGATGTGGGCCTTCC		
TBC1D3	GAAACGCAGGGTGTGTGGCC	1293bp	66°C
	CTGGTGACCAGGCCTGCAGTCC		

TRPM6	TTTTGTCCCTGAATGCTGGC	492bp	58°C
	CCTACCTCAGACCATGTGACA		
AMER1	GGAATGCCATGGGTGATGCC	465bp	59°C
	GGAGGACGAAGAGGAGGAAG		
FAM72B	ACTCTGGTAGGAGGGTCACTTA	568bp	57°C
	TAGGGTCAGACCTCAAACATGG		
GPR42	GCATTGAACGCTTCCTGAGT	458bp	59°C
	CGTCACGTAGATCCTCCACA		
KRIT1	TCTCATCACATTCCACCCCA	923bp	59°C
	CCCTGTCCTTGACCTACCTG		
LMTK3	GTGCTCACTTGAACCACGAA	657bp	58°C
	AGAAGAGGAGGAGGAGGAGG		
NOL3	GACTCGGGACTGCTGTTGGA	363bp	57°C
	GGGAGTGACTGAGGCTCAAT		
USP17L3	CTTGCCTCTATGGAAGCCAGA	623bp	61°C
	TGGGTGGGGTGTTAGCGGC		
ZNRF3	GAGGAGCGAGGATAAAGGGG	912bp	56°C
	GGGGTTCACAGGAGCTCTAG		

RPTN	GGTTAAAGTGAGAGGCTTGTCC	400bp	56°C
	GAAAGGCAGAACTCCCACCA		
SLC50A1	GGCTGAGTTATGGGGCTTTG	270bp	58oC
	AGGCTGGCTTCATAGAGGAA		
TARS2	CCCTACCTGCTTTCCTTCCA	250bp	56°C
	GAGGGCGAGTGTATCTGTGA		
ANAPC2	TACGATCCAGTGAGAGGCTC	359bp	58°C
	CTGAGCAAGGCGGTGAAGAT		
CACNA1A	ATCGTCACTCTCGCTGTAGG	623bp	57°C
	ACGTGTCCTATTCCCCTGTG		
CDKL5	GACACTGCAGCCCAATGAAA	452bp	60°C
	GTTGCTGGGAAGAAAAGGGG		
DDHD1	CAAATGCAGCCCGGGTTC	324bp	62°C
	TCGGGTCGGGATTTGAAAGA		
DNM1	ATGTACCACGCACTGAAGGA	979bp	58°C
	CAAGTGCTGGAAAGAAGGGG		
HEPACAM	CGCATGCTCATACACGTTCA	534bp	58°C
	TCCAGCAAGCTCTAACACCA		

TH: 27243

ROR2	GGCATGGAGACCTGTTTGTG	294bp	58°C
	GGTGGAGAGTGGGTTGGTAG		
TAF1	GACTAGGAGTGGGAAGAGGC	444bp	57°C
	TCACAGCTGAGCCCACTTAT		
ZBTB6	GAATGATTATGTGTAAGCCGACA	392bp	56°C
	ATCTCAACCGACACATCCGA		

Results and Discussion

Mutations in *AFF2* gene underlie Fragile XE syndrome, a form of non-syndromic X-linked cognitive disability.

3.1. Family MR200

Family MR200 affected with Intellectual disability was recruited from (Ghrini) FR Tank District of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.1). The normal cousin couple III:8 and III:9 have five children, out of which two (IV:2 and IV:3) shows ID and the other three members (IV:1, IV:4 and IV:5) were normal. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.1.1. Clinical Features

Affected individuals in family MR200 exhibited clinical manifestations of intellectual disability including problem in interaction with the other people of the society. They have long term memory loss. They show aggression in their action, learning disability. They have limited personal skills and have minimal language disability. Phenotypic representation is given in (Figure 3.2).

3.1.2. Molecular Genetic Studies

Due to the involvement of different genes in causing ID (table 2.8), microsatellite marker for each genes were used for linkage studies and genotype for affected members and normal members of the family MR200. In this family, seven DNA samples (III:6, III:8, III:9, IV:1, IV:2, IV:3 and IV:5), including two affected and five normal individuals were genotyped. In family linkage was not established with known genes. After family was excluded from linkage, the genomic DNA was subjected to whole Exome sequencing.

3.1.3. Whole Exome Sequencing and Variants Filtration

Affected individual (IV: 2) from family MR200 was subjected to Exome sequencing. Total variants 40000 were found in data of the affected individual, 39980 variants occurring in intronic/UTRs/synonymous were excluded. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of inheritance i.e. homozygous or compound

heterozygous. Following filtering of variants four deleterious sequence variants were found to be present in different genes are selected. One was located in AFF2 (Chr:X:g.148059506G>T (hg19 build); NM_001169123.1:c.3418G>T; p.Asp1140Tyr), gene considered an unlikely candidate to cause this condition. Second was located in CTSA (Chr:20:g.44520261_44520263del (hg19 build); NM_001167594.2:c.108_110del; p.Leu37del), third one is present in FUT2 (Chr:19:g.49206451G>A (hg19 build); NM_000511.5:c.238G>A; p.Ala80Thr), fourth one is present in DMD (Chr:X:g.31792222A>G (hg19 build); NM_004006.2:c.7397T>C; p.Met2466Thr). Table 3.1 shows filtered variant with amino acid change as a result of four variants in MR200.

3.1.4. Segregation Analysis

Co-segregation of selected variants in all affected and normal members were checked. Gene variants of FUT2, CTSA and DMD were not co-segregated in the family. Among the selected gene variants, the variant [NM_001170628.1:c.2371G>T; Chr:X: 148059506G>T (GRCh37)] in AFF2 was found to co-segregate with diseased phenotype within the family MR200 as shown in (Figure 3.3). Two affected individuals (IV: 2 and IV3) were carrying the mutant allele in Exon 17 of gene where the normal member (III:9) were heterozygous for the allele and the remaining members (III:6, III:8, IV:1 and IV:5) are wild type. This variant leads to a substitution of asparagine with tyrosine and at the evolutionary conserved position 1140 (p.Asp1140Tyr) according to the UCSC Human Genome (GRCh37/ hg19). The p.Asp1140Tyr variant is not listed in the Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org/).

3.1.5. *In silico* Structural Analysis

In order to understand the biological role of proteins deep study about structure and function has key role. Although computational strategy is not equivalent to experimental methods of structure determination in providing high-resolution, but still provide useful knowledge that can be used to understand biological mechanism.

To predict secondary and three dimensional protein models, Psipred (McGuffin *et al.*, 2000) 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). Pockets on three-dimensional structures of proteins were identified

using CASTp server (Binkowski *et al.*, 2003; Dundas *et al.*, 2006). Meta SNP (Capriotti *et al.*, 2013), I-Mutant2.0 (Emidio *et al.*, 2005), and PREDICT SNP (Bendl *et al.*, 2014) were utilized to estimate effect of mutation on stability of protein and to determine whether the mutation has an impact on normal function of protein or not. This structure analysis has been carried out in the present study using computational tools to study target proteins and their behavior. A secondary structure analysis show that after mutation structure features gets effected i.e. number of alpha (α)-helices, beta (β) strands/sheets and coils alter. Figure 3.4 shows representation of secondary structure features for AFF2, respectively. In figure 3.4 Part A is graphical comparative analysis of wild and mutant type depicting difference of structure elements and Part B and C is the sequence representation of wild and mutant type proteins, respectively with amino acids participating in structure elements have been highlighted. In case of AFF2, amino acids involved in point mutations i.e. Aspartic acid (D)/Tyrosine (Y), respectively were involved in making coil structure. In this protein, it has been observed that apart from difference in number of alpha helices, beta sheets and coils, there is difference in amino acids too which are involved in making these structure elements in wild and mutant types.

Sequences of genes were subjected to protein BLAST at NCBI as a prerequisite of protein modeling to find out the related proteins in database. No suitable templates were available to perform homology modeling approach hence alternative approach i.e. threading to model genes was adopted. Quality of predicted models is estimated through a parameter i.e. C-score which is confidence score. Three dimensional structures modeled for wild and mutant proteins are shown in Figure 3.5. Maximum percentage of residues in the favored region in evaluation process proved reliability of predicted models. There is no change in length of protein in case of missense mutations but the single residue difference (highlighted in part B of Figure 3.5) also has huge impact and involved in altering conformation and function of protein. Point mutations of AFF2 were also analyzed employing various tools to further check effect and stability of proteins. Protein stability decreased after substitution mutation according to calculations of I-Mutant2.0 which is Protein Stability Changes calculator. Meta SNP employs various predictors such as PANTHER, PhD-SNP, SIFT and SNAP to estimate implications of mutation on normal protein. Predict SNP also stated mutations as disease causing. Results of

all predictors Predict SNP, I Mutant and Meta SNP are shown in Figure 3.6 Part A, B and C, respectively.

Proteins exhibit specific conformations due to their secondary and tertiary structure. Binding pocket or active site is region where specific molecules bind, called ligands which can be any chemical or small interacting protein. To perform a specific function within body as per interactomics, proteins interact with other partners and this interaction takes place at these active sites which in turn depicts chemical change / reaction. In proteins, active site is normally hydrophobic pocket that includes side chain atoms. Indicating molecules that bind to target protein and interaction study is helpful for 3D structure's role in docking as well as for drug designing. Number of pockets and amino acids involve in each pocket is specific to every protein structure hence also used for discrimination, interaction and mutational analysis. Count of interaction sites was examined for the target proteins and results collected depict changes in number of pockets. Normal structures of AFF2 were observed with a total number of pockets as 210 respectively. The number of active sites remained same in case of point mutations in AFF2 but amino acids contributing these active sites got changed. This change shows implication of mutation on overall conformation of protein which eventually alters interaction and function of proteins.

3.1.6. *In silico* Interaction Analysis

From STITCH analysis shown in (Figure 3.7) AFF2 have strong interaction with NUMR2 protein. NUMR2 (Neuromedin U) is a protein like molecule (neuropeptide) used by nerve cells to communicate with each other. Its expression is high in the central nervous system and upper gastrointestinal tract. Recently, it has been reported that a particular receptor for neuromedin U is GPR66/FM-3 (NmU-R1). Expressional study of human NmU-R2 confirmed that it's mRNA have high expression in human CNS tissues. On the basis of these findings, we conclude that NmU-R2 is a novel neuromedin U subtype receptor that is likely to induce CNS-specific neuromedin U effects.

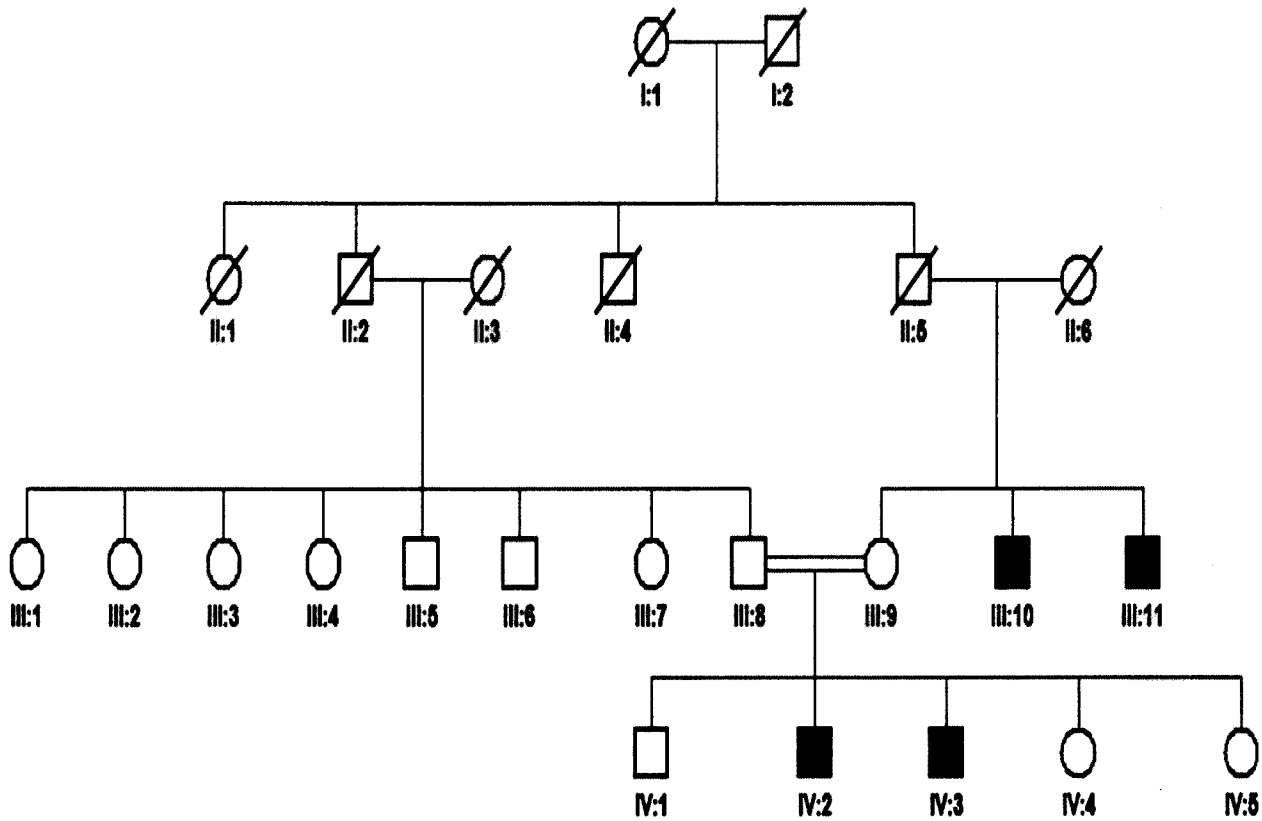


Figure 3.1: Pedigree of family MR200. Females are represented by circles and males are represented by squares. Affected individuals are represented by filled symbols while unaffected individuals are represented by clear symbols. Deceased members are represented by lines passes through symbols. A double line represents consanguineous marriages.

Figure 3.1: Phenotypic representation of affected individuals (IV: 3 and IV: 2) of family MR200.

Table 3. 1: List of filtered variants identified in the Exome sequencing data of the proband (IV: 2) from family MR200.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
ChrX:148059506	*AFF2*	c.5418G>T	Hom	missense	p.Asp1140Tyr
Chr20:44520261_44520263	*CTSA*	c.108_110del	Hom	in-frame deletion	p.Leu37del
Chr19:49906451	*FUT2*	c.238G>A	Hom	missense	p.Ala80Thr
ChrX:31792222	*DMD*	c.7397T>C	Hom	missense	p.Met2466Thr

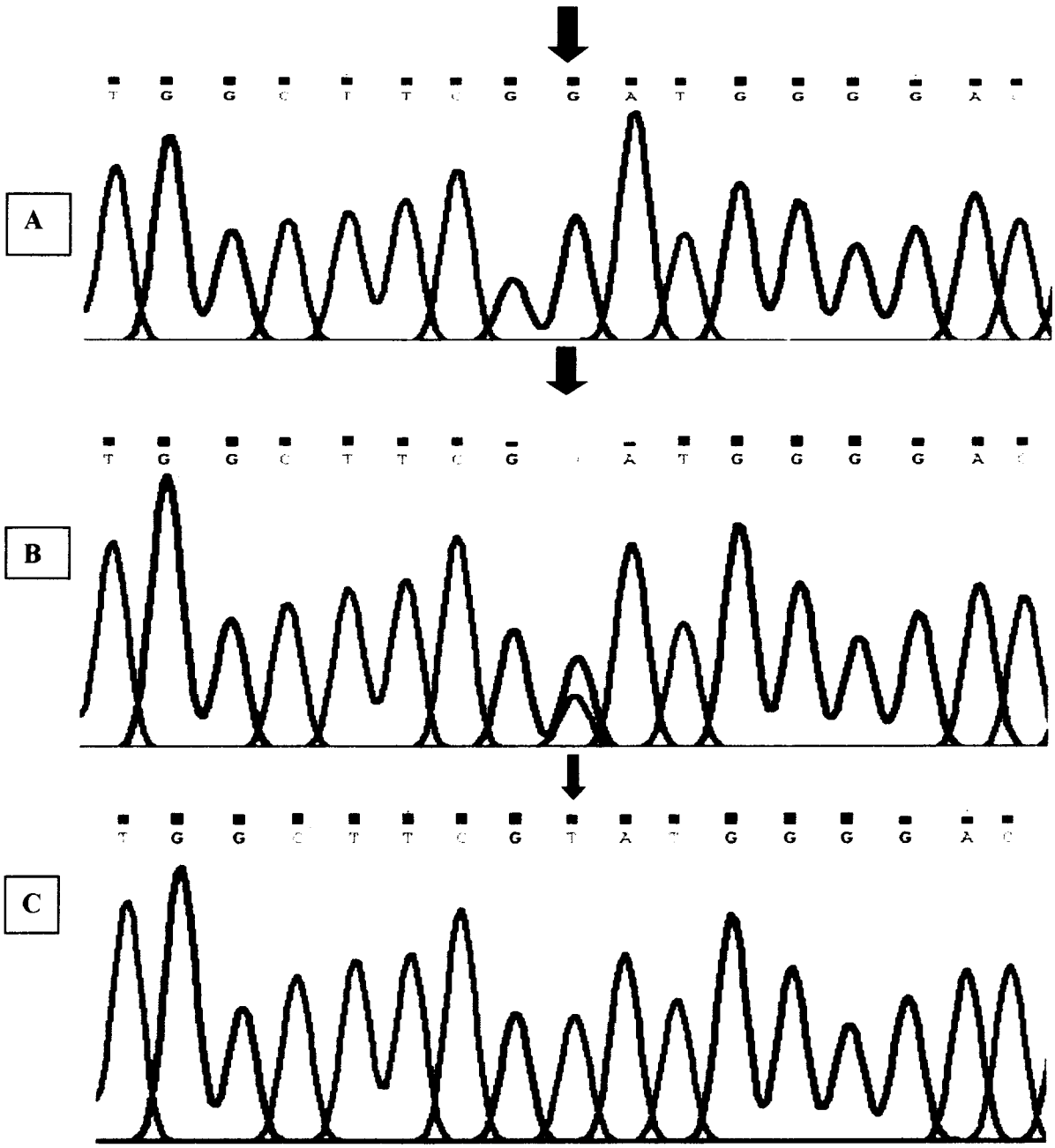
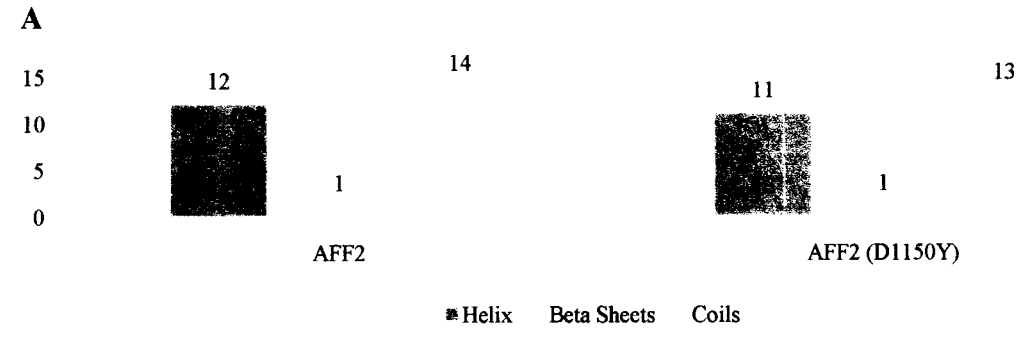
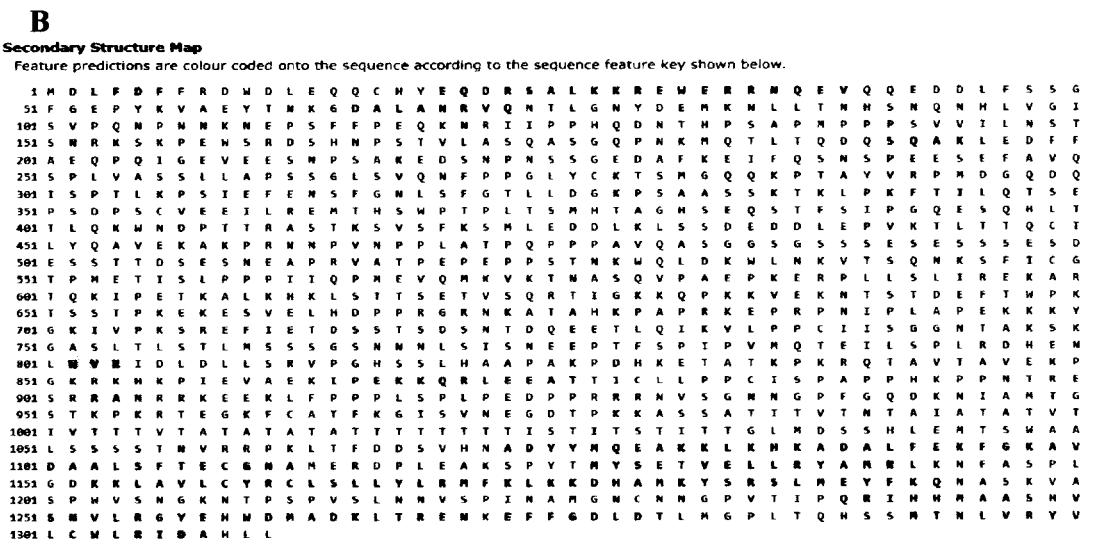


Figure 3.3: Chromatogram of variants c.3418G>T, p.Asp1140Tyr in **(A)** Wild type, **(B)** Carrier and **(C)** Affected individual.





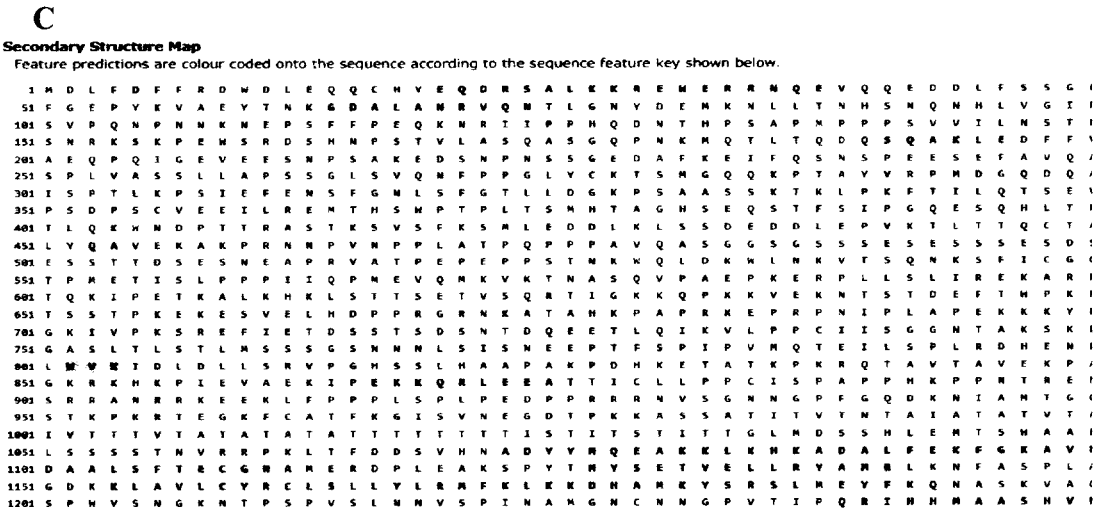
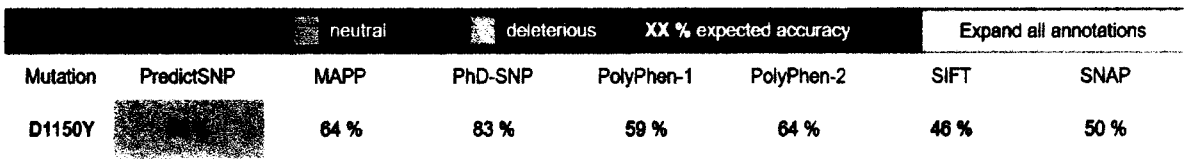


Figure 3.4: Secondary structure feature elements predicted for AFF2 protein **A)** Comparison of wild and mutant types, **B)** Sequence representation for wild type, **C)** Sequence representation for mutant type. In part B and C amino acids with purple, yellow and black color are representing α-helices, beta sheets and coils, respectively. Numbers on each side are positions of amino acids.

Figure 3.5: 3D Protein Models: A) Normal AFF2 protein structure shown in secondary type display style, **B)** Overlay of normal and mutant AFF2 structures highlighting the mutation, i.e. substitution of ASP (purple color) with TYR (red color) at position 1150. Proteins were shown in Rounded Ribbon Display Style with Green and cornflower blue color patterns representing wild-type and mutant species, respectively.

A



B

Position	WT	NEW	Stability	RI	pH	T
1150	D	Y	Decrease	1	7.0	25

C

Mutation	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP	RI
D1150Y	Disease 0.638	Neutral 0.150	Neutral 0.060	Neutral 0.430	Neutral 0.192	6

Figure 3.6: Prediction results for mutation effect analysis of AFF2 (D1150Y): **A)** Predict SNP results, **B)** I-Mutant results, **C)** Meta SNP results. RI is Reliability Index.

STITCH analysis

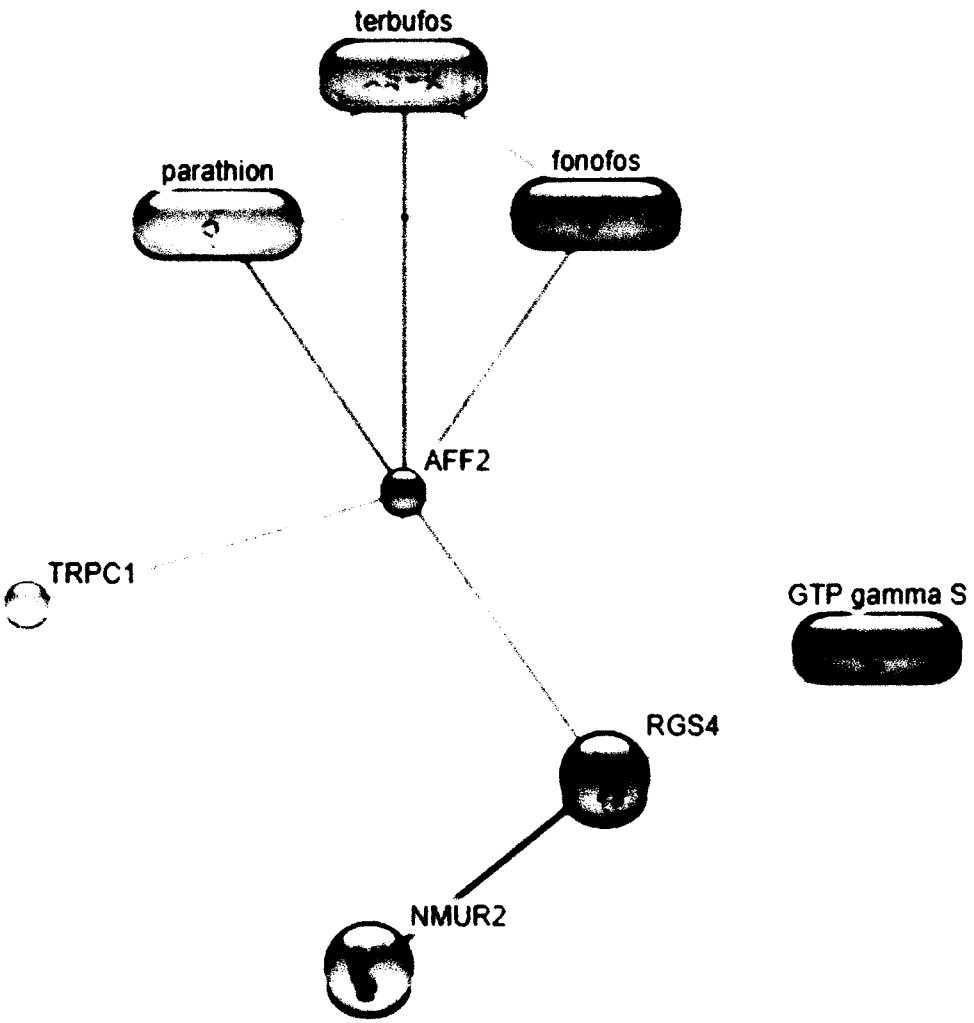


Figure 3.7: Stronger links are represented by thicker lines. Protein-protein interactions are shown in blue, chemical-protein interaction in green interaction between chemicals in red.

3.1.5. Discussion

In this research study, family MR200 from district Tank (KPK) Pakistan having individuals affected with intellectual disability was investigated. Sequence analysis revealed a novel c.3418G> T missense mutation in 17th exon of the AFF2 gene and co-segregated with diseased phenotype in the family. Polyphen2 and Mutation Taster shows missense mutation as damaging.

The AFF2 gene is involved in the production of a protein present in the cell nucleus, but its function is poorly understood. This protein has been reported to act as a transcription factor that binds to specific regions of DNA and helps regulate the activity of many other genes, but these genes are unknown.. (Bensaid et al., 2009).

Fragile syndrome is a condition that affects cognitive function and thinking ability. In most of the cases the individuals affected with Fragile XE syndrome have mild ID. People affected with Fragile XE syndrome have limited cognitive function and learning disabilities. The most common symptoms for people with Fragile XE syndrome are poor writing skills, slurred speech, hyperactivity and inattention. Autistic behaviors such as intense interest in a particular subject, repetitive behaviors and hand flapping are shown by some affected individuals. In Fragile XE syndrome, unlike some other features, cognitive function remained stable and did not decrease with age. In the current study, it was determined that mutations in AFF2, are responsible for mental retardation (Handt et al., 2014).

Computational analysis using different bioinformatics tools and software's helped in successfully exploring target proteins and facilitated mutational analysis in a cost and time effective way. Structure modeling not only predicted features in a protein sequence such as alpha helixes, beta sheets and coils but also revealed information about amino acids involved in making these structure elements. Three dimensional models revealed how conformation of amino acids in protein changes due to mutation. This change is involved in disruption of proteins own overall globular conformation and its interaction with other interacting proteins thereby leading to abnormal functionality. Point mutation was also observed in effecting protein structure in case of AFF2 although it was a single base change. Effect of mutation was also checked using various tools which showed protein as disease causing and a decrease in its stability. This picture of mutational effect through three dimensional structural and mutational study reports provides insight into the normal and mutant forms of protein, its inetractomics

and biological mechanism. This analysis also helps in understanding experimental findings collected from wet lab process. This also portrayed design of future experiments, by leading to the development of latest medicines with the help of drug designing and interaction studies. In this chapter, a novel Fragile X E syndrome arising due to AFF2 mutation is described.

Mutations in *RPTN* gene cause a new genetic disorder of Intellectual disability.

3.2. MR212

Family MR212 having individuals affected with Intellectual disability was recruited from (Rega) Buner District of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.8). The consanguineous couple III:8 and III:9 have four children, out of which three (IV:1, IV:2 and IV:3) shows Intellectual disability and the individual (IV:4) was normal. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.2.1. Clinical Features

Affected individuals of the family MR212 have intellectual disability. They have problem in social behavior. They have long term memory loss. They have learning and walking disability. They have very limited personal skills and have language disability. Information collected from family members shows that they have born normally.

3.2.2. Molecular Genetic Studies

Family MR212 has been screen for known genes causing ID (Table 2.8). Microsatellite marker for each gene was used for linkage studies and genotype for affected individuals and normal individuals of the family MR212. In this family, genotyping was performed for three affected and four normal individuals, including DNA samples (III:8, III:10, III:11, IV:1, IV:2, IV:3 and IV:4). In family linkage was not established with known genes. After family was excluded from linkage, the whole Exome sequencing was planned to identify the causative variants.

3.2.3. Whole Exome Sequencing and Variants Filtration

Affected individual (IV: 2) from family MR212 was subjected to Exome sequencing. Variants present in exome data were first checked for previously reported causing intellectual disability. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of

inheritance i.e. homozygous or compound heterozygous as shown in (Table 3.2). Only two new potentially harmful gene variants located in homozygous regions shared between four affected cases detected by complementary Affymetrix 6.0 single nucleotide microarray analyzes after variant screening [(chr:1:108,281,720-147,011,783 [hg19], and chr1:147,283,701-167,615,368 [hg19])], One was located in RPTN (Chr:1:g.152128983 (hg19 build); NM_001122965.1:c.592G>T; p.Glu198*), gene considered an unlikely candidate Intellectual disability. Second was located in SLC50A1 (Chr:1:g.155109390T>C (hg19 build); NM_001122839.1:c.245T>C; p.Leu82Pro), third one is present in TARS2 (Chr:1:g.150470022G>A (hg19 build); NM_025150.4:c.1037G>A; p.Arg346His), fourth one is present in CAPZA1 (Chr:1:g.113197143C>G (hg19 build); NM_006135.2:c.276C>G; p.Asn92Lys).

3.2.4. Segregation Analysis

Co-segregation of selected variants (Table 3.2) in all affected and normal members was checked. Gene variants of *SLC50A1*, *TARS2* and *CAPZA1* were not co-segregated in the family. Among the selected gene variants, the only variant [NM_001122965.1:c.592G>T; Chr:1:g.152128983C>A (GRCh37)] in *RPTN gene* was found to co-segregate and in consistence with phenotypes within the family. Three affected individuals (IV: 1, IV: 2 and IV3) were carrying the mutant allele in homozygous state where the other normal family members (III: 8, III:10, III:11 and IV:4) were heterozygous for the allele. According to the UCSC Human Genome (GRCh37/ hg19), this variant leads to stop gain at the evolutionary conserved position 198 (p.Glu198*). In the Genome Aggregation Database (gnomAD), this variant (p.Glu198*) was not listed.

3.2.5. *In silico* Analysis

Psipred (McGuffin et al., 2000) and I-Tasser (Roy et al., 2010; Zhang 2008) were used for the predication of secondary and three dimensional protein models. Predicted models reliability was checked using RAMPAGE (Lovell et al., 2002). Models were visualized via UCSF Chimera (Pettersen et al., 2004).

In order to understand the biological role of proteins deep study about structure and function has key role. Although computational strategy is not equivalent to experimental method of structure determination in providing high-resolution but still provide useful knowledge that

can be used to understand biological mechanism. Linear sequence of protein keeps in it all information about secondary structure features and under conformational chemistry adopts specific globular three dimensional (3D) shapes called tertiary structure. This structure analysis has been carried out in the present study using computational tools to study target proteins and their behavior. A secondary structure analysis show that after mutation structure features gets effected i.e. number of alpha (α)-helices, beta (β) strands/sheets and coils alter. Figure 3.10 shows representation of secondary structure features for RPTN, respectively. In Figures 3.10 Part A is graphical comparative analysis of wild and mutant type depicting difference of structure elements and Part B and C is the sequence representation of wild and mutant type proteins, respectively with amino acids participating in structure elements have been highlighted. Sequences of genes were subjected to protein BLAST at NCBI as a prerequisite of protein modeling to find out the related proteins in database. No suitable templates were available to perform homology modeling approach hence alternative approach i.e. threading to model genes was adopted. Quality of predicted models is estimated through a parameter i.e. C-score which is confidence score. Three dimensional structures modeled for wild and mutant proteins are shown in Figure 3.11. Maximum percentage of residues in the favored region in evaluation process proved reliability of predicted models. RPTN mutation has truncated the protein hence the difference in structure is obvious and prominent (Figure 3.11 Part B).

Proteins have specific forms due to their secondary and tertiary structure. A binding pocket or active region is an area where specific molecules, called ligands, bind which can be any chemical or small interacting protein. To perform a specific function within body as per interactomics, proteins interact with other partners and this interaction takes place at these active sites which in turn depicts chemical change / reaction. In proteins, active site is normally hydrophobic pocket that includes side chain atoms. Indicating molecules that bind to target protein and interaction study is helpful for 3D structure's role in docking as well as for drug designing. Number of pockets and amino acids involve in each pocket is specific to every protein structure hence also used for discrimination, interaction and mutational analysis. Count of interaction sites was examined for the target proteins and results collected depict changes in number of pockets. Normal structures of RPTN were observed with a total number of pockets as 126 respectively. A huge difference in pocket number has been observed in RPTN as length

of protein was effected after mutation i.e. 24 pockets. This change shows implication of mutation on overall conformation of protein which eventually alters interaction and function of proteins.

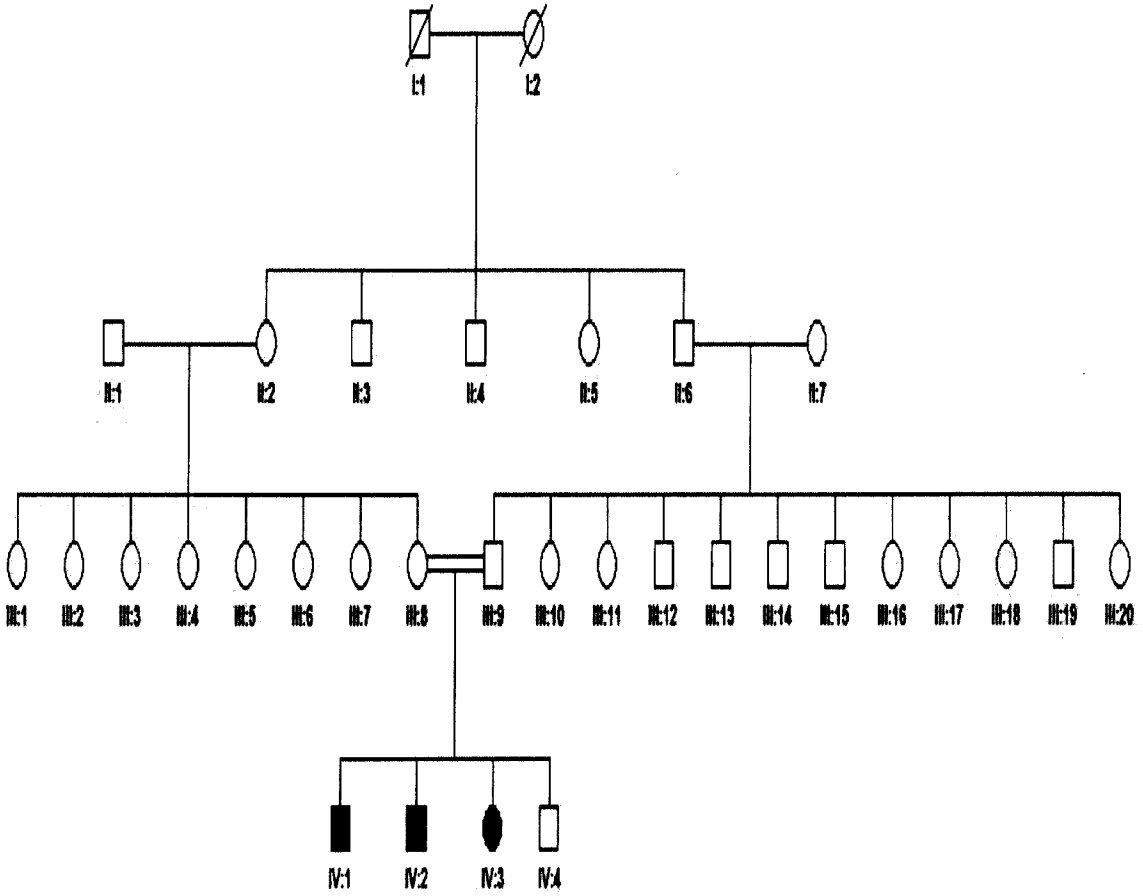


Figure 3.8: Pedigree of family MR212. Females are represented by circles and males are represented by squares. Affected individuals are represented by filled symbols while unaffected individuals are represented by clear symbols. Deceased members are represented by lines passes through symbols. A double line represents consanguineous marriages.

Table 3. 2: List of filtered variants identified in the Exome sequencing data of the proband (IV: 2) from family MR212.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr1:152128983	*RPTN*	c.592G>T	Hom	Stop gain	p.Glu198*
5109390	*SLC50A1*	c.245T>C	Hom	missense	p.Leu82Pro
Chr1:150470022	*TARS2*	c.1037G>A	Hom	missense	p.Arg346His
Chr1:113197143	*CAPZA1*	c.276C>G	Hom	missense	p.Asn92Lys

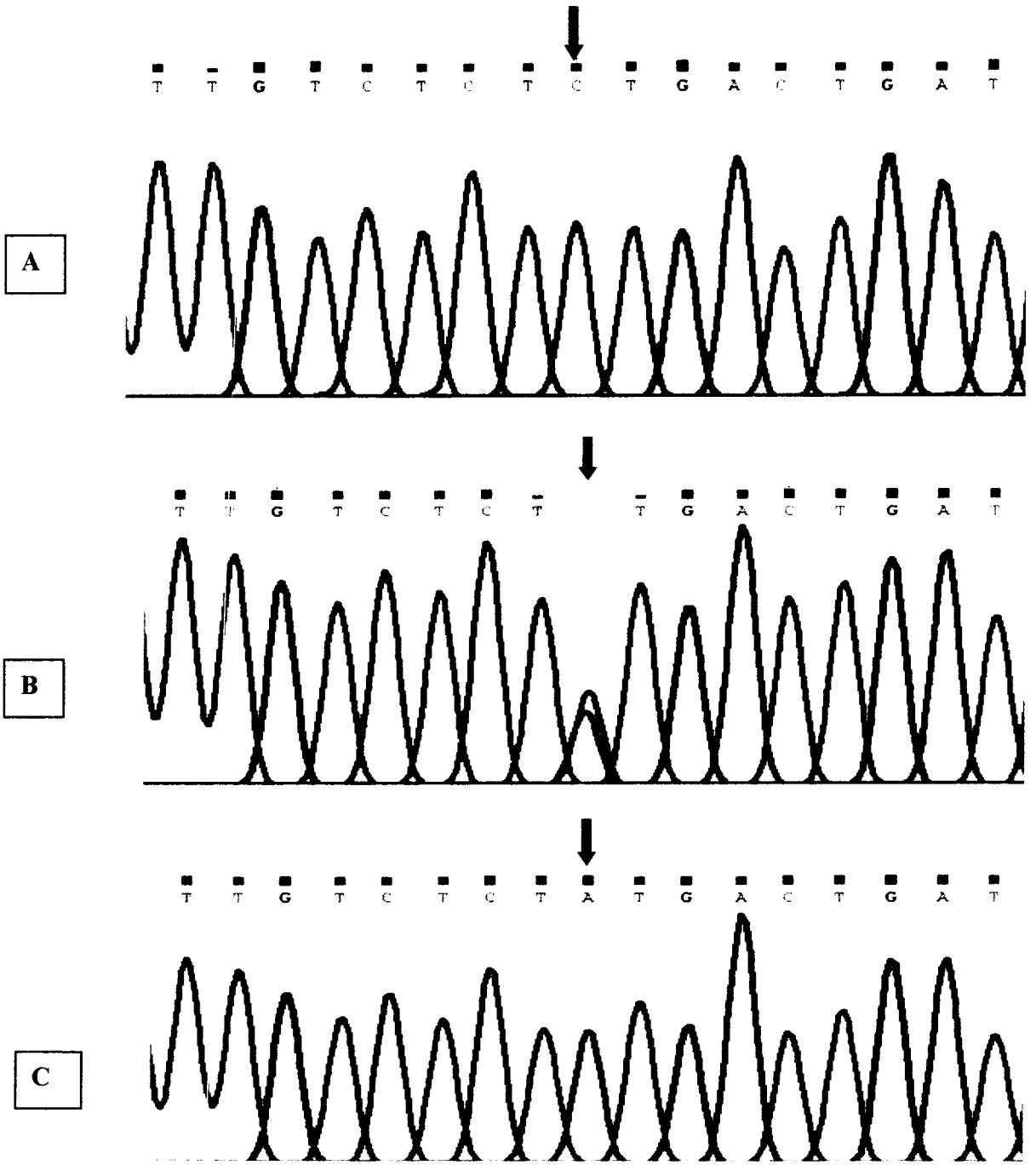
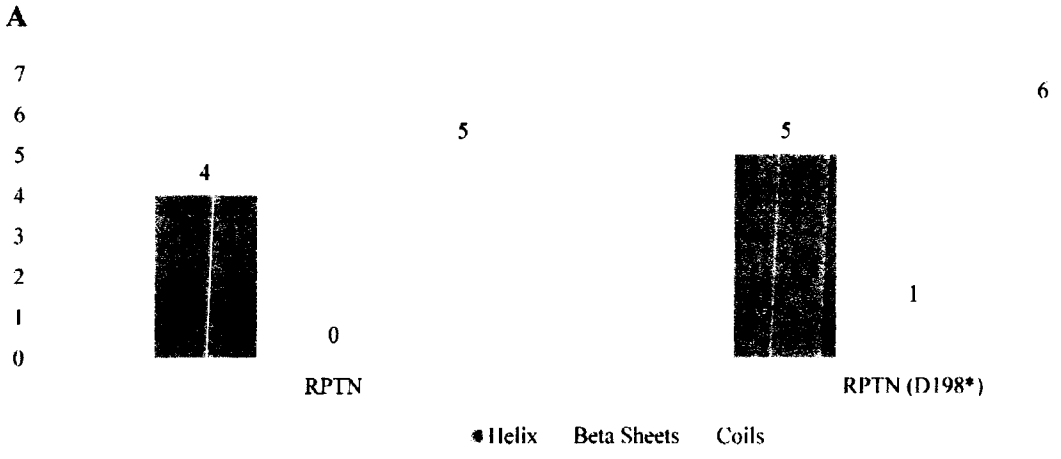


Figure 3.9: Chromatogram of variants c.592G>T, p.Glu198* in **(A)** Wild type, **(B)** Carrier and **(C)** Affected individual (IV: 2) in exon 3 of *RPTN* gene



B

Secondary Structure Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.

```
  1 M A Q L L N S I L S V I D V F H K Y A K G N G D C A L L C K E E L K Q L L L A E F G D I L Q R P N D  50
 51 P E T V E T I L N L L D Q D R D G H I D F H E Y L L L V F Q L V Q A C Y H K L D N K S H G G R T S Q 100
101 Q E R G Q E G A Q D C K F P G N T G R Q H R Q R H E E E R Q N S H H S Q P E R Q D G D S H H G Q P E 150
151 R Q D R D S H H G Q S E K Q D R D S H H S Q P E R Q D R D S H H N Q S E R Q D K D F S F D Q S E R Q 200
201 S Q D S S S G K K V S H K S T S G Q A K N Q G H I F A L N R C E K P I Q D S H Y G Q S E R H T Q Q S 250
251 E T L G Q A S H F N Q T N Q Q K S G S Y C G Q S E R L G Q E L G C G Q T D R Q G Q S S H Y G Q T D R 300
301 Q D Q S Y H Y G Q T D R Q G Q S S H Y S Q T D R Q G Q S S H Y S Q P D R Q G Q S S H Y G Q M D R K G 350
351 Q C Y H Y D Q T N R Q G Q G S H Y S Q P N R Q G Q S S H Y G Q P D T Q D Q S S H Y G Q T D R Q D Q S 400
401 S H Y G Q T E R Q G Q S S H Y S Q M D R Q G Q G S H Y G Q T D R Q G Q S S H Y G Q P D R Q G Q N S H 450
451 Y G Q T D R Q G Q S S H Y G Q T D R Q G Q S S H Y S Q P D K Q G Q S S H Y G K I D R Q D Q S Y H Y G 500
501 Q P D G Q G Q S S H Y G Q T D R Q G Q S F H Y G Q P D R Q G Q S S H Y S Q N D R Q G Q S S H Y G Q T 550
551 D R Q G Q S S H Y G Q T D R Q G Q S Y H Y G Q T D R Q G Q S S H Y I Q S Q T G E I Q G Q N K Y F Q G 600
601 T E G T R K A S Y V E Q S G R S G R L S Q Q T P G Q E G Y Q N Q G Q G F Q S R D S Q Q N G H Q V W E 650
651 P E E D S Q H H Q H K L L A Q I Q Q E R P L C H K G R D W Q S C S S E Q G H R Q A Q T R Q S H G E G 700
701 L S H W A E E E Q G H Q T N D R H S H E S Q E G P C G T Q D R R T H K D E Q N H Q R R D R Q T H E H 750
751 E Q S H Q R R D R Q T H E D K Q N R Q R R D R Q T H E D E Q N H Q R
```

C

Secondary Structure Map

Feature precictions are colour coded onto the sequence according to the sequence feature key shown below.

```
  1 M A Q L L N S I L S V I D V F H K Y A K G N G D C A L L C K E E L K Q L L L A E F G D I L Q R P N D  50
 51 P E T V E T I L N L L D Q D R D G H I D F H E Y L L L V F Q L V Q A C Y H K L D N K S H G G R T S Q 100
101 Q E R G Q E G A Q D C K F P G N T G R Q H R Q R H E E E R Q N S H H S Q P E R Q D G D S H H G Q P E 150
151 R Q D R D S H H G Q S E K Q D R D S H H S Q P E R Q D R D S H H N Q S E R Q D K D F S F D Q S
```

Figure 3.10: Secondary structure feature elements predicted for RPTN protein **A)** Comparison of wild and mutant types, **B)** Sequence representation for wild type, **C)** Sequence representation for mutant type. In part B and C amino acids with purple, yellow and black color are representing α-helices, beta sheets and coils, respectively. Numbers on each side are positions of amino acids.

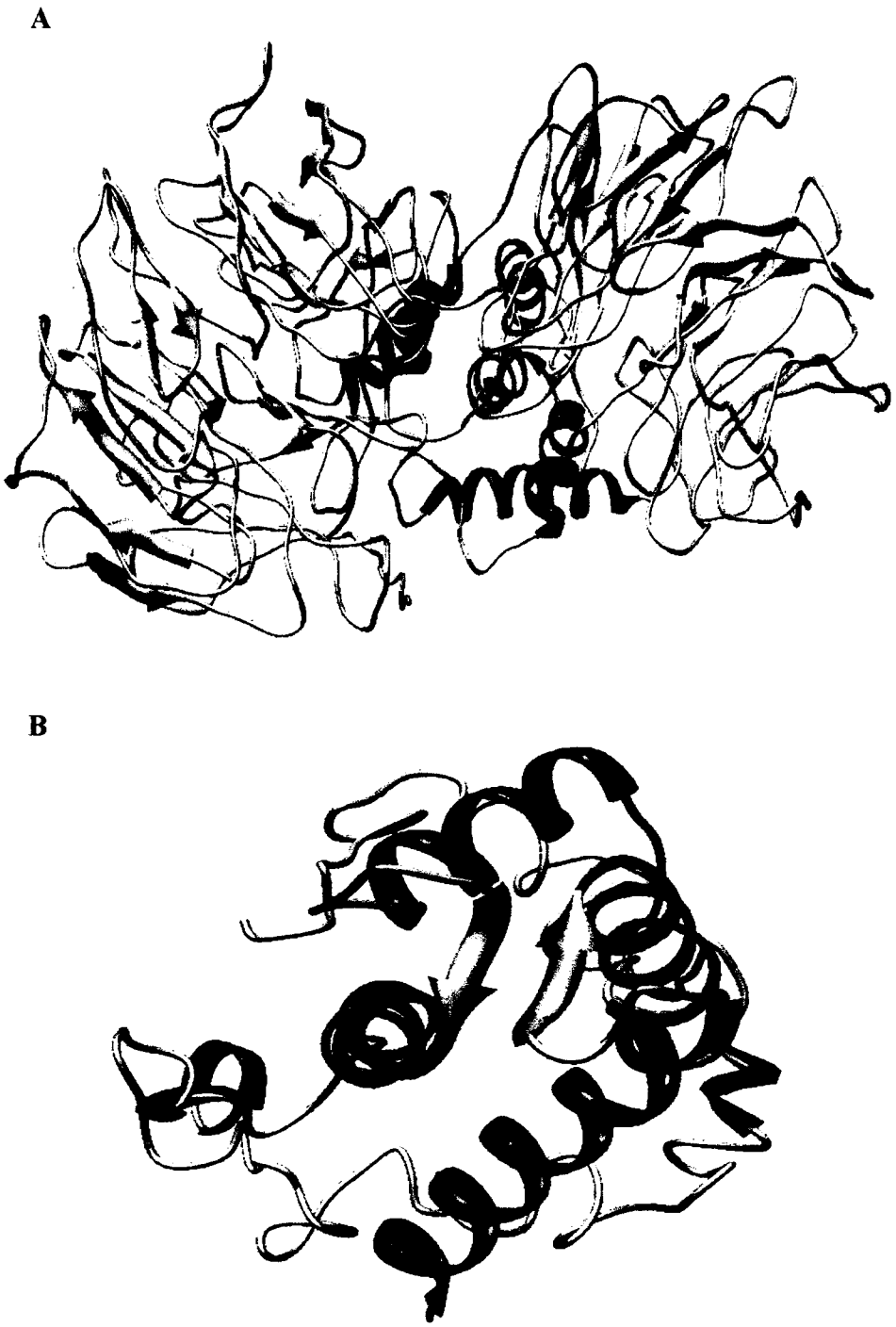


Figure 3.11: 3D Model for RPTN: A) Normal protein structure B) Mutant protein structure. Models have been shown in secondary type display style with red cylinders, yellow arrow heads and grey turns representing alpha helix, beta sheets and coils, respectively in the structure.

STITCH analysis

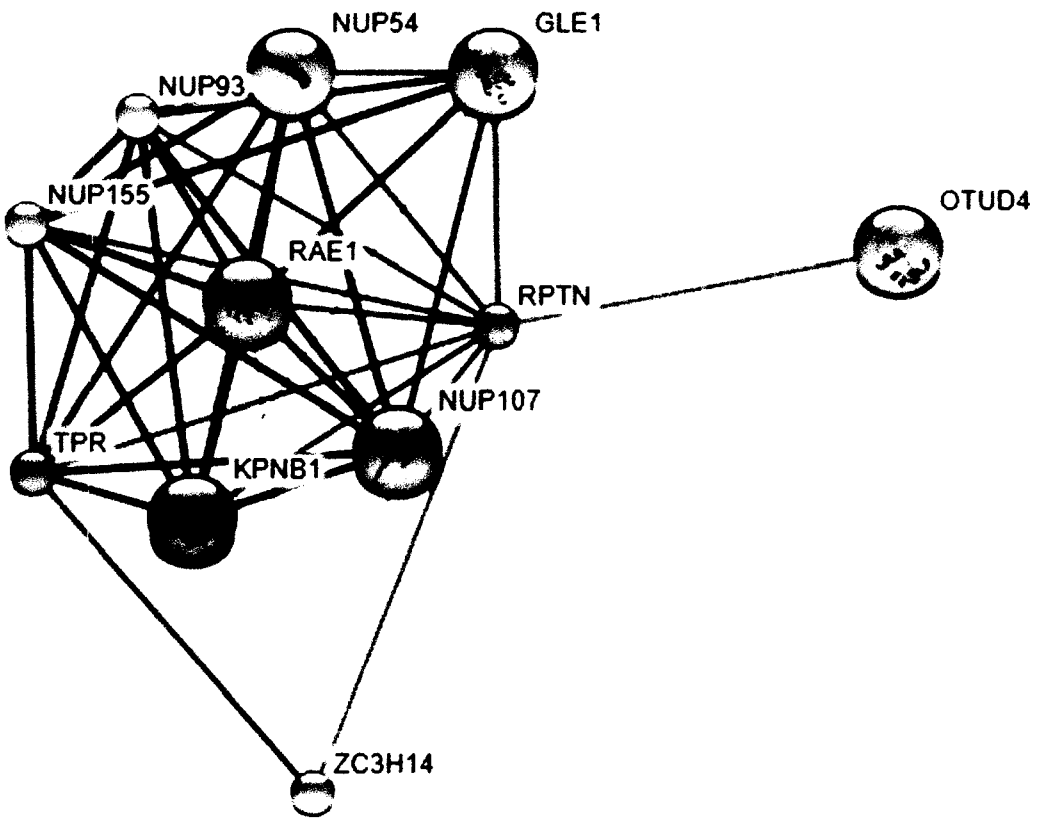


Figure 3.12: Stronger links are represented by thicker lines. Protein-protein interactions are shown in blue, chemical-protein interaction in green interaction between chemicals in red.

3.2.6. Discussion

In this study family with Intellectual disability was evaluated at molecular level. Classical linkage analysis was carried for exclusion study. Subsequently, complete exome sequencing identified a pathogenic RPTN (new gene) mutation with stop gain c.592G>T; p.Glu198*. Mutation is not found in 1000 genome.

Repetin (RPTN) protein is a part of S100 family and is been reported that repetin protein is expressed in the normal epidermis. It has been shown that repetin has high expression in both human and mouse brain. In order to investigate the expression of repetin in neuropsychiatric disorders, repetin in the serum levels of individuals affected with schizophrenia (n = 88) or bipolar disorder (n = 34) and in chronic psychostimulant users (n = 91) was determined. Individuals affected with bipolar disorders, schizophrenia or psychostimulant shows diminish RPTN level in serum in comparison study with healthy individuals (n = 115) or age-matched controls (n = 92) ($p < 0.0001$. The results show that RPTN plays an important role in emotional and mental abilities. Decreased serum RPTN levels suggest a possible association with the pathogenesis of schizophrenia and bipolar disorder. (Wang et al., 2015).

Computational analysis using different bioinformatics tools and software's helped in successfully exploring target proteins and facilitated mutational analysis in a cost and time effective way. Structure modeling not only predicted features in a protein sequence such as alpha helixes, beta sheets and coils but also revealed information about amino acids involved in making these structure elements. Three dimensional models revealed how conformation of amino acids in protein changes due to mutation.

The highest percentage of residues in the recipient region in the evaluation process confirms the reliability of the assumed models. RPTN mutation has truncated the protein hence the difference in structure is obvious and prominent. In proteins, active site is normally hydrophobic pocket that includes side chain atoms. Indicating molecules that bind to target protein and interaction study is helpful for 3D structure's role in docking as well as for drug designing. Number of pockets and amino acids involve in each pocket is specific to every protein structure hence also used for discrimination, interaction and mutational analysis. Count of interaction sites was examined for the target proteins and results collected depict changes in number of pockets. A normal structure of RPTN was observed with a total number of pockets 126, respectively. A huge difference in pocket number has been observed in RPTN as length

of protein was effected after mutation i.e. 24 pockets. This change shows implication of mutation on overall conformation of protein which eventually alters interaction and function of proteins. This picture of mutational effect through three dimensional structural and mutational study reports provides insight into the normal and mutant forms of protein, its inetractomics and biological mechanism. This analysis also helps in understanding experimental findings collected from wet lab process. This also portrayed design of future experiments, by leading to the development of latest medicines with the help of drug designing and interaction studies. Further expressional study is required to whether mutation in this gene is responsible for the diseased phenotype.

Further research of patients and animal models of these conditions will provide a better understanding of the recurring role of repetin in neural structure and functioning, and thus provide insight into future treatment. Development of effective treatments for patients suffering from these debilitating conditions.

Mutations in *ROR2* gene underlie a new genetic disorder of Intellectual disability.

3.3. MR295

Family MR295 having individuals affected with Intellectual disability was recruited from (Malakand) province KPK of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.13). The normal couple II: 5 and II: 6 have two children (III: 3 and III: 4) affected with Intellectual disability. We have total seven individual blood samples (II: 1, II: 2, II: 6, III: 1, III: 2, III: 3 and III: 4) show willingness to participate in this research study. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.3.1. Clinical Features

Affected individuals of the family MR295 have intellectual disability. They have problem in Emotional response. They have memory loss. They have learning and difficult in walking. They have very limited personal skills, need care and have language disability.

3.3.2. Molecular Genetic Studies

Family MR295 has been screen for known genes causing ID. Microsatellite marker for each gene was used for linkage studies and genotype for affected individuals and normal individuals of the family MR295. In this family, seven DNA samples (II:1, II:2, II:6, III:1, III:2, III:3 and III:4), including three affected and four normal individuals were genotyped. In family linkage was not established with known genes. After family was excluded from linkage, the whole Exome sequencing was planned to identify the causative variants.

3.3.3. Whole Exome Sequencing and Variants Filtration

Affected individual (III: 3) from family MR295 was subjected to Exome sequencing. Variants present in exome data were first checked for previously reported causing intellectual disability. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of

inheritance i.e. homozygous or compound heterozygous as shown in (Table 3.3). After filtering for variants, there were only two variants in the gene located in the homozygous region shared between the two affected cases, which were detected by single-nucleotide microarray scans. [(chr:9:88,279,039-96,906,751 [hg19], and chr:9:118,665,465-135,139,901 [hg19])], one of which co-segregated with the disease phenotype as determined by dideoxy sequence analysis. One was located in ROR2 (Chr:9:94488948 (hg19 build); NM_004560.3:c.1261C>G; p.Leu421Val), gene considered an unlikely candidate Intellectual disability. Second was located in ZBTB6 (Chr:9:125673127 (hg19 build); NM_006626.5:c.1225A>G; p.Ser409Gly. Table 3.3 shows the amino acid change as a result of these two variants.

3.3.4. Segregation Analysis

Co-segregation of selected variants (Table 3.3) in all affected and normal members were checked. Gene variant of ZBTB6 was not co-segregated in the family. Among the selected gene variants, the variant [NM_004560.3:c.1261C>G; Chr9:g.94488948G>C (GRCh37)] in ROR2 gene was found to co-segregate with diseased phenotype within the family MR295 shown in (figure). Two affected individuals (III: 3 and III:4) were carrying the mutant allele in homozygous state where the other normal family members (II:1, II:2, III:1 and III:2) were heterozygous for the allele. According to the human UCSC genome (GRCh37/hg19), this variant leads to a substitution at the evolutionarily conserved position 421 (p.Leu421Val).

3.3.5. *In silico* Analysis

Psipred (McGuffin et al., 2000) and I-Tasser (Roy et al., 2010; Zhang 2008) were used for the prediction of secondary and three dimensional protein models. In order to understand the biological role of proteins deep study about structure and function has key role. Although, computational strategy is not equivalent to experimental method of structure determination in providing high-resolution but still provide useful knowledge that can be used to understand biological mechanism. Linear sequence of protein keeps in it all information about secondary structure features and under conformational chemistry adopts specific globular three dimensional (3D) shapes called tertiary structure. This structure analysis has been carried out in the present study using computational tools to study target proteins and their behavior. Secondary structure analysis show that after mutation structure features gets effected i.e.

number of alpha (α)-helices, beta (β) strands/sheets and coils alter. Figure 3.16 shows representation of secondary structure features for ROR2 respectively. In each of these figures Part A is graphical comparative analysis of wild and mutant type depicting difference of structure elements and Part B and C is the sequence representation of wild and mutant type proteins, respectively with amino acids participating in structure elements have been highlighted. In case of ROR2, amino acids involved in point mutations i.e. Leucine (L)/Valine (V), respectively was involved in making coil structure. In each protein, it has been observed that apart from difference in number of alpha helices, beta sheets and coils, there is difference in amino acids too which are involved in making these structure elements in wild and mutant type.

Sequences of genes were subjected to protein BLAST at NCBI as a prerequisite of protein modeling to find out the related proteins in database. No suitable templates were available to perform homology modeling approach hence alternative approach i.e. threading to model genes was adopted. Quality of predicted models is estimated through a parameter i.e. C-score which is confidence score. Three dimensional structures modeled for wild and mutant proteins are shown in Figure 3.17. There is no change in length of protein in case of missense mutations but the single residue difference (highlighted in part B of Figure 3.17) also has huge impact and involved in altering conformation and function of protein. Point a mutation of ROR2 was also analyzed employing various tools to further check effect and stability of proteins. Protein stability decreased after substitution mutation according to calculations of I-Mutant2.0 which is Protein Stability Changes calculator. Meta SNP employs various predictors such as PANTHER, PhD-SNP, SIFT and SNAP to estimate implications of mutation on normal protein. Predict SNP also stated mutations as disease causing. Results of all predictors Predict SNP, I Mutant and Meta SNP are shown in Figure 3.18 Part A, B and C, respectively.

Proteins have specific confirmation due to their secondary and tertiary structure. The binding pocket or active site is the region to which a special molecule called a ligand attaches and can be any chemical or small protein that interacts with it. To perform a specific function within body as per interactomics, proteins interact with other partners and this interaction takes place at these active sites which in turn depicts chemical change / reaction. In proteins, the active site is usually a hydrophobic sheath containing side chain atoms. Indicating molecules that bind to

target proteins and studying their interactions contribute to the role of 3D structures in docking and drug design. Number of pockets and amino acids involve in each pocket is specific to every protein structure hence also used for discrimination, interaction and mutational analysis. Count of interaction sites was examined for the target proteins and results collected depict changes in number of pockets. Normal structure of ROR2 was observed with a total number of pockets as 107, respectively. The number of active sites remained same in case of point mutations ROR2 but amino acids contributing these active sites got changed. This change shows implication of mutation on overall conformation of protein which eventually alters interaction and function of proteins.

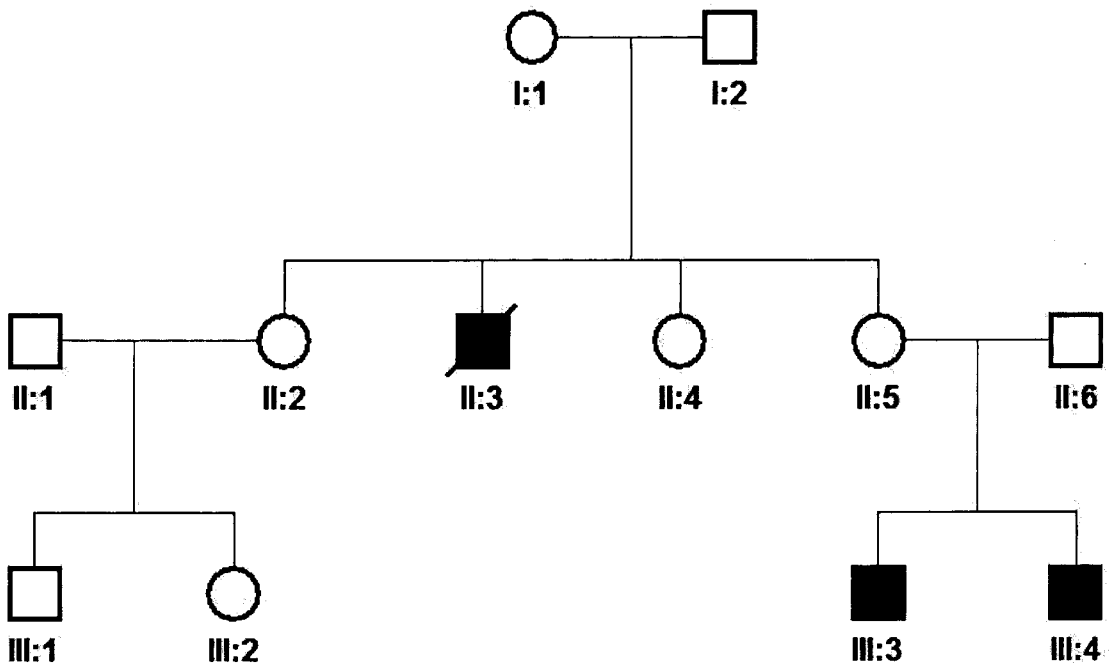


Figure 3.13. Pedigree of family MR295. Females are represented by circles and males are represented by squares. Affected individuals are represented by filled symbols while unaffected individuals are represented by clear symbols. Deceased members are represented by lines passes through symbols.

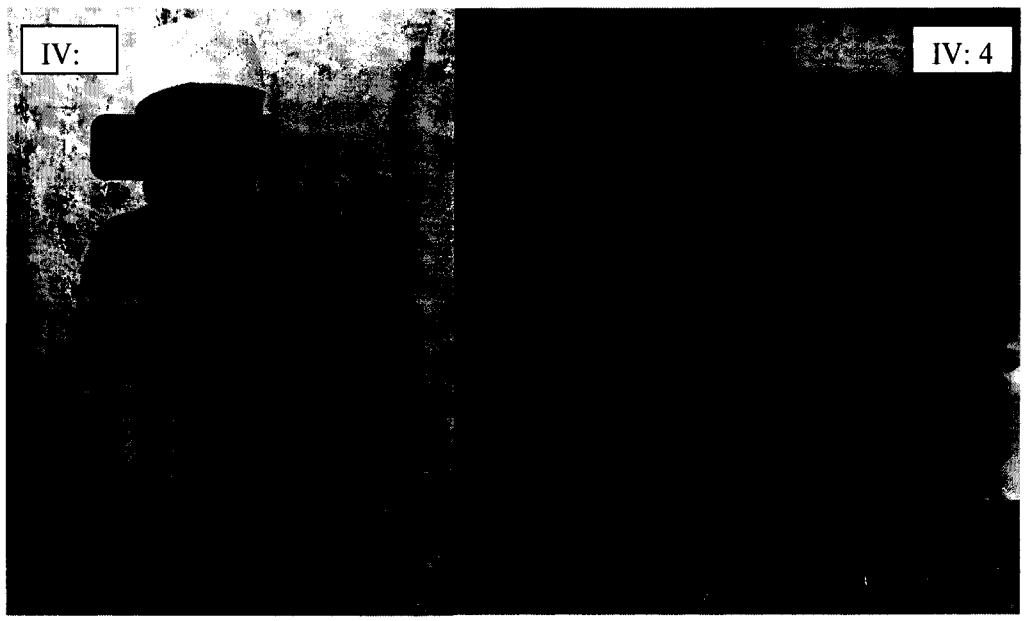


Figure 3.14. Phenotypic appearance of affected individual (III: 3 and III: 4) from family MR295.

Table 3. 3: List of filtered variants identified in the Exome sequencing data of the proband (III: 3) from family MR295.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr9:94488948	*ROR2*	c.1261C>G	Hom	missense	p.Leu421Val
Chr9:125673127	*ZBTB6*	c.1225A>G	Hom	missense	p.Ser409Gly

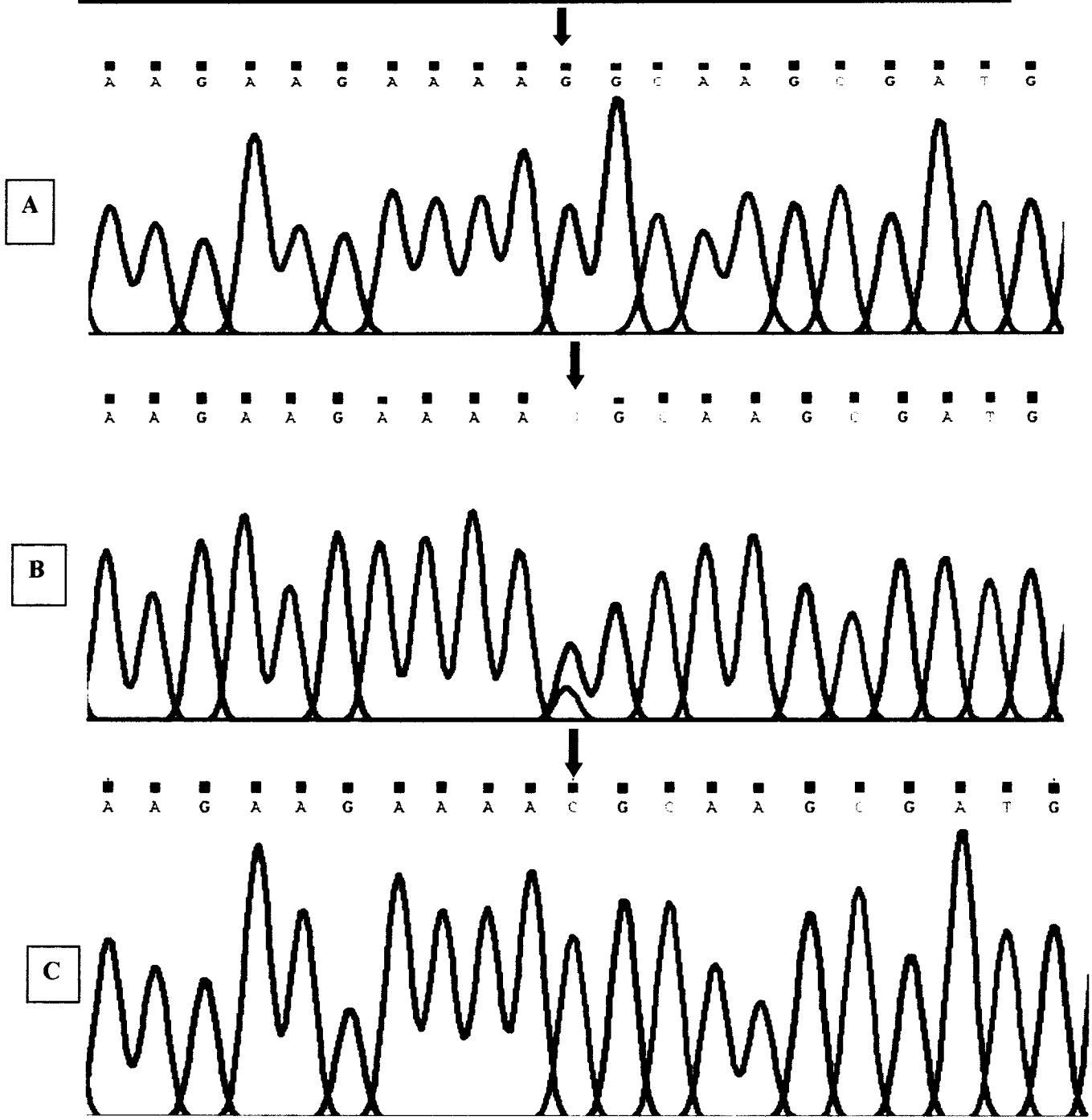
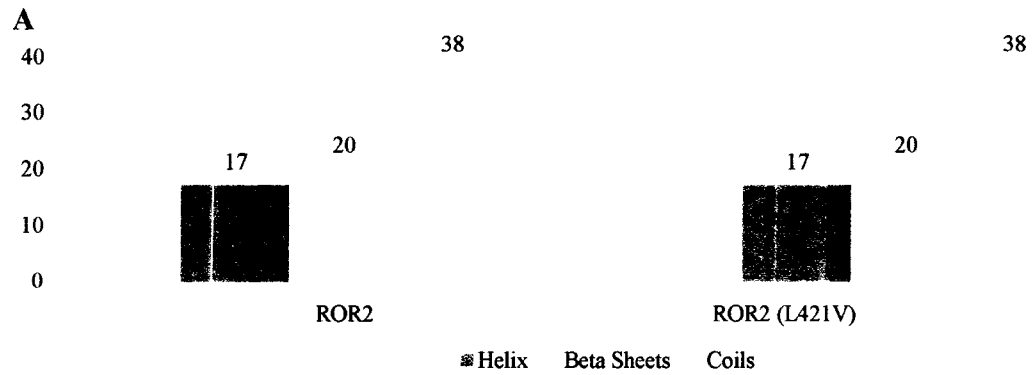


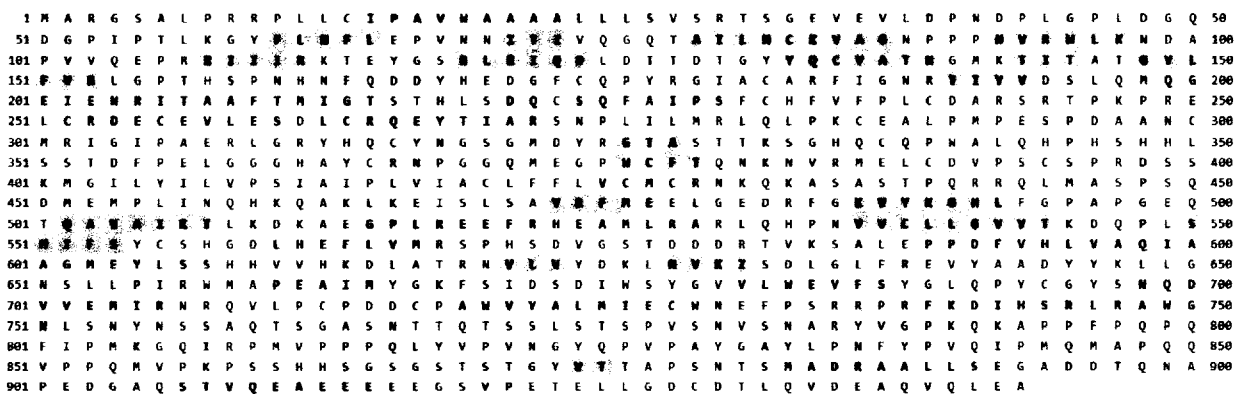
Figure 3.15. Chromatogram of variants c.1261C>G, p.Leu421Val in **(A)** Wild type, **(B)** Carrier and **(C)** Affected individual (IV: 3) in exon 8 of *ROR2* gene.



B

Secondary Structure Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.



C

Secondary Structure Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.

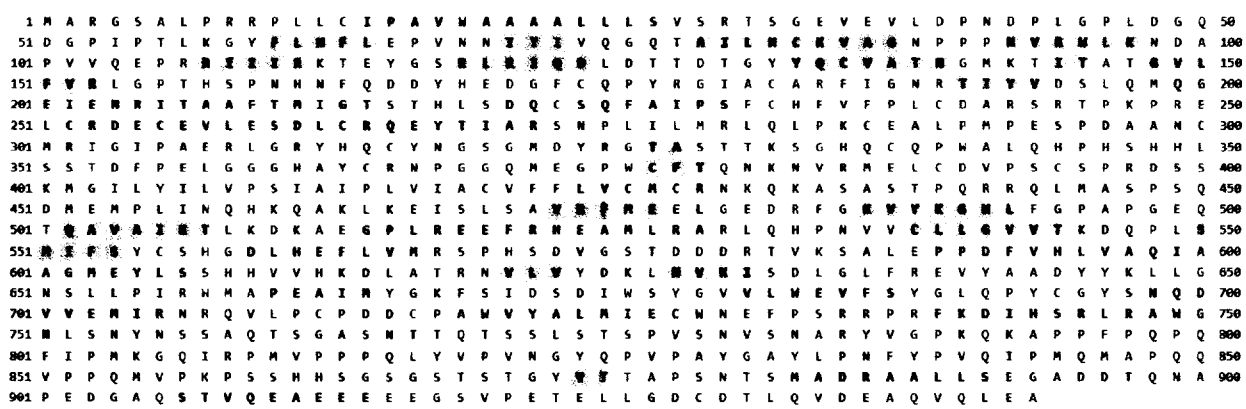


Figure 3.16: Secondary structure feature elements predicted for ROR2 protein **A)** Comparison of wild and mutant types, **B)** Sequence representation for wild type, **C)** Sequence representation for mutant type. In part B and C amino acids with purple, yellow and black color are representing α-helices, beta sheets and coils, respectively. Numbers on each side are positions of amino acids.

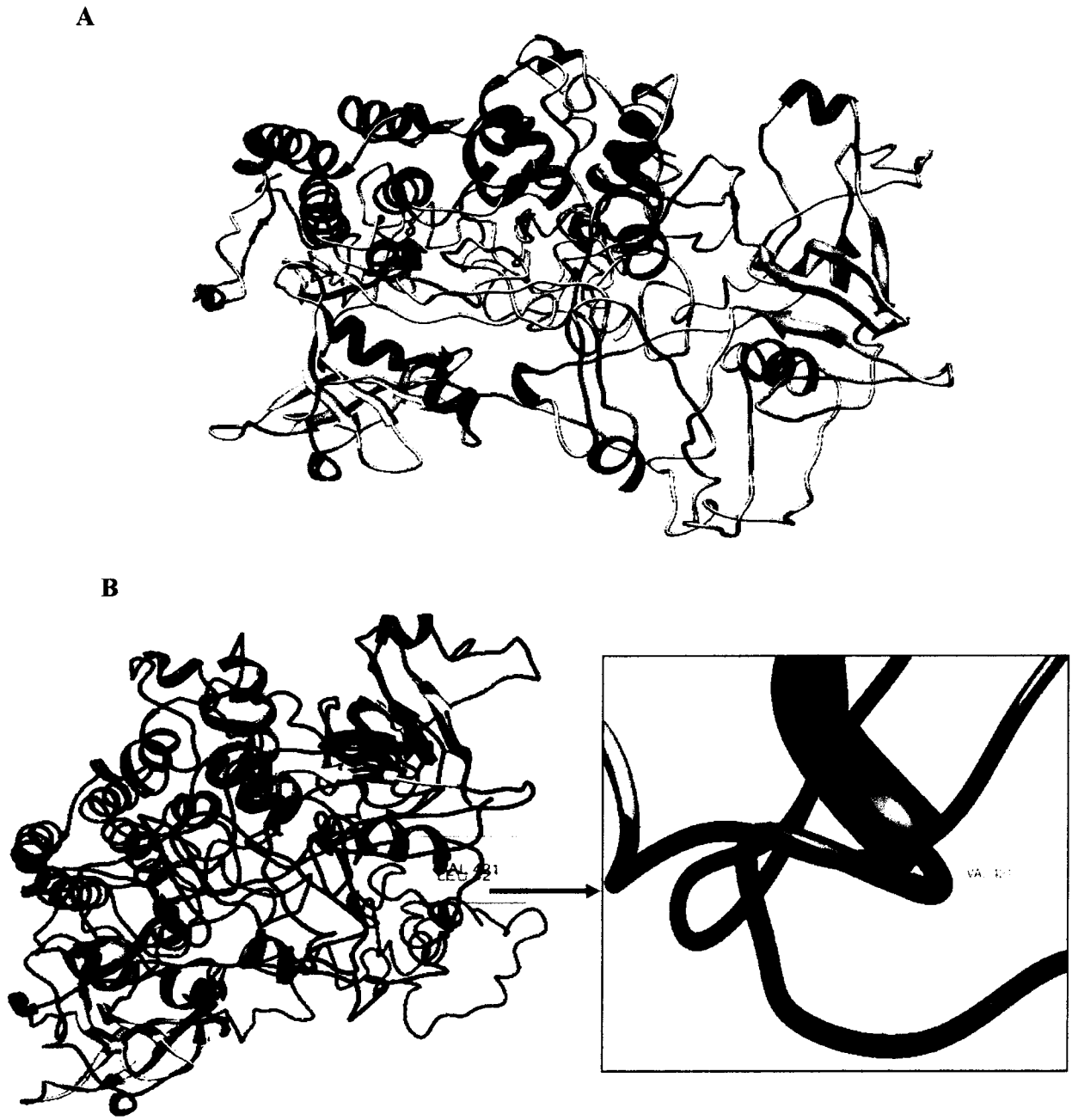
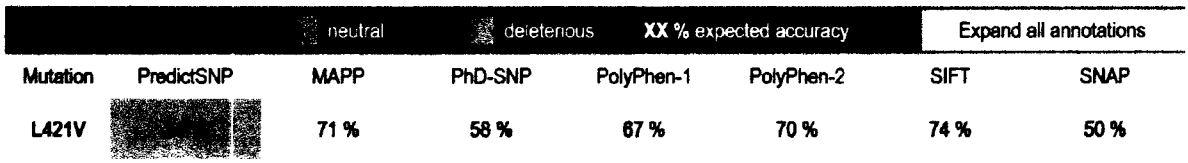


Figure 3.17: 3D Protein Models: **A)** Normal ROR2 protein structure shown in secondary type display style, **B)** Superimposed normal and mutant ROR2 structures highlighting mutation i.e. substitution of LEU with VAL at position 421. Proteins have been shown in Rounded Ribbon Display Style with Green and cornflower blue color representing wild and mutant types respectively. Normal protein residue (LEU) has been shown in magenta color and mutant protein residue (VAL) has been shown in red color.

A



B

C

Mutation	PANTHER	PhD-SNP	SIFT	SNAP	Meta-SNP	RI
L421V	Neutral 0.226	Neutral 0.231	Disease 0.010	Disease 0.625	Neutral 0.383	2

Figure 3.18: Prediction results for mutation effect analysis of ROR2 (L421V). **A)** PredictSNP results, **B)** I-Mutant results, **C)** Meta SNP results. RI is Reliability Index.

STITCH Analysis

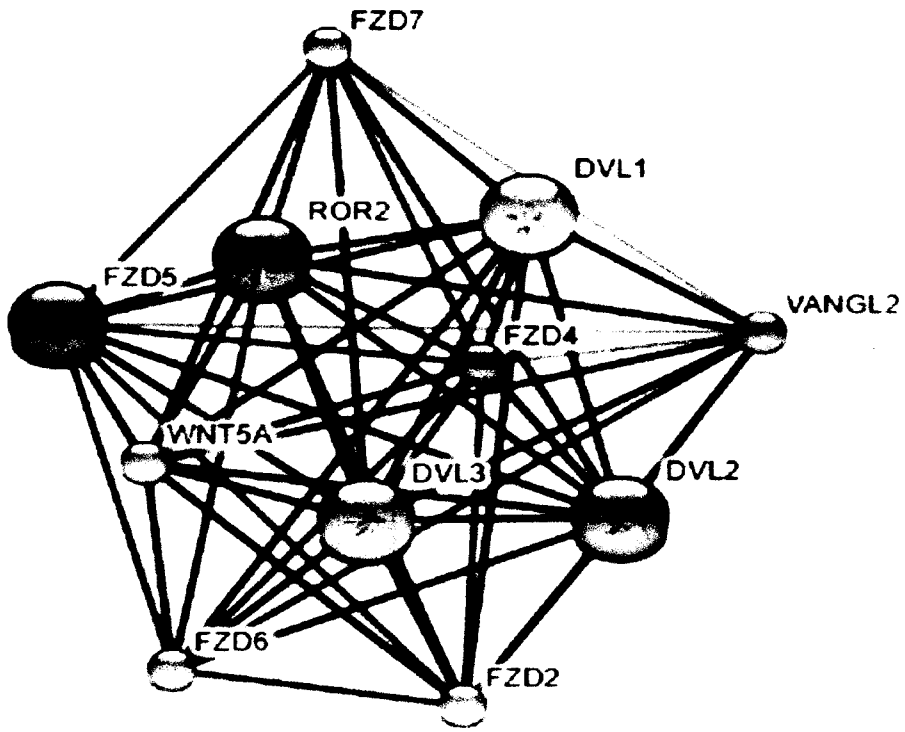


Figure 3.19: Protein-protein interactions are shown in grey, chemical-protein interactions in green and interactions between chemicals in red.

3.3.6. Discussion

In the current research study, family affected with intellectual disability from Malakand (KPK) district of Pakistan was investigated and a missense mutation c.1261C>G in the exon 8 of ROR2 was co-segregated in the affected individuals (III: 3 and III: 4) and in consistence with phenotype.

ROR2 is 943 amino acids protein. The ROR family of tyrosine kinases make up a small subfamily of receptor tyrosine kinases (RTKs) and are characterized by a unique conserved domain structure. STITCH analysis shows ROR2 protein interaction with other protein (Figure 3.16), which are greatly expressed in human brain. Protein like VANGL2, which plays an important role in brain development. In silico analysis was carried out to find the affects of missense mutation on ROR2 protein. Psipred (McGuffin et al., 2000) and I-Tasser (Roy et al., 2010; Zhang 2008) were used for prediction of secondary and three dimensional protein models. A secondary structure analysis show that after mutation structure features gets effected i.e. number of alpha (α)-helices, beta (β) strands/sheets and coils alter. Figure 3.17 shows representation of secondary structure features for ROR2 respectively.

In ROR2, amino acids involved in point mutations i.e. Leucine (L)/Valine (V) were involved in making coil structure. In this protein, it has been observed that apart from difference in number of alpha helices, beta sheets and coils, there is difference in amino acids too which are involved in making these structure elements in wild and mutant types. There is no change in length of protein in case of missense mutations but the single residue difference (highlighted in part B of Figure 3.18) also has huge impact and involved in altering conformation and function of protein. Point mutations of ROR2 were also analyzed employing various tools to further check effect and stability of proteins. To perform a specific function within body as per interactomics, proteins interact with other partners and this interaction takes place at these active sites, which in turn describes the chemical change / reaction. In protein, the active region is usually a hydrophobic pocket containing atoms in the side chain Indicating molecules that bind to target protein and interaction study is helpful for 3D structure's role in docking as well as for drug designing. Number of pockets and amino acids involve in each pocket is specific to every protein structure hence also used for discrimination, interaction and mutational analysis. Count of interaction sites was examined for the target proteins and result collected depicts changes in number of pockets. A normal structure of ROR2 was observed

with a total number of pockets as 107, respectively. The number of active sites remained same in case of point mutations in ROR2 but amino acids contributing these active sites got changed. This change shows implication of mutation on overall conformation of protein which eventually alters interaction and function of proteins thereby leading to abnormal functionality. Point mutation was also observed in effecting protein structure in case of ROR2 although it was a single base change.

This analysis also helps in understanding experimental findings collected from wet lab process. This also portrayed design of future experiments, by leading to the development of latest medicines with the help of drug designing and interaction studies. In this chapter, ID arising due to ROR2 mutation is described, it remains an interesting possibility that more pathogenic mutations in the ROR2 gene may be detected in more severely affected patients showing more significant ID phenotypic overlap.

3.4. MR205

Family MR205 having individuals affected with Intellectual disability was recruited from district Tank (KPK) Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.20). We have total four individual blood samples (V:6, V:7, V:8 and V:9) out of which two individuals (V:6 and V:8) shows intellectual disability while the two members (V:7 and V:9) are normal. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.4.1. Clinical Features

Affected individuals in family MR205 exhibited clinical manifestations of intellectual disability including problem with Emotional response. They have long term memory loss. They have learning disability. They have very limited personal skills, need care and have language disability. They has problem with Repetitive behaviors, Nonverbal communication and Intelligence.

3.4.2. Whole Exome Sequencing and Variants Filtration

Affected individual (V: 6) from family MR205 was subjected to Exome sequencing. Variants present in exome data were first checked for previously reported causing intellectual disability. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of inheritance i.e. homozygous or compound heterozygous as shown in (Table 3.4). Following filtering of variants five variants were found to be present in genes. One was located in TBC1D3 (Chr:17:.36287705G>A (hg19 build); NM_001123391.3:c.234G>A; p.Trp78*), gene considered an unlikely candidate Intellectual disability. Second was located in HRNR (Chr:1:152190060 (hg19 build); NM_001009931.2:c.4045T>C; p.Tyr1349His. third one is located in DRP2 (Chr:X:100511153 (hg19 build); NM_001939.2:c.2293C>T; p.Arg765Trp), compound heterozygous given in (Table 3.5) fourth one is located in MSRA (Chr:8:10285751del (hg19 build); NM_001135670.2:c.517delGA>G; p.Ser173Alafs*36); Chr:8:10285753_10285754insG (hg19 build); NM_001135670.2:c.519_520insG; p.Lys174Glufs*21), fifth one is located in SSC5D (Chr19:56000753_56000754ins81(hg19

build);
NM_001144950.1:c.85_86ins81(TCTCGATGATCTGCTGGAGTATCCCTACGGCGGCGT CGAGCAGACCGCCAGCCAACTGCTGCCGCTGAGCATCGCCTATCCG); p.Gly28_Cys29insSerArg*); (Chr:19:56000745_56000746insGGCCGCTGCCTA (hg19 build); NM_001144950.1:c.77_78insGGCCGCTGCCTA; p.Pro26_His27insAlaAlaAlaTyr); (Chr:19:56000748del (hg19 build); NM_001144950.1:c.80delCA>C; p.His27Leufs*121). Sixth one is located in SUN3 (Chr:7:48058122G>T (hg19 build); NM_001284350.1:c.122C>A; p.Ser41*); (Chr:7:48058129G>T (hg19 build); NM_001284350.1:c.115C>A; p.Pro39Thr). Seven one is present in ANKIB1 (Chr:7:91936742G>T (hg19 build); NM_019004.1:c.258G>T; p.Leu86Phe); (Chr:7:91936750G>T (hg19 build); NM_019004.1:c.266G>T; p.Gly89Val). Eight one is present in TRPM6 (Chr:9:77357522T>C (hg19 build); NM_017662.4:c.5155A>G; p.Met1719Val); (Chr:9:77357542G>A (hg19 build); NM_001177310.1:c.5120C>T; p.Ala1707Val). Nine one is present in NPIPB13 (Chr:16:30236689G>A (hg19 build); NM_001321892.1:c.1078C>T; p.Pro360Ser); (Chr:16:30236722G>A (hg19 build); NM_001321892.1:c.1045C>T; p.Pro349Ser).

3.4.3. Segregation Analysis

Co-segregation of selected homozygous variants (Table 3.4) in all affected and normal members was checked. No gene variant was co-segregated in affected individuals. Then co-segregation of selected compound heterozygous variants (Table 3.5) was checked in all affected individuals of the family MR205. No gene variants were co-segregated in affected individuals in family MR205. Further NGS is required to find the causes of disease in this family on the availability of funds in future.

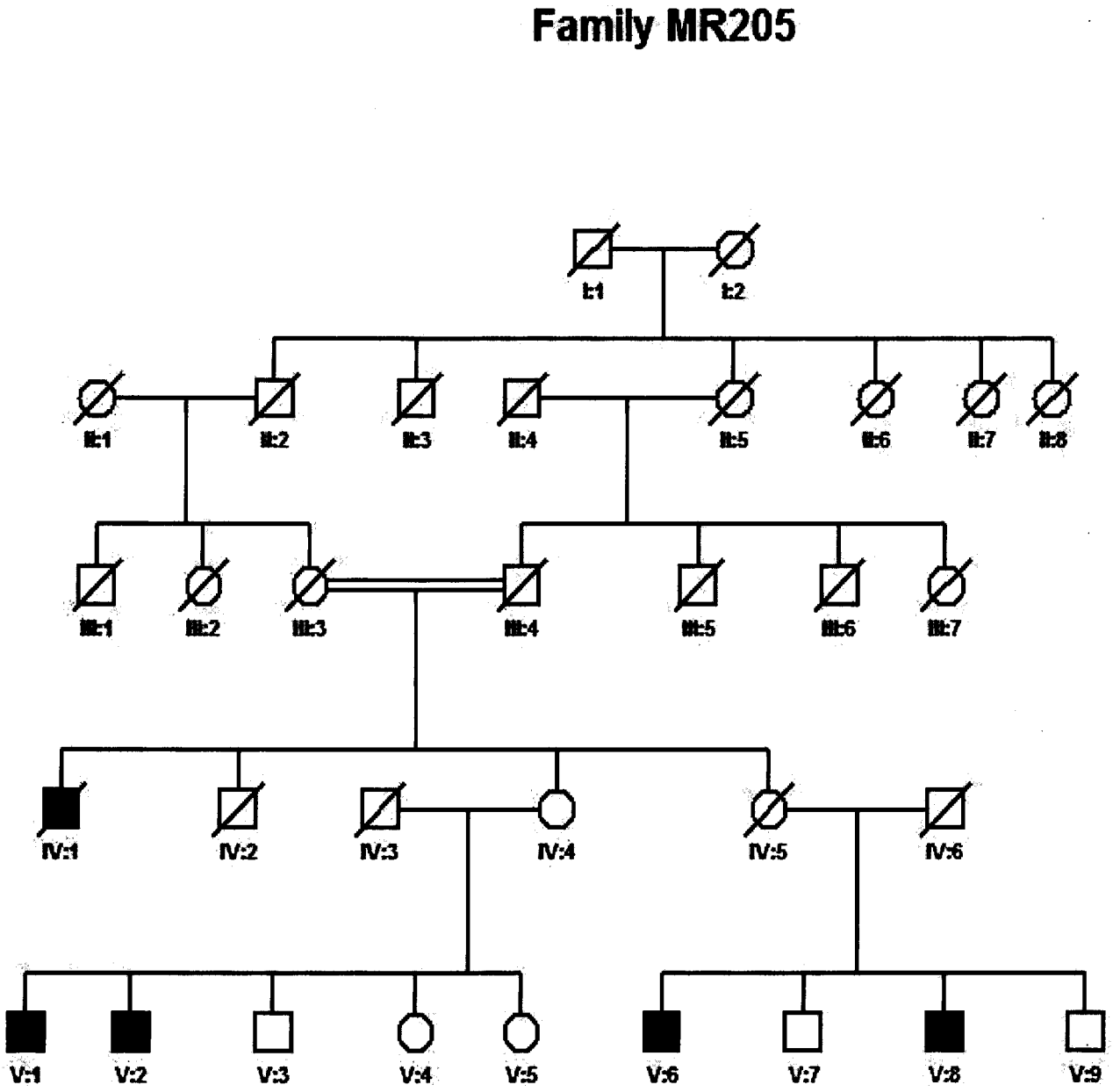


Figure 3.20: Pedigree of family MR205. Females are represented by circles and males are represented by squares. Affected individuals are represented by filled symbols while unaffected individuals are represented by clear symbols. Deceased members are represented by lines passes through symbols. A double line represents consanguineous marriages.

Figure 3.21: Phenotypic appearance of affected individual (V: 6 and V: 8) from family MR205.

Table 3. 4: List of filtered homozygous recessive variants identified in the Exome sequencing data of the proband (V: 6) from family MR205.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr17:36287705	*TBC1D3*	c.234C>A	Hom	Stop gain	p.Trp78*
Chr1:152190060	*HRNR*	c.4045T>C	Hom	missense	p.Tyr1349His
ChrX:100511153	*DRP2*	c.2293C>T	Hom	missense	p.Arg765Trp

Table 3.5: List of filtered compound heterozygous variants identified in the Exome sequencing data of the proband (V: 6) from family MR205.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr7:48058122	*YUY3*	c.12C>A	Het	Stop gain	p.Ser41*
Chr7:48058129		c.5G>A		missense	p.Pro39Thr
Chr7:91936742	*ANKIB1*	c.258G>T	Het	missense	p.Leu86Phe
Chr7:91936750		c.266G>T		missense	p.Gly89Val
Chr9:77357522	*TRPM6*	c.5155A>G	Het	missense	p.Met1719Val
Chr9:77357542		c.5135C>T		missense	p.Ala1712Val
Chr16:.30236689	*NPIPB13*	c.1078C>T	Het	missense	p.Pro360Ser
Chr16:30236722		c.1045C>T		missense	p.Pro349Ser
Chr8:10285751del	*MSRA*	c.517GA>G	Het	Frame	p.Ser173Alafs*36
Chr8:10285753_10285754insG		c.519C>CG		shift	p.Lys174Glufs*21
				Frame	
				shift	
Chr19:56000748del	*SSC5D*	c.80CA>C	Het	Frame	p.His27Leufs*121
Chr19:56000753_56000754ins81		c.85_86ins81		shift	p.Gly28_Cys29insSerArg*
				Stop gain	

3.5. MR208

Family MR208 having individuals affected with Intellectual disability was recruited from district Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.22). We have total six individual blood samples (V: 6, V: 7, V: 8, V: 9, IV: 1 and IV: 2) out of which two individuals (V: 8 and IV: 1) has affected with intellectual disability, while the other four members (V: 6, V: 7, V: 9 and IV: 2) are normal. Normal members show willingness to participate in this research study. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.5.1. Clinical Features

Affected individuals in family MR205 exhibited clinical manifestations of intellectual disability including problem with body use. They have severe memory loss. They have problem with learning, Nonverbal communication. They have problem with intelligence. Affected individual (V: 8) have difficulty in walking. They have very limited personal skills, need care. Information collected from family members shows that they have born normally.

3.5.2. Whole Exome Sequencing and Variants Filtration

Affected individual (V: 8) from family MR208 was subjected to Exome sequencing. Variants present in exome data were first checked for previously reported causing intellectual disability. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of inheritance i.e. homozygous or compound heterozygous as shown in (Table 3.6). Following filtering of variants, eight gene variants were present. One was located in USP17L3 (Chr:8:7835477C>A (hg19 build); NM_001256871.1:c.31G>T; p.Glu11*), gene considered an unlikely candidate Intellectual disability gene. Second was located in FAM72B (Chr:1:120842024A>G (hg19 build); NM_001100910.1:c.202A>G; p.Lys68Glu). Third one is located in MADCAM1 (Chr:19:501743T>C (hg19 build); NMNM_130760.2:c.742T>C; p.Ser248Pro), fourth one is located in GPR42 (Chr:19:35862811G>T (hg19 build); NM_001348195.1:c.550G>T; p.Val184Leu), fifth one is located in LMTK3

(Chr:19:49002060G>C (hg19 build); NM_001080434.1:c.2353C>G; p.Pro785Ala), Sixth one is located in ZNRF3 (Chr:22:29279833C>T (hg19 build); NM_001206998.1:c.79C>T; p.Arg27Trp). Seven one is present in AMER1 (Chr:X:63410151A>G (hg19 build); NM_152424.3:c.3016T>C; p.Ser1006Pro). Eight one is present in CNTNAP3B (Chr:9:43915527G>A (hg19 build); NM_001201380.2:c.3610G>A; p.Ala1204Thr).

3.5.3. Segregation Analysis

Co-segregation analysis of selected variants given in (Table 3.6) will checked on the availability of fund in the future.

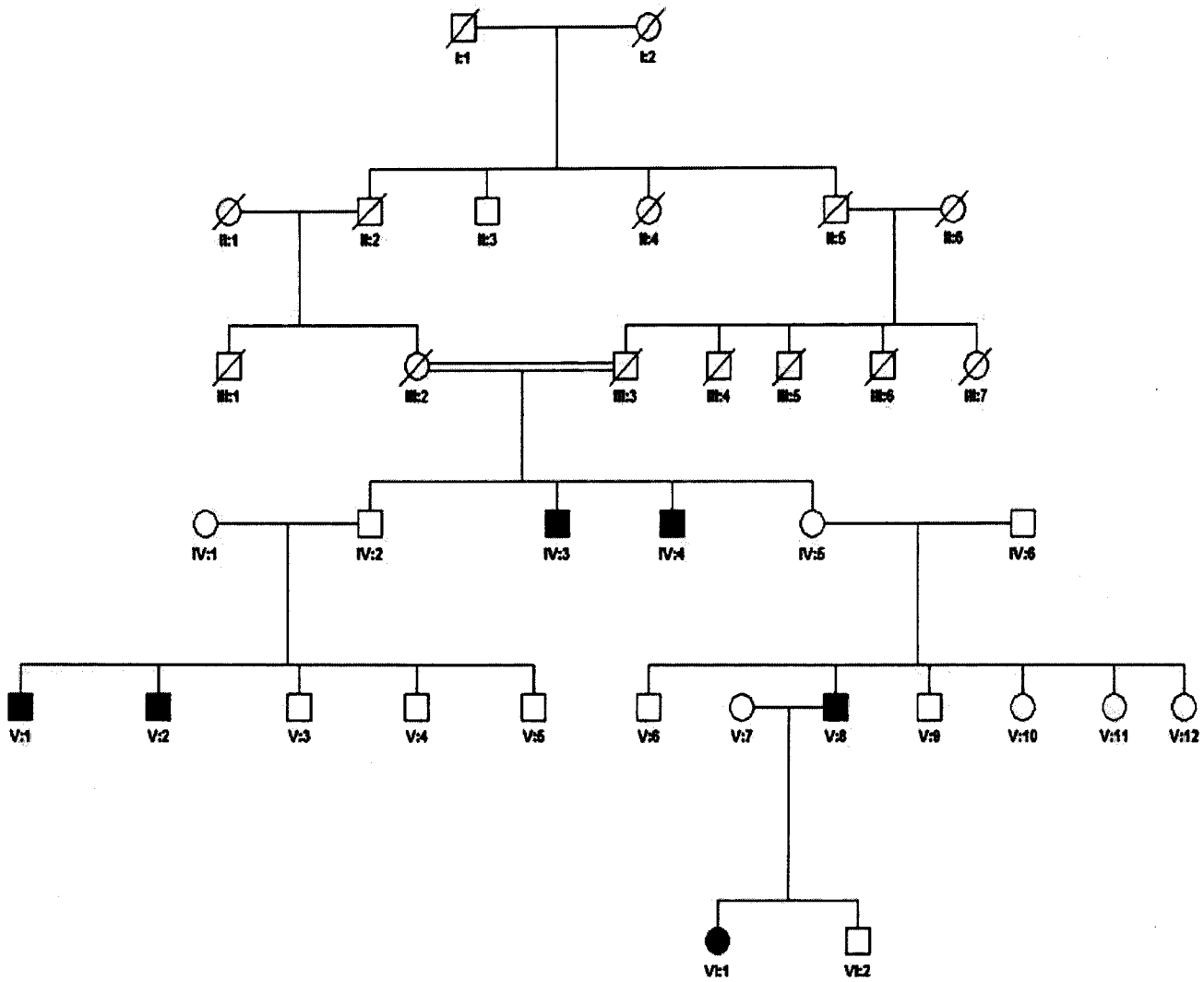


Figure 3.22: Pedigree of family MR208. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

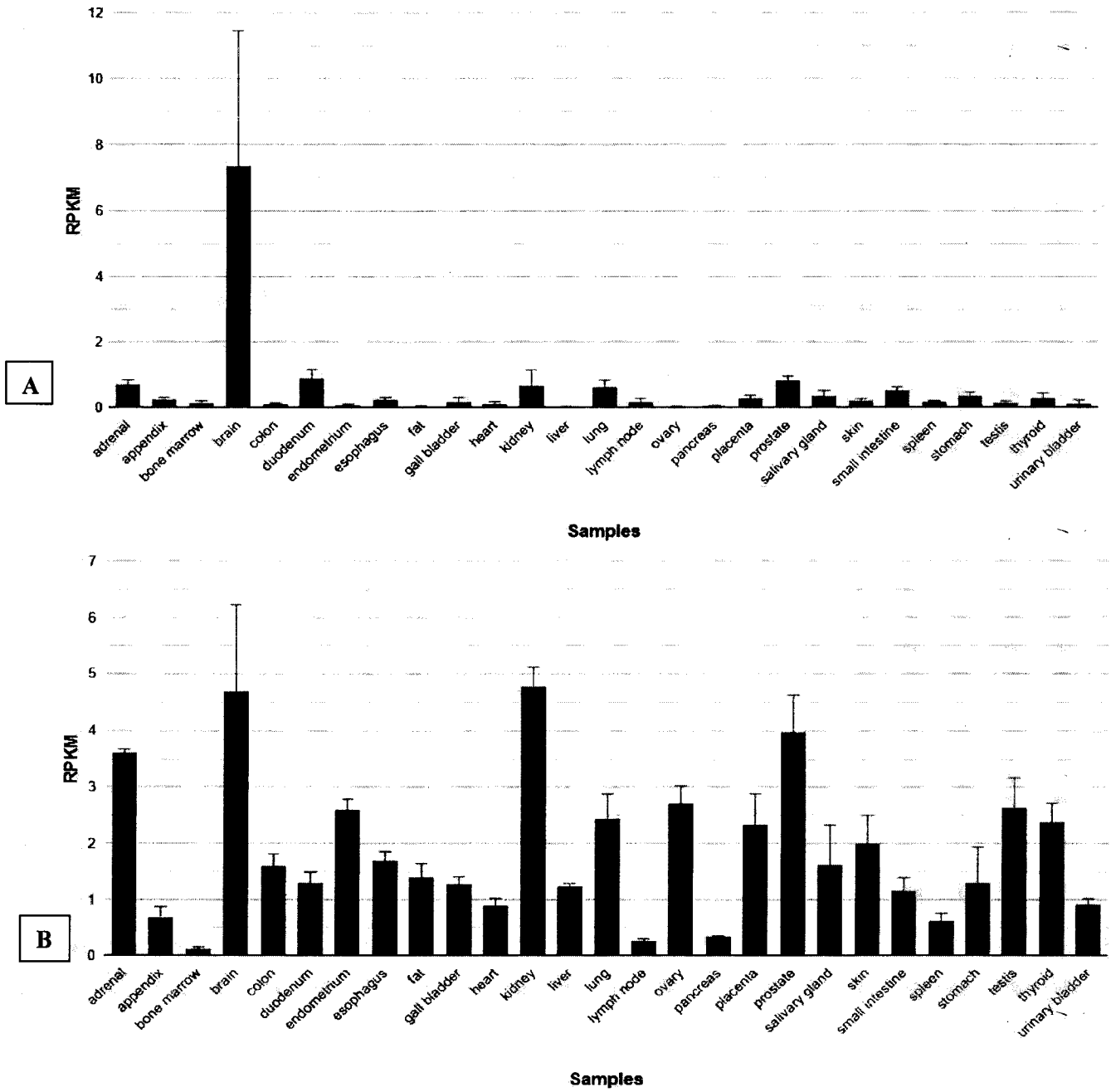


Figure 3.23: (A) represents *LMTK3* gene expression in brain, **(B)** represents *ZNRF3* gene expression in brain. (www.ncbi.nlm.nih.gov).

Table 3. 6: List of filtered homozygous recessive variants identified in the Exome sequencing data of the proband (V: 8) from family MR208

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr[illegible]:7835477[illegible]	*USP[illegible]*	c.[illegible]G>T	Hom	Stopgain	p.Glu[illegible]
Chr1:120842024	*FAM72B*	c.202A>G	Hom	missense	p.Lys68Glu
Chr19:501743	*MADCAM1*	c.742T>C	Hom	missense	p.Ser248Pro
Chr19:35862811	*GPR42*	c.550G>T	Hom	missense	p.Val184Leu
Chr19:49002060	*LMTK3*	c.2353C>G	Hom	missense	p.Pro785Ala
Chr22:29279833	*ZNRF3*	c.79C>T	Hom	missense	p.Arg27Trp
ChrX:63410151	*AMER1*	c.3016T>C	Hom	missense	p.Ser1006Pro
Chr9:43915527	*CNTNAP3B*	c.3610G>A	Hom	missense	p.Ala1204Thr

3.6. MR206

Family MR206 having individuals affected with Intellectual disability was recruited from district Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.24). We have total four individual blood samples (III: 2, III: 3, III: 4 and IV: 1) out of which two individuals (III: 2 and IV: 1) has affected with intellectual disability, while the other two members (III: 3 and III: 4) are normal. Normal members show willingness to participate in this research study. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.6.1. Clinical Features

Affected individuals in family MR206 exhibited clinical manifestations of intellectual disability including problem with body use. They have problem with their thinking skills. They have problem with learning, relationships with family. They have problem with keeping balance. They have very limited personal skills, need care. They have problem with using Objects. Information collected from family members shows that they have born normally.

3.6.2. Whole Exome Sequencing and Variants Filtration

Affected individual (III: 2) from family MR206 was subjected to Exome sequencing. Variants present in exome data were first checked for previously reported causing intellectual disability. All the variants present in exome data are first filter by setting allele frequency or MAF<0.001, predicted protein structure i.e. non-sense, missense, frame-shift and splice site variants were chosen and further filtered based on their compatibility with the autosomal recessive mode of inheritance i.e. homozygous or compound heterozygous as shown in (Table 3.7). Following filtering of variants, four gene variants were present. One was located in RHBDD2 (Chr:7:75517464C>T (hg19 build); NM_001346188.1:c.484C>T; p.Gln162*), gene considered an unlikely candidate Intellectual disability gene. Second was located in USP17L7 (Chr:8:11991270C>G (hg19 build); NMNM_001256869.1:c.249G>C; p.Gln83His). Third one is located in PDZD3 (Chr:11:119059569T>A (hg19 build); NM_024791.3:c.1238T>A; p.Ile413Asn), fourth one is located in GRIPAP1 (Chr:X:48838215C>T (hg19 build); NM_020137.4:c.1669G>A; p.Glu557Lys).

3.6.3. Segregation Analysis

Co-segregation analysis of selected variants given in (Table 3.7) will checked on the availability of fund in the future.

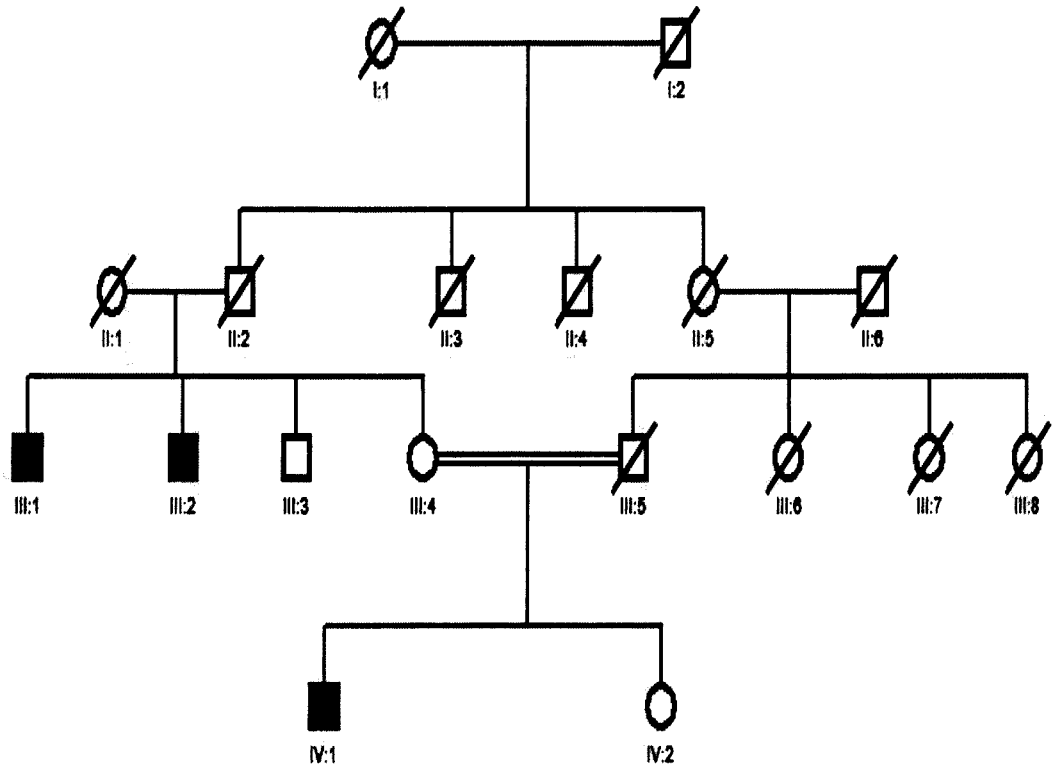
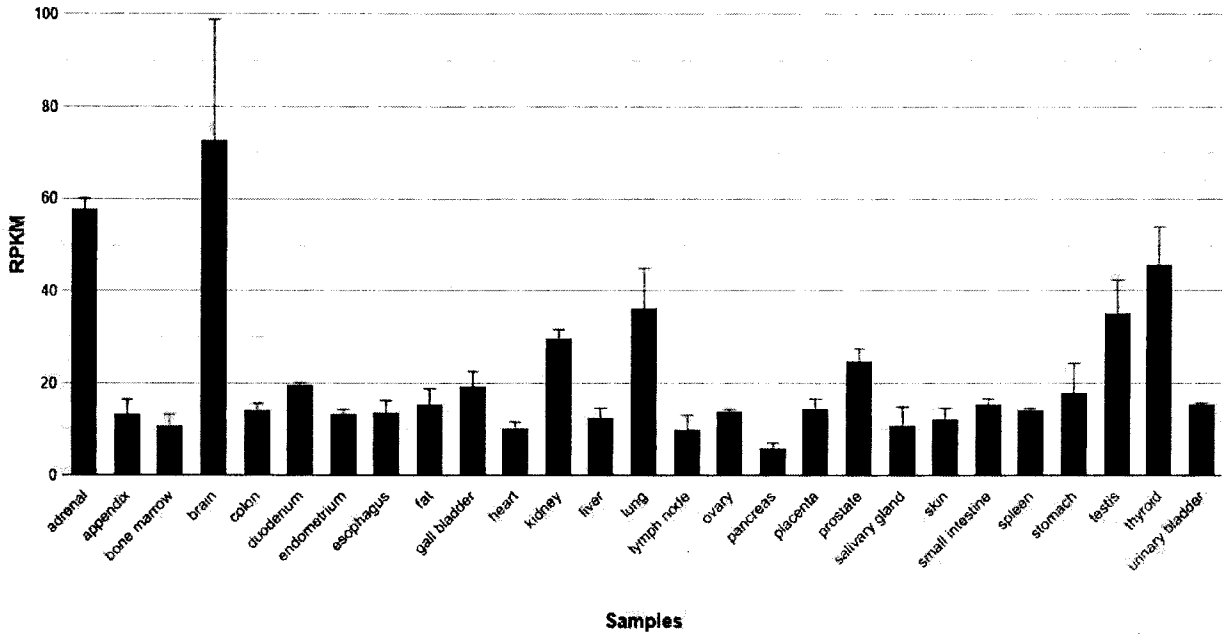


Figure 3.24: Pedigree of family MR206. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

A



B

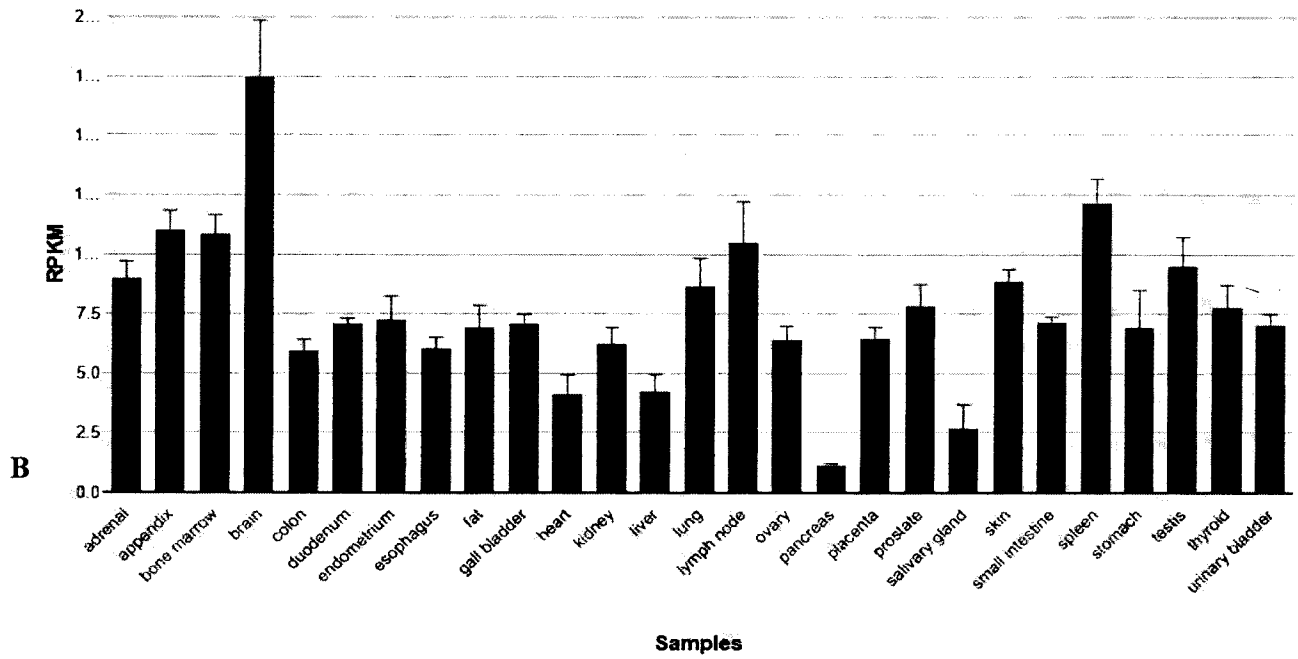


Figure 3.25: (A) represents *RHBDD2* gene expression in brain, **(B)** represents *GRIPAP1* gene expression in brain.(www.ncbi.nlm.nih.gov).

Table 3. 7: List of filtered homozygous recessive variants identified in the Exome sequencing data of the proband (IV: 6) in family MR205.

Position (hg19 build)	Gene	Substitution	Zyg-osity	Effects	Amino acid change
Chr7:75517464	*RHBDD2*	c.484C>T	Hom	Stop/gain	p.Gln162*
Chr8:11991270	*USP17L7*	c.249G>C	Hom	missense	p.Gln83His
Chr11:119059569	*PDZD3*	c.1238T>A	Hom	missense	p.Ile413Asn
ChrX:48838215	*GRIPAP1*	c.1669G>A	Hom	missense	p.Glu557Lys

3.6.4. Discussion

Family MR205 from district Tank (KPK) having intellectual disability was investigated. Whole exome sequencing was carried out in this family and variants were selected on the basis of homozygous recessive pattern of inheritance shown in (Table 3.4), but no variant was co-segregated among the selected gene variants. After that compound heterozygous variants were selected as shown in (Table 3.5), but no variant was co-segregated among the selected gene variants. NGS is required in the future to find the exact causes of diseased phenotypes in the selected family.

Family MR208 having members affected with intellectual disability was investigated. The DNA of proband (III: 2) was subjected to whole exome sequencing and variants were selected from exome data on the basis of homozygous recessive pattern of inheritance as shown in the (Table 3.6). Co-segregation will be checked in the affected individuals in the selected list of variants in the future on the availability of funds.

Family MR206 having intellectual disability from district Tank (KPK) Pakistan was investigated. Variants were selected from exome file on the basis homozygous recessive pattern of inheritance given in (Table 3.7). Most of the variants selected on the basis of their expression in brain and homozygosity. Co-segregation will be checked among the selected variants in the future.

3.7. MR201

Family MR201 having individuals affected with Intellectual disability was recruited from district Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.26). We have total six individual blood samples (IV: 6, V: 3, V: 4, V: 5, V: 6 and V: 7) out of which four individuals (V: 4, V: 5, V: 6 and V: 7) has affected with intellectual disability, while the other two members (IV: 6, and V: 3) are normal. Normal members show willingness to participate in this research study. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.7.1. Clinical Features

Affected individuals in family MR206 exhibited clinical manifestations of intellectual disability. They have problem with body use. They have problem with their thinking skills. They have problem with learning, social relationships with family members and other people of the society. They have problem with keeping balance. They have very limited personal skills, need care. They have problem with using Objects and Emotional response.

3.7.2. Microarray Analysis and Variants Selection

Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (family MR201). Proband (V: 6) share common homozygous region (chr16:59,003,011-74,060,285) total of 15,057,275 bp, comprising total 187 genes. Out of 187 genes, 46 genes (table 3.8) based on their expression in brain were selected for screening in the selected family to find the exact mutation. Proband (V: 6) share another homozygous (chr11:89,070,497-95,304,834) region of 6,234,338 bp contained 92 genes. 7 genes (table 3.9) based on their expression in brain were selected for screening in the selected family MR201 to find exact pattern of disease phenotypes. Proband (V: 6) third homozygous (chr7:83,519,060-93,643,511) region of 10,124,452 bp contained 76 genes. 4 genes (table 3.10) were selected based on their expression in brain for further screening. GRM3 consider the candidate gene in this family, as this gene has most expression in brain and involved in schizophrenia risk (Saini et al., 2017).

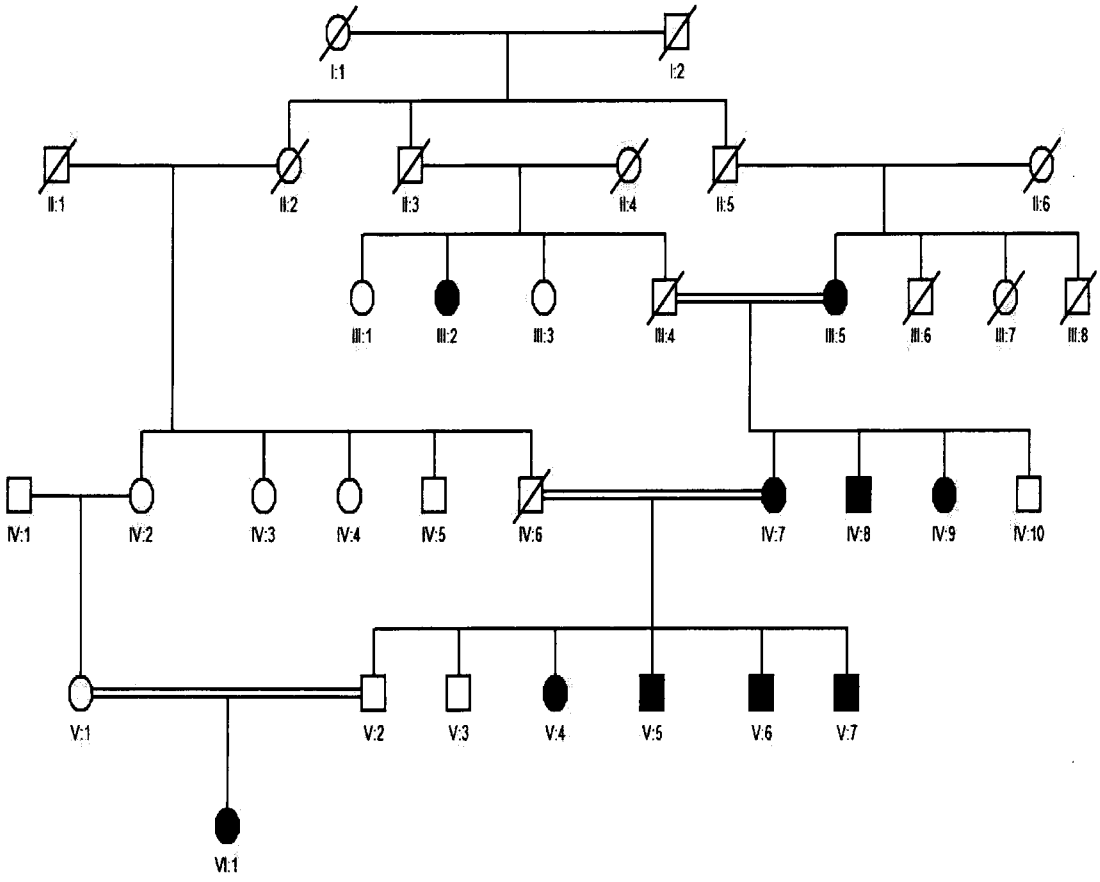


Figure 3.26: Pedigree of family MR201. Pedigree of family MR208. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

Table 3. 8: List of filtered candidate gene selected from microarray mapping data in family MR201.

Position (GRCh37/hg19 build)	Gene	Protein
Chr16: 61,681,146-62,070,939	*GDH8*	799aa
Chr16: 71,859,680-71,895,336	*ZNF821*	412aa
Chr16: 70,375,978-70,439,237	*ST3GAL2*	*350aa*
Chr16: 70,252,295-70,289,543	*AARS*	*968aa*
Chromosome 16: 68,843,604-69,085,180	*TANGO6*	*1094aa*
Chromosome 16: 68,530,090-68,576,072	*ZFP90*	*636aa*
Chromosome 16: 68,310,974-68,358,563	*PRMT7*	*692aa*
Chromosome 16: 67,884,805-67,888,855	*NRN1L*	*165aa*
Chromosome 16: 67,842,082-67,844,195	*THAP11*	*314aa*
Chromosome 16: 67,674,531-67,719,421	*GFOD2*	*112aa*
Chromosome 16: 67,562,407-67,639,185	*CTCF*	*727aa*
Chromosome 16: 67,438,014-67,481,237	*ATP6V0D1*	*351aa*
Chromosome 16: 67,389,809-67,393,535	*TPPP3*	*176aa*
Chromosome 16: 67,277,510-67,289,499	*PLEKHG4*	*1191aa*
Chromosome 16: 67,227,103-67,229,278	*TMEM208*	*103aa*

Table 3. 9: List of filtered candidate gene selected from microarray mapping data in family MR201.

Position (GRCh37/hg19 build)	Gene	Protein
Chromosome 11: 94,377,311-94,401,419	*C1QB83*	423aa
Chromosome 11: 94,128,928-94,181,972	*PANX1*	426aa
Chromosome 11: 93,784,227-93,814,963	*MED17*	*651aa*
Chromosome 11: 93,729,948-93,784,391	*TAF1D*	*278aa*
Chromosome 11: 93,661,639-93,730,358	*CEP295*	*2601aa*
Chromosome 11: 93,144,171-93,197,964	*SLC36A4*	*504aa*
Chromosome 11: 92,352,096-92,896,470	*FAT3*	*4557aa*
Chromosome 11: 90,201,160-90,223,364	*CHORDC1*	*322aa*

Table 3. 10: List of filtered candidate gene selected from microarray mapping data in family MR201.

Position (GRCh37/hg19 build)	Gene	Protein
Chromosome 7: 86,645,913-86,804,884	[illegible]	879aa
Chromosome 7: 84,995,553-85,186,855	*SEMA3D*	777aa
Chromosome 7: 84,983,556-84,984,506	*HNRNPA1P8*	*951aa*
Chromosome 7: 83,955,777-84,492,724	*SEMA3A*	*771aa*

3.8. MR293

Family MR293 having individuals affected with Intellectual disability was recruited from district Malakand (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.27). We have total five individual blood samples (IV: 2, V: 1, V: 2, V: 3 and V: 4) out of which four individuals (V: 1, V: 2, V: 3 and V: 4) has affected with intellectual disability, while the other one members (IV: 2) is normal. Normal members show willingness to participate in this research study. Phenotypic examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.8.1. Clinical Features

Affected individuals in family MR293 exhibited clinical manifestations of intellectual disability. They have problem with body use. They have problem with their thinking skills. They have problem with learning, social relationships with family members and other people of the society. They have problem with keeping balance. They have very limited personal skills, need care. They have problem Emotional response.

3.8.2. Molecular Genetic Studies

Family MR293 has been screen for known genes causing ID. Microsatellite marker for each gene was used for linkage studies and genotype for affected individuals and normal individuals of the family. In this family, five DNA samples (IV: 2, V: 1, V: 2, V: 3 and V: 4), including four affected and one normal individuals were genotyped. In family linkage was not established with known genes. After family was excluded from known linkage, Single-nucleotide polymorphism (SNP) microarray genotyping was planned to find the causative variants.

3.8.3. Microarray Analysis and Variants Selection

Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR293). Affected individuals (V: 1, V: 2, V: 3 and V: 4) share common homozygous region (chr:15:38,657,924-54,623,246) total of 15,965,323 bp, contained total 246 genes. 8 genes (table 3.11) based on their expression in brain were selected for screening in the selected family to find the exact mutation. Affected individuals (V: 1, V: 2, V: 3 and V: 4) share

another homozygous (chr:8:30,023,540-42,001,906) region of 11,978,367 bp contained 97 genes. 6 genes (table 3.12) based on their expression in brain were selected for screening in the selected family MR201 to find exact pattern of disease phenotypes.

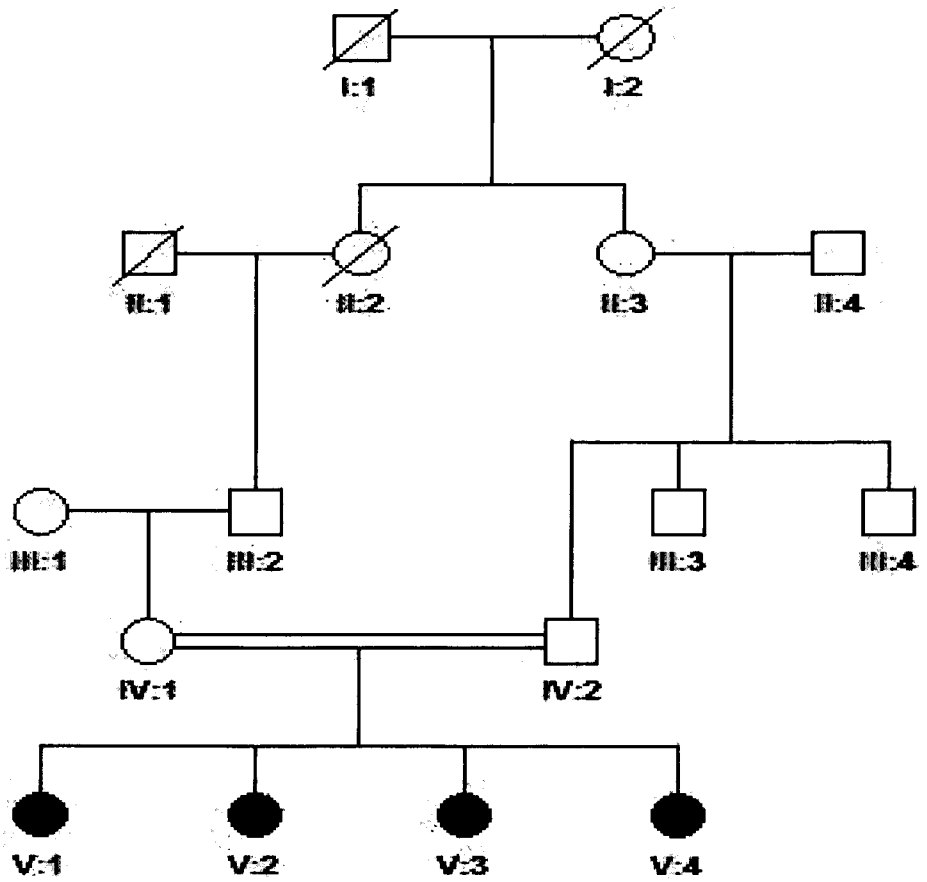


Figure 3.27: Pedigree of family MR293. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

Table 3. 11: List of filtered candidate gene selected from microarray mapping data (chr15:38,657,924-54,623,246) in family MR293.

Position (GRCh37/hg19 build)	Gene	Protein
Chromosome 15: 43,870,761-44,195,252	*ERMD5*	336aa
Chromosome 15: 42,738,734-42,920,809	*TTBK2*	1244aa
Chromosome 15: 42,575,659-42,720,981	*STARD9*	*4700aa*
Chromosome 15: 41,517,176-41,544,269	*RPAP1*	*1393aa*
Chromosome 15: 40,594,020-40,664,342	*KNL1*	*2342aa*
Chromosome 15: 40,035,739-40,039,188	*SRP14*	*136aa*
Chromosome 15: 39,799,032-39,920,892	*GPR176*	*514aa*
Chromosome 15: 38,454,127-38,487,710	*FAM98B*	*433aa*

Table 3. 12: List of filtered candidate gene selected from microarray mapping data (chr8:30,023,540-42,001,906) in family MR293.

Position (GRCh37/hg19 build)	Gene	Protein
Chromosome 8: 37,736,160-37,759,101	*EIF4EBP2*	239aa
Chromosome 8: 33,591,332-33,600,106	*DUSP26*	211aa
Chromosome 8: 33,473,386-33,513,601	*TTI2*	*508aa*
Chromosome 8: 30,995,802-31,033,715	*PURG*	*347aa*
Chromosome 8: 30,774,457-30,814,314	*PPP2CB*	*309aa*
Chromosome 8: 30,638,600-30,646,064	*SMIM18*	*95aa*

3.9. MR296

Family MR296 having individuals affected with Intellectual disability was recruited from district Mardan (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.28). We have total three individual blood samples (II: 2, II: 8 and III: 1) out of which two individuals (II: 8 and III: 1) has affected with intellectual disability, while the other one members (II: 2) is normal. Normal members show willingness to participate in this research study. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.9.1. Clinical Features

Affected individuals in family MR296 exhibited clinical manifestations of intellectual disability. They have problem with their thinking skills. They have problem with learning, social relationships with family members and other people of the society. They have very limited personal skills, need care. They have problem with Emotional response.

3.9.2. Microarray Analysis and Variants Selection

Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR296). Affected individuals (II: 8 and III: 1) share single common homozygous region (chr11:49,779,717-51,274,692) total of 1,494,976 bp, contained total 18 genes. 3 genes (table 3.13) based on their expression in brain were selected for screening in the selected family to find the exact mutation.

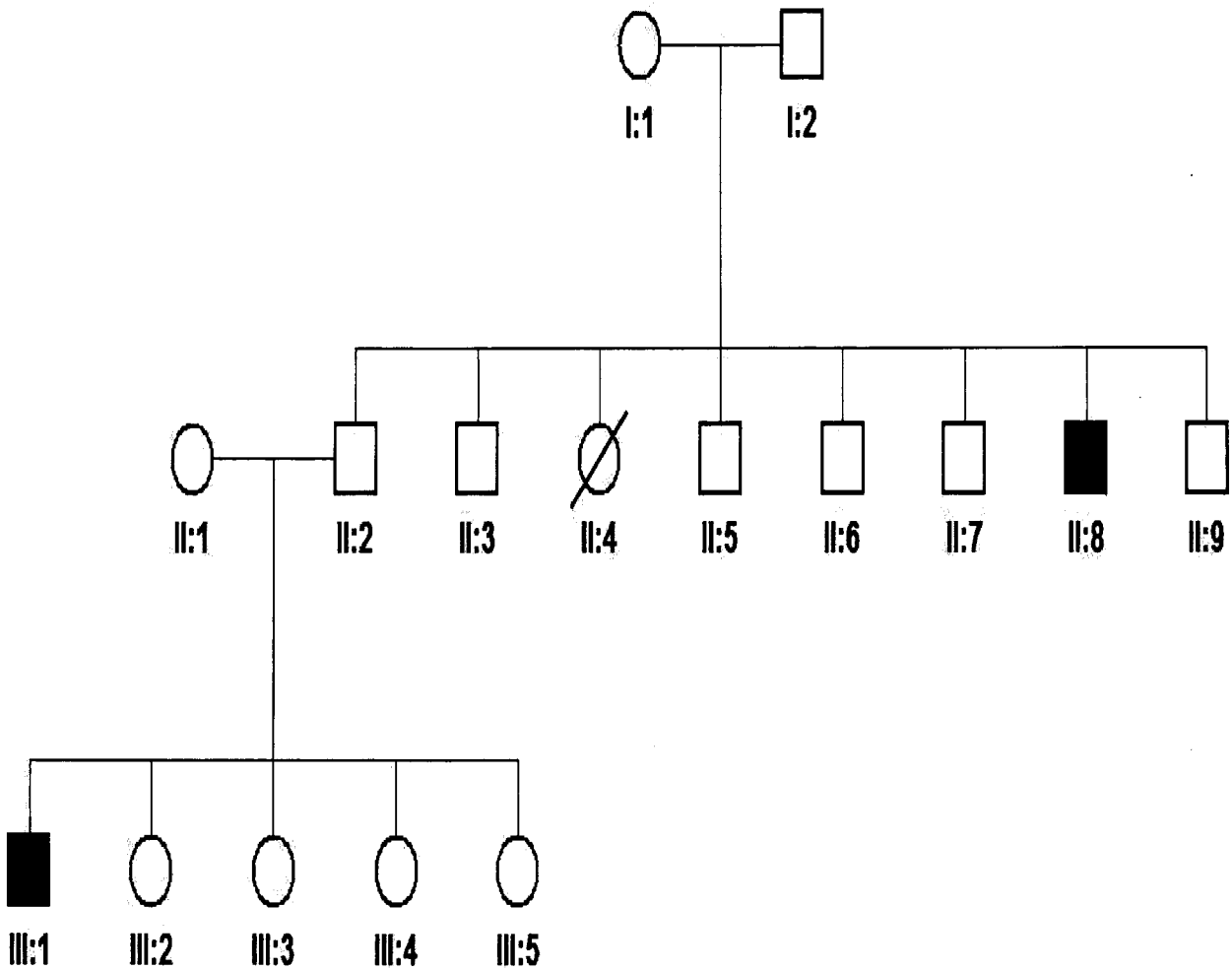


Figure 3.28: Pedigree of family MR296. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

Table 3. 13: List of filtered candidate gene selected from microarray mapping data (chr11:49,779,717-51,274,692) in family MR296.

Position (GRCh37/hg19 build)	Gene	Protein
Chromosome 11: 49,915,013-49,915,644	*OR4A49P*	N/A
Chromosome 11: 49,898,267-49,899,186	*OR4A1P*	N/A
Chromosome 11: 49,558,546-49,810,419	*AC136759.1*	*N/A*

3.9.4. Discussion

Family MR201 having intellectual disability was recruited from district Tank (KPK) Pakistan. Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR201). Proband (V: 6) share common homozygous region (chr:16:59,003,011-74,060,285) total of 15,057,275 bp, comprising total 187 genes. Out of 187 genes, 46 genes (Table 3.8) based on their expression in brain were selected for screening in the selected family to find the pattern of inheritance of mutation with diseased phenotype. Proband (V: 6) share second homozygous (chr:11:89,070,497-95,304,834) region of 6,234,338 bp contained 92 genes. 7 genes (Table 3.9) based on their expression in brain were selected for screening in the selected family MR201 to find exact pattern of disease phenotypes. Proband (V: 6) third homozygous (chr:7:83,519,060-93,643,511) region of 10,124,452 bp contained 76 genes. 4 genes (Table 3.10) were selected based on their expression in brain for further screening. GRM3 consider the candidate gene in this family, as this gene has most expression in brain and involved in schizophrenia risk (Saini et al., 2017). Marker saturation is performed using existing and newly created microsatellite markers. A multipoint linkage analysis was performed with Simwalk2275 on an autosomal recessive genetic model with full access using the estimated allele frequency of 0.0003.

Family MR293 having individuals affected with Intellectual disability was recruited from district Malakand (KPK) of Pakistan. In this study family with Intellectual disability was evaluated at molecular level. Family was excluded by classical linkage analysis. After family was excluded from linkage, the genomic DNA was subjected to Single-nucleotide polymorphism (SNP) microarray genotyping. Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR293). Affected individuals (V: 1, V: 2, V: 3 and V: 4) share common homozygous region (chr:15:38,657,924-54,623,246) total of 15,965,323 bp, contained total 246 genes. 8 genes (Table 3.11) based on their expression in brain were selected for screening in the selected family to find the exact pattern of mutation. Affected individuals (V: 1, V: 2, V: 3 and V: 4) share second homozygous (chr:8:30,023,540-42,001,906) region of 11,978,367 bp contained 97 genes. 6 genes (Table 3.12) based on their expression in brain were selected for screening in the selected family MR201 to find causative gene.). Marker saturation will be performed with existing and newly developed microsatellite markers.

Family MR296 having individuals affected with Intellectual disability was recruited from district Mardan (KPK) of Pakistan. Microarray genotyping was performed using Illumina Human CytoSNP-12v2.0 in family (MR296). Affected individuals (II: 8 and III: 1) share single common homozygous region (chr:11:49,779,717-51,274,692) total of 1,494,976 bp, contained total 18 genes. 3 genes (Table 3.13) based on their expression in brain were selected for screening in the selected family to find the causative gene. Microsatellite marker for each genes will be designed and check linkage in each genes in the future.

Co-segregation analysis of *AFF2* and *RPTN* variants in six families

3.10. MR202

Family MR202 having individuals affected with Intellectual disability was recruited from district Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.29). We have total four individual blood samples (IV: 3, IV: 4, IV: 5 and IV: 6) out of which two individuals (IV: 5 and IV: 6) has affected with intellectual disability, while the other two members (IV: 3 and IV: 4) are normal. Normal members show willingness to participate in this research study. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.10.1. Clinical Features

Affected individuals in family MR202 exhibited clinical manifestations of intellectual disability including problem with long term memory. Affected individual (IV: 5) is very aggressive in action. They have problem with object use. They have problem with learning, social relationships with family members and other people of the society. They have very limited personal skills, need care. They have problem with Emotional response.

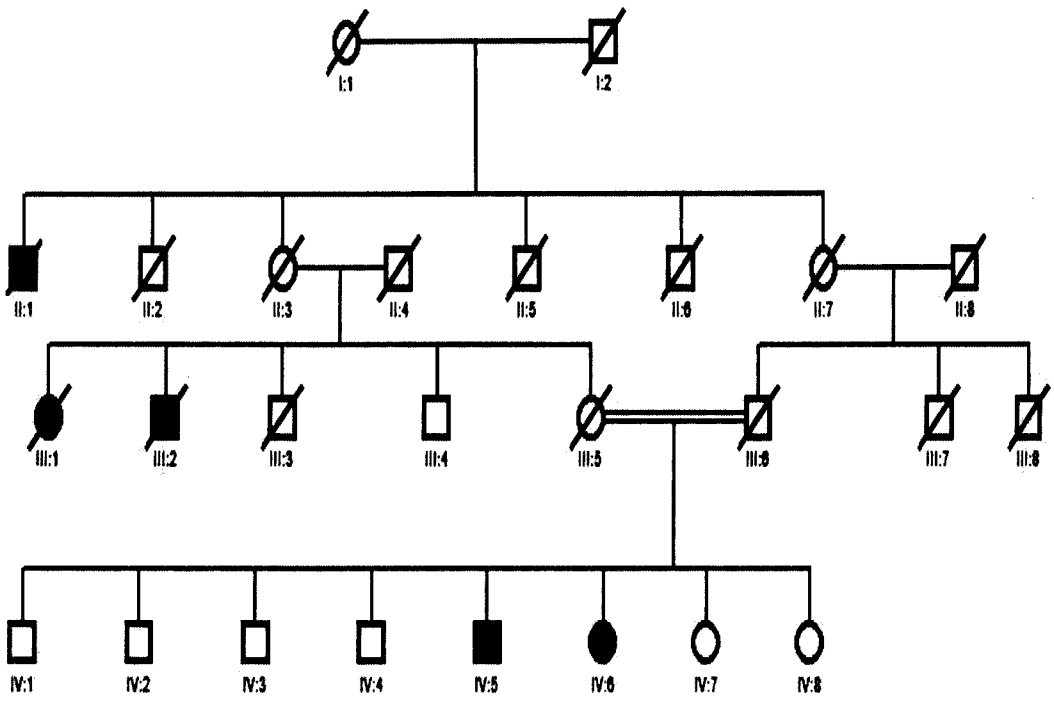


Figure 3.29: Pedigree of family MR202. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.11. MR203

Family MR203 having individuals affected with Intellectual disability was recruited from district Swabi (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.30). We have total five individual blood samples (III: 4, III: 5, IV: 1, IV: 2 and IV: 3) out of which two individuals (IV: 2 and IV: 3) has affected with intellectual disability, while the other three members (III: 4, III: 5 and IV: 1) are normal. Normal members show willingness to participate in this research study. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.11.1. Clinical Features

Affected individuals in family MR203 exhibited clinical manifestations of intellectual disability having problem with language. Affected individual (IV: 3) is very aggressive in action. They have problem with learning, social relationships with family members and other people of the society. They have very limited personal skills, need care. They have problem in nonverbal communication. Also, they have repetitive behavior.

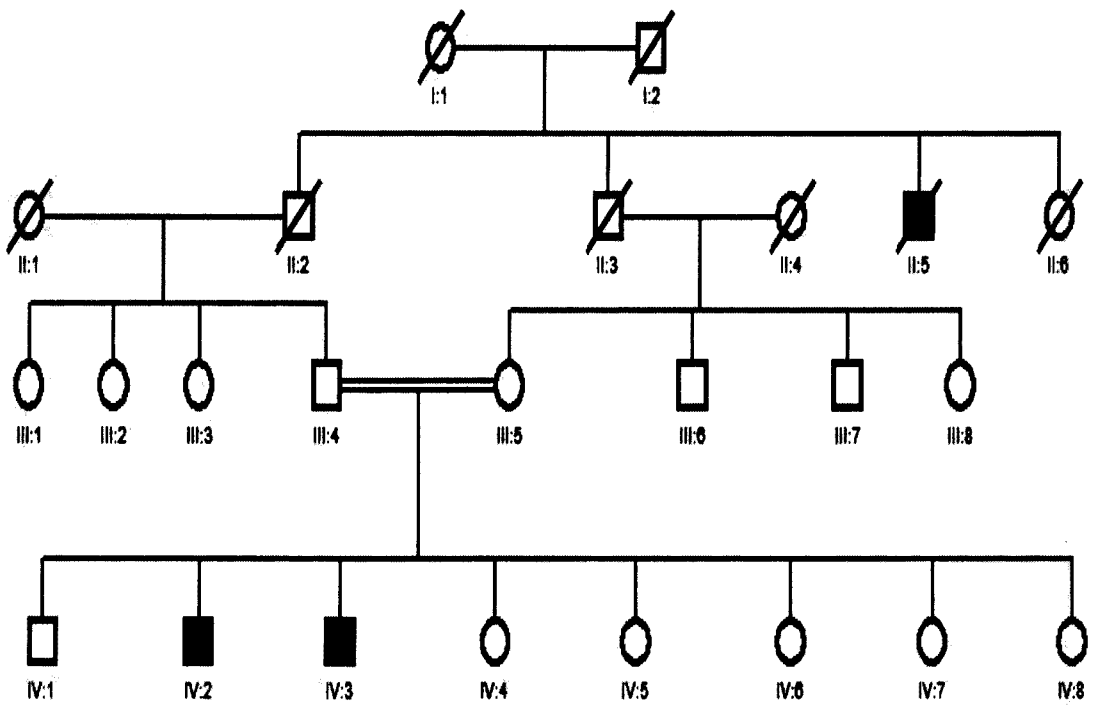


Figure 3.30: Pedigree of family MR203. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.12. MR207

Family MR207 having individuals affected with Intellectual disability was recruited from district FR Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.31). We have total four individual blood samples (V: 2, V: 4, V: 5, and VI: 1) out of which two individuals (V: 2 and V: 5) has affected with intellectual disability, while the other two members (V: 4, and VI: 1) are normal. Normal members show willingness to participate in this research study. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.12.1. Clinical Features

Affected individuals in family MR207 exhibited clinical manifestations of intellectual disability including problem with language. They have problem with learning, social relationships with family members and other people of the society. They have problem in verbal as well as nonverbal communication. They have repetitive behavior. They have also a problem with intelligence.

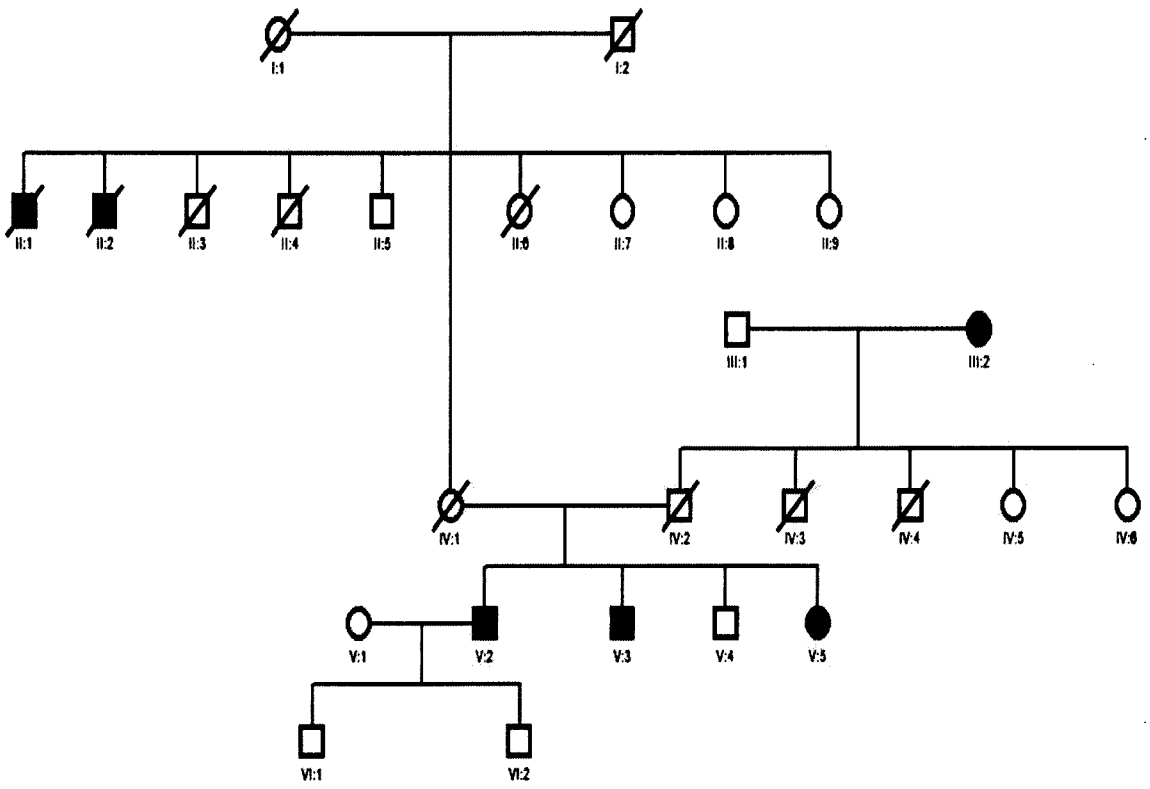


Figure 3.31: Pedigree of family MR207. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.13. MR209

Family MR209 having individuals affected with Intellectual disability was recruited from village station ketch district Tank (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.32). We have total three individual blood samples (III: 3, IV: 1 and IV: 2) out of which two individuals (III: 3 and IV: 1) has affected with intellectual disability, while the other one member (IV: 2) is normal. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.13.1. Clinical Features

Affected individuals in family MR209 exhibited clinical manifestations of intellectual disability having problem with language. They have problem with learning, social relationships with family members and other people of the society. They have problem in verbal as well as nonverbal communication. They have repetitive behavior. They has problem with intelligence. They have problem with body use, object use, response and adaptation to change.

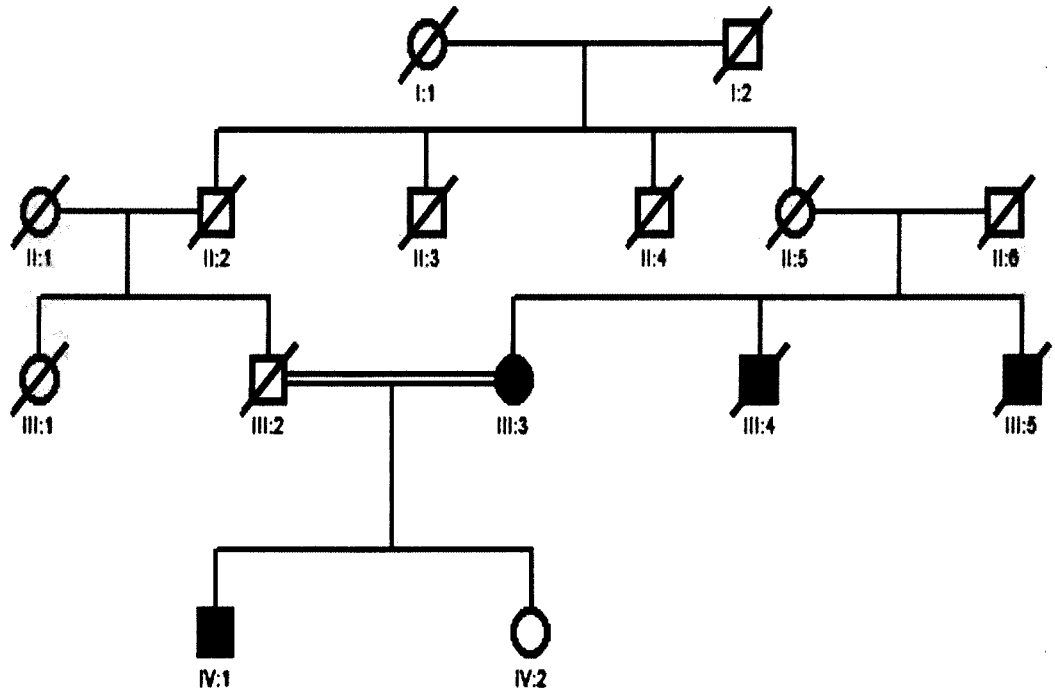


Figure 3.32: Pedigree of family MR209. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.14. MR213

Family MR213 having individuals affected with Intellectual disability was recruited from village markhani maidan district Dir lower (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.33). We have total six individual blood samples (III: 4, III: 5, IV: 1, IV: 2, IV: 3 and IV: 4) out of which three individuals (IV: 1, IV: 2 and IV: 3) has affected with intellectual disability, while the other three members (III: 1, III: 2 and IV: 4) are normal. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.14.1. Clinical Features

Affected individuals in family MR213 exhibited clinical manifestations of intellectual disability including problem with language. They have problem with learning, social relationships with family members and other people of the society. They have problem in verbal as well as nonverbal communication. They have repetitive behavior. They has problem with intelligence. They have problem with body use, object use, response and adaptation to change.

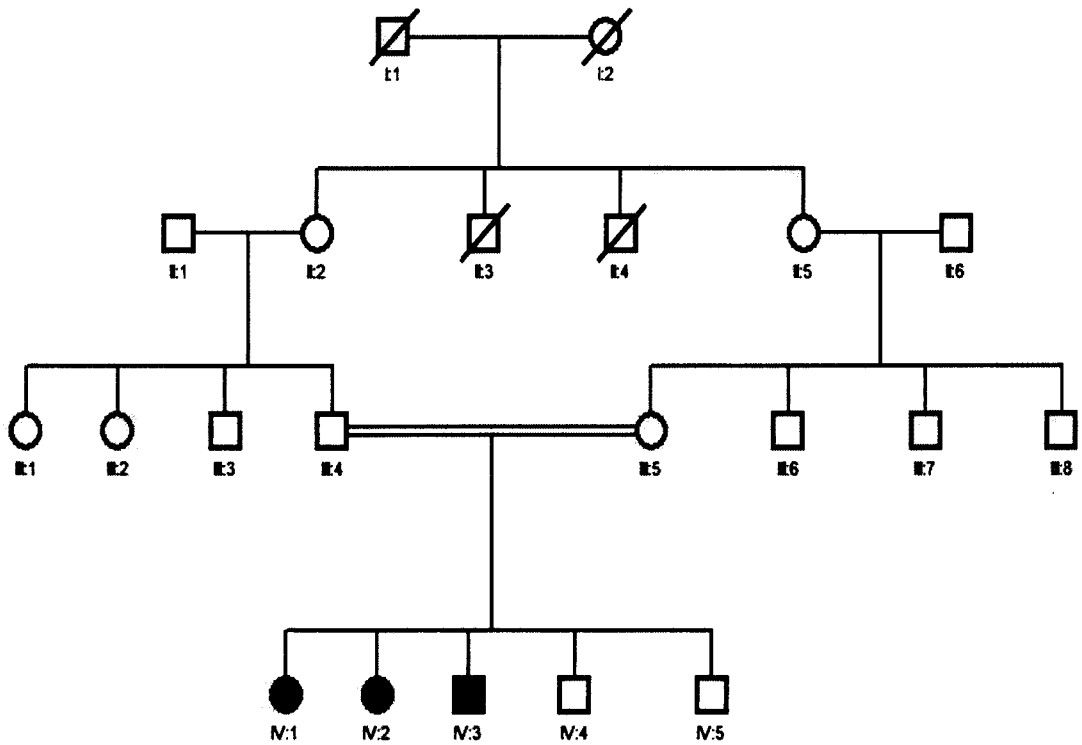


Figure 3.33: Pedigree of family MR213. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.15. MR214

Family MR214 having individuals affected with Intellectual disability was recruited from village markhani maidan district Dir lower (KPK) of Pakistan. Information was collected from normal family members and pedigree was drawn as shown in (Figure 3.34). We have total four individual blood samples (III: 3, IV: 1, IV: 2, and IV: 3) out of which two individuals (IV: 1, and IV: 2) has affected with intellectual disability, while the other two members (III: 3 and IV: 3) are normal. Examination of the affected individuals was performed by Phsychiatric at local government hospital.

3.15.1. Clinical Features

Affected individuals in family MR214 exhibited clinical manifestations of intellectual disability having problem with language. They have problem with learning, social relationships with family members and other people of the society, problem in verbal as well as nonverbal communication, repetitive behavior, as well as have problem with intelligence, have problem with body use, object use, response and adaptation to change.

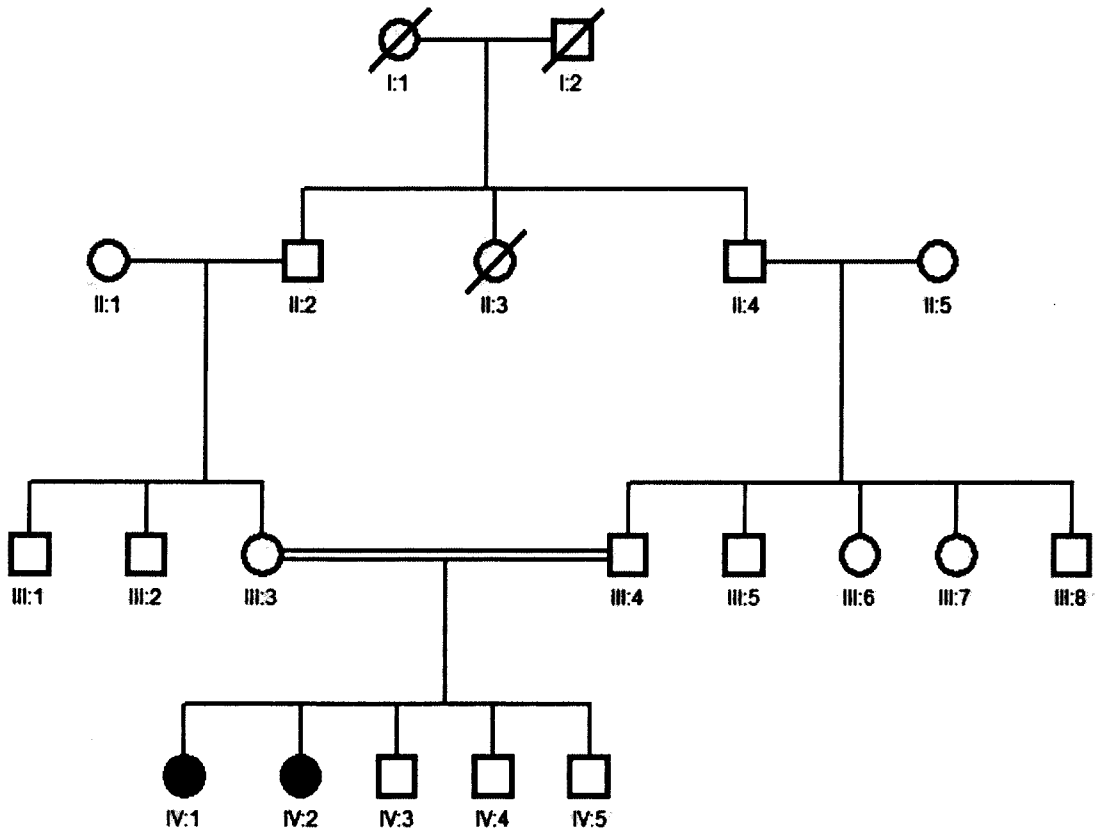


Figure 3.34: Pedigree of family MR214. Consanguineous marriages are represented by a double line. Circles represents female and squares represent males. Filled symbols represent affected individuals while clear symbols represent unaffected individuals. Lines passes through symbols represent deceased members.

3.16. Co-segregation analysis

DNA of the affected individual from family MR202 (IV: 3 and IV: 4), MR203 (IV: 2 and IV: 3), MR207 (V: 2 and V: 5), MR209 (III: 3 and IV: 1), MR213 (IV: 1, IV: 2 and IV: 3), MR214 (IV: 1 and IV: 2) were subjected to Sanger sequencing for *AFF2* and *RPTN* mutation. But no pathogenic variant was identified, after the exclusion *AFF2* and *RPTN* genes in the selected families; whole exome sequencing will be carried out in the future on the availability of funds to find the causative gene/variants.

3.17. Future Perspective

The goals of the bigger assignment are to define the genomic basis of hereditary disorders which occur in the Pakistani people and to transform the discoveries of this investigation into straight investigative and therapeutic advantages. These discoveries have formed the basis of a widespread molecular genetic investigations and treatment.

The scientific developments in genetics have promptly improved the swiftness and scope of diagnostic genetic testing. The influence of this is particularly noteworthy in the arena of neuro-developmental conditions, where disorders display important phenotypic and genetic heterogeneity. Next generation sequencing expertise permit the screening of several potentially causative genes in parallel, which means that recognition of the accurate phenotype of an individual is no longer fully necessary.

In recent years a growing numbers of genes involved in conditions of human brain development and growth have been discovered. These studies offer extraordinary prospects to identify the molecules associated with the normal brain growth and what happens when their cellular functions are disturbed. The next challenge is to use this information to develop new and advanced treatments targeted directly to avoid the recognized cellular deficiency. The potential of genomics to permit delivery of such cures appears to be progressively possible to renovate the practice of medicine. Further findings will help in drug designing. In future, our research help to enhance he public awareness to reduce the consanguineous marriages in the population which will result in the reduction of inherited disease.

Using this research study, we can predict the interaction of a mutated protein with other proteins and how it affects the pathway and leads to disease.

Our mutated sequence can be analyzed by protein expression in a variety of expression systems. This research will help in the future prenatal diagnosis and gene therapy of related diseases.

References

1. AAMR (2005). Definition of intellectual disability, American Association on Mental Retardation.
2. Agha Z, Iqbal Z, Azam M, Vissers LE, Gilissen C, Ali SH, Raiz M, Veltman JA, Pfundt R, et al. Exom sequencing identifies three novel candidate gene implicated in intellectual disability. PLoS one. 2014 Nov 18;9(11);e 112687.
3. Ahmed I, Buchert R, Zhou M, Jiao X, Mittal K, Sheikh TI, Scheller U, Vasli N, Rafiq MA, et al. Mutation in DCPS and EDC3 in autosomal recessive intellectual disability indicate a crucial role for mRNA decapping in neurodevelopment. Hum Mol Genet. 2015 jun 1; 24(11):3172-80.
4. Ahuja AS., Thapar A and Owen MJ. Genetics of mental retardation. Indian J. Med. Sci. 2005, 59: 407-417.
5. Al-Dosari, M. S., Al-Owain, M., Tulbah, M., Kurdi, W., Adly, N., Al-Hemidan, A., Masoodi, T. A., Albash, B., Alkuraya, F. S. Mutation in MPDZ causes severe congenital hydrocephalus. J. Med. Genet. 2013, 50: 54-58.
6. APA (1994). Diagnostic and Statistical Manual of Mental Disorders IV (ed 4). Washington, DC American Phsychiatric Association.
7. Bahn S, Volk B and Wisden W. Kainate receptor gene expression in the developing rat brain. J. Neurosci. 1994, 14: 5525–5547.
8. Basel-Vanagaite L, Attia R, Yahav M, Ferland RJ, Anteki L, Walsh CA, Olender T, Straussberg R, Magal N, Taub E, Drasinover V, Alkelai A, Bercovich D, Rechavi G, Simon AJ and Shohat M. The CC2D1A, a member of a new gene family with C2 domains, is involved in autosomal recessive nonsyndromic mental retardation. J. Med. Genet. 2005, 43: 203–210.
9. Basel-Vanagaite L. Clinical Approaches to Genetic Mental Retardation. The Israel Medical Association Journal. 2008, 10: 821–826.
10. Basel-Vanagaite L. Genetics of autosomal recessive non-syndromic mental retardation: Recent advances. Clin. Genet. 2007, 72: 167–174.

11. Bennett R L, K A Steinhaus, S B Uhrich, C K O'Sullivan, R G Resta, D Lochner-Doyle, D S Markel, V Vincent, and J Hamanishi. Recommendations for standardized human pedigree nomenclature. Pedigree Standardization Task Force of the National Society of Genetic Counselors. Am J Hum Genet. 1995 Mar; 56(3): 745–752.
12. Bensaid M, Melko M, Bechara EG, Davidovic L, Berretta A, Catania MV, Gecz J, Lalli E, Bardoni B. FRAXE-associated mental retardation protein (FMR2) is an RNA-binding protein with high affinity for G-quartet RNA forming structure. Nucleic Acids Res. 2009 Mar; 37(4):1269-79. doi: 10.1093/nar/ gkn1058. Epub 2009 Jan 9.
13. Bernkopf M, Webersinke G, Tongsook C, Koyani CN, Rafiq MA, Ayaz M, Muller D, Enzinger C, Aslam M, Naeem F, Schmidt K and 7 others. Disruption of the methyletransferase-like 23 gene METTL23 causes mild autosomal recessive intellectual disability. Hum Mol Genet. 2014 Aug 1;23(15):4015-23.
14. Bhalla K, Luo Y, Buchan T. Beachem M Guzauskas GF, Ladd S, Bratcher SJ, Schroer RJ, Balsamo J, DuPont BR, Lilien J, srivastava AK. Alterations in cDH15 and KIRREL3 in patients with mild to severe intellectual disability. Am J Hum Genet. 2008; 83:703-713.
15. Bittles A.H. and Black M.L.. Consanguineous Marriage and Human Evolution: Annual Review of Anthropology. 2010. 39: 193-207
16. Castellvi-Bel S and Mila M. Genes responsible for nonspecific mental retardation. Mol. Genet. Metab. 2001, 72: 1104–1108.
17. Chelly J, Khelfaoui M, Francis F, Cherif B and Bienvenu T. Genetics and pathophysiology of mental retardation. European Journal of Human Genetics. 2006, 14: 701–713.
18. Derijard, B., Hibi, M., Wu, I.-H., Barrett, T., Su, B., Deng, T., Karin, M., Davis, R. J. JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. Cell. 1994, 76: 1025-1037.
19. des Portes V, Boddaert N, Sacco S, Briault S, Maincent K, Bahi N, Gomot M, Ronce N, Bursztyn J Adamsbaum C, Zilbovicius M, Chelly J, Moraine C. Specific

39. Khan MA, Rafiq MA, Noor A, Hussain S, Flores JV, Rupp V, Vincent AK, Malli R and 11 others. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am J Hum Genet. 2012 May 4;90(5):856-63.
40. Law R, Dixon-Salazar T, Jerber J, Cai N, Abbasi AA, Zaki MS, Mittal K, Gabriel SB, Rafiq MA, Khan V, et al (2014). Biallelic truncating mutations in FMN2, encoding the actin-regulating protein Formin 2, cause nonsyndromic autosomal recessive intellectual disability. Am J Hum Genet. 2014 Dec 495(6):721-8.
41. Lenski C, Abidi F, Meindl A, Gibson A, Platzer M, Frank Kooy R, Lubs HA, Stevenson RE, Ramser J, Schwartz CE. Novel truncating mutations in the polyglutamine tract binding protein 1 gene (PQBP 1) cause renpenning syndrome and X-linked mental retardation in another family with microcephaly. Am J Hum Genet. 2004; 74:777-80.
42. Lisik MZ, Sieron AL. X-Linked mental retardation. Med Sci Monit. 2008; 14; RA 221-29.
43. Lovell CS, Davis WI, Arendall BW, Bakker WIP, Word MJ, Prisant GM, Richardson SJ, Richardson CD, Structure validation by Calpha geometry: phi,psi and Cbeta deviation.*Proteins: Structure, Function & Genetics* 2002, 50: 437-450.
44. Majnemer, A. (2012). Measures for Children with Developmental Disabilities: An ICF-CY Approach.(London, UK: Mac Keith Press).
45. 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(2); 419-436.
46. McGuffin LJ, Bryson K, Jones DT: The PSIPRED protein structure prediction server.Bioinformatics 2000, 4: 404-5.
47. Mehdi S. Q. R. Qamar Q. Ayub S. Khaliq A. Mansoor M. Ismail M. F. Hammer P. A. UnderhillL. L. Cavalli-Sforza. The Origins of Pakistani Populations, Evidence from Y chromosome Markers: Genomic Diversity: Applications in Human Population Genetics. 1999; 83-90.
48. Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, & Naeem F. Identification of mutations in TRAPPC9, which encodes the NIK-and IKK-β-

binding protein, in nonsyndromic autosomal-recessive mental retardation. The American Journal of Human Genetics. 2009, 85(6), 909-915.

49. Mir A, Sritharan K, Mittal K, Vasli N, Araujo C, Jamil T,& Mahmood H. Truncation of the E3 ubiquitin ligase component FBXO31 causes non-syndromic autosomal recessive intellectual disability in a Pakistani family. Human genetics. 2014, 133(8), 975-984.
50. Molinari F, Foulquier F, Tarpey PS, Morelle W, Boissel S, Teague J, Edkins S, Futreal PA, Stratton MR, Turner G, Matthijs G, Gecz J, Munnich A and Colleaux L. Oligosaccharyltransferase-subunit mutations in nonsyndromic mental retardation. Am. J. Hum. Genet. 2008, 82: 1150-1157.
51. Molinari F, Rio M, Meskenaite V, Encha-Razavi F, Auge J, Bacq D, Briault S, Vekemans M, Munnich A, Attie-Bitach T, Sonderegger P and Colleaux L. Truncating neurotrypsin mutation in autosomal recessive nonsyndromic mental retardation. Science. 2002, 298: 1779–1781.
52. Musante L, Ropers HH. Genetics of recessive cognitive disorders: Trends in Genetics. 2014, 30(1): 32-39.
53. Opitz JM. Historiography of the causal analysis of mental retardation. Speech to the 29th Annual Gatlinburg Conference on Research and Theory in Mental Retardation, Gatlinburg, TN. 1996
54. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12.
55. Philips, A. K., Pinelli, M., de Bie, C. I., Mustonen, A., Maatta, T., Arts, H. H., Wu, K., Roepman, R., Moilanen, J. S., Raza, S., Varilo, T., Scala, G., Cocozza, S., Gilissen, C., van Gassen, K. L. I., Jarvela, I. Identification of C12orf4 as a gene for autosomal recessive intellectual disability. Clin. Genet. 2017, 91: 100-105.
56. Rafiq MA, Kuss AW, Puettmann L, Noor A, Ramiah A, Ali G& Khan MA. Mutations in the alpha 1, 2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. The American Journal of Human Genetics. 2011, 89(1), 176-182.

41. Khan MA, Rafiq MA, Noor A, Hussain S, Flores JV, Rupp V, Vincent AK, Malli R and 11 others. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am J Hum Genet. 2012 May 4;90(5):856-63.
42. Law R, Dixon-Salazar T, Jerber J, Cai N, Abbasi AA, Zaki MS, Mittal K, Gabriel SB, Rafiq MA, Khan V, et al (2014). Biallelic truncating mutations in FMN2, encoding the actin-regulating protein Formin 2, cause nonsyndromic autosomal recessive intellectual disability. Am J Hum Genet. 2014 Dec 495(6):721-8.
43. Lenski C, Abidi F, Meindl A, Gibson A, Platzer M, Frank Kooy R, Lubs HA, Stevenson RE, Ramser J, Schwartz CE. Novel truncating mutations in the polyglutamine tract binding protein 1 gene (PQBP 1) cause renpenning syndrome and X-linked mental retardation in another family with microcephaly. Am J Hum Genet. 2004; 74:777-80.
44. Lisik MZ, Sieron AL. X-Linked mental retardation. Med Sci Monit. 2008; 14; RA 221-29.
45. Lovell CS, Davis WI, Arendall BW, Bakker WIP, Word MJ, Prisant GM, Richardson SJ, Richardson CD, Structure validation by Calpha geometry: phi,psi and Cbeta deviation.*Proteins: Structure, Function & Genetics* 2002, 50: 437-450.
46. Majnemer, A. (2012). Measures for Children with Developmental Disabilities: An ICF-CY Approach.(London, UK: Mac Keith Press).
47. 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(2); 419-436.
48. McGuffin LJ, Bryson K, Jones DT: The PSIPRED protein structure prediction server.Bioinformatics 2000, 4: 404-5.
49. Mehdi S. Q. R. Qamar Q. Ayub S. Khaliq A. Mansoor M. Ismail M. F. Hammer P. A. UnderhillL. L. Cavalli-Sforza. The Origins of Pakistani Populations, Evidence from Y chromosome Markers: Genomic Diversity: Applications in Human Population Genetics. 1999; 83-90.
50. Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, & Naeem F. Identification of mutations in TRAPPC9, which encodes the NIK-and IKK-β-

binding protein, in nonsyndromic autosomal-recessive mental retardation. The American Journal of Human Genetics. 2009, 85(6), 909-915.

51. Mir A, Sritharan K, Mittal K, Vasli N, Araujo C, Jamil T,& Mahmood H. Truncation of the E3 ubiquitin ligase component FBXO31 causes non-syndromic autosomal recessive intellectual disability in a Pakistani family. Human genetics. 2014, 133(8), 975-984.
52. Molinari F, Foulquier F, Tarpey PS, Morelle W, Boissel S, Teague J, Edkins S, Futreal PA, Stratton MR, Turner G, Matthijs G, Gecz J, Munnich A and Colleaux L. Oligosaccharyltransferase-subunit mutations in nonsyndromic mental retardation. Am. J. Hum. Genet. 2008, 82: 1150-1157.
53. Molinari F, Rio M, Meskenaite V, Encha-Razavi F, Auge J, Bacq D, Briault S, Vekemans M, Munnich A, Attie-Bitach T, Sonderegger P and Colleaux L. Truncating neurotrypsin mutation in autosomal recessive nonsyndromic mental retardation. Science. 2002, 298: 1779–1781.
54. Musante L, Ropers HH. Genetics of recessive cognitive disorders: Trends in Genetics. 2014, 30(1): 32-39.
55. Opitz JM. Historiography of the causal analysis of mental retardation. Speech to the 29th Annual Gatlinburg Conference on Research and Theory in Mental Retardation, Gatlinburg, TN. 1996
56. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12.
57. Philips, A. K., Pinelli, M., de Bie, C. I., Mustonen, A., Maatta, T., Arts, H. H., Wu, K., Roepman, R., Moilanen, J. S., Raza, S., Varilo, T., Scala, G., Cocozza, S., Gilissen, C., van Gassen, K. L. I., Jarvela, I. Identification of C12orf4 as a gene for autosomal recessive intellectual disability. Clin. Genet. 2017, 91: 100-105.
58. Rafiq MA, Kuss AW, Puettmann L, Noor A, Ramiah A, Ali G& Khan MA. Mutations in the alpha 1, 2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. The American Journal of Human Genetics. 2011, 89(1), 176-182.

59. Rafullah R, Aslamkhan M, Paramasivam N, Thiel C, Mustafa G, Wiemann S, Schlesner M. Wade RC, Rappold GA, Berkel S. Homozygous missense mutation in the LMAN2L gene segregates with intellectual disability in large consanguineous Pakistani family. J Med Genet. 2016 Feb;53(2): 138-44.

60. Rauch, A., Hoyer, J., Guth, S., Zweier, C., Kraus, C., Becker, C., Zenker, M., Huffmeier, U., Thiel, C., Ruschendorf, F., *et al.* Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. Am J Med Genet. 2006, A 140, 2063-2074.

61. Reif R, Sales S, Hettwer S, Dreier B, Gisler C, Wolfel J, Luscher D, Zurlinden A, Stephan A, Ahmed S, Baici A, Ledermann B, Kunz B and Sonderegger P. Specific cleavage of agrin by neurotrypsin, a synaptic protease linked to mental retardation. FASEB. J. 2007, 21: 3468–3478.

62. Roy A, Kucukural A, Zhang Y: I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 2010, 5, p. 725.

63. Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

64. Sherr, E.H., Michelson, D.J., Shevell, M.I., Moeschler, J.B., Gropman, A.L., and Ashwal, S. Neurodevelopmental disorders and genetic testing: current approaches and future advances. Ann Neurol. 2013, 74, 164-170.

65. Shevell, M. I. Office evaluation of the child with developmental delay. Seminars in pediatric neurology. 2006, 13, 256-261.

66. Shevell, M.I. Diagnostic approach to developmental delay. In Current management in child neurology, B. Maria, ed. (Shelton, Connecticut, BC Decker Inc.). 2009, pp 292-299.

67. Shevell, M.I. Present conceptualization of early childhood neurodevelopmental disabilities. Journal of child neurology. 2010, 25, 120-126.

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

69. Tinel, A., Tschopp, J. The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. Science. 2004, 304: 843-846.

70. Topper S, Ober C, & Das S. Exome sequencing and the genetics of intellectual disability. Clinical genetics. 2011, 80(2), 117-126.

71. Uyguner O, Kayseroli H, Li Y, Karaman B, Nurnberg G, Hennies HC, Becker C, Nurenberg P, Basan S, Apak MY and Wollnik B. A new locus for autosomal recessive non-syndromic mental retardation maps to 1p21.1-p13.3. Clin. Genet. 2007, 71: 212-219.

72. van Bokhoven, H. Genetic and epigenetic networks in intellectual disabilities. Annual review of genetics. 2011, 45, 81-104.

73. Vissers Lisenka E. L. M , Christian Gilissen and Joris A. Veltman. Genetic studies in intellectual disability and related disorders: Nature Reviews Genetics. 2016. 17: 9-18.

74. Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Glaser D, Liebscher v. Meitinger T, Strom TM. Copy-number variations measured by single-nucleotide polymorphism oligonucleotide arrays in patients with mental retardation. Am J Hum Genet. 2007; 81;768-779.

75. Wang s, Ren H, Xu J, Yu Y, Han S, Qiao H, Cheng S, Xu C, An S, Ju B, Yu C. Diminished serum repetin levels in patients with schizophrenia and bipolar disorder. Scientific Report. 2015; 5: 7975.

76. Winnepenninckx B, Rooms L, Kooy RF. Mental retardation: A review of the genetic causes, Brit J Develop Disab, 2003: 49:29-44.

77. Zhang Y: I-TASSER server for protein 3D structure prediction.*BMC* Bioinformatics 2008, 9: 40.

Research Papers published from this Thesis

ISSN: 1971-8571 Vol 13, No. 2 December 2021

International Journal of Computational Intelligence in Control

Whole Exome Analysis in Consanguineous Pakistani Families Determined ROR2 and RPTN as Novel Candidate Genes to be involved in Autosomal Recessive Non-Syndromic Intellectual Disability

Iftikhar Ahmed[1,2], Muhammad Muzammal[1], Sumra Wajid Abbasi[3], Muzammil Ahmad Khan[2], Asif Mir[1]

[1]Department of Biological Sciences, International Islamic University, Islamabad Pakistan.
[2]Gomal Centre of Biochemistry and Biotechnology, Gomal University, Dera Ismail Khan, 29050, Khyber Pakhtunkhwa, Pakistan.
[3]Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Punjab, Pakistan.

#Correspondence should be addressed to:-
Dr. Asif Mir, Associate Professor, Department of Biological Sciences, International Islamic University, Islamabad Pakistan.
Email ID: Asif.mir@iiu.edu.pk

Running Head: Novel candidate gene involved in Non-Syndromic Intellectual Disability

Statements and Declarations

-Funding

The current study was supported through an HEC Pakistan, funded by International Research Support Initiative Program (IRSIP) awarded to Iftikhar Ahmed.

-Conflict of interest/ Competing interests

None declare by all authors

-Data availability Statement:

The reference sequence data was obtained from UCSC genome browser (http://genome.ucsc.edu/). The patient's data (sequence, photographs, and pedigrees) is stored in the password-protected computer of Lab and is available upon request.

-Acknowledgment.

We are grateful to the volunteer family for their valuable participation in the present biochemical genetics study. The current data has not been published anywhere, except presented in the M.Phil. and Ph.D. Thesis of few students that are already on-board in this manuscript.

-Author's contribution.

All the authors have read, edited and approve the final version of manuscript. Experiments and data analysis (Iftikhar Ahmed and Asif Mir), *In silico* analysis (Muhammad Muzammal,

The institutional ethical review board of International Islamic University Islamabad, Pakistan (No: BI&BT)/FBAS-2018-3591), approved the present molecular study.

-Consent to Publish

The patient's guardians have given their consent to publish their clinical information and photographs.

Whole Exome Analysis in Consanguineous Pakistani Families Determined ROR2 and RPTN as Novel Candidate Genes to be involved in Autosomal Recessive Non-Syndromic Intellectual Disability

Iftikhar Ahmed[1,2], Muhammad Muzammal[1], Sumra Wajid Abbasi[3], Muzammil Ahmad Khan[2], Asif Mir[1].

1. Department of Biological Sciences, International Islamic University, Islamabad Pakistan.
2. Gomal Centre of Biochemistry and Biotechnology, Gomal University, Dera Ismail Khan, 29050, Khyber Pakhtunkhwa, Pakistan.
3. Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Punjab, Pakistan.

#Correspondence should be addressed to:-
Dr. Asif Mir, Associate Professor, Department of Biological Sciences, International Islamic University, Islamabad Pakistan.

...cy Families Page 126

Pak J Med Res
Vol. 60, No. 2, 2021

Original Article

A Novel Hemizygous Variant in the *AFF2* Gene Causing Fragile XE (FRAXE) Syndrome: First Report from Pakistan

Iftikhar Ahmed[1], Muhammad Ilyas[1], Gaurav V Harlalka[2,3], Asif Mir[1]
Department of Biological Science, International Islamic University, H-10, Islamabad, College of Medicine and Health, University of Exeter[2], Exeter, Devon, UK, Department of Pharmacology, Rajarshi Shahu College of Pharmacy[3], Malvihir, Buldana, Maharashtra, India.

Abstract

Background: Fragile XE (FRAXE) is an X-linked recessive condition that affects 1 in 50,000 of new born males with intellectual disability (ID). It is characterized by mild intellectual disability (ID), speech delay cognitive impairment, and in some cases with phenotypes of Autism Spectrum disorder (ASD).
Methodology: In this study, a family was investigated with two male siblings having neuro developmental delay. Whole exome sequencing analysis (WES) was carried out to identify the pathogenic variant. Sanger sequencing was performed in normal and affected family members and co-segregation analysis was done.
Results: Two probands were affected in a family diagnosed with intellectual disability. A novel hemizygous variant (c.3348G>T, p.Asp1150Tyr) in *AFF2* gene was identified as the causal variant cause in affected individuals. This variant was novel from Pakistani population.
Conclusion: In this study, a novel hemizygous variant (c.3348G>T, p.Asp1150Tyr) identified in *AFF2*. These findings paved the way for further studies on genetic and clinical spectrum of rare X-linked recessive disease involved in ID.
Key words: *AFF2*, hemizygous, intellectual disability, neurological disorders.

Introduction

Cognition is the result of cellular, biological and multiple molecular events in the nervous system. Minor defect in any of these events can results in intellectual disability or cognitive impairment.[1,2] It can also be termed as neuro-development disorder as it results from defect in synapse formation.[3,4]

Overall, in general population, the prevalence of intellectual disability (ID) is 2-3%.[5,6] Its clinical and genetic heterogeneity make the diagnosis challenging for scientists and physicians.[7,8] More than 900 genes are reported to cause intellectual disability so far.[9,10]

The X-linked intellectual disability is a heterogeneous group of genetic disorders. There are more than 141 genes linked with disease located on x-chromosome.[11] Its prevalence is high in males as x-chromosome genes contribute to cognition. Among x-linked intellectual disability, Fragile X syndrome (FXS) is common and characterized by moderate to severe disability (OMIM - 309548).[12] There is high repeat expansion of a CGG in *FMR1* gene which cause methylation, that ultimately halts the production of *FMR1* protein leads to (FXS) syndrome.[13]

The *AFF2* gene (also as *FMR2* gene) cause non-specific x-lined intellectual disability with prevalence as 1/25,000 to 100,000 in new born male. Micro deletion in *AFF2* genes leads to Fragile XE (FRAXE) syndrome [14]. It is characterized by mild to moderate intellectual disability.[15]

In FRAXE syndrome, learning, thinking ability and cognitive function, affected badly. Also there is delay in speech; hyperactivity, poor writing skills, and very short attention span are common symptoms of people affected with this syndrome. It

Corresponding Author:
Asif Mir
Department of Biological Sciences
International Islamic University, Islamabad.
Email: asif.mir@iiu.edu.pk

Received: 13 September 2020, Accepted: 30 April 2021,
Published: 08 July 2021

Authors Contribution

IA, GVH & AM conceptualized the project and drafting, revision & writing of manuscript. IA & MI did the literature search and data collection. IA, MI & AM performed the statistical analysis.

turnitin **Turnitin Originality Report**

PhD Thesis by Iftikhar Ahmed

From Thesis MS (Thesis)

Processed on 26-Feb-2022 21:36 PKT
ID: 1771416951
Word Count: 16408

Similarity Index	Similarity by Source	
17%	Internet Sources:	12%
	Publications:	7%
	Student Papers:	8%

sources:

1 3% match (Internet from 14-Oct-2020)
https://www.researchsquare.com/article/rs-42138/v1

2 1% match (student papers from 12-May-2015)
Class: Quick Submit
Assignment: Quick Submit
Paper ID: 540788865

3 1% match (student papers from 05-Mar-2017)
Submitted to Higher Education Commission Pakistan on 2017-03-05

4 1% match (student papers from 18-Dec-2016)
Submitted to Higher Education Commission Pakistan on 2016-12-18

5 1% match (student papers from 16-Jul-2018)
Submitted to Higher Education Commission Pakistan on 2018-07-16

6 1% match (Internet from 21-Mar-2017)
http://www.oatext.com/Insight-to-FBXO31-Novel-Mutation-p.Cys283Asn-causing-Non%E2%80%91Syndromic-Autosomal-Recessive-Intellectual-Disability-using-Computational-Methods.php

7 1% match (Internet from 06-Mar-2020)
http://irigs.iiu.edu.pk:64447/gsdl/collect/00electron/tmp/T14807E2015BIMS.html

8 1% match (Internet from 20-Sep-2017)
http://prr.hec.gov.pk/Thesis/310S.pdf

9 1% match (publications)
Myriam Srour, Afnan AlHakeem, Michael Shevell. "Global developmental delay and intellectual disability", Elsevier BV, 2020

10 1% match (Internet from 02-Dec-2020)
https://jbiomedsci.biomedcentral.com/articles/10.1186/s12929-018-0481-x

11 < 1% match (student papers from 11-Sep-2015)
Submitted to Higher Education Commission Pakistan on 2015-09-11

12 < 1% match (student papers from 07-Mar-2017)
Class: Quick Submit
Assignment: Quick Submit
Paper ID: 781168822

< 1% match (student papers from 11-May-2011)

13 Submitted to Higher Education Commission Pakistan on 2011-05-11

14 < 1% match (student papers from 29-May-2017)
Submitted to Higher Education Commission Pakistan on 2017-05-29

15 < 1% match (student papers from 04-Jul-2012)
Submitted to Higher Education Commission Pakistan on 2012-07-04

16 < 1% match (student papers from 07-Jan-2016)
Submitted to Higher Education Commission Pakistan on 2016-01-07

17 < 1% match (student papers from 05-Apr-2010)
Submitted to Higher Education Commission Pakistan on 2010-04-05

18 < 1% match (student papers from 10-May-2013)
Submitted to Higher Education Commission Pakistan on 2013-05-10

19 < 1% match (student papers from 02-Feb-2012)
Submitted to Higher Education Commission Pakistan on 2012-02-02

20 < 1% match (student papers from 28-Aug-2020)
Submitted to Higher Education Commission Pakistan on 2020-08-28

21 < 1% match (student papers from 10-Nov-2009)
Submitted to Higher Education Commission Pakistan on 2009-11-10

22 < 1% match (student papers from 19-Dec-2016)
Submitted to Higher Education Commission Pakistan on 2016-12-19

23 < 1% match (student papers from 28-Jun-2015)
Submitted to Higher Education Commission Pakistan on 2015-06-28

24 < 1% match (student papers from 27-May-2021)
Submitted to Higher Education Commission Pakistan on 2021-05-27

25 < 1% match (student papers from 21-Dec-2009)
Submitted to Higher Education Commission Pakistan on 2009-12-21

26 < 1% match (student papers from 09-Apr-2010)
Submitted to Higher Education Commission Pakistan on 2010-04-09

27 < 1% match (student papers from 30-Sep-2010)
Submitted to Higher Education Commission Pakistan on 2010-09-30

28 < 1% match (student papers from 10-Jan-2011)
Submitted to Higher Education Commission Pakistan on 2011-01-10

29 < 1% match (student papers from 06-Sep-2012)
Submitted to Higher Education Commission Pakistan on 2012-09-06

30 < 1% match (student papers from 28-Jun-2013)

Submitted to Higher Education Commission Pakistan on 2013-06-28

31 < 1% match (student papers from 17-Nov-2009)
Submitted to Higher Education Commission Pakistan on 2009-11-17

32 < 1% match ()
Gaurav V. Harlalka, Anna Lehman, Barry Chioza, Emma L. Baple et al. "Mutations in (GM2 synthase) underlie a new disorder of ganglioside biosynthesis", Brain

33 < 1% match ()
Shuai Wang, Huixun Ren, Jie Xu, Yanjun Yu et al. "Diminished serum repetin levels in patients with schizophrenia and bipolar disorder", Scientific Reports

34 < 1% match ()
Maxime Sasseville, Nancy Côté, Christine Guillemette, François J Richard. "New insight into the role of phosphodiesterase 3A in porcine oocyte maturation", BioMed Central

35 < 1% match ()
Neil V Morgan, Bryndis Yngvadottir, Mary O'Driscoll, Graeme R Clark et al. "Evidence that autosomal recessive spastic cerebral palsy-1 (CPSQ1) is caused by a missense variant in ", Brain Communications

36 < 1% match ()
Siying Lin, Gaurav V. Harlalka, Abdul Hameed, Hadia Moattar Reham et al. "Novel mutations in associated with autosomal recessive anophthalmia/microphthalmia, and review of the literature", BMC Medical Genetics

37 < 1% match (student papers from 24-Jan-2012)
Submitted to University of Newcastle upon Tyne on 2012-01-24

38 < 1% match (Internet from 24-May-2019)
http://www.pragatipublication.com/assets/uploads/doc/b25c3-612-622.13602.pdf

39 < 1% match (Internet from 04-Jan-2020)
https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awz382/5679762?searchresult=1

40 < 1% match (Internet from 04-Apr-2021)
https://academic.oup.com/brain/article/140/4/940/3058658

41 < 1% match (publications)
M. Srour, M. Shevell. "Genetics and the investigation of developmental delay/intellectual disability", Archives of Disease in Childhood, 2013

42 < 1% match (publications)
Anita Rauch. "Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation", American Journal of Medical Genetics Part A, 10/01/2006

43 < 1% match (Internet from 11-Mar-2021)
https://archive.org/stream/aet2858.0001.001.umich.edu/aet2858.0001.001.umich.edu_djvu.txt

44 < 1% match (Internet from 27-Dec-2020)
https://archive.org/stream/in.ernet.dli.2015.206067/2015.206067.The-Indian_djvu.txt