Molecular Analysis of Lysosomal Storage Disorders in Pakistani Population

By

Rutaba Gul

Department of Zoology

Faculty of Biological Sciences

Quaid-I-Azam University, Islamabad, Pakistan

2023

Molecular Analysis of Lysosomal Storage Disorders in Pakistani Population

A dissertation in the partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Zoology

by

Rutaba Gul

Department of Zoology

Faculty of Biological Sciences

Quaid-I-Azam University, Islamabad, Pakistan

2023

In the name of Allah Almighty, Most Gracious, Most Merciful

DEDICATED TO:

THE HOLY PROPHET HAZRAT MUHAMMAD (PBUH) AND HIS FAMILY WHO ENLIGHTENED OUR SOULS WITH THE ESSENCE OF FAITH IN ALLAH

&

ARE SOURCE OF GUIDANCE AND KNOWLEDGE FOR HUMANITY AS A WHOLE.

Author's Declaration

I **Ms. Rutaba Gul** hereby state that my PhD thesis titled "Molecular Genetic Analysis of Lysosomal Storage Disorders from Pakistani Population" is my own work and has not been submitted previously by me for taking any degree from Quaid-i-Azam University, Islamabad, Pakistan.

At any time if my statement is found to be incorrect even after my Graduate the University has the right to withdraw my Ph.D. degree.

Ms. Rutaba Gul

Date: 04.01.2024

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis "Molecular Genetic Analysis of Lysosomal Storage Disorders from Pakistani Population" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Quaid-i-Azam University towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/ cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of Ph.D degree, HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Student / Author Signature: 

Name: **Ms. Rutaba Gul**

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled "Molecular Genetic Analysis of Lysosomal Storage Disorders from Pakistani Population" was conducted by **Ms. Rutaba Gul** under the supervision of **Dr. Sabika Firasat**. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Department of Zoology of Quaid-i-Azam University, Islamabad in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Field of Molecular Biology.

Student Name: **Ms. Rutaba Gul** Signature:_____

Examination Committee:

a) External Examiner 1: Signature:_____

Dr. Mazhar Qayyum
Professor (Retd.)
Department of Zoology, Wildlife and Fisheries, PMAS Arid Agriculture University,
Rawalpindi

b) External Examiner 2: Signature:_____

Dr. Shamim Akhter
Professor and Chairperson
Department of Zoology, Wildlife and Fisheries, PMAS Arid Agriculture University,
Rawalpindi

Supervisor Name: **Dr. Sabika Firasat** Signature:_____

Name of HOD: **Prof. Dr. Amina Zuberi** Signature:_____

Date: 04.01.2024

CHAIRPERSON
Department of Zoology
Quaid-i-Azam University

DECLARATION

I, hereby declare that the work presented in this thesis is the result of my own effort and research work, carried out in Molecular Biology Lab, Department of Zoology, Quaid I Azam University Islamabad, Pakistan. This thesis is my own composition and no part of it has been presented for any degree previously, nor does it contain, without proper acknowledgement or reference, any material from the published resources to the best of my knowledge.

Rutaba Gul

CONTENTS

CHAPTER 1: Introduction

Chapter 2: Materials and Methods

Chapter 3: Results

List of Figures

List of Tables

Annexures

LIST OF ABBREVIATIONS

(NH4)2SO4	Ammonium sulfate
°C	Degree centigrade
µL	Micro liter
3MC	Malpuech-Michels-Mingarelli-Carnevale syndrome
6MWT	6 minute walk test
∞	Infinity
A	Amperes
A	Adenine
ABC	ATP binding cassette
ACE	Angiotensin converting enzyme
ACMG	American college for medical genetics
AFR	Africa
AJK	Azad jammu and Kashmir
Ala	Alanine
ALT	Alanine aminotransferase
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
AST	Aspartate triaminase
ATP	Adenosine triphosphate
AV/VA	Articuloventricle/Ventriculoarticlular
BBB	Blood brain barrier
BD	Becton Dickinson

BEACH	Beige and Chediak-Higashi domain
BEC	Bioethical committee
BLAT	Blast like alignment tool
Blood CP	Blood complete picture
BMT	Bone marrow transplant
bp	Base pair
BP	Blood pressure
BS	Strongly benign
C	Cytosine
C6S	Chondroitin 6 sulfate
CBC	Complete blood count
CHS	Chediak Higashi syndrome
CK	Creatin kinase
cm	Centimeter
CNS	Central nervous system
COVID-19	Corona virus disease of 2019
Cys	Cystine
DBS	Dry blood spot
DNA	Deoxyribo nucleic acid
dNTP	Deoxynucleoside triphosphates
EAS	East asian ancestry
EC	Enzyme code
ECG	Electro cardiogram
EDTA	Ethylene diamine tetra acetate

EJC	Exon exon junction complex
ERT	Enzyme replacement therapy
ESR	Erythrocyte sedimentation rate
EUR	European
F	Female
fathmm	Functional analysis through hidden markov models
FBS	Faculty of biological sciences
FGF	Fibroblast growth factors
fl	Femto liter
fs	Frame shift
G	Guanine
g/dL	Gram per deciliter
GAGs	Glycosamino glycans
GD	Gaucher diseases
gDNA	Genomic DNA
GEF	Guanine nucleotide exchange factor
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GMAF	Global minor allele frequency
gnomAD	Genome aggregation database
GS	Griscelli syndrome
GTP	Guanosine triphosphate

HCl	Hydrochloric acid
HCT	Hematocrit
HEC	Higher education commission
HGMD	Human genome mutation database
His	Histidine
HLH	Hemophagocytic lymphohistiocytosis
HOPE	Have y/our protein explained
HPF	High power field
HPLC	High performance liquid chromatography
HS	Heparan sulfate
HSCT	Hematopoitic stem cell transplant
ICT	Islamabad capital territory
ID	Intellectual disability
IEM	Inborn errors of metabolism
IGF1	Insuline like growth factor
Ile	Isoleucine
IOPD	Infantile onset pompe disease
IQ	Intelligent quotient
IRSIP	International research support initiative program
I-TASSER	Iterative Threading ASSEmbly Refinement
IUERT	Intra uterine enzyme replacement therapy
kb	Kilobase
Kg	Kilogram
KPK	Khyber Pakhtunkhwa

KS	Keratin sulfate
LCMS	Liquid chromatography mass spectrometry
Leu	Leucine
LOF	Loss of function
LROs	Lysosome related organelle
LSD	Lysosomal storage disorder
Lys	Lysine
m	Months
M6P	Mannose 6-phosphate
M6PR	Mannose 6-phosphate receptor
MAF	Minor allele frequency
MCH	Mean corpuscle hemoglobin
MCHC	Means corpuscle hemoglobin concentration
MCV	Mean corpuscle volume
Met	Methionine
mg/dL	Mili gram per deciliter
MgCl	Magnesium chloride
mL	Milliliter
Mm	Millimeter
mm/hr	Millimeter per hour
mmol/L	Millimole per liter
MRI	Magnetic rosanance imaging
mRNA	Messenger ribonucleic acid
MSUD	Maple syrup urine disease

NaCl	Sodium chloride
NFκB	Nuclear factor kappa B
ng/mL	Nanogram per microliter
NGS	Next generation sequencing
NMD	Non sense mediated decay
OMIM	Online mendelian inheritance in man
OSA	Obstructive sleep apnea
PANTHER	Protein analysis through evolutionary relationships
PCR	Polymerase chain reaction
PCT	Premature termination codon
PE	Paired end
pg	Picograms
pH	Potential of hydrogen
PHD-SNP	Predictor of human deleterious single nucleotide polymorphisms
Phe	Phenylalanine
PJL	Punjabi population
PK	Protein kinase
PM	Moderate pathogenic
PMut	Python module
POLYPHEN-2	Polymorphism Phenotyping v2
PP	Probably pathogenic
Pro	Proline
PROVEAN	Protein variation effect analyzer

PS	Strongly pathogenic
PTC	Premature termination codon
PVS	Very strong pathogenic
RBC	Red blood cells
RBG	Randome blood glucose
RDW-CV	Red cell distribution width
RNA	Ribonucleic acid
ROH	Runs of homozygosity
RPM	Rounds per minute
SapC-DOPS	Sphingolipid activator protein C and dioleylphosphatidylserine
SARS-CoV-2	Severe acute respiratory syndrome coronavirus
SAS	South Asian superpopulation
SDS	Sodium dodecyle sulfate
Ser	Serine
SIFT	Sorting intolerant from tolerant
SNP&GO	Single nucleotide polymorphisms & gene ontology
SNPs	Single nucleotide polymorphisms
SRT	Substrate reduction therapy
STRING	Search tool for the retrieval of interacting genes/proteins
SUMF1	sulfatase modifying factor 1
T	Thymine
Taq	Thermus aquaticus
TBE	Tris boric acid EDTA
TE	Tris EDTA

Thr	Threonine
TIM	Triosephosphate isomerase
TLC	Thin layer chromatography
Tm	Melting temperature
TMS	Tendem mass spectrometry
TNE	Tris NaCl EDTA
Trp	Tryptophan
Tyr	Tyrosine
U/L	Unit per liter
UDP	Uridine diphosphate
UTRs	Untranslated regions
V	Volts
Val	Valine
WBC	White blood cells
WD-40	Tryptophan aspartic acid 40 repeats domain
WGS	Whole genome sequencing
WT	Wild type
y	Years

ACKNOWLEDGEMENT

I have no words to express my deepest words of gratitude to **Almighty Allah**, whose blessings can't be even counted. He is the most compassionate and merciful, who knows what was there in the universe, what is and what will be. He has enabled me to elucidate a drop from existing ocean of knowledge. I offer my countless salutations to the **Holy Prophet Muhammad (PBUH)**, who enabled me to recognize our creator and whose moral and spiritual teaching enlightened my heart and flourished my thoughts to achieve higher ideas of life.

My deepest gratitude is to my supervisor**, Dr. Sabika Firasat**, Associate Professor, Department of Zoology, Faculty of Biological sciences, Quaid-i-Azam University Islamabad Pakistan whose expertise, understanding, and patience, added considerably to my research experience. I am thankful to her for her inspiration, reassurance, and counseling, especially for creating and maintaining a friendly environment throughout my study. I have been amazingly fortunate to have an advisor like her.

Huge lap of gratitude to the **Higher Education Commission, Islamabad, Pakistan** for providing me the opportunity in the form of **International Research Support Initiative Program (IRSIP)** to complete my research project with high impact.

I am grateful to **Prof. Dr. Torben Hansen**, Human physiology and Metagenomics group, Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark for providing me an opportunity to carryout whole genome sequencing and NGS data analysis in his lab without which this work would have not been possible. I am also thankful to **Mikkel Schubert, Asmat Ullah, Elionora Peña Lozano, Anne Catherina Baun Thuesen, Tina Hvidtfeldt Lorentzen, Anette Marianne Prior Gjesing, Lise Ryborg** and CBMR for helping me with sample preparation, data analysis and other technical assistance. I also want to expand my gratitude to whole Hansen group, UCPH members for helping me in a way or the other to make my stay memorable at Copenhagen.

I owe an immense debt of gratitude to **Dr. Mulazim Hussain**, Pediatrician, Childrens Hospital, Pakistan Institute of Medical Sciences, Islamabad, for his tiring efforts for patients recruitments, clinical evaluation and sample collection. Without him this study would not be so easy.

My heartiest obligations to all the participating families and their immense cooperation for sampling is the key of success for this study.

I feel it obligatory to offer thanks to my lab fellows **Muhammad Tufail, Fatima Akbar, Kiran Fatima, Samra Amir, Bushra Gul** and **Dil Nawaz** for their cooperation and moral support.

Many friends have helped me stay sane through this time. I want to specially mention in **Hina Ahmed Nawaz**, **Ramsha Mehboob, Dr. Saman Illyas**, **Hafsa Illyas, Dr. Amna Abdul Rahim, Dr. Sara Latif** who were there whenever I need them.

I would like to thanks all of the technical staff of the department of Zoology during the conduct of this research work. Specially like to mention **Mr. Anwar** and **Mr. Naeem** for always been there to help.

Most importantly, none of this would have been possible without the love and patience of my Family. I would like to express my heart-felt gratitude to **my Father Mirza Khan** and **my sweetest Mother Kaneez Fatima, my adorable sister-in-laws Azizah Razak** and **Sarah Noor, my supporting pillars my brothers Muhammad Umar Khan, Muhammad Usman Khan and Adnan Munir Khan, my fresh pill my nephew M. Farzam Khan and my aunt Fakhr un Nisa,.** Without your help and support, I can't be able to do this whole work and write this thesis or even complete this thesis. I would also like to mention my roommates **Naila Akbar** and **Nida Nayab** for bearing my frustrations and support me like sisters. I would also like to expand my gratitude to **Fatima Shehzad, Gul Abeer Khan, Dr. Omar Afzal, Dr. Syed Farhat, Rabail Kiani** and **Dr. Kiran Afshan.**

And last but not least, I have no words to express my feelings and my gratitude to the **APK** whose limitless moral support, affection and care are actually the main pillars of support to completion of this whole study. I feel so much lucky to have you by my side in every up and down of my research and my life.

Finally, I would like to thank everybody who was important to the successful completion of this thesis, as well as expressing my apology that I could not mention personally one by one.

RUTABA GUL

List of Publications

The research work presented in the current dissertation resulted in the publication of the following articles.

1. **Gul, R**., Firasat, S., Hussain, M., Afshan, K., & Nawaz, D. (2019). *IDUA* gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families. Congenit. Anom, 60, 1-2.
2. **Gul, R**., Firasat, S., Hussain, M., Tufail, M., Ahmad, W., & Afshan, K. (2021). Neurological manifestations in Pakistani lysosomal storage disorders patients and molecular characterization of Gaucher disease. Genetika, 53(3), 1017-1029.
3. **Gul, R**., Firasat, S., Hussain, M., Akbar, F., Fatima, K., Sabar, F., Afshan, K. Mutations in the gene encoding Alpha-L-Iduronidase identified in Mucopolysaccharidosis-I segregating families from Pakistan. *Pakistan Journal of Zoology*. Under-review (April 2022).
4. **Gul, R.,** Firasat, S., Shubert, M., Ullah, A., Pena, E., Thuesen Anne C.B., Hussain, M., Stæger, Frederik F., Gjesing, Anette P., Albrechtsen, A., Hansen, T. Identifying the Genetic Causes of Phenotypically Diagnosed Pakistani Mucopolysaccharidoses Patients by Whole Genome Sequencing. *Frontiers in Genetics*. Revision submitted (29-January 2023).

ABSTRACT

Inborn errors of metabolism (IEM) are considered previously as group of disorders caused by the metabolic pathways' enzymes deficiency; however, recent classification precisely label them as any disorder with biochemical pathway impairment. Among IEM, Lysosomal Storage Disorders (LSDs) are a vast group, each subtype of which is caused by genetic mutation/s of gene/s encoding lysosomal protein. Most LSDs are inherited in an autosomal recessive manner thus, current study was performed to explore molecular genetics of LSDs in affected families recruited from consanguineous Pakistani population. The study was approved from Bioethical Committee of Quaid-i-Azam University, Islamabad and followed the rules mentioned in Declaration of Helsinki.

A total of 45 families have been recruited from different regions of Pakistan affected with various sub-types of LSDs. After DNA extraction from blood samples, initially nine families affected with Mucopolysaccharidosis type I (MPS I) and Gaucher disease were put through Sanger's sequencing of all coding exons, intron-exon boundaries and un-translated regions of *IDUA* and *GBA* gene respectively to identify disease causing variants. Data analysis of *IDUA* regions revealed segregation of a disease causing variant in each of MPS-1 affected family including a prevalent missense variant p.Leu490Pro segregating with disease phenotype in three un-related families, a single nucleotide deletion p.His262Thr*fs**55, a 14 nucleotide deletion p.Asn190His*fs**204 and a non-sense variant p.Glu486Ter. A previously reported disease causing missense variant p.Leu483Pro of *GBA* gene was identified in all three Gaucher disease segregating families. Remaining thirty-six families were subjected to Whole Genome Sequencing (WGS) at Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark, during a six month research stay under International Research Support Initiative Program (IRSIP) funded by the Higher Education Commission, Islamabad, Pakistan. WGS data analysis identified pathogenic variants in five MPS-1 affected families in *IDUA* gene, seven MPS-IV type A affected families in *GALNS* gene, five MPS-III type A affected families in *SGSH* gene, one MPS-III type B affected family in *NAGLU* gene, one Pompe disease affected family in *GAA* gene, three MPS-II affected families in *IDS* gene, one Hereditary Fructose Intolerance affected family in *ALDOB* gene, one

Intellectual Disability affected family in *TRAPPC4* gene, one Mingarelli, Malpeuch, Michels and Carnevale (3MC) syndromes affected family in *MASP1* gene, one Schimke-Immuno Osseous Dysplasia affected family in *SMARCAL1* gene, one Alkuraya-Kucinskas Syndrome affected family in *KIAA1109* gene, one Intellectual Disability affected family in *HERC1* gene, one Noonan Syndrome affected family in *RRAS2* gene, one Cerebral Ataxia affected family in *VWA3B* gene, one Mucolipidosis II/III affected family in *GNPTG* gene, one Biotinidase affected family in *BTD* gene, three Niemann-Pick disease type C affected families in *NPC1* gene, two Niemann-Pick disease type A/B affected families in *SMPD1* gene, one Pycnodysostosis affected family in *CTSK* gene, one Griscelli Syndrome type 2 affected family in *RAB27A* gene and a novel candidate gene (*ABCA5*) for LSD-like phenotypes. Identification of *ABCA5* mutation segregating with disease phenotype in this study is the first report of *ABCA5* involvement in LSD in humans. *ABCA5* has previously been associated with symptoms strongly related with LSD in animal models. Pathogenicity of all identified disease-causing variants was also checked using different *In-Silico* tools.

In conclusion, out of 45 families, each of which was initially recruited as affected with a subtype of LSD based on clinical protocol followed by clinicians, forty-one families were diagnosed at molecular basis. The remaining four families were not solved due to poor quality of the available DNA samples for WGS and unfortunately re-sampling is not possible due to the death of the patients. Among forty-one solved families, LSDs were confirmed in thirty-two families and based on molecular results four families were affected with genetic disorders other than LSDs e.g., Hereditary Fructose Intolerance, Intellectual Disability, Schimke Immuno Osseous Dysplasia and Cerebral Ataxia. An interesting finding was that one family was inheriting two LSDs i.e., MPS III type A and Pompe disease, furthermore, five families were inheriting more than one genetic disorder. This study highlights the importance of searching for homozygous pathogenic variants through WGS in highly inbred populations like our Pakistani population and may help in reducing the burden of early morbidity and mortality caused by these rare disorders by paving the path for future diagnostics.

1. INTRODUCTION

Inborn errors of metabolism (IEM) are considered previously as class of disorders caused by the deficiency of enzymes involved in metabolic pathways; however, recent studies precisely label them as any disorder with biochemical pathway impairment as inborn errors of metabolism (El-Hattab, 2015; Ferreira *et al.,* 2019). Inborn errors of metabolism have been categorized into three types according to their pathophysiology: first include intermediate metabolic pathway causing toxication adjacent to metabolic blockage examples include defects of urea cycle and maple syrup urine disease (MSUD), second leads to deficiency of energy for example mitochondrial respiratory chain defects and third type results in obstruction of anabolism or catabolic process of complex molecules in various cell organelles like lysosomal storage disorders (Leonard and Morris, 2006).

Lysosomes were firstly discovered by Christian de Duve and his colleagues in early 1960's, as a cell organelle and function as digestive unit for cell and also recycle the macromolecules. The environment of lysosome is highly acidic, ~ 4.5 which is due to the presence of ATP driven proton pump, V-ATPase along with the occurrence of pantheon luminal hydrolases and ion channels, which collectively make the organelle perfect for the breakdown of the macromolecules including polysaccharides, lipids and proteins. According to literature, lysosomal biogenesis is dependent upon the coordination of biosynthesis of lysosomal proteins and endosome-lysosome trafficking (Figure 1.1). Lysosomal proteins are firstly synthesized and modified in the lumen of endoplasmic reticulum and then transported to Golgi complex through vesicles. The tag used for the lysosomal enzymes is mannose-6-phosphate which interact with mannose-6-phosphate receptor (M6PR) and whole complex is then enclosed in a clathrin-coated vesicle at the trans-Golgi network to early endosome (Figure 1.1). The delivery of these newly synthesized proteins of lysosomes follow direct and indirect pathways. In direct pathway proteins are sorted out at trans Golgi network and deliver directly into early endosome while in indirect pathway proteins are carried to the plasma membrane and then through endocytosis they are transferred into early endosome and then to late endosome which is further fused into lysosome. Some proteins not destined for lysosomes are transferred back

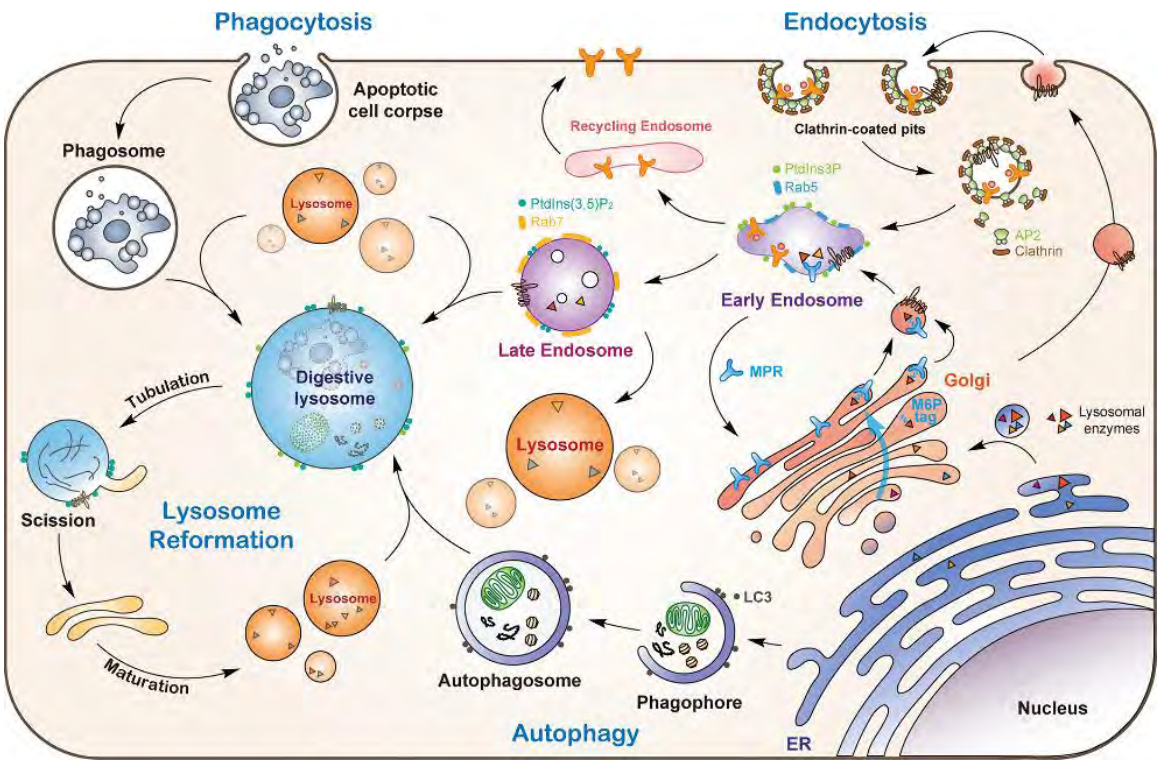


Figure 1.1: Lysosomal biogenesis through direct and indirect pathways and phagocytosis and its degradation through lysosome and reformation of lysosome (Adapted from Yang and Wang, 2021).

to plasma membrane or Golgi network via retromer-dependent or independent pathways (Luzio *et al.,* 2014; McNally and Cullen, 2018; Yang and Wang, 2021). Lysosomal enzymes are encoded by nuclear genome and are key to maintain the internal homeostasis of cell. Disruption of lysosomal proteins including enzymes which result in the accumulation of substrate in the lumen, there are some non-enzymatic proteins like integral membrane proteins, trafficking proteins and proteins participating in post-translational modifications leading to malfunctioning of lysosomal cellular processes like endocytosis, autophagy, pH regulation, Ca^{2+} homeostasis and exocytosis, leads to *lysosomal storage disorders* (LSDs) (Figure 1.2) (Platt *et al.,* 2012; Scerra *et al.,* 2022). Depending upon the number of enzymatic proteins and non-enzymatic proteins, there are more than 60 subtypes of lysosomal storage disorders mostly inherited as autosomal recessive except for Danon disease, Fabry disease and Hunter syndrome (MPS II), which are inherited as X-linked disorders. Most of the LSDs show progressive severity of phenotype and early morbidity and mortality in the children (Table 1.1) (De Duve, 2005; Fuller *et al.,* 2006;

Forgac, 2007; Linhart and Elliott, 2007; Coutinho *et al.,* 2012; Settembre *et al.,* 2013; Luzio *et al.,* 2014; Nair *et al.,* 2019). LSDs affect different cells and organs depending upon the type of compounds accumulated in those cells e.g., Mucopolysacchraidoses, Gaucher disease etc. Despite of being highly interrelated, LSDs differ variably because of accumulation of different compounds in specific tissues e.g., sub-types involving central nervous system like Niemann-Pick disease and non-CNS involving sub-types like Morquio syndrome (Marques and Saftig, 2019). In few severe cases accumulation can occur in almost every organ of the body e.g., Hurler phenotype of Mucopolysaccharidosis type I. Early onset of the phenotypes progresses rapidly than the late onset or attenuated forms. Literature is also available regarding acquired form of LSDs, in which an increased lipid diet intake can leads to the LSDs even when there is no mutation in the genes (Ferraz *et al.,* 2016; Rampanelli *et al.,* 2018).

The incidence rate for overall LSDs is very rare affecting 1:7000-1:8000 live births. There is no universal diagnostic procedure to detect LSDs. However, emerging advancements and techniques for treating various LSDs make it compulsory for any LSD to be diagnosed on time. Tests used to diagnose any certain sub-type of LSD depends upon chronological symptoms, patients' age at the time of presentation and clinical signs. There are a number of diagnostic methods available for LSDs. The first step is analysis of urine for respective accumulated macromolecules, that usually include glycosaminoglycans (GAGs) by using high-pressure liquid chromatography (HPLC). But it is used for definite group of LSDs and showed potent risk of false-negative results specially in patients with severe or attenuated sub-types like MPS III and MPS IV (Mabe *et al.,* 2004; Meikle *et al.,* 2004). For enzymatic levels measurement from fibroblasts or leukocytes by using dry blood spot (DBS) analysis through tendem mass spectrophotometry (TMS), this method is very helpful and accurate for diagnosing enzyme levels to narrow down the sub-types of LSDs involving deficiency of particular enzyme. DBS is further followed by genetic testing for differential diagnosis. The genetic testing for enzymatic/non-enzymatic proteins either through whole genome sequencing or whole exome sequencing are proved to be the best diagnostic procedures as they are more accurate methods to find out the pathogenic variants. Genetic testing also helps in screening of carriers that will further help the family to lower the risks for future off-springs by avoiding cousin marriages.

Once the index case has been diagnosed with any LSD condition, the fetus at risk for LSD can easily be diagnosed through prenatal diagnosis for certain condition. Which make it easy to treat the condition at very early stages of life (Meikle *et al.,* 2004; Kingma *et al.,* 2015; Mehta and Winchester, 2022). Next generation sequencing is a latest technology which helps in advancing genetic make up for exploring and diagnosing genetic disorders at such a fast pace to save early morbidity and mortality in coming generations (Torshizi et al., 2018; Pruneri et al., 2021). Next-generation sequencing (NGS) technologies have revolutionized genetic diagnosis since their inception. These advanced techniques, including whole-exome sequencing (WES) and whole-genome sequencing (WGS), offer comprehensive and efficient approaches to identify genetic variants responsible for various inherited disorders. NGS allows for the rapid and cost-effective analysis of a patient's entire genetic makeup, enabling the detection of both common and rare genetic mutations (Hu et al., 2021). This has significantly improved diagnostic accuracy, especially for conditions with complex genetic etiologies. Additionally, NGS has shortened the diagnostic odyssey for many patients and their families, leading to earlier interventions and personalized treatment strategies (Schuler et al., 2022). Newborn screening for inborn errors of neonates including inborn errors of metabolism through DBS followed by genetic screening through NGS has incredibly reduced the false positive results during screening (Huang et al., 2022).

To date the accessible treatment options for LSDs, include enzyme replacement therapy (ERT), where the required enzyme is injected in the body to cope with the needs of cells. But there is a limitation for exogenous enzymes as it is not capable of crossing blood brain barrier due to large in size, thus affecting less in all those sub-types where there is involvement of central nervous system. Other options in-clinical trial phase includes substrate reduction therapy (SRT), pharmacological chaperon therapy (PCT), bone marrow transplant (BMT), hematopoietic stem cell transplant (HSCT) and gene therapy. For the majority of the treatment options in clinical trial phase including BMT, HSCT, early diagnosis is very crucial during the first year of birth of the patient (Del Grosso *et al.,* 2022).

In Pakistani population the rate of consanguinity is really very high, prevalence of genetic diseases is also high but there on time diagnosis is very low. The reason observed during current study for late diagnosis is late presentation due to lack of awareness and access to advanced clinical facilities. For early diagnosis newborn screening also lacks in here including heel prick for detection of inborn errors of metabolism (IEM). Also, non-specific clinical tests locally are not enough and accurate for diagnosing such complicated disorders with overlapping phenotypes. Dry blood spot analysis used for diagnosing LSDs is currently not available in any public and private sector hospital within country so, performing such tests from any international facility is way too expensive to afford by a common person from low- and middle-income families. The goal of current study is to evaluate the landscape of LSDs in Pakistan and ease the specific and on time diagnosis by founding the founder mutations or highly prevalent variants to reduce the cost and time also helps in drawing a genetic map of different sub-types of LSDs, their prevalence and designing their therapies.

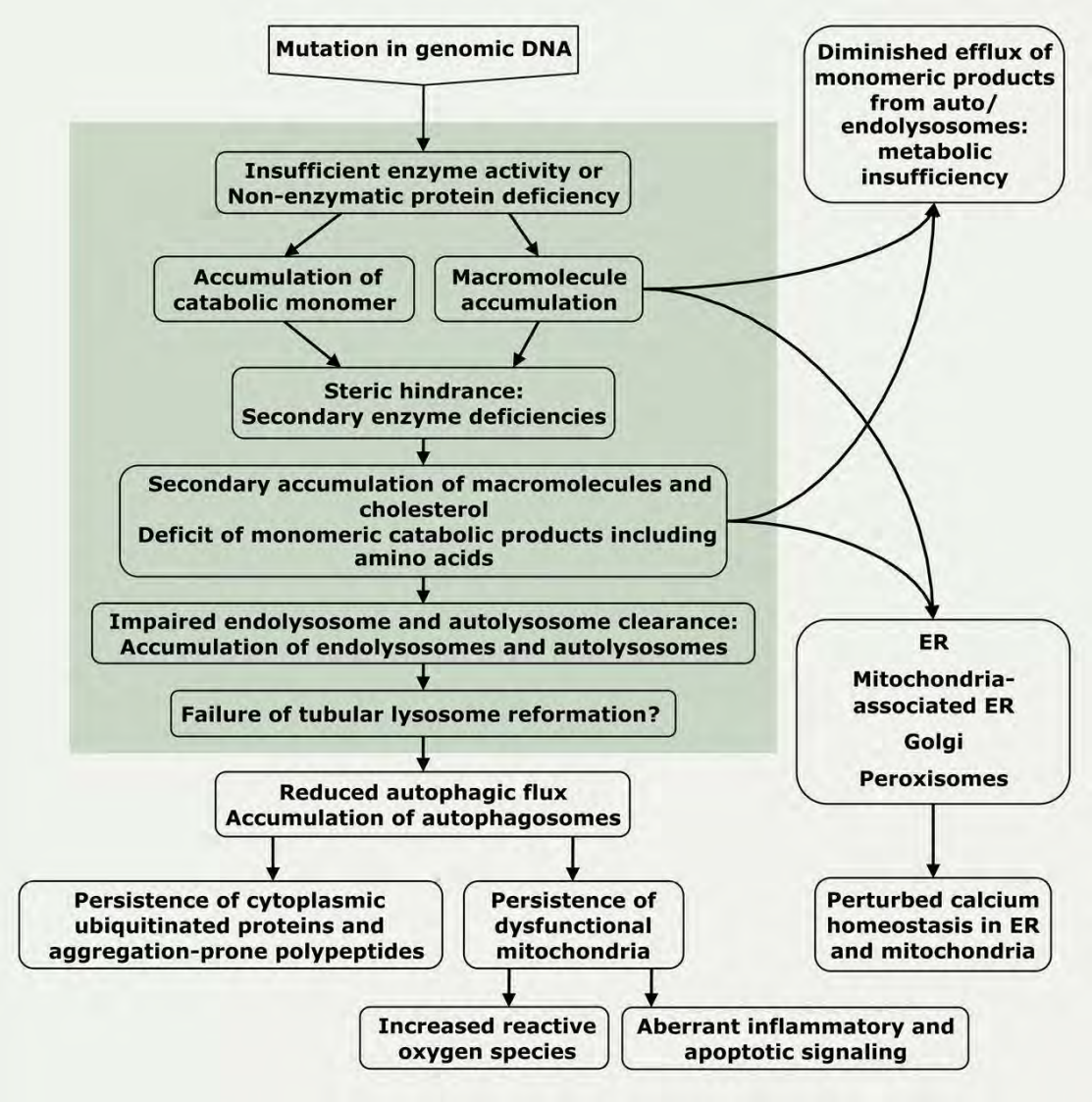


Figure 1.2: Hypothetical cascade of events in lysosomal storage disorders pathology (Adapted from Platt *et al.*, 2012).

Table 1.1: Some majorly classified subtypes of Lysosomal Storage Disorders both enzymatic and non-enzymatic and genes involved in them.

S. no.	CATEGORY AS PER SUBSTRATE	DISEASES	GENES
Enzymatic Lysosomal defects			
1.	Sphingolipids	Tay-Sachs Disease Sandhoff's disease GM1 gangliosidosis Fabry's disease Gaucher disease Farber's disease Niemann-Pick disease type A/B Multiple sulfatase deficiency	*HEXA* *HEXB* *GLB1* *GLA* *GBA* *ASAH1* *SMPD1* *SUMF1*
2.	Mucopolysaccharidosis (Glycosaminoglycans)	MPS-I, Hurler, Scheie MPS-II, Hunter MPS-IIIA, Sanfilippo A MPS-IIIB, Sanfilippo B MPS-IIIC, Sanfilippo C MPS-IIID, Sanfilippo D MPS-IV A, Morquio A MPS-IV B, Morquio B MPS-VI, Maroteaux-Lamy MPS-VII, Sly MPS-IX	*IDUA* *IDS* *SGSH* *NAGLU* *HGSNAT* *GNS* *GALNS* *GLB1* *ARSB* *GUSB* *HYAL1*
3.	Glycogen	Pompe disease	*GAA*
4.	Oligosaccharide/Glycopeptides	α-Mannosidosis β-Mannosidosis α-Fucosidosis Schindler-Kanzaki disease Aspertylglucosaminuria	*MAN2B1* *MANBA* *FUCA1* *NAGA* *AGA*
5.	Peptides	Pycnodysostosis	*CTSK*
6.	Monosaccharides/Amino acid monomers	Cysteinosis	*CTNS*
7.	Multiple enzyme deficiencies/Mucolipidoses	Mucolipidosis-I, Sialidosis Mucolipidosis-II Mucolipidosis-III	*NEU1* *GNPTAB* *GNPTG*
8.	Acid Anhydrides	Griscelli type 1 Griscelli type 2 Griscelli type 3 Chediak-Higashi syndrome	*RAB27A* *MLPH* *MYO5A* *LYST*
Non- Enzymatic Lysosomal defects			
9.	Transporter	Niemann-Pick disease type C1/2 Salla disease	*NP-C1/2* *SLC17A5*
10.	Activator defects	Danon disease Galactosialidosis	*LAMP2* *CTSA*

1.1. Mucopolysaccharidosis type I

Mucopolysaccharidosis I (MPS I, OMIM#252800), an autosomal recessive disease and is a subtype of lysosomal storage disorders (LSDs) caused by the deficiency of an enzyme α-L-iduronidase (IDUA, EC 3.2.1.76, OMIM#252800) (Kiely *et al.,* 2017). MPS I is a rare disease effecting 1:100,000 live births (Beck *et al.,* 2014). Patients with MPS I has IQ less than 70 in 25% of cases who remained untreated till 2-3 years of age (Shapiro *et al.,* 2015).

Three sub types of MPS 1 based on disease severity include Hurler (MPS I-H, OMIM#607014) the most severe neuropathic form with life expectancy of less than 10 years, Hurler-Scheie (MPS I-H/S, OMIM#607015) intermediate form with mild or no neurological involvement and patients may survive till third decade of life, Scheie (MPS I-S, OMIM#607016) mildest form of all having late onset and normal life span with available therapies (Muenzer *et al.,* 2009; Bertola *et al.,* 2011; Ngiwsara *et al.,* 2018). IDUA enzyme is involved in the degradation of two glycosaminoglycans (GAGs) that are heparan sulfate and dermatan sulfate (Beck *et al.,* 2014). IDUA enzyme has three important domains: a TIM barrel domain accommodating the catalytic site, a fibronectin-like domain and a β-sandwich domain. IDUA structure is bound to analogs of iduronate exemplify the Michaelis complex and expose a B conformation in the intermediate of glycosyl-enzyme, which propose to hold on a double displacement reaction involving the general acid/base Glu182 and the nucleophilic Glu299. Unexpectedly, the N-glycan is required for enzymatic activity and is attached to Asn372 and interacts with analogs of iduronate in the active site (Bie *et al.,* 2013). Malfunctioning of IDUA enzyme leads to accumulation of GAGs which further activates the TLR4 pathway and leads to activation of NFκB which start producing cytokines and activates the innate immune system (Simonaro *et al.,* 2005; Metcalf *et al.,* 2010). There is difference in the structure and sulfation pattern of GAGs which is actual pathophysiological event behind MPS-I bone deformities because distribution of growth factors depends on the sulfation pattern of GAGs (Kingma and Jonckheere, 2021). The IDUA enzyme is encoded by gene *IDUA* with 13 introns and 14 exons, which is located on chromosome 4p16.3 and encodes a protein of 653 amino acids (Scott *et al.,* 1991, 1992c, 1995).

To date there are more than 250 mutations have been reported for *IDUA* gene which include missense, nonsense, deletions, insertions and splicing variants (Human Gene Mutation Database, accessed on 5 March 2022). According to a survey in 1998, 60% of all marriages were consanguineous and among them, more than 80% were first cousin marriages (Hussain and Bittles, 1998). A study conducted from Pakistan (January 2011 to October 2014) for determining the frequency of inherited metabolic disorders showed carbohydrate disorders 51%, followed by lysosomal storage disorders 32.7%, organic acid and energy metabolism defects 10%, amino acid disorders 3.3% and miscellaneous 2.2%. In addition, 97% of cases had consanguinity among the parents (Cheema *et al.,* 2016).

In previous few studies UK based Pakistani families have been diagnosed with MPS I phenotype and genotype, revealed recurrent missense variant p. Leu490Pro, and the same variant was also reported from the Indian families and also a novel non-sense variant p. Glu486Ter was reported from Pakistan (Beesley *et al.,* 2001; Vijay *et al.,* 2005; Arora *et al.,* 2007; Uttarilli *et al.,* 2016; Gul *et al.,* 2020; Pakistan Mutation Database, 2020).

Diagnosis for MPS I include initial phenotypic symptoms, but it is very challenging to make accurate diagnosis on initial symptoms like hernia and respiratory tract infection during early childhood. Therefore, other options involve a sensitive but nonspecific quantification test of urinary GAGs analysis in which extra dilution of urine can leads to false negative results (Muenzer *et al.,* 2009). Second and the most accurate method in practice is analysis of enzyme levels in dry blood spot, cultured fibroblast or peripheral blood leukocytes either through tandem mass spectrometry or fluorimetric assay (Lin *et al.,* 2013; Scott *et al.,* 2013).

Treatment options are enzyme replacement therapy (ERT) with Iaronidase (Aldurazyme®), but these are the exogenous enzymes not capable of crossing blood brain barrier (BBB) so is not effective in case of neurocognitive functions. Hematopoietic stem cell transplantation (HSCT) is effective in early age of the patient with age limit below 2 years because of reproduction of microglial cells and enzyme restoration in brain. In addition, in early age the severity of disease phenotype is less so the outcomes for the treatment will show fruitful results. Bone marrow transplant

(BMT) in MPS I patients showed mean intelligent quotient of 76 in almost every group of patients (Krivit *et al.,* 1995; Aldenhoven *et al.,* 2015; Shapiro *et al.,* 2015; Sun, 2018). Some of the treatment options in their clinical trials include gene therapy, genome editing, small molecule therapy, anti-inflammatory therapy, and some other novel therapies (Ou *et al.,* 2019; Schuh *et al.,* 2020; Kingma and Jonckheere, 2021).

1.2. Mucopolysaccharidosis type II

Mucopolysaccsaccharidosis type II or Hunter syndrome (OMIM#309900) is a X-linked recessive sub type of lysosomal storage disorders (LSDs). Hunter syndrome is caused by the scarcity of enzyme iduronate-2-suphatase (IDS, EC 3.1.6.13, OMIM#300823). This is ubiquitously expressed hydrolase and is involved in the removal of two sulfate units from heparan sulfate and dermatan sulfate. It affects 1 in 100,000 live births in male children while few rare cases of females were also reported (Neufeld, 2001, Muenzer, 2004).

In a study, it was shown that in zebra fish heparan sulfate sulfotransferases, involved in modification of sulfation pattern of heparan sulfate, are entangled in cardiac valve formation, limb development and somite specification (Walsh and Stainer, 2001; Bink *et al.,* 2003; Norton *et al.,* 2005). In a separate study, indecisive character of chondroitin sulfate and dermatan sulfate prosteoglycans in development has been shown. However, there is an increasing supporting evidence for the role of chondroitin sulfate and dermatan sulfate in the development of vertebrate brain and motor nerve outgrowth (Bernhardt and Schachner, 2000; Sugahara and Mikami, 2007). Proteoglycans derived from chondroitin sulfate and dermatan sulfate bind to multiple morphogens like family of fibroblast growth factors (FGF) and help in regulation of several general biological processes like cellular division and specific biological processes like neurogenic activity (Sugahara *et al.,* 2003). Active IDS enzyme to get maximum catalytic activity need biotransformation of cysteine residues to formyl glycine. Sulfatase-modifying factor 1 (SUMF1) enzyme helps IDS to get this biotransformation (Demydchuk *et al.,* 2017; Gleitz *et al.,* 2018; Wada *et al.,* 2020). Loss of function of IDS enzyme leads to the excessive accumulation of GAGs in various organs, which further leads to the multiple organ failure including mental retardation, hepatosplenomegaly, stiff joints and many more (Muenzer, 2004;

Baehner *et al.,* 2005; Lin *et al.,* 2009; Demydchuk *et al.,* 2017). The enzyme IDS is encoded by gene *IDS* which is mapped on chromosome X long arm q28. It has 9 coding exons which transcribe into 550 residue protein. In *IDS* gene, 20 kb downstream of the transcribing region there is a pseudogene, *IDS2*. It has homology with exon 2, 3, intron 2, 3, and 7 (Bondeson *et al.,* 1995). *IDS2* shows 100% homology with exon 3 while 96% with other sequences (Bondeson *et al.,* 1995). Hunter syndrome is categorized into two main categories; first one with severe and early onset of the disease in which phenotypes appear between 2 to 4 years of age and patient shows behavioral and progressive intellectual disability by affecting the peripheral and central nervous system and death occur in childhood. Second one is milder and have late onset of the disease with normal or partial involvement of the neural cognition and survival period is up to adulthood (Young and Harper, 1983; Li *et al.,* 1999). Clinical profiles of the patients showed few common and prevalent symptoms like coarse facial features, hepatosplenomegaly, skeletal deformities, joint contractures, cardiovascular and respiratory disorders, delay in achieving milestones and intellectual disabilities (Demydchuk *et al.,* 2017).

The limited studies from Pakistan regarding the genetics of Hunter syndrome have not explained much of the population specific prevalent variants. A Pakistani group published a study of Hunter syndrome screening three Pakistani families and identified 2 novel variants (c.114_117dupCGTT and c.1006 + 1G>C) and a formerly reported nonsense variant p. Gln389Ter (Jonsson *et al.,* 1995; Zubaida *et al.,* 2019).

Diagnosis for MPS II involve urinary GAGs analysis in which extra dilution of urine can leads to false negative results (Muenzer *et al.,* 2009). Second method in practice is analysis of enzyme levels in dry blood spot, cultured fibroblast, or peripheral blood leukocytes either through tandem mass spectrometry or fluorometric assay (Lin *et al.,* 2013; Scott *et al.,* 2013).

Possible treatment options include enzyme replacement therapy (ERT) through idursulfase (Elaprase®) and idursulfase Beta (Hunterase®). Hematopoietic stem cell transplantation (HSCT) was not successful at initial stages in MPS II patients. Therapeutic iduronate 2-sulfatase production from novel single stranded RNA virus (Sendai) vector is in clinical trial phase (Guffon *et al.,* 2009; Nakanishi

and Otsu, 2012; Aldenhoven *et al.,* 2015; Sano *et al.,* 2019; BdO *et al.,* 2021; Ohira *et al.,* 2021).

1.3. Mucopolysaccharidosis type III

Mucopolysaccharidosis type III or Sanfilippo syndrome, discovered 60 years ago, is a sub type of lysosomal storage disorders with autosomal recessive mode of inheritance. The ratio of Sanfilippo syndrome ranges from subtype to geographical location but estimated to affect 1 in 70,000 live births (Sanfilippo, 1963; Valstar *et al.,* 2008; Fedele, 2015; Khan *et al.,* 2017).

MPS III has been further classified into four categories MPS III type A (OMIM#252900), MPS III type B (OMIM#252920), MPS III type C (OMIM#252930) and MPS III type D (OMIM#252940). The deficiency of enzyme *N*-sulfoglucosamine sulfohydrolase (EC 3.10.1.1, OMIM#605270) lead the way to the phenotype of MPS III type A and is encoded by gene *SGSH*, which is located on chromosome 17q25.3 and has 8 exons translating into a 502-residue protein (Scott *et al.,* 1995). To date there have been 155 variants reported in *SGSH* gene (HGMD professional 2021.2, assessed on 18 September 2021), MPS III type B is caused by the deficiency of enzyme *N*-acetyl-α-glucosaminidase (EC 3.2.1.50, OMIM#609701) which is encoded by gene *NAGLU*, on chromosome 17q21.1 and has 6 exons which produce 743 amino acid protein (Zhao *et al.,* 1996) with more than 160 variants in HGMD professional 2021.2 assessed on 18 September 2021, MPS III type C is caused by the deficiency of enzyme α-glucosaminide *N*-acetyle transferase (EC 2.3.1.78, OMIM#610453) encoded by gene *HGSNAT*, localized on chromosome 8p11.1. This gene possess 18 exons which translate into a 635 amino acid protein (Fan *et al.,* 2006; Hrebicek *et al.,* 2006) with 68 variants reported in HGMD professional 2021.2 assessed on 18 September 2021 and MPS III type D is caused by insufficiency of enzyme *N*-acetylglucosamine 6-sulfatase (EC 3.1.6.14, OMIM#607664) which is encoded on gene *GNS*, located on chromosome 12q14.3, this gene consists of 14 exons translated into 552 residue protein (Robertson *et al.,* 1988) with reports of 25 variants only in HGMD professional 2021.2, assessed on 18 September 2021. Sanfilippo syndrome is caused by the excessive accumulation of heparan sulfate (HS) (Bame and Rome. 1986 a,b). All these four afore mentioned enzymes are involved in the

degradation of heparan sulfate (HS) chains, so deficiency of any of these enzymes leads to the accumulation of heparan sulfate and ultimately to Sanfilippo phenotype. Neurological manifestations are also involved in the patients of Sanfilippo syndrome, which includes the delayed behavioral and developmental milestones starting from early years of life (Shapiro *et al.,* 2017; Ozkinay *et al.,* 2021). In most severe forms, there are decrease of 9 to 15 points in the developmental quotient (Shapiro *et al.,* 2016). Prevalence for each type varies from population to population, as subtype A is common in northern part of Europe and sub-type B is common in southern part of Europe, subtype C is in general less common and subtype D is the rarest form (Zelei *et al.,* 2018; Beneto *et al.,* 2020). Heparan sulfate (HS) is involved in the central nervous system (CNS) development, which explains the pathology of neurological manifestations in the children affected from Sanfilippo syndrome having dementia from childhood (De Pascuale and Pavone, 2019). Recent studies have confirmed a direct connection between Sanfilippo and Parkinson's disease. Patients affected with Sanfilippo mutations, have higher chances for developing Parkinson's disease (Winder *et al.,* 2012). General clinical phenotypes presented by all subtypes of Sanfilippo syndrome include central nervous system deterioration in first decade of life including cortical atrophy, hyperactivity, motor deterioration, aggressive behavior, delayed learning, sleeping apnea, severe mental retardation and some somatic manifestation including hepatosplenomegaly, hirsutism, dysphagia, joint stiffness, hypoacusia, hypertrichosis, skeletal deformities, and loss of speech. Death usually occurs in second or third decade of life, but there are some exceptions up to fourth and fifth decade (Neufeld, 2001; Ruijter *et al.,* 2003; Valstar *et al.,* 2011; Andrade *et al.,* 2015).

There was no previous study conducted from Pakistan related to Sanfilippo cases to the best of my knowledge. However, some of the international groups have published some research revealing different variants from Pakistani born patients/parents in different parts of the world. In literature, the novel or reported variants of Sanfilippo syndrome from Pakistani patients include a novel deletion causing Sanfilippo syndrome type D, c.1169delA in exon 10, which causes a pre mature termination codon (Beesly *et al.,* 2003), a novel non sense variant leading to

Sanfilippo syndrome type C, p. Ser296Ter in exon 10 and splice site variant c.744-2A>G in intron 7 (Feldhammer *et al.,* 2009 a,b).

Diagnosis for MPS III involve either the urinary heparan sulfate excretion measurement or the enzyme level measurements from dry blood spot (DBS) through tendem mass spectrometry (TMS) or fluorimetry (Seker Yilmaz *et al.,* 2020).

Not a single therapeutic approach has been found successful against reverting the symptomatology of Sanfilippo syndrome and neuronal cognitions caused by it. However, some of the approaches tested commonly include delivering enzyme replacement therapy (ERT) into central nervous system (CNS), pharmacological chaperon therapy (PCT), substrate reduction therapy (SRT), gene therapy and human stem cell transplantation (HSCT) (Valstar *et al.,* 2010; Beneto *et al.,* 2020). Some other approaches used on different cell and animal models include coenzyme Q_{10} (Matalonga *et al.,* 2014), over expression of TFEB (Lotfi *et al.,* 2018), master regulator of lysosomal biogenesis (Bajaj *et al.,* 2019) and modified RNAs use to revert aberrant splicing processes (Matos *et al.,* 2014).

1.4. Mucopolysaccharidosis type IV

Mucopolysaccharidosis type IV, also termed Morquio syndrome, firstly discovered by a pediatrician Luis Morquio (1867-1935). It has autosomal recessive mode of inheritances (Morquio, 1929). It is a rare disorder with prevalence of 1 in 40,000 live births (Nelson *et al.,* 2003). Clinical profile includes pectus carinatum, short stature, genu valgum, forearm deformity, hip dysplasia, scoliosis/kyphosis, atlantoaxial instability, dental deformities, odontoid hypoplasia, corneal clouding, obstructive or restrictive lung disease, hearing loss and cardiac arrhythmias. There is no neuronal cognition in Morquio patients. At the mean height at age of 18 years for males is 122.4±21.5 cm and in females it is 113.1±22.6 cm (Montano *et al.,* 2008).

There are two types of Morquio syndrome, clinically difficult to diagnose on the basis of phenotype, type A (OMIM#253000) being the most severe one, caused by the scarcity of enzyme N-acetylgalactosamine-6-sulfatase (GALNS, EC 3.1.6.4, OMIM#612222) (Hendriksz *et al.,* 2013; Sun *et al.,* 2018). Type B (OMIM#253010), which is less severe and rare, is caused by the deficiency of enzyme β-d-galactosidase

(GLB1, EC 3.2.1.23, OMIM#611458) is also allelic with GM1 gangliosidosis (Chudley and Chakravorty, 2002; Stockler-Ipsiroglu *et al.,* 2021). GALNS enzyme is encoded by the *GALNS* gene, which is located on chromosome 16q24.3 and consists of 14 exons and translated into 522 amino acid protein (Masue *et al.,* 1991; Masuno *et al.,* 1993). GLB1 enzyme is encoded by *GLB1* gene, which is mapped on chromosome 3p21.33. It has 16 exons, which are translated into 677 amino acid protein (Brunetti-Pierri and Scaglia, 2008). GALNS enzyme is involved in the degradation of chondroitin-6-sulfate (C6S) and keratin sulfate (KS), while GLB1 enzyme is involved in the degradation of keratin sulfate only (Celik *et al.,* 2021). GALNS is phosphorylated with 6-mannose phosphate which helps to transport the protein to lysosome through endocytic pathway. GALNS is activated through post-translational modification at Cys79 into formyl glycine aldehyde (Dahms *et al.,* 1989; Dierks *et al.,* 1998; Rivera-Colon *et al.,* 2012). The enzyme GALNS is a member of sulfatase family including 10 different human sulfatases showing close similarity of 20-30% at amino acid residue level and conservation of Cys79 catalytic site in all of them (Schmidt *et al.,* 1995). Generally, sulfatases are involved in the catalysis of hydrolysis of ester bonds from broad spectrum of substrates (Rivera-Colon *et al.,* 2012). GLB1 enzyme is also involved in the hydrolyzation of terminal beta galactosyl residues of GM1 gangliosides and glycoproteins (Okada and John, 1968). Thus, in absence of beta galactosidase enzyme there is accumulation of polysaccharide substrate with non-reducing end, glycans without O- and N- linkage and beta-linked galactosyl residues (Wolfe *et al.,* 1974; Lawrence *et al.,* 2019). Rising evidence but from very limited studies performed for Morquio type B suggested that these elevations of intracellular GAGs trigger the inflammation inside the cell through apoptosis of cells of connective tissues, break up of cartilage tissues and increasing levels of pro-inflammatory cytokines. These up levels of cytokines in the cells cause the pathophysiology of pain and inflammation in the patients suffering from Morquio syndrome type B. The thickness of articular cartilage was also observed more in animal models of Morquio type B with hyperplastic synovium (Simonaro *et al.,* 2005). The location of mutant variants describes the pathogenesis of the *GLB1* clinical presentation and phenotype, ranging from skeletal abnormalities in Morquio type B to neurological manifestations in GM1 gangliosidosis. Both the allelic forms share some phenotypes in common i.e there are mild skeletal abnormalities observed in

GM1 gangliosidosis and mild neuronal cognitions in Morquio syndrome type B (Abumansour *et al.,* 2020).

From Pakistan, few studies previously performed on some cases of Morquio syndrome, reported five novel variants p. Phe216Ser, p. Met38Arg, p. Ala291Ser, p. Glu121Arg*fs**37, p. Tyr294Ter*fs* and two reported variants p. Pro420Arg, p. Arg386Cys (Tomatsu *et al.,* 2004; Morrone *et al.,* 2014; Ullah *et al.,* 2017; Zubaida *et al.,* 2018).

Initial diagnosis for MPS IV (A and B) involves the clinical presentation and radiological findings (X-Ray) of the body. Further followed by urinary excretion or liquid chromatography-mass spectrometry (LC-MS/MS) based method for KS measurements from dry blood spot (DBS) or urine and molecular genetic analysis (Piraud *et al.,* 1993; Martell *et al.,* 2011; Peracha *et al.,* 2018; Chien *et al.,* 2020; Lin *et al.,* 2020).

Treatment options for MPS IV include enzyme replacement therapy (Elosulfase alpha, VIMIZIM®) which reduces the keratan sulfate when assessed through 6 minute walk test (6MWT) and is working for MPS IV type A but not for MPS IV type B, molecular chaperon therapy, substrate reduction therapy, hematopoietic stem cell transplantation (HSCT), in-vivo adeno-associated viral gene therapy, anti-oxidant therapy, inhibiting protein aggregation, stop-codon read through therapy and anti-inflammatory therapy (Hendriksz *et al.,* 2014; Akyol *et al.,* 2019; Seker Yilmaz *et al.,* 2020). In a recent study, recombinant human GALNS enzyme (rhGALNS) was infused with hydrogel like polyethylene glycol to deliver the exogenous enzyme successfully into the fibroblasts for sustainable and longer release but this might be the expensive and frequently required treatment (Jain *et al.,* 2020).

1.5. Gaucher Disease

Gaucher disease (GD) was first discovered by a medical doctor names Gaucher in 1882. He noticed fluid filled spaces inside the macrophages which is now called as ‘Gaucher bodies’ (Gaucher, 1882). The fatty substance stored in the macrophages was a glycolipid known as glucocerebroside. Which is accumulated by the deficiency of the enzyme glucocerebrosidase (GCase, EC 3.2.1.45,

OMIM#606463) (Figure 1.3) in the lysosomes of macrophages, primarily in liver, spleen, osteoclasts, brain, bone marrow, less often in lungs, kidneys, heart, conjunctivae, and skin. In normal functioning cell, saposin C presents the substrate glucocerebroside to the glucosidase, beta, acid (GBA), which activates the enzyme. This enzyme then causes the hydrolytic breakdown of glycosylceramide into glucose and ceramide (Orvisky *et al.,* 2002; Hoops *et al.,* 2007). GD affects 1:40,000–100,000 live births. It has an autosomal recessive mode of inheritance. The most common phenotypes of GD patients include supranuclear gaze palsies, auditory processing defect, ataxia, occasional seizures, myoclonus, hepatosplenomegaly, anemia, and abdominal distention (Hughes *et al.,* 2019).

The variation of phenotypes in the GD patients with same genotype has been under debate since a decade ago, for example, one of the most prevalent variant p. Leu444Pro in homozygous state had severely affected autistic children with seizures and also successful students with brilliant academic record and mild symptoms (Sidransky *et al.,* 2003). There are three main types of GD: most common form in Western world is type 1 non-neuronopathic (GD1, OMIM#230800) which can be differentiated from other two types on the basis of un-involvement of central nervous system (CNS) during early stages, type 2 acute neuronopathic (nGD2, OMIM#230900) which typically don't have any ethnic predilection and has very early onset of central nervous system manifestations in the patient starting at the age of 6 month or before, and cause death of the affected individual within 2 years of life but some exceptions also lye in case with proper medical care and support and live up to 4 years, and type 3 chronic neuronopathic (nGD3, OMIM#231000) with slight later onset as nGD2, showing central nervous system cognitions after months or years of birth with slightly slow progressions than nGD2 (Goker-Alpan *et al.,* 2003; Daykin *et al.,* 2021; Motta *et al.,* 2021). In 5% of GD patients, an extra pyramidal disease develops, which resembles Parkinson's disease (Bultron *et al.,* 2010; Schapira, 2015). GCase is encoded by gene *GBA*, present on chromosomal locus 1q21 consists of 11 exons which translated into 536 amino acid protein. To date there are 459 variants reported in Human Genome Mutation Data Base for causing Gaucher disease in humans. From *GBA* gene, 16 kb downstream lye a pseudo gene (*GBAP*) which has more than 96% exonic homology with *GBA*. The region of chromosome with *GBA*

gene is a gene rich region with 7 genes and 2 pseudogenes within 85 kb of chromosome 1q. Between *GBA* and *GBAP* gene lye another pseudogene metaxin (*MTX1*), which also transcribe with *GBA*. These pseudogenes seem to be produced by the duplication of this region (Long *et al.,* 1996; Winfield *et al.,* 1997; Martínez-Arias *et al.,* 2001; HGMD Professional 2021.2, Accessed on 15-10-2021). In all three sub-forms, the complexity and heterogeneity of genotype-phenotype makes it difficult to diagnose them solely on the basis of genotype except the presence of p. Asn409Ser allele make the exclusion of nGD involvement (Schiffmann *et al.,* 2020). In non-European population like Taiwan, China, Japan, Pakistan, Egypt, Korea and India nGD3 is the most prevalent form of disease (Eto and Idda, 1999; Choy *et al.,* 2007; Abdelwahab *et al.,* 2016).

From Pakistan few published studies revealed the disease causing alleles p. Leu444Pro/Leu483Pro (Naz *et al.,* 2018; Gul *et al.,* 2021), g.9551G>A (no aa change), p. Ile379Phe, p. Glu381Asp, p. Arg385Lys, c.1170C>A (no aa change), p. Glu395His, c.1152G>A (no aa change), p. Ala400Pro (Sattar *et al.,* 2021).

Diagnosis of GD involve measurement of activity of β-glucosidase in leucocytes or fibroblasts and biopsies of spleen, liver and bone marrow showing striated cytoplasm on Leishman staining or pink sheets in tissue sections stained with hematoxylin and eosin (H&E) (Thomas *et al.,* 2013; Hughes *et al.,* 2019). With advancement in technology next generation sequencing (NGS) is also a commonly used method in order to get better insight into genotype-phenotype correlation (Hagege *et al.,* 2017; Abou Tayoun and Heather, 2019; Han and Lee, 2020).

Treatment options for GD include enzyme replacement therapy, which was developed firstly in 1990s after extraction of enzyme (alglucerase) from human placenta and now a number of pharmaceuticals in the world are providing with the exogenous recombinant GCase including Cerezyme®, Sanofi-Genzyme, Vpriv®, Shire, authorized in 2010 and Elelyso®, Pfizer (Goker-Alpan, 2011; Stirenemann *et al.,* 2017). But ERT in case of neuronopathic GDs is not effective because of the inability of exogenous enzymes to cross blood brain barrier (BBB). So other options like pharmacological chaperon therapy, intrathecal gene therapy are in their advanced clinical trials (Schiffmann *et al.,* 2020).

Recent COVID-19 pandemic hits many people with severe infection particularly the ones already infected or suffering from some under-laying diseases/infections. In a study, it was reported that the patients of GD have high levels of angiotensin-converting enzyme (ACE) beside other elevated biomarkers, although the relation between ACE-2 receptors and Covid-19 infection is not clear, but luckily no GD patient have been reported with severe infection or death with Covid-19 (Zimran *et al.*, 2020)

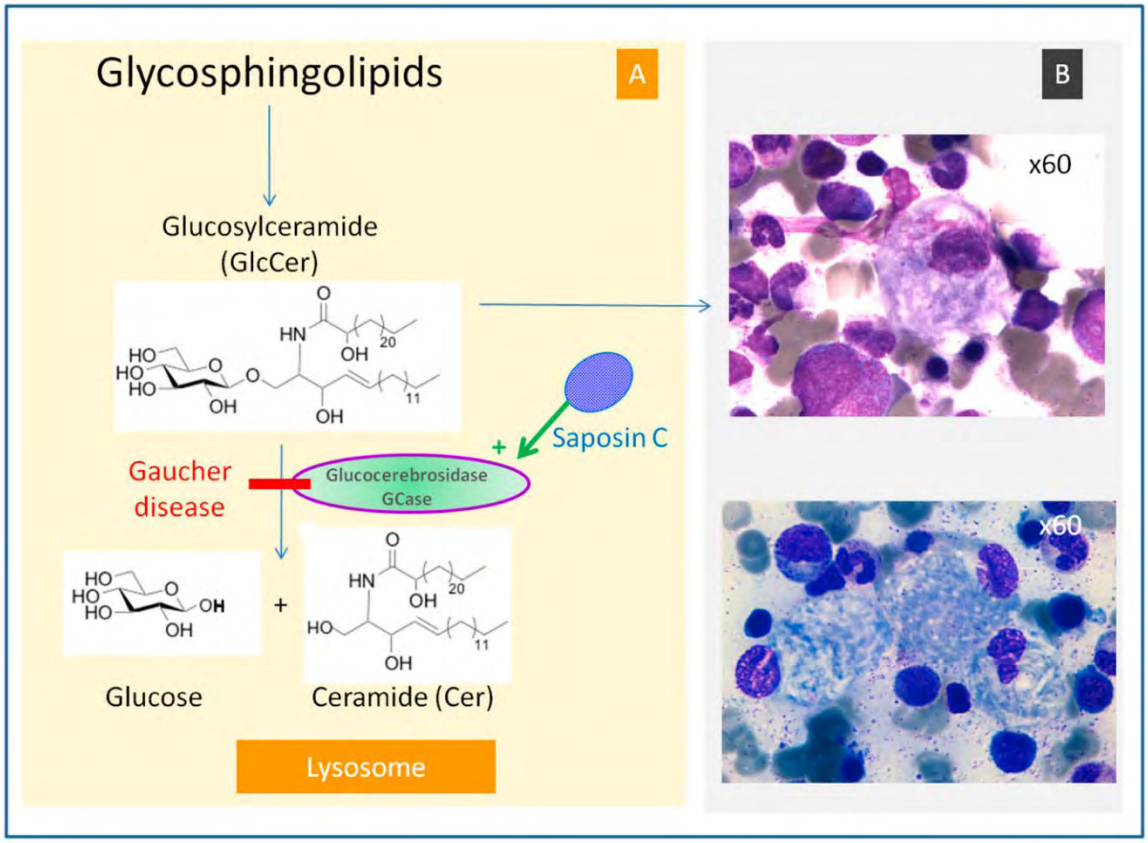


Figure 1.3: Glucoceramide (GlcCer) hydrolysis by enzyme glucocerebrosidase (GCase) inside lysosomal cavity. (A) Saposin actiates GCase and its deficiency leads to the accumulation of substrate GlcCer (B) Which further form aggregates inside macrophages which appears as "crumples tissue paper" (Courtesy by Stirnemann *et al.*, 2017).

1.6. Niemann-Pick Disease

Niemann-Pick disease (NPD) was first discovered by Albert Niemann and Ludwig Pick in late 1920's. NPD is heterogeneous group of autosomal recessive disorders which is lysosomal lipid storage disease. Prevalence of NPD varies from 1: 250,000 to 1:120,000 in different sub-types (Abghari *et al.,* 2019). The most common clinical profiles include hepatosplenomegaly, sphingomyelin storage in reticuloendothelial tissues and involved neurological cognition in some but not all cases (Vanier, 2010).

NPD is further classified into four categories, type A (no enzyme) (OMIM#257200) which has severe and early central nervous system deterioration, type B (partial enzyme) (OMIM#607616) which has chronic visceral involvement but no nervous system malfunctioning, type C and D have moderate or slow sub-acute nervous system involvement and mild visceral storage (Croker, 1961). Type A and B have showed mutations in the gene *SMPD1* which encode an enzyme acid sphingomyelinase (EC 3.1.4.12, OMIM#607608), which causes the accumulation of sphingomyelin. *SMPD1* gene is located on chromosome 11p15.4 preferably expressed from maternal chromosome i.e paternally imprinted. *SMPD1* gene has 6 exons which translated into 631 amino acid proteins (Schuchman *et al.,* 1992; Simonaro *et al.,* 2006; Henry *et al.,* 2013). In types A and B NPD large, lipid-laden foam cells are present in the spleen, lymph nodes, liver, bone marrow, adrenal cortex, and lung airways (Niemann and Link, 1933; Iaselli *et al.,* 2011). Almost half of infants affected with NPD-A have cherry red spot in macula (Schuchman, 2001; McGovern *et al.,* 2006). In type B, there is no clear involvement of central nervous system, but profound hepatosplenomegaly was found accompanying the liver failure (Hollak *et al.,* 2012). Type C is caused by two genes *NPC1* (95%) (OMIM#607623) chromosome 18q11-q12, *NPC2* (5%) (OMIM#601015) chromosome 14q24.3. The products of these genes are not involved directly in the metabolism but help in trafficking of the lipids. Cholesterol, which normally exits the endocytic pathway and go to the endoplasmic reticulum, accumulates in the lysosomes in this scenario, which leads to NPD type C (Newton *et al.,* 2018). *NPC1* has 25 exons encoding a 1278 amino acid protein located in late endosome, which is intracellular cholesterol transporter 1, which by binding cholesterol to its N terminus and has 13 transmembrane domains (Kwon *et al.,* 2009).

NPC2 has 5 exons encoding a 151 amino acid protein localized in the lumen of lysosomes (Vanier and Millat, 2003). It is suggested that NPC1 protein is an efflux pump for cholesterol and maintains the lipid (cholesterol) contents of the membranes and interact with NPC2 protein. Although knowledge for NPC1 and NPC2 is not complete and clear yet (Scott and YA, 2004; Trinh *et al.,* 2018; Winkler *et al.,* 2019). Most pathogenic variants in NP-C are missense comprising of 70-80% of total variants reported so far while less common (20-30%) are splicing, frameshift or pre-mature stop codon causing mutations (Dardis *et al.,* 2020). Niemann-Pick disease type D is only related to a certain ethnic group Acadian couple who married in ~1700, today present in Nova Scotia. NPD type D is also caused by the same *NPC1* gene, suggested that these two disorders are allelic variants (Winsor and Welch 1978; Greer *et al.,* 1997).

From Pakistan only study that reported variants from our population are p. Trp573Ser, p. His423Tyr, p. Arg443Ter, p. Arg498His, compound heterozygote with p. Gly247Asp, p. Arg498His in *SMPD1* gene (Cheema *et al.,* 2020).

NPD is difficult to diagnose through clinical profiles and phenotypic presentation due to quite overlapping features. Therefore, biochemical and genetic screening is important (McGovern *et al.,* 2017).

Treatment options involve: 1) enzyme replacement therapy, recently introduced for NPD type B by introducing olipudase alpha is in the final clinical trials and showing promising results by alleviating the symptoms without any markable side effects and 2) substrate reduction therapy (Barton *et al.,* 1991; Cox *et al.,* 2000; Wasserstein *et al.,* 2018; Thurberg *et al.,* 2020). While other treatment options in clinical trials include chaperon therapy (Han *et al.,* 2020), gene therapy (Massaro *et al.,* 2020), SapC-DOPS nanovesicles (Saposin C, phospholipid dioleoylphosphatidylserine) (Tylki-Szymańska *et al.,* 2011), novel modified amino acid N-acetyl-L-leucine for treating NPD particularly NPD type C is also in progress and in utero enzyme replacement therapy (IUERT) (Goasdoué *et al.,* 2017; Fields *et al.,* 2021).

1.7. Pycnodysostosis

Pycnodysostosis (OMIM#265800, Greek: pycnos = dense; dys = defective; osteon= bone), is a rare autosomal recessive disease with elevated density of bones and is caused by the deficiency of an enzyme which is member of papain-like cysteine protease family, called cathepsin K (CTSK, OMIM#601105) which was reported firstly by Maroteaux and Lamy in 1962 (Maroteaux and Lamy, 1962 a, b; Bizaoui *et al.,* 2019). It affects 1-1.7 per million and only ~200 patients suffering from pycnodysostosis have been reported in the medical literature with 54 pathogenic variants. Interestingly, all of the pathogenic variants in pycnodysostosis patients in *CTSK* gene are loss-of-function mutations that leads to absence of cathepsin K enzyme (Li *et al.,* 2009; Bertola *et al.,* 2010; Khan *et al.,* 2010; Turan, 2014; LeBlanc and Savarirayan, 2020; HGMD, Professional 2021.2, Accessed on 17-Oct-2021).

There are 11 members of cathepsins in humans (B, C, F, H, K, L, O, S, V, W and Z), differentiated based on structure, catalytic functions and type of substrate they act upon (Turk *et al.,* 2001). This huge family of protease have a conserved active site with a triad of residues Cys-Asn-His (Vasiljeva *et al.,* 2007). CTSK is the only protease that is involved in the degradation and cleavage of triple helix and telopeptide of type 1 collagen and matrix protein during resorption (Garnero *et al.,* 1998; Dai *et al.,* 2020). Clinical presentations of patients include short stature, frequent fractures, respiratory problems, strabismus, obstructive sleep apnea (OSA) (>60%) and 48% of patients with OSA between age range of 5 to 10 years may require non-invasive ventilation , stubby hands and feet with dysmorphic nails and also dental issues (increased caries , hypodontia, enamel hypoplasia and malocclusion), neuronal cognition is generally not involved in these patients with exception of mild psychomotor abnormalities in 30% of affected individuals (Turan, 2014; Otaify *et al.,* 2018; Bizaoui *et al.,* 2019). The fracture rate in pycnodysostosis patients is really high on average up to 0.2 fractures per year and first fracture within first decade of life (Bizaoui *et al.,* 2019). Incomplete remodeling of osteoclasts, fracture healing ability is delayed in these patients and surgical fixation is also another complicity due to sclerotic bones and narrow medullary canal which causes intraoperative iatrogenic fracture (Grewal *et al.,* 2019). Reported average height in adults pycnodysostosis patients in males is <150 cm and in females is 130-134 cm (Bizaoui *et al.,* 2019).

CTSK is encoded by gene *CTSK*, which is mapped on chromosome 1q21, with 8 exons translated into 329 amino acid protein. And *CTSK* predominantly expressed in the osteoclasts where the number of osteoclasts remain same as healthy individuals but the region of demineralized bones surrounding those osteoclasts (Everts *et al.,* 1985; Gelb *et al.,* 1995; Polymeropoulos *et al.,* 1995). In 50% of the patients, growth hormones are deficient and almost all of them have low levels of IGF1. When growth hormones are administered to such patients the insuline like growth factors 1 (IGF1) level restored satisfactorily (Rothenbuler *et al.,* 2010).

Few studies from Pakistan had previously shown the genetic variants from cohort including p. Ala277Val, p. Gly243Glu and p. Arg46Trp (Naeem *et al.,* 2009; Khan *et al.,* 2010; Kausar *et al.,* 2017).

Diagnosis of CTSK involve clinical presentations like short stature and hands and feet structure, radiological findings confirm the increased bone density and recurrent multiple fractures and finally the genetic testing (Mujawar *et al.,* 2009; Arman *et al.,* 2014; Baglam *et al.,* 2017). But pycnodysostosis is confusing with other osteopetrosis syndromes so molecular genetic testing is the only proven way to exclude other osteopetrosis syndromes overlapping (Sayed *et al.,* 2021).

There is no cure or treatment strategy currently available for CTSK patients other than regular dental hygiene and tooth extraction and avoid fractures, however novel therapies are in trials to find some better option for treating CTSK (Fonteles *et al.,* 2007). Early hematopoietic stem cell transplant (HSCT) could be one of the successful therapies for pycnodysostosis patients (Stark and Savarirayan, 2009). In a few reports, the treatment of some patients with growth hormones also had shown some betterment in growth and attaining height near to a normal length (Verma and Singh, 2020; Sayed *et al.,* 2021).

1.8. Chediak-Higashi Syndrome

Chediak-Higashi syndrome (CHS, OMIM#214500) with atypical large granules in leukocytes and was firstly reported by a Cuban pediatrician Bagues Cesar, and later by Chediak and Higashi (Bagues-Cesar, 1943; Chediak, 1952; Higashi, 1954). CHS is an exceptionally rare autosomal recessive disorder with less than 500

cases reported in the previous research studies, showing multiple degrees of oculocutaneous albinism, neutropenia, pyogenic infections, late neurological manifestations and many more. About 50-85% of cases showed accelerated phase called as hemophagocytic lymphohistiocytosis (HLH), caused by the enlargement of lysosomes and lysosome related organelles (LROs) leading to cytotoxicity of thymus (T) and natural killer (NK) cells, which if left untreated can be lethal (Kaplan *et al.,* 2008; Lozano *et al.,* 2014; Sharma *et al.,* 2020).

The gene involve in causing this disorder is *CHS1* also termed as *LYST* (OMIM#606897) which is 87.9% homologous with *Lyst*, a murine gene for beige that is a developed mouse model for human *CHS1* gene, mapped on chromosome 1q42-44 with total 53 exons among them 51 are coding and translated into a protein of 3801 amino acid (Barbosa *et al.,* 1996; Nagle *et al.,* 1996). Which encodes a lysosomal trafficking regulator whose exact function is not known yet, but it is thought to involve in the regulation of lysosomal-related organelle size, fission, and secretion (Perou *et al.,* 1996; Durchfort *et al.,* 2012). LYST is highly expressed in the bone marrow, spleen, cerebellum, and thymus. Its carboxy termini contains three important domains: Pleckstrin homology domain (PH), Beige and Chediak-Higashi domain (BEACH) and tryptophan aspartic acid-40 (WD-40) domain (Nagle *et al.,* 1996; Sharma *et al.,* 2020). Patients with CHS were diagnosed during the first decade of their life due to multiple organs and system failure. Mostly death occur during the first decade of their life. Hypopigmentation is also seen in the CHS patients particularly in hair, eyes and skin. Hairs appears to be like blonde, gray, or white with a distinguishing metallic or silvery sheen where pigment clumps are scattered throughout hair shaft but are not properly transferred to the epithelial cells, causing hypopigmentation. Eyes have nystagmus, strabismus, photophobia and decreased visual acuity. Iris and retina have lower levels of pigmentation (Zalickson *et al.,* 1967; Kaya *et al.,* 2011; Ho *et al.,* 2013; Lozano *et al.,* 2014; Sharma *et al.,* 2020). Position of LYST is important to determine the number and size of lytic granules in the natural killer cells (Gil-Krzewska *et al.,* 2016). To the best of my knowledge there was not a single study reported from Pakistani population related to Chediak-Higashi syndrome.

Diagnosis for Chediak-Higashi syndrome is made through laboratory evaluation, imaging studies, histological findings (Sánchez-Guiu *et al.,* 2014).

Previously, the criterion of oculocutaneous albinism was one of the flags to identify the CHS, which now is challengeable because some patients have no signs of oculocutaneous albinism (Introne *et al.,* 2016).

The best known and effective treatment option for HLH or accelerated phase is hematopoietic stem cell transplant (HSCT), although the efficiency of HSCT for delaying or stopping the progressiveness of the disease is not reported yet (Eapen *et al.,* 2007; Kaplan *et al.,* 2008; Jackson *et al.,* 2013). According to recent research, an advanced technique known as haploidentical HSCT (haplo-HSCT), has now evolved with more effective outcomes against CHS (Sachdev *et al.,* 2021).

In a most recent study during the pandemic of COVID-19, a case of a 6-week-old infant was reported with SARS-CoV-2 infection and lateral diagnosed with HLH phase of Chediak-Higashi syndrome. Despite of immunosuppression with HLH-2004, the infant recovered from the SARS-CoV-2 infection suggesting that the patients of HLH are more prone to the infection, but early accurate diagnosis can be of real help to prevent the mortality due to infection (Lange *et al.,* 2021).

1.9. Griscelli Syndrome

Griscelli syndrome (GS), was discovered by Griscelli in 1978, is an autosomal recessive disorder (Griscelli *et al.,* 1978). The most common clinical profiles of the patients showed partial albinism, hepatosplenomegaly, cytopenia, immunodeficiencies followed by frequent fever, silvery grey hair and bronze skin upon exposure to sun (Griscelli *et al.,* 1978; Castaño-Jaramillo *et al.,* 2002).

Formerly, termed as Elejalde syndrome, it is caused by the mutations of any member of *MYO5A-RAB27A-MLPH* tripartite complex. Each gene is responsible for causing each type of griscelli syndrome i.e griscelli syndrome type 1 (GS1, OMIM#214450), griscelli syndrome type 2 (GS2, OMIM#607624) and griscelli syndrome type 3 (GS3, OMIM#609227) (Pastural *et al.,* 1997; Huizing *et al.,* 2008; Mohammadzadeh *et al.,* 2015). Two genes involved in causing GS are mapped on chromosome 15; *RAB27A* (OMIM#603868) mapping on chromosome 15q15-q21.1 having total 7 exons with 5 coding exons translating into 221 amino acid protein and *MYO5A* (OMIM#160777), located on chromosome 15q21 with 42 exons producing

1880 amino acid long protein and on chromosome 2. While the third member of tripartite complex is *MLPH* (OMIM#606526) located on chromosome 2q37.3, comprises of total 16 exons and among them 15 exons coded the 600 amino acid protein (Ensembl Genome Browser 104, Accessed on 19-Oct-2021). In GS1, there is hypopigmentation accompanied by severe neurological impairment, caused by mutations in gene myosin 5A (*MYO5A*) that encode an organelle motor protein Myosin Va which has an important role in neuron functioning (Langford *et al.,* 1998; Bahadoran *et al.,* 2001; Gironi *et al.,* 2019). GS2 has same albinism with immune defect that leads to hemophagocytic syndrome (HS). It is caused by the deficiency of a small GTPase protein Rab27a encoded by gene *RAB27A*, which is involved in the regulation of intracellular regulated secretory pathway (Ménasché *et al.,* 2000; Pastural *et al.,* 2000; Wilson *et al.,* 2000; Stinchcombe *et al.,* 2001). The third type is caused by the defect in gene melanophilin (*MLPH*), which interacts with the *RAB27A* gene as an effector protein (GTPase) for protein trafficking and has expression in melanocytes and lymphocytes but not in the cells of brain. For pigment distribution in the melanocytes *RAB27A* interacts with *MYO5A* and *MLPH* which clearly explains why it causes the hypopigmentation of melanocytes (Matesic *et al.,* 2001; Hume *et al.,* 2002; Strom *et al.,* 2002; Ménasché *et al.,* 2003; Krzewski and Cullinane, 2013; Shah *et al.,* 2016). GS2 represents 0.1% of all inborn errors of immunity and 2.2% of all disorders related to dysregulations (El-Helou *et al.,* 2019). There is not a single published study from Pakistan or Pakistani patients until now.

Diagnosis for Griscelli syndrome can be made through laboratory and microscopic analysis of hair shaft which show uneven small and large clumps of pigment. And also, skin biopsy shows irregular and dense pigment granules trapped in the basal epidermal layer and not transferred to the keratinocytes (Ridaura-Sanz *et al.,* 2018). The GS patients also have immune deficiency with low levels of immunoglobulins and isohaemagluttinins (Griscelli *et al.,* 1978; Castano-Jaramillo *et al.,* 2021). And finally, the genetic testing for GS patients (Çağdaş *et al.,* 2012).

Only treatment option for GS2 is only bone marrow transplant (Ménasché *et al.,* 2003), but no other effective treatment option is currently available or in practice for GS patients except for some medications for HLH phase (Marsh and Haddad, 2018; Kim *et al.,* 2019; Henter *et al.,* 2020).

1.10. Pompe Disease

Pompe disease (OMIM#232300) was firstly discovered by a Dutch pathologist, Johannes Cassianus Pompe while performing an autopsy of a seven-month-old child. The disease is also termed glycogen storage disease type II (GS II). The phenotypes include severe myopathies of cardiac and skeletal muscles (Pompe, 1932). There are 1:8684-138,000 live births worldwide (Spada *et al.,* 2018).

Pompe disease is caused by the deficiency of enzyme acid alpha-glucosidase (GAA, OMIM#606800) involved in the degradation of glycogen inside lysosome. Accumulation of glycogen in muscle cells leads to lethal myopathies. GAA enzyme is encoded by gene *GAA* located on chromosome 17q25, comprising of 20 exons translated into 952 long amino acids protein (Hoefsloot *et al.,* 1990). Enzyme acid alpha-glucosidase was discovered firstly in 1963 by a Belgian biochemist Henry-Gery Hers. Dr. Hers was the first who triggered the search for lysosomal storage disorders by predicting that like Pompe disease other deposition disorders can also be explained on the principle of insufficiency/absence of other lysosomal enzymes. Pompe disease is the distinctive disease in LSD group because it was first among all sub-types to be documented as storage disorder (Hers, 1963; Kohler *et al.,* 2018). Proper folding and glycosylation of GAA in endoplasmic reticulum is critical for transport to Golgi body where it attaches a mannose 6-phosphate (M6P), a targeting signal for lysosome. The phosphorylated M6P then adds the binding of enzyme with M6P receptor (M6PR). This bounded enzyme is then pinched off from the Golgi in an endosome and transported to the lysosome where enzyme dissociates from receptor and enters lysosome while receptor cycle back to carry other enzymes. After synthesis GAA undergo extensive post-translational modifications i.e., cleaved proteolytically at carboxyl and amino termini, enroute from rough endoplasmic reticulum to lysosome adding the efficiency of enzyme by making it catalytically active (Hermans *et al.,* 1993; Wisselaar *et al.,* 1993; Moreland *et al.,* 2005). There are two subsets of Pompe disease, one with severe cardiac myopathies, obstruction of left ventricular outflow, muscle weakness, hypotonia and respiratory distress, with age of onset less than 12 months and without proper medical care patients could not survive above one year of age termed as classic infantile onset Pompe disease (IOPD), while second subset have all similar features except for severe cardio myopathies termed as non-classic IOPD

(Van den Hout *et al.,* 2003; Kishnani *et al.,* 2006). The reported variants from Pakistani population include p. Arg854Gln and p. Asn925His for causing IOPD (Ullah *et al.,* 2020).

Diagnostic tools for Pompe disease patients include serum level measurements for creatin kinase (CK), aminotransferases (AST), alanine aminotransferases (ALT), lactate dehydrogenase (LDH), urinary glucose tetra saccharide (Glc_4), chest x-ray, electrocardiogram (ECG), echocardiography (Echo), magnetic resonance imaging (MRI), muscle biopsy and leukocytes enzyme levels measurements (Kishnani *et al.,* 2006; Burton *et al.,* 2017).

Treatment options used for Pompe disease patients include enzyme replacement therapy (ERT), substrate reduction therapy (SRT), genetic suppression of autophagy, stimulation of lysosomal exocytosis and gene therapy (Douillard-Guilloux *et al.,* 2008; Douillard-Guilloux *et al.,* 2010; Hahn *et al.,* 2014; Settembre and Ballabio, 2014; Kohler *et al.,* 2018).

1.11. Mucolipidosis type II/III

Mucolipidoses (ML), a very rare disorder affecting 2.5 to 10 individuals per 1,000,000 live births and inherited as autosomal recessive disorder (Maroteaux and Lamy 1966; Cathey *et al.,* 2010).

It has three sub-types depending upon the sub-unit of enzyme that lacks its functional activity. ML II (OMIM#252500) also known as cell disease, ML III also termed as pseudo-Hurler dystrophy has ML III α/β (OMIM#252600), ML III γ (OMIM#252605) (Leroy and Martin 1975; Raas-Rothschild *et al.,* 2000, 2012; Cathey *et al.,* 2010). Mutations in any of the above-mentioned sub-types reduced the activity of hexameric membrane bounded enzyme UDP-N-acetyl glucosamine-1-phosphotransferase (GlcNAc-PTase) comprise of three sub-units namely α2, β2 and γ2 (Reitman and Kornfeld, 1981; Raas-Rothschild *et al.,* 2000). The membrane bound α/β sub-units are encoded by *GNPTAB* gene (OMIM#607840) located on chromosome 12q23.3 having 21 exons encoding 1256 amino acids and soluble γ-sub-unit is encoded by gene *GNPTG* (OMIM#607838) located on chromosome 16p13.3 with 11 exons translated into a 305-residue protein. The soluble γ subunit binds with

α subunit directly and enhances its enzymatic activity to transport M6P (Di Lorenzo *et al.,* 2018). Inside Golgi bodies, GlcNAc-PTase is responsible for the first step of phosphorylation of enzyme-bounded mannose residues into mannose-6-phosphate. M6P serves as the recognition marker for the newly synthesized lysosomal proteins. If this marker is absent, the newly synthesized proteins are released in the cytoplasm and wasted out without performing their functions to degrade the substrates resulting in their accumulation leading to multiple organ failure (Reitman and Kornfeld, 1981). Phenotypes of all three sub-types vary ranging from severe to mild in ML III α/β and milder in ML III γ. Some died in early childhood while some survive to the adulthood with skeletal deformities (Maroteaux and Lamy, 1966; Raas-Rothschild *et al.,* 2000; Cathey *et al.,* 2010; Liu *et al.,* 2014; Yang *et al.,* 2017; Tüysüz *et al.,* 2018). Some of the common clinical profiles found in all patients of ML II/III include craniosynostosis, osteopenia, rickets, neonatal hyperparathyroidism, deformed long tubular bones, joint stiffness, thoracic deformity, kyphosis, dysplasia of hip bone, clubfeet, joint contractures, waddling gait, carpal tunnel syndrome and spinal deformity. Most of the times ML III is misdiagnosed as juvenile arthritis or progressive pseudo rheumatoid arthropathy because of the presenting clinical features (Pohl *et al.,* 2010; Alegra *et al.,* 2014; Yang *et al.,* 2017; Oussoren *et al.,* 2018).

Some studies published pathogenic variants causing ML from Pakistani population including p. Gln9Ter, p. Ala387Val in *GNPTAB* gene in compound heterozygous form and a four base pair deletion c.478_479insTAGG (p. Ala160*fs*) in *GNPTG* gene (Khan *et al.,* 2018; Khan *et al.,* 2020).

Clinically ML II/III can be diagnosed through enzyme level measurements from plasma, dried blood and cultured media of fibroblasts. Another diagnostic tool is single-chain antibody fragments bind and identify M6P residues through western blotting (Pohl *et al.,* 2010; Alegra *et al.,* 2014; Velho *et al.,* 2014).

Currently there is no treatment available for treating ML II/III (α, β, γ). Like other rare LSDs e.g., Mucopolysaccharidoses (MPS), natural history and data collected from various ethnicities will help to identify therapies for future to treat such complex and rare disorders (Oussoren *et al.,* 2018).

1.12. Aim and Objectives

The aim of this study was to perform molecular analysis of lysosomal storage disorders in Pakistani population.

The objectives of this study are:

1. Recruitment of approximately 50 LSD affected families from local population through expert pediatricians, diagnosed on clinical basis.

2. Molecular genetic screening of enrolled families using Sangers and Next Generation Sequencing to identify causative variants.

3. In-silico analysis of identified variants to explore their effect on structure/function of protein and to establish genotype-phenotype correlation.

4. To provide genetic diagnosis to affected families to help patients to get accurate treatment (if any) and genetic counseling to affected families.

2. MATERIALS AND METHODS

❖ SUMMARY

To perform this study, following were the major steps followed to complete the study:

1. Field Work

This step includes patients' recruitments on the basis of their presenting clinical profiles, phenotypes and basic medical evaluations through tests like blood complete picture, radiological findings, bone marrow test, urine testing etc.

2. Laboratory Work

Firstly, DNA was extracted. The quantity and quality of the extracted DNA was measured through NanoDrop. Nine families were selected to subject to Sanger sequencing based on availability of the facility, their data analysis, *In-Silico* analysis and the data was published (Gul *et al.,* 2020; Gul *et al.,* 2021).

Remaining 36 families were subjected to Whole Genome Sequencing during International Research Support Initiative Program (IRSIP) sponsored by the Higher Education Commission (HEC), Pakistan for six months stay at the University of Copenhagen Denmark, and their data analysis, *In-Silico* analysis were performed and publications are in progress.

2.1. Ethical Approvals

This study was approved by the Institutional Review Board (IRB)/Bioethics Committee (BEC) of Quaid-i-Azam University, Islamabad, Pakistan with reference to meeting number BEC-FBS-QAU2019-198 (Annexure 1 {1.4}).

2.2. Patients' Recruitment

The child physician helped to diagnose each type with respective diagnostics techniques based on radiological and biochemical analysis. All patients were informed before taking their details, pictures and blood samples, and written informed consents

were obtained to participate and publish their required details as per Rules of Helsinki (World Helsinki Association, 2013) (Annexure 1 {1.1, 1.2, 1.3}). Detailed pedigrees were drawn from the gathered information provided by the parents/guardians of the patients by using HaploPainter 1.043®. Pictures of the patients with specific diagnostic features and medical history containing all the helpful data were also collected. 5-10 ml of peripheral blood samples were collected from patients and their accompanying parents/guardians/siblings in a 10 ml vacutainer tubes (Becton Dickinson and Company (BD), USA, vacutainer K2 EDTA 18 mg). All collected blood samples were stored at 4°C in the Lab of Molecular Biology, Department of Zoology, Quaid-i-Azam University, Islamabad, Pakistan, till DNA extraction.

2.2.1 Inclusion and Exclusion Criteria

All families have been diagnosed through clinical examination and some basic available tests like X-ray, complete blood count (CBC), bone marrow test (BMT), computer tomography (CT) scan, magnetic resonance imaging (MRI) were performed and for few families enzyme analysis was also performed (Table 3.6). The families recommended as strong candidates for LSDs after clinical confirmation by expert child physician have been included in this study.

2.3. DNA Extraction

The protocols used for the extraction of genomic DNA (gDNA) from collected blood samples were modified and mixed protocols of Proteinase K-Phenol-Chloroform methods (Miller *et al.,* 1988; Sambrook and Russel, 2006). The steps followed during the experiments were:

- **Day 1**

❖ The blood samples were thawed and poured in a 50 mL falcon tubes (SPL Lifesciences).

❖ A washing buffer was prepared by using Tris-Ethylene Diamine Tetra Acetate (EDTA) (TE) buffer (10mM Tris HCl, 2mM EDTA, pH 8.0). Total 40 mL volume was made by pouring this washing buffer in the falcon tube containing blood sample (5 mL blood + 35 mL TE buffer and vice versa).

❖ Tubes were centrifuged (BECKMAN COULTER Allegra X-30R Benchtop Centrifuge, USA) at 3000 rounds per minute (RPM) for 20 minutes at 20° C. The supernatant was discarded by using the glass pipettes with a manual sucker up to maximum of 20 mL out of 40 mL so that the white blood cells (WBCs) pellet should not be disturbed.

❖ Again 20 mL of TE washing buffer was added to make final volume of 40 mL and did a second centrifugation at 3000 rpm for 20 minutes at 20° C, to collect the remaining suspended white blood cells (WBCs) and washed away the red blood cells (RBCs). This time up to 30 mL of supernatant was discarded by using glass pipettes and manual sucker.

❖ These washing steps were repeated for three to four times to eliminate any traces of RBCs and clearing out the pellet of WBCs.

❖ In last washing step all of the supernatant was discarded and in the pellet 25 µL of proteinase K (PK) (20 mg/ mL), 50 µL of 20% sodium dodecyl sulfate (SDS) and 3 mL of Tris+NaCl+EDTA (TNE) buffer (10 mM Tris HCl, 400 mM NaCl, 2 mM EDTA) was added and kept incubated at 37° C for overnight digestion.

- **Day 2**

❖ On second day of gDNA extraction, the pellet was completely dissolved out so, for precipitation of proteins 1 ml of 6 molar (M) NaCl was added and after vigorous shaking the tubes were placed immediately on ice for 15 minutes.

❖ A mixture of phenol:chloroform:isoamylalcohol with a proportion of 25:24:1 was added to the falcon tubes in a volume of 1 mL each.

❖ Another round of centrifugation was performed for 20 minutes at 3000 rpm and 20° C to separate the layers of precipitated proteins, lipids, salts, cellular debris in the lower and gDNA in upper layer.

❖ Upper layer was separated in another 50 mL falcon tube carefully. Added approximately 5 mL of chilled isopropanol to the separated layer and mixed thoroughly to precipitate out the gDNA thread. The tubes were then centrifuged again at 3000 rpm for 20 minutes at 20° C to settle down the suspended gDNA thread into a pellet at the bottom of the falcon tube.

❖ Supernatant was completely discarded off and added 70% ethanol to washout any remaining traces of solvents used previously. Centrifuged at 3000 rpm, 20° C for 20 minutes.

❖ Discarded the ethanol and let the pellet in falcon air dry at room temperature for 3 to 4 hours. Once ethanol was completely dried, dissolving Tris buffer (0.2 mM) was added in a suitable quantity ranging from 300-500 µL and incubated at 37° C for a whole night to get a suspension with completely dissolved gDNA.

- **Day 3**

❖ On the last day of gDNA extraction, the dissolved gDNA was heat shocked at 70° C in a water bath for 1 hour to deactivate any remaining nucleases and let it cool at room temperature. After a short spin of 1000 rpm for 2 minutes at 20° C, the stock gDNA was shifted to screw caps (1.5 mL Axygen™, USA).

❖ Extracted DNA was quantified by using nanodrop (Titertek, BERTHOLD, Germany) and gel electrophoresis. Then stock samples were stored at -20° C for future use.

2.3.1. Agarose Gel Electrophoresis for gDNA Quality and Quantity Assurance

In order to check the quality and concentration of the extracted gDNA, agarose gel electrophoresis was performed as follow:

❖ 1X Tris-Boric acid-EDTA (TBE) buffer was prepared from the stock solution of 10X TBE buffer (1 mM of Tris borate, 0.5 mM EDTA, pH 8.0) by taking 5 mL of 10X TBE buffer and adding 45 mL of distilled water and made a final volume of 50 mL 1X TBE buffer.

❖ For gDNA 1% of agarose gel was prepared by putting 0.5 g of agarose (Ultrapure ™ Agarose-1000, ThermoFisher Scientific, USA), in 50 mL of 1X TBE buffer in a 100 mL conical glass flask and heated it in a microwave oven for 3 minutes to melt the gel powder and prepared a suspension.

❖ After taking out from microwave, 3 µL of 1% ethidium bromide was added and mixed thoroughly.

❖ The suspension was poured into a gel casting apparatus with combs to make wells for sample loading, gradually and slowly to avoid any bubble formation (Hinder the gDNA running). And allowed gel to solidify at room temperature for 30 minutes.

❖ When gel solidified, combs and stoppers were removed carefully and placed the gel inside of gel tank (CS-3000V, Cleaver Scientific Limited, UK) filled with 1X TBE buffer.

❖ For gDNA samples loading, 5 µL of loading dye (Bromophenol blue 0.25%, sucrose 40%) was used to locate our samples during run. 5 µL of sample was mixed in each 5 µL of loading dye and loaded the whole mixture in each well.

❖ As, amount of florescence is proportional to amount of DNA present. Therefore, in one well a DNA standard (with known DNA concentration) was also loaded to compare each test sample with standard for visual assurance of DNA quantity.

❖ Samples were run at 120 Volts, 120 Amperes for thirty minutes.

❖ Once gel running was completed, the gel was removed from the tank and placed on the stage of gel documentation system (Biometra, Gottingen, Germany) to visualize the samples under ultraviolet radiation with complete care and protection.

2.4. Genotyping

When all samples were quantified then next step was the genotyping of all samples to identify the disease-causing variants. For genotyping several steps has been performed as mentioned in detail below:

2.4.1 The Sanger Sequencing

Nine families were selected for polymerase chain reaction (PCR) based amplification and Sanger sequencing in Molecular Biology Lab at Quaid-I-Azam University, Islamabad, Pakistan, based on their monogenic status, which include six MPS I and three Gaucher disease families because for each disease a single gene is reported i.e *IDUA* and *GBA*, respectively. Beside this, both genes were not too long to scan through Sanger sequencing as Sanger sequencing involves individual exon scanning through forward and reverse primers which is time taking process.

2.4.2 Primer Designing and PCR Amplification

For primer designing, the whole sequence of *IDUA* and *GBA* genes including Untranslated regions (3', 5' UTRs), introns and exons, were obtained through Ensembl Genome Browser (https://asia.ensembl.org/index.html) and the transcripts used for primer designing for *IDUA* was ENST00000247933.9 and for *GBA* was ENST00000327247.9, for getting the best primers Primer 3 Web (Version 4.1.0) was

used (https://primer3.ut.ee/), Then blast like alignment tool (BLAT) (http://genome.ucsc.edu/cgi-bin/hgBlat) was applied in order to confirm the specificity of each primer set to respective gene as it is used to check the single hit for each set of primer. For checking the product size of primer, melting temperature and GC contents, the *In-silico* PCR (https://genome.ucsc.edu/cgi-bin/hgPcr) was performed. In order to check the properties of the primers including their melting temperatures, Primer stats was performed (https://www.bioinformatics.org/sms2/pcr_primer_stats.html). The primers set used for *IDUA* are given below in the table 2.1 and *GBA* are given in table 2.2.

Table 2.1: List of primers used for the amplification of exons of *IDUA* gene.

Primer no.	5'-3' sequence	Primer length (bp)	Product size (bp)	GC contents	Tm
1 F	CAACTCACATAGCACCAACG	20	532	50	57.8
1 R	TCTCTGAAGCTCTCACAGTGC	21		52	58.5
2 F	GGCTTGAACGTGTGTGTCA	19	274	52	59.3
2 R	TCCCATCTGTGCCTCTGTAA	20		50	59.2
3,4 F	CATACCAGGCCTTCATAGGG	20	708	55	59.4
3,4 R	CCAACCTATCCCTTGTCACC	20		55	59.3
5 F	TGAGGCGAGATTCACCTGT	19	288	52.6	59.4
5 R	CGAGCAGGCATCGTAGTAGTT	21		52.4	59.6
6,7 F	CAGACCACCACGACTTTGAC	20	841	55	59.2
6,7 R	GGTAGCTCAGGAAGGCATTG	20		55	59.8
8 F	CGGACACCCCCATTTACA	18	582	55.6	60.2
8 R	TCCCCTTGGTGAAGGAGTC	19		57.9	60
9,10 F	CTGGGGACTCCTTCACCAA	19	821	57.9	61.1
9,10 R	ACGTGCACCAGCAAAAGC	18		55.6	61
11, 12 F	TACCTGGACAACGGGCTCT	19	872	57.89	60.7
11,12 R	CAGCCTCATTCCCTCAAGC	19		57.89	60.9
13,14 F	GCCTTGTTCTTGGCCTGA	18	592	55.56	59.9
13, 14 R	TCAACCCGAGACTGACCTGT	20		55	60.7

A: adenine, bp: base pair, F: forward, GC: guanine cytosine, R: reverse, T: thymine, Tm: melting temperature

Table 2.2: List of primers used for the amplification of exons of *GBA* gene.

Primer no.	5'-3' sequence	Primer length (bp)	Product size (bp)	GC contents	Tm
1,2 F	GGGAAGCCGGAATTACTTG	19	850	52.63	59.51
1,2 R	GAGCCAAAATTGCACCACT	19		47.37	58.71
3,4 F	ACCGTGTTCAGTCTCTCCTAGC	22	652	54.5	62.9
3,4 R	GACAGAATGGGCAGAGTGAG	20		55	59.9
5,6 F	AGGAGCCCAAGTTCCCTTT	19	806	52.63	60
5,6 R	CTGATGGAGTGGGCAAGATT	20		50	60
7 F	AGGCTGTTCTCGAACTCCTG	20	496	55	59.6
7 R	GGGAATGGTGCTCTAGGAATC	21		52.38	59.92
8 F	GTTGCATTCTTCCCGTCACC	20	367	55	62.8
8 R	CTGGACAGGAAGGGCTTCTG	20		60	62.3
9 F	CTCTCCCACATGTGACCCTTA	21	399	52.3	59.9
9 R	GCCTCCATGGTGCAAAAGGGG	21		52.9	59.1
10, 11 F	GCAGAAAAGCAGGGTCAGTG	20	575	55	60.98
10, 11 R	TGCTGTGCCCTCTTTAGTCA	20		50	59.59

A: adenine, bp: base pair, F: forward, GC: guanine cytosine, R: reverse, T: thymine, Tm: melting temperature

For amplification of both genes primers through polymerase chain reaction (PCR) the conditions used were described as follow:

- The reaction mixture was prepared in a 200 µL PCR tubes (Axygen, USA) containing a total volume of 25 µL.
- Reagents used were from SOLIS BIODYNE KIT, Estonia.
- Mixture contained 2 µL of extracted DNA sample, 10X Taq buffer $(NH_4)_2SO_4$ 2.5 µL, 2.5 mM $MgCl_2$ 2.5 µL, 10 mM dNTPs 2.5 µL, 10pmol/ µL of forward and reverse primers 0.37 µL each, 5U/ µL Taq Polymerase 0.3 µL and at the end total volume was raised to 25 µL by putting ultra-pure PCR water 14.46 µL.
- The mixture was vortexed and short spinned to mixed and collect all ingredients well. PCR tubes were then placed in thermo cycler (BioRad T100) to amplify the required exons of both genes. The specific thermal cycling used for each primer pair of *IDUA* and *GBA* are mentioned in table 2.4 and 2.5 respectively.
-

Table 2.3: Thermal cycling conditions for *IDUA* gene primers pairs.

S. no	Primer pairs	Initial denaturation °C	First 10 cycles of amplification °C	Second 30 cycles of amplification °C	Extension °C	Hold ∞
1.	1 F 1 R	95	57.5	57.5	72	25°C
2.	2 F 2 R	95	68	58	72	25°C
3.	3,4 F 3,4 R	95	68	58	72	25°C
4.	5 F 5 R	95	68	58	72	25°C
5.	6,7 F 6,7 R	95	70.5	60.5	72	25°C
6.	8 F 8 R	95	67	57	72	25°C
7.	9,10 F 9,10 R	95	65	69.5	72	25°C
8.	11,12 F 11,12 R	95	66	65	72	25°C
9.	13,14 F 13,14 R	95	69	59	72	25°C

F: forward, R: reverse

Table 2.4: Thermal cycling conditions for *GBA* gene primers pairs.

S. no.	Primers pairs	Initial denaturation °C	First 10 cycles of amplification °C	Second 30 cycles of amplification °C	Extension °C	Hold ∞ °C
1.	1,2 F 1,2 R	95	68	58	72	25
2.	3,4 F 3,4 R	95	68	58	72	25
3.	5,6 F 5,6 R	95	68	58	72	25
4.	7 F 7 R	95	68	58	72	25
5.	8 F 8 R	95	68	58	72	25
6.	9 F 9 R	95	68	58	72	25
7.	10,11 F 10,11 R	95	68	58	72	25

F: forward, R: reverse

2.4.2. Agarose Gel Electrophoresis for PCR Products

The amplified products were quantified once again through agarose gel electrophoresis as follows:

- 1X Tris-Boric acid-EDTA (TBE) buffer was prepared from the stock solution of 10X TBE buffer (1 mM of Tris borate, 0.5 mM EDTA, pH 8.0) by taking 5 mL of 10X TBE buffer and adding 45 mL of distilled water to make a final volume of 50 mL 1X TBE buffer.
- For amplified DNA 2% of agarose gel was prepared by putting 1 g of agarose (Ultrapure ^{TM} Agarose-1000, ThermoFisher Scientific, USA), in 50 mL of 1X TBE buffer in a 100 mL conical glass flask and heated it in a microwave oven for 3 minutes to melt the gel powder and prepared a suspension.
- After taking out from microwave, 3 µL of 1% ethidium bromide was added and mixed thoroughly.
- The suspension was poured into a gel casting apparatus with combs to make wells for sample loading, gradually and slowly to avoid any bubble formation (Hinder the PCR product running). And allowed gel to solidify at room temperature for 30 minutes.
- When gel solidified, combs and stoppers were removed carefully and placed the gel inside of gel tank (CS-3000V, Cleaver Scientific Limited, UK) filled with 1X TBE buffer.
- For amplified PCR samples loading, 5 µL of loading dye (Bromophenol blue 0.25%, sucrose 40%) was used to locate our samples during run. 5 µL of sample was mixed in each 5 µL of loading dye and loaded the whole mixture in each well.
- Loaded samples were run at 120 volts (V), 120 ampere (A) for 30 minutes.
- To check fragment size for each PCR product a 1 kb ladder was also loaded in first well and visualization of photograph confirmed the desired size fragment in each sample.

Once gel running was completed, the gel was removed from the tank and placed on the stage of gel documentation system (Biometra, Gottingen, Germany) to visualize the samples under ultraviolet (UV) radiation with complete care and protection.

2.4.3. Purification and Sanger Sequencing of Amplified Products

Before performing Sanger sequencing, all amplified products were purified to eliminate all unused reagents. For purification WizPrep™ Gel/PCR Purification Mini Kit-WIZBIO, Korea, was used. The steps followed are mentioned below:

- GP binding buffer 100 µL was added to each PCR tube (Axygen, USA), and mixed thoroughly and left at room temperature for 10 minutes.
- This mixture was then transferred to spin columns labelled in advance and centrifuged (BECKMAN COULTAR, Microfuge 16 Centrifuge, USA) at 13000 rpm for 2 minutes.
- Columns were washed with 350 µL washing buffer and centrifuged again. The supernatant was discarded. Same step of washing was repeated again to wash out any impurity left. For removing all washing buffer out of spin columns, a spin of 13000 rpm for additional 2 minutes was performed.
- Spin columns were shifted to 1.5 mL labelled Eppendorfs ®(Autoclaved).
- Elution buffer 30 µL was poured in the spin column and centrifuged for 3 minutes at 13000 rpm, to elute the purified amplified product from spin columns into 1.5 mL Eppendorfs.
- In order to confirm the quantity of purified product, 5 µL of product was loaded onto 2% agarose gel and performed electrophoresis.

After purification the products along with their respective forward and reverse primers were sent for commercial Sanger Sequencing.

2.4.4. Data Analysis and Variants Identification

The sequenced data when received was analyzed through BioEdit 7.2. and the identified variants were confirmed through Mutation Taster2 (https://www.mutationtaster.org/). Further variants were studied through *In-silico* analysis to confirm the pathogenicity of the identified variants by using tools like Sorting Intolerant from Tolerant (SIFT) (https://sift.bii.a-star.edu.sg/), Protein Variation Effect Analyzer (PROVEAN) (http://provean.jcvi.org/index.php), Polymorphism Phenotyping v2 (PolyPhen-2) (http://genetics.bwh.harvard.edu/pph2/), Have (y)our Protein Explained (HOPE)

(https://www3.cmbi.umcn.nl/hope/), Iterative Threading ASSEmbly Refinement (I-TASSER) (https://zhanggroup.org/I-TASSER/), Mutation Prediction Loss of Function (MutPred LOF) (http://mutpred2.mutdb.org/mutpredlof/). I-Mutant2.0 (https://folding.biofold.org/i-mutant/i-mutant2.0.html), Protein Analysis Through Evolutionary Relationships (PANTHER) (http://www.pantherdb.org/), Single Nucleotide Polymorphisms & Gene Ontology (SNPs&GO) (https://snps-and-go.biocomp.unibo.it/snps-and-go/index.html), Predictor of Human Deleterious Single Nucleotide Polymorphisms (PHD-SNP) (https://snps.biofold.org/phd-snp/phd-snp.html), Mutation Assessor (http://mutationassessor.org/r3/), Python module (PMut) (http://mmb.irbbarcelona.org/PMut/), Prediction of Protein Stability Changes for Single Site Mutations from Sequences (MUpro) (http://mupro.proteomics.ics.uci.edu/), Functional Analysis through Hidden Markov Models (fathmm) (http://fathmm.biocompute.org.uk/), Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (https://string-db.org/), ConSurf (https://consurfdb.tau.ac.il/), and Genome Aggregation Database (gnomAD) (https://gnomad.broadinstitute.org/).

2.5. Next Generation Sequencing

The use of NGS is one of the best methods to diagnose complicated and rare disorders like LSDs. During this study 36 families affected with various suspected LSDs sub-types, were subjected to Whole Genome Sequencing to identify their pathogenic variants in either already reported genes or in novel genes.

2.5.1. Whole Genome Sequencing and Variant Calling

Whole genome sequencing was performed by using MGISEQ-T7 (DNBSEQTM), a recently developed sequencing platform (2019) by MGI Tech and Complete Genomics, that uses combined probe anchor synthesis and DNA nanoball for generating short reads at a very large scale. The read length for MGISEQ-T7 is 100 bp paired-end (PE). It has rough coverage of ×30. MGISEQ-T7 has showed the lowest percent of low-quality reads i.e 4.2%, moderate levels of K-mer errors i.e 6.39% and highest ratio of duplication with a percentage of 3.04% (Kim *et al.,* 2020).

Read mapping and genotype calling were performed using an open-source pipeline implemented in PALEOMIX (Schubert *et al.,* 2014). The modified pipeline

and detailed instructions are available at https://github.com/Hansen-Group/Gul2022. Briefly, pre-analyses quality assurance was performed using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC v1.10 (Ewels *et al.,* 2016). Read pairs were subsequently processed using fastp v0.20.1 (Chen *et al.,* 2018) to trim adapter sequences, remove reads with low complexity, and merge and correct overlapping read pairs (with a minimum overlap of 11 bp). Reads were mapped using BWA mem v0.7.17-r1188 (Li, 2013) against the hg38 human reference genome distributed as part of the GATK (McKenna *et al.,* 2010) resource bundle (https://gatk.broadinstitute.org/hc/en-us/articles/360035890811-Resource-bundle), including alternative and decoy contigs. Mapped reads were post-processed using samtools v1.11 (Danecek *et al.,* 2021) and the 'bwa-postalt.js' script included in the BWA-kit which post-processes alignments mapping to alternative contigs such as HLA variants. Duplicate, unmapped, and orphaned reads, and secondary and supplementary alignments were filtered. The resulting BAMs were recalibrated using GATK ApplyBQSR trained using dbSNP release 151 (Sherry *et al.,* 2001) for known variable sites.

Genotypes (GVCFs) were called for each sample using GATK HaplotypeCaller, merged, and called using GATK GenotypeGVCFs. Variant recalibration was carried out separately for SNPs and indels using GATK ApplyVQSRNode, trained using positive lists from the GATK resource bundle. Features used in the calibration were ExcessHet, DP, MQ, QD, SOR, FS, ReadPosRankSum, MQRankSum, and BaseQRankSum (only for SNPs). We kept only variables within the upper 98.0 tranche for indel and 99.6 tranche for SNPs. The resulting VCFs were annotated using VEP v104 (McLaren *et al.,* 2016) using the Ancestral Allele, ExACpLI, GERP Conservation Scores, and LOFTEE v1.0.3 (Karczewski *et al.,* 2020, https://github.com/konradjk/loftee) plugins. Additional custom annotation was included as described in the github repository.

2.4.6. Data Analysis and Variant Prioritization

All variants were evaluated as per ACMG guidelines (Richard *et al.,* 2015). Variants that were classified to be pathogenic, likely pathogenic or uncertain significance according to ACMG guidelines were prioritized.

Runs of homozygosity (ROH)-based inbreeding coefficient (F) was calculated using PLINK (v1.90b6) on a dataset including both affected and unaffected individuals as well as 77 additional individuals with Pakistani origin that are part of an ongoing LSDs study (unpublished) for better allele frequency estimates. Results are only shown for the individuals in present study. For ROH calculation, only variants with minor allele frequency (MAF>5%), and no missingness were used. ROHs were defined as having more than 100 variants, a total length larger than 1MB, on average at least 1 variant per 50 kB, consecutive variants less than 1 MB apart, and at most 1 heterozygous call per window of 50 variants. The total length of ROHs was then divided by the total length of the genome (3 GB).

Principal component analysis was performed on by first merging with the 1000 genomes project (https://www.nature.com/articles/nature15393). Non-overlapping sites and sites with a MAF difference between LSD-samples and the 1000 genomes South Asian super population of more than 25% were removed. The PCA was calculated using PLINK -pca function on remaining overlapping sites with overall MAF > 0.05, no missingness, and LD-pruned (window size of 1MB, step size of 1, R2-threshold of 0.8).

Non-synonymous variants including missense, nonsense, frameshift and indels in coding and splice sites changes in non-coding regions with a MAF of less than 0.01 in large, outbred populations (gnomAD exomes of South Asian population) were considered as putative causal candidates. Further prioritization of these rare variants was based on segregation with disease phenotypes within respective family. As a first phase for WGS data analysis, already reported genes causing MPS were screened for pathogenic variants and during second phase, unsolved families were checked for variants in other LSDs related genes.

In autosomal recessive disorders, variants that were homozygous in affected individuals and in heterozygous state in their parents were prioritized based on recessive inheritance pattern of the disease and consanguinity among parents revealed by their pedigrees. The lists of heterozygous variants in each family were searched for compound heterozygous or digenic inheritance. In X-linked disorders, pathogenic variants on X-chromosome segregated from the mother were extracted. In families

where genotypes of parents were available, exomes of each family were analyzed for dual molecular diagnosis. As a reference population, more than a 100 healthy control samples from the general Pakistani population were analyzed for the prevalence of potentially causal variant as per ACMG guidelines.

3. RESULTS

3.1. Summary

During this study from December 2017 to December 2019, total forty-five families affected with various types of LSDs had been recruited from different regions of Pakistan including Punjab, Khyber Pakhtunkhwa, Balochistan, Islamabad Capital Territory and Azad Jammu and Kashmir (Table 3.1, Figure 3.1). Six families affected with MPS type I and three families affected with Gaucher disease have been selected to scan through Sanger Sequencing at ***Lab of Molecular Biology, Department of Zoology, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan***. While thirty-six families have been subjected to Whole Genome Sequencing at ***Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark*** through International Research Support Initiative Program (IRSIP) provided by the **Higher Education Commission (HEC), Pakistan** (Annexure 2 {2.1, 2.2}).

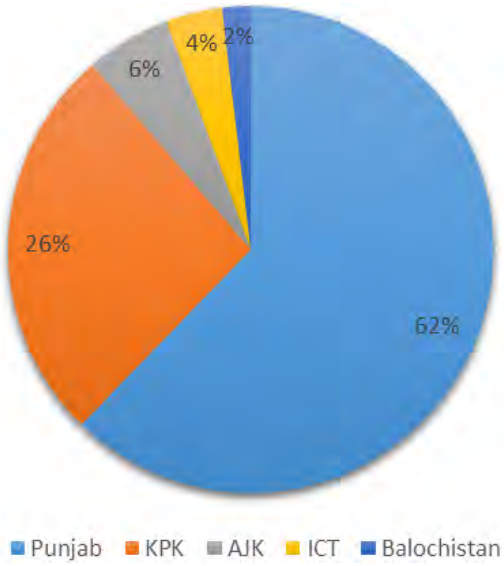


Figure 3.1: Pie chart showing the gathered families from different regional areas of Pakistan (Punjab, Khyber Pakhtunkhwa (KPK), Islamabad Capital Territory (ICT), Balochistan) and Azad & Jammu Kashmir (AJK).

Principal component analysis of the present LSDs families merged with the 1000 genomes projects showed the LSDs families clustering with the South Asian (SAS) superpopulation (Figure 3.2). The SAS population includes a Punjabi population (PJL) verifying the genetic ancestry of the LSDs families.

Table 3.1: Total number of identified diseases through phenotypes and genotype in the families collected during this study from December 2017 to December 2019.

S. No.	Type of diseases	Collected families	Sanger sequencing	Next Generation Sequencing
1.	Mucopolysaccharidosis type I	18	6	12
2.	Mucopolysaccharidosis type II	2	N/A	2
3.	Mucopolysaccharidosis type III	6	N/A	6
4.	Mucopolysaccharidosis type IV	8	N/A	8
5.	Gaucher disease	6	3	3
6.	Niemann-Pick disease	2	N/A	2
7.	Pycnodysostosis	1	N/A	1
8.	Griscelli syndrome	1	N/A	1
9.	Chediak-Higashi syndrome	1	N/A	1

N/A: not applicable

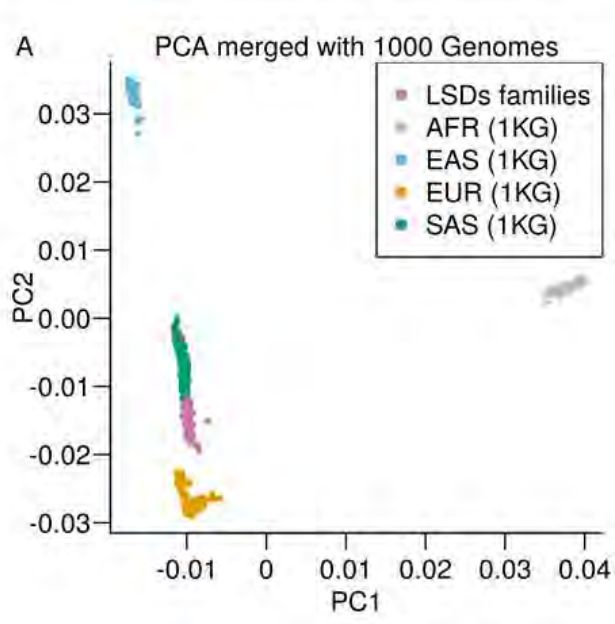


Figure 3.2: Principal component analysis (PCA) of the LSDs families merged with the 1000 Genomes project colored according to super population: African Ancestry (AFR), East Asian Ancestry (EAS), European Ancestry (EUR), and South Asian Ancestry (SAS).

3.2. RESULTS OF FAMILIES SUBJECTED TO SANGER SEQUENCING

3.2.1. Study of Clinical Aspects and Genetic Screening of Six Mucopolysaccharidosis type I Families.

Out of eighteen collected families, initially recruited on the basis of their presenting clinical phenotypes, six were subjected to Sanger Sequencing (Table 3.2). Parental consanguinity or non-consanguinity is mentioned in figure 3.3. Sequencing data analysis revealed a novel variation, which is a non-sense variation NM_000203.5 (*IDUA*); c.1456G>T, (p.Glu486Ter*) (Gul *et al.,* 2020) (Figure 3.4A) in pedigree A which is pathogenic as per American College of Medical Genetics (ACMG) guidelines (Richard *et al.,* 2015), and second is a reported missense variant NM_000203.5 (*IDUA*); c.1469T>C, (p.Leu490Pro) in pedigree C, E and F (Tieu *et al.,* 1995) (Figure 3.4B) which too is pathogenic, thirdly a novel fourteen nucleotide deletion NM_000203.5 (*IDUA*); c.568_581delAACGTCTCCATGAC, (p.Asn190His*fs**204) (Figure 3.5A) in pedigree B, which is a pathogenic variant as per ACMG classification and lastly another reported deletion NM_000203.5 (*IDUA*); c.784delC, (p.His262Thr*fs**55) (Figure 3.5B) (Al-Jasmi *et al.,* 2012) in pedigree D which is also a pathogenic variant (Table 3.3). There are also some benign variants identified and then classified as per American College Medical Genetics (ACMG) guidelines during the sequence data analysis (Table 3.4).

The conservation analysis in multiple orthologues showed highly conserved pattern of the amino acids during fourteen nucleotide and single nucleotide deletions (Figure 3.6 a,b). Furthermore, modeling of wild type and both frameshift mutant variants of the IDUA protein identified in pedigree B and pedigree D families respectively was performed using I-Tasser (Figure 3.7).

Table 3.2: Clinical details, age and cast of six unrelated patients with Mucopolysaccharidosis type I subjected to Sanger Sequencing.

Pedigree ID	Patient ID	Age	Cast	Clinical Profile
A	IV-1	1.5y	Pakhtoon	Coarse facial features, joint contractures, aggressive behavior, umbilical hernia, abdominal distention, hepatosplenomegaly, delay in achieving milestones, excessive hair growth, flat nasal bridge, gibbus formation
B	IV-2	1.5y	Sayed	Delay in achieving milestone, abdominal distention, speech delay, hernia, gibbus formation, short stature, coarse facial features, joint contractures
C	V-2	6y	Chalotra	Delay in achieving milestone, abdominal distention, aggressive behavior, umbilical hernia, short stature, respiratory/ear infection, coarse facial features, joint contractures, dysostosis/multiplex, hepatosplenomegaly, mild IQ,
D	V-3	6m	Jutt	Delay in achieving milestone, short stature, chest problem, coarse facial features, joint contractures, J shaped sella, inferior beaking of lumber vertebrae, oar shaped ribs, subtle opening of proximal ends of metacarpal bones
E	IV-2	8y	Mughal	Delay in achieving milestone, abdominal distention, hepatosplenomegaly, aggressive behavior, speech delay, short stature, corneal opacity, joint contractures, coarse facial features
F	IV-1	3y	Awan	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, umbilical hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding, hyper cellular fragment and trails in bone marrow RBC = 1.77×10^{6}/ul (3.5-5.5), Hemoglobin = 5.5 g/dL (13-18), HCT = 17.3% (35-50), MCV = 98 fl (76-96), MCHC = 31.7 g/dL (32-36), Platelets = 69×10^{3}/ul (150-400),
	IV-2	10.5 y		Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding, cranio facial disproportion due to enlarged calvarium, J shaped sella, oar shaped ribs, hypoplastic L2 vertebrae with posterior displacement ALP = 226 U/L (upto 135), Red cell count = 5.7 million/ul (3.5-5.5), Hemoglobin = 8.7 g/dL (12-15), HCT = 33.5 fl (35-55), MCV = 58.9 fl (75-100), MCH = 15.3 pg (25-35), Reticulocytes =

				0.7% (2-6), RDW-CV = 20.1% (11-15), Neutrophils = 56.9% (60-70), Anisocytsosis ++, Microcytes ++, Hypochromia ++, Poikilocytsosis ++, Serum ferritin = 17.5 ng/ml (24-336)
	IV-4	13y		Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, umbilical hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding
	IV-3	11y		Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding, heart anomaly

ALP: alkaline phosphatase, HCT: hematocrit, IQ: intelligent quotient, MCV: mean corpuscle volume, MCH: mean corpuscle hemoglobin, MCHC: mean corpuscle hemoglobin concentration, RBC: red blood cells, RDW-CV: red cell distribution width, y: years

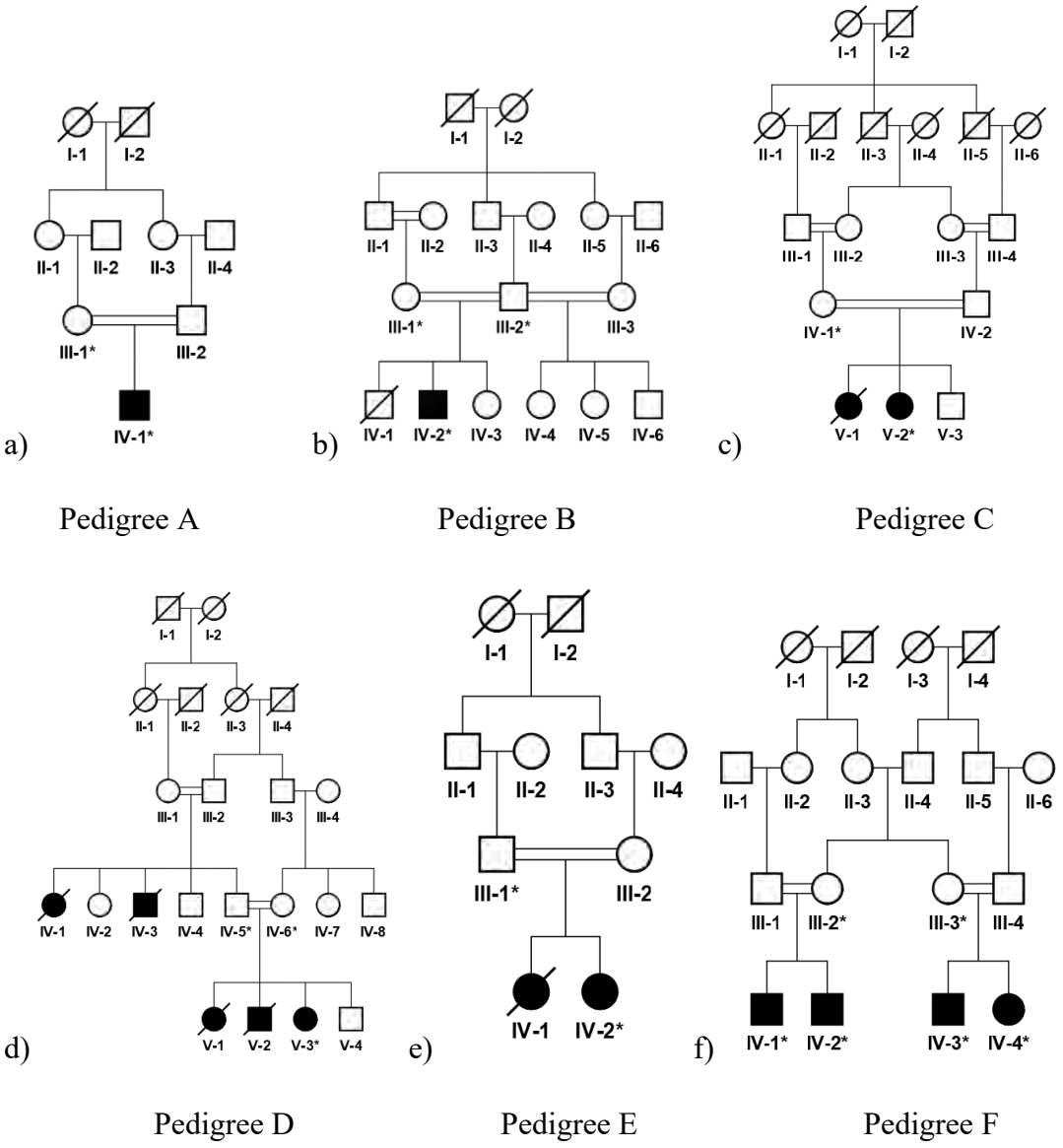


Figure 3.3: Pedigrees of families inheriting MPS I phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual.

Table 3.3: Patient's characteristics along with identified pathogenic variants of *IDUA* gene as predicted by Mutation Taster.

Family ID	Gender	Age at enrollment	Nucleotide change	Protein change	ACMG classification	Status/exon
A	Male	1.5y	c.1456G>T	p.G486Ter	Pathogenic PVS1, PP5, PM2	Homozygous/ 10
B	Male	1.5y	c.568_581del AACGTCTCCAT GAC	p.N190H*fs** 204	Pathogenic PVS1, PM2, PP5	Homozygous/ 5
C	Female	6y	c.1469T>C	p.L490P	Pathogenic PP5, PS3, PP3, PM2	Homozygous/ 10
D	Female	6m	c.784delC	p.H262T*fs** 55	Pathogenic PVS1, PP5, PM2	Homozygous/ 6
E	Female	8y	c.1469T>C	p.L490P	Pathogenic PP5, PS3, PP3, PM2	Homozygous/ 10
F	Male	3y	c.1469T>C	p.L490P	Pathogenic PP5, PS3, PP3, PM2	Homozygous/ 10

A: adenine, C: cytosine, G: guanine, m: months, PM: moderate pathogenic, PP: possibly pathogenic, PS: strong pathogenic, PVS: very strong pathogenic, T: thymine, y: years

Table 3.4: Benign variants identified during Sanger Sequencing of MPS-I families.

S. No	Nucleic Acid change	Amino acid change	ACMG interpretation Strongly benign (BS1)	Global Minor Allele Frequency (GMAF) ClinVar	Exon/ Intron	dbSNP ID
	Exonic variants					
1	c.24C>A	No change	BS1	0.39257	Exon 1	rs11248061
2	c.60G>A	No change	BS1	0.39437	Exon 1	rs10902762
3	c.99T>G	H33Q	BS1	0.10803	Exon 1	rs10794537
4	c.314G>A	R105Q	BS1	0.16873	Exon 3	rs3755955
5	c.543T>C	No change	BS1	0.21985	Exon 5	rs6815946
6	c.942G>C	No change	BS1	0.21865	Exon 7	rs6830825

7	c.1164G>C	No change	BS1	0.21865	Exon 8	rs6836258
8	c.1081G>A	A361T	BS1	0.21765	Exon 8	rs6831280
9	c.1230C>G	No change	BS1	0.21785	Exon 9	rs115790973
10	c.1232T>C	V411A	N/A	N/A	Exon 9	Current study
11	c.1292C>T	A431V	N/A	N/A	Exon 9	Current study
12	c.1360G>A	V454I	BS1	0.21865	Exon 9	rs73066479
13	c.1364C>G	A455G	N/A	N/A	Exon 9	Current study
14	c.1654A>T	T552S	N/A	N/A	Exon 12	Current study
15	c.1786A>C	S596R	N/A	N/A	Exon 13	Current study
	Intronic variants					
16	g.15984T>C	No change	BS1	0.21885	Intron 9/10	rs115134980
17	g.15989C>G	No change	N/A	N/A	Intron 9/10	Current study

A: adenine, ACMG: American college of medical genetics, BS: strongly benign, C: cytosine, G: guanine, N/A: not applicable, T: thymine

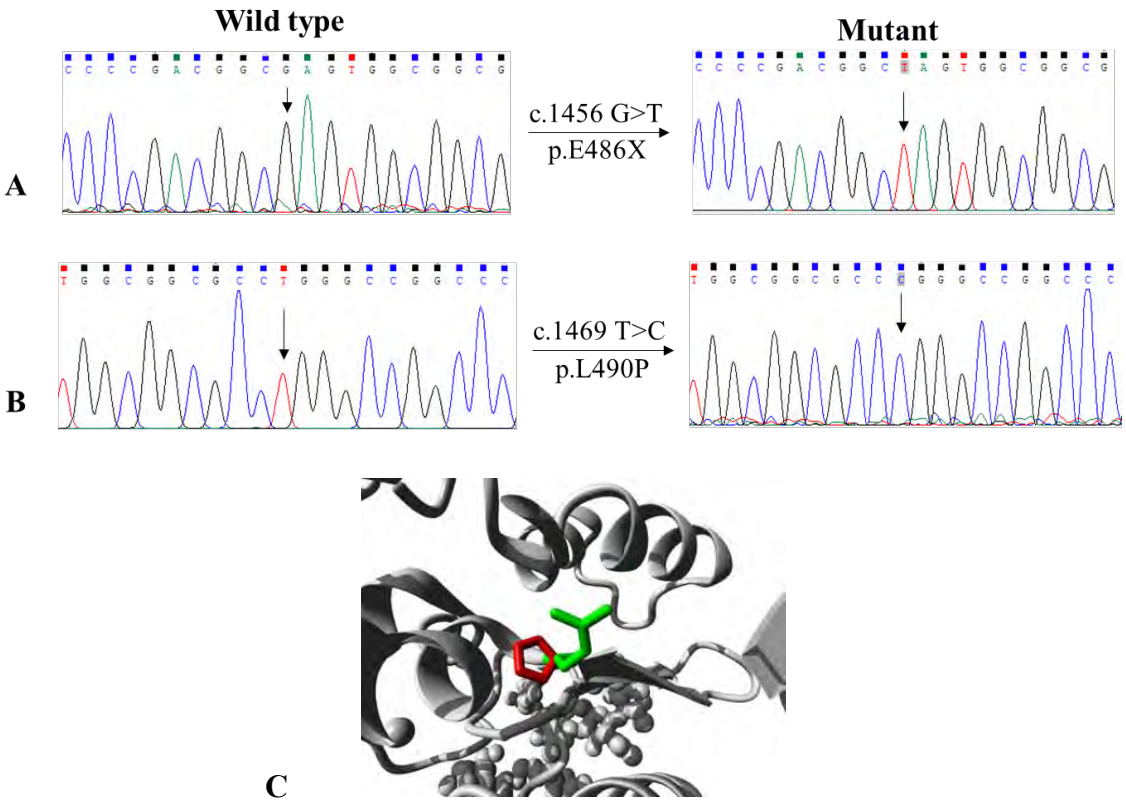


Figure 3.4: (A) DNA sequence chromatogram showing novel variant c.1456G>T causing a (p.Glu486Ter*) in affected individual (right) and unaffected control (left). (B) DNA sequence chromatogram showing previously reported missense variant c.1469T>C causing a (p.Leu490Pro) in affected individual (right) and unaffected control (left). (C) *In-silico* prediction of (p.Leu490Pro) as generated by HOPE. Wild type amino acid residue is colored as green while mutant is colored as red.

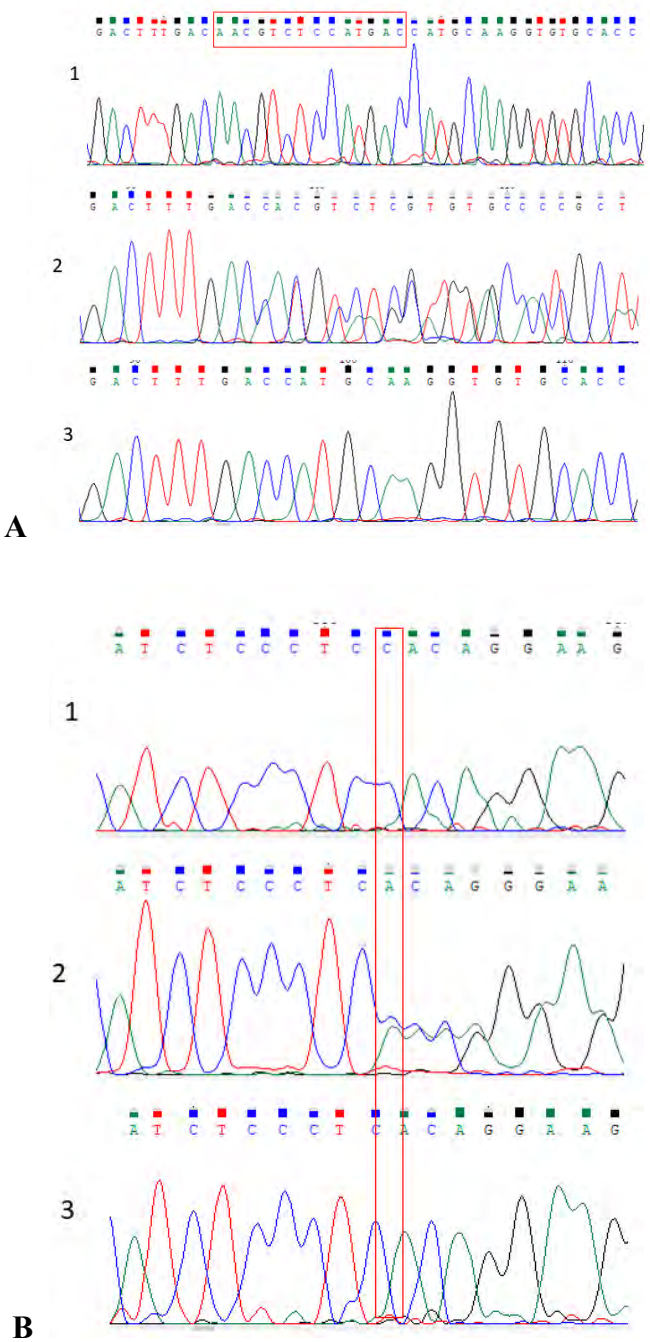
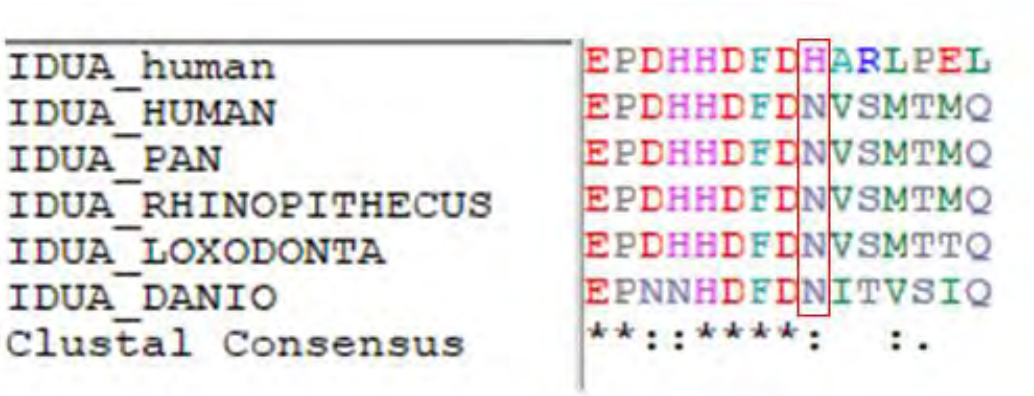


Figure 3.5: (A) Chromatograms showing the novel deletion of fourteen nucleotides i.e., c.568_581delAACGTCTCCATGAC leading to p.N190H*fs204, (i) Homozygous wild type (ii) Heterozygous carrier (iii) Homozygous mutant. (B) Chromatograms showing the novel deletion of cytosine nucleotide i.e., c.784delC causing p.H262T*fs**55. (i) Homozygous wild type (ii) Heterozygous carrier (iii) Homozygous mutant.**

a



b

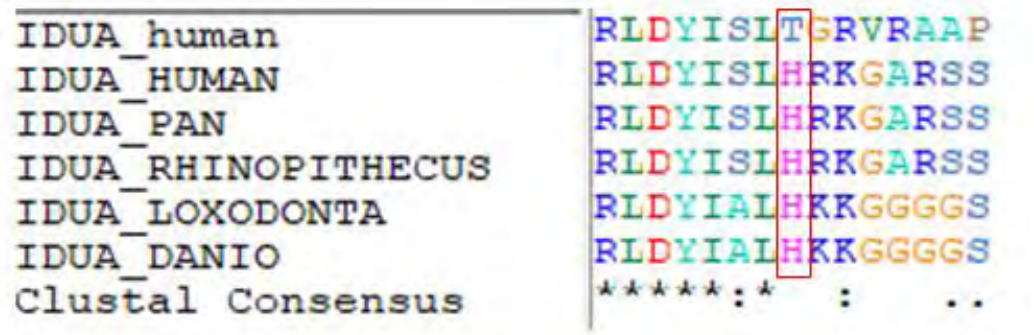


Figure 3.6: Conservation of amino acid residues altered by a) novel fourteen nucleotide deletion i.e., c.568_581delAACGTCTCCATGAC resulting in p.N190H*fs204 b) deletion of cytosine i.e., c.784delC causing p.H262T*fs**55 among other IDUA orthologs.**

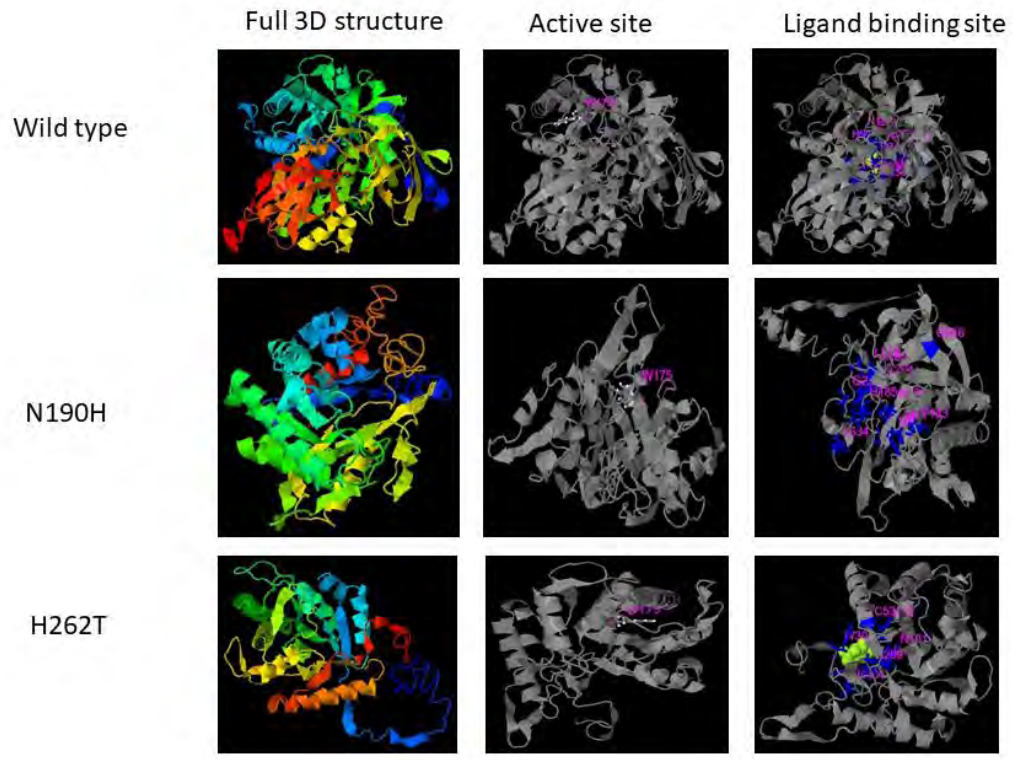


Figure 3.7: Protein structure prediction using I-TASSER software. a) wild type protein b) the truncated IDUA protein caused by 14 nucleotide deletions leading to p.N190Hfs*204 and c) the truncated IDUA protein caused by cytosine nucleotide deletion leading to p.H262Tfs*55 as compared to wild type protein. The tertiary structure of wild type and both mutants, the predicted active site of enzyme and the predicted ligand binding sites of the WT and mutant proteins are shown as indicated.

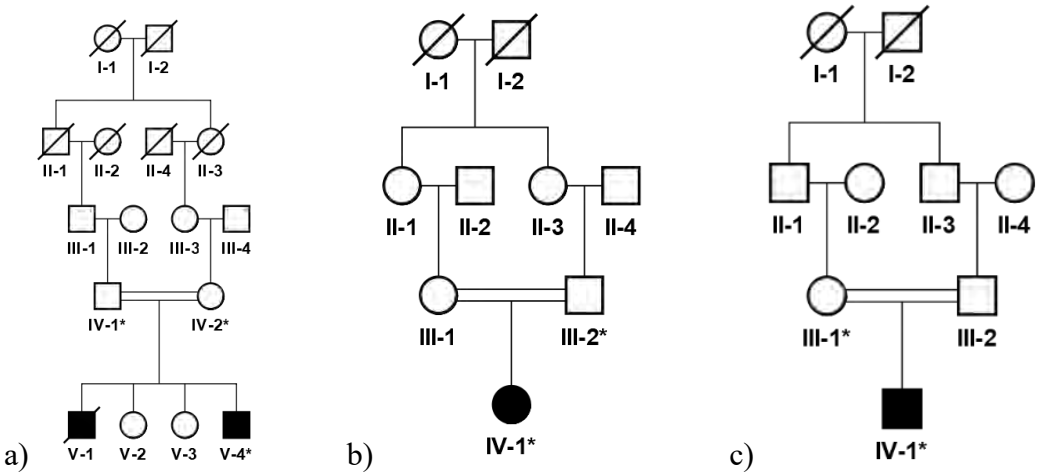
3.2.2. Study of Clinical Aspects and Genetic Screening of Three Gaucher Disease (GD) Families.

During this study total six families affected with Gaucher disease on the basis of presenting clinical profiles have been recruited (Table 3.5). Parental consanguinity is shown in Figure 3.8. Three were subjected into the Sanger sequencing for *GBA* gene which revealed a missense variant NM_000157.4 (*GBA*); c.1448T>C, (p. Leu483Pro) (Table 3.6) (Figure 3.9) in all of the three explored families (Tsuji *et al.,* 1987; Gul *et al.,* 2021).

Table 3.5: Clinical details, age and caste of three unrelated patients with Gaucher disease subjected to Sanger Sequencing.

Pedigree no.	Patient ID	Age	Cast	Clinical profile
A	V-4	4y	Chodhary	Chronic diarrhea, abdominal distention, hepatosplenomegaly, anemia, histiocytes with wrinkled tissue paper appearance Hemoglobin = 11.9 g/dL (12-15), RDW-CV = 16.7% (11.6-15), Neutrophils = 45.6% (60-70), Lymphocytes = 45% (30-40), Monocytes = 6.3% (0-5), ALP = 279 U/L (upto 135), AST = 70 U/L (upto 35), Sodium = 132 mmol/L (135-145)
B	IV-1	13m	Pathan	Macrocephaly, coarse facial features, abdominal distention, hepatosplenomegaly, everted umbilical, gum hypertrophy, seizures, paralysis, gross ascites WBC = 16.6×10^3/ul (4-11), RBC = 3.2×10^6/ul (4-6), Hemoglobin = 7.4 g/dL (11.5-17.5), HCT = 23.5 fl (36-54), MCV = 73 fl (76-96), MCH = 23 pg (27-33), MCHC = 31.5 g/dL (33-35), Platelets = 16×10^3/UL (150-450), Neutrophil = 35% (40-75), Lymphocytes = 60% (20-45), ALT = 52 U/L (10-50)
C	IV-1	1y	Cheema	Abdominal distention, hepatosplenomegaly, anemia, delay in achieving milestones

ALP: alkaline phosphatase, ALT: alkaline transferase, AST: aspartate aminotransferase, HCT: hematocrit, m: months, MCV: mean corpuscle volume, MCH: mean corpuscle hemoglobin, MCHC: mean corpuscle hemoglobin concentration, RBC: red blood cells, RBG: random blood glucose, RDW-CV: red cell distribution width, WBC: white blood cells, y: years



Pedigree A Pedigree B Pedigree C

Figure 3.8: Pedigrees of families inheriting Gaucher disease phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual.

Table 3.6: Demographic, biochemical and molecular genetics diagnosis details of proband of three Gaucher disease family subjected to Sanger Sequencing.

Clinical data		**Pedigree A**	**Pedigree B**	**Pedigree C**
Gender		M	F	M
Age at diagnosis/ age at examination		± 1 year/ 6.5 years	3 months/ 13 months	6 months/ 1 year
Region		Punjab	KPK	Punjab
Treatment		ERT	N/A	ERT
Splenectomy		Yes	N/A	N/A
GCase activity		0.63 nmol/ml/hr	N/A	0.5 µmol/l/h
RBCs		4.55 million/µl	3.22×10^{-6}/µl	4.33 10^6/uL
ALT		22 U/L	52 U/L	16 U/L
WBCs		N/A	16.6×10^{-3}/µl	4.8 10^3/uL
Age at diagnosis		5 months	3 months	14 months
Hemoglobin		11.9 g/dL	7.4g/dL	13.4 g/dL
Platelets		220×10^{-3}/µl	16×10^{-3}/µl	105 10^3/uL
Genotype	**DNA mutation**	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C
	Protein change	p.L483P/ p.L483P	p.L483P/ p.L483P	p.L483P/ p.L483P

M: male, F: female, N/A: not available, ERT: enzyme replacement therapy

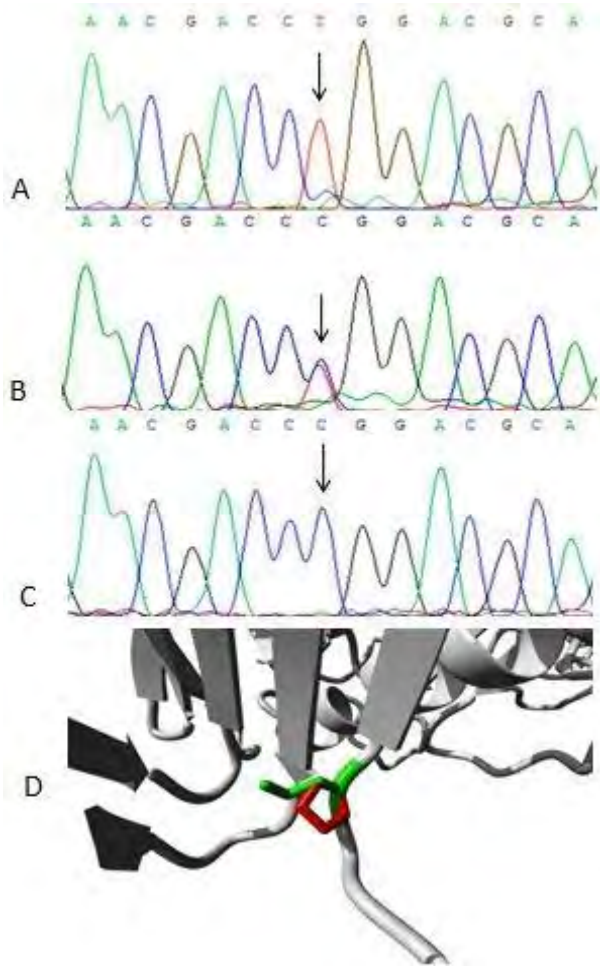


Figure 3.9: (A) Sequence chromatogram of normal allelic sequence. (B) Sequence chromatogram of heterozygous carrier for c.1448T>C. (C) Sequence chromatogram of homozygous mutant for c.1448T>C in affected individuals of Gaucher affected families. (D) Structural effect of the amino acid substitution as shown by HOPE analysis. Mutated amino acid is shown in red color while wild type amino acid is shown in green color.

3.3. RESULTS OF FAMILIES SUBJECTED TO WHOLE GENOME SEQUENCING

Thirty-six families were subjected to Next Generation Sequencing (NGS) to find the genetic basis of disease phenotype in each of these analyzed families. Whole Genome Sequencing was performed at Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark during six months stay sponsored and funded by the Higher Education Commission (HEC), Islamabad, Pakistan under International Research Support Initiative Program (IRSIP). Details of NGS data analysis as per disease subtype is given below:

3.3.1. Study of Clinical Aspects and Genetic Screening of Twelve Mucopolysaccharidosis type I Families.

As described in section 3.3.1, out of total 18 collected MPS 1 segregating families, six were subjected to Sanger Sequencing at Molecular biology lab, Department of Zoology, Faculty of Biological Sciences, Quaid-I-Azam University, Islamabad, Pakistan, and twelve were subjected to NGS. Interestingly among these twelve clinically diagnosed as MPS I families by local Physician (Table 3.7), six families were segregating disease causing variant in MPS-1 causing reported gene i.e., *IDUA* gene, NM_000203.5 (*IDUA*); c.1172-1173insGCTGCTGGC, (p.Gly391insLeuLeuAla) in family B, NM_000203.5 (*IDUA*); c.1469T>C, (p.Leu490Pro) in family E, F, G, I and K. However, six families (A, C, D, H, J, L) despite of initial clinical diagnosis as MPS I, didn't show any disease-causing variant in *IDUA* gene. Their detailed genotypes are mentioned as: NM_000512.5 (*GALNS*); c.871G>A, (p.Ala291Thr) in family A; NM_000199.5 (*SGSH*); c.220G>A, (p.Arg74Cys) and NM_000152.5 (*GAA*); c.1930G>T, (p.Ala644Ser) in family C; NM_172232.4 (*ABCA5*); c.26569G>A, (p.Arg857Cys) in family D; NM_000035.4 (*ALDOB*); c.264G>T, (p.Asp88Glu) in family H; NM_016146.6 (*TRAPPC4*); c.454+3A>G and NM_003922.4 (*HERC1*); c.5941G>A, (p.Val1981Ile) in family J; and NM_014140.4 (*SMARCAL1*); c.1810C>T, (p.Gln604Ter*) and NM_012250.6 (*RRAS2*); c.439C>T, (p.Arg147Trp) in family L. Five families were identified with multigenic inheritances affected with both *IDUA* and other gene or affected with both other genes instead of *IDUA*. Their genotypes are mentioned as: NM_000199.5

(*SGSH*); c.220G>A, (p.Arg74Cys) and NM_000152.5 (*GAA*); c.1930G>T, (p.Ala644Ser) in family C; NM_000203.5 (*IDUA*); c.1469T>C, (p.Leu490Pro) and NM_001384125.1 (*KIAA1109*); c.1297A>G, (p.Thr4033Ala) in family I, NM_016146.6 (*TRAPPC4*); c.454+3A>G and NM_003922.4 (*HERC1*); c.5941G>A, (p.Val1981Ile) in family J; NM_000203.5 (*IDUA*); c.1469T>C, (p.Leu490Pro) and NM_139125.4 (*MASP1*); c.C73T, (p.Arg25Ter*) in family K and NM_014140.4 (*SMARCAL1*); c.1810C>T, (p.Gln604Ter*) and NM_012250.6 (*RRAS2*); c.439C>T, (p.Arg147Trp) in family L. All the variants were found in homozygous state in affected individuals while heterozygous in their parents. Phenotypically normal siblings were either wild type or carrier of the respective variants in each family. Detailed family history and consanguinity are mentioned in figure 3.10. Runs of homozygosity-based F-coefficient of the affected individuals showed that the high rate of consanguinity in MPS I families was also reflected in the genetic data (Figure 3.11) and close to 0 for the non-consanguineous MPS 2 affected individuals.

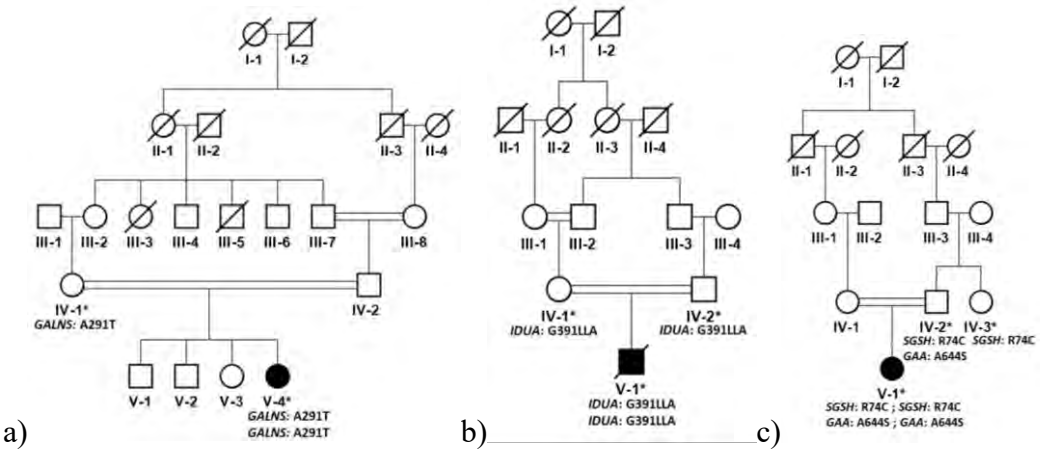
These variants were predicted deleterious/disease causing/pathogenic by different mutation effect prediction tools including MutationTaster, Polyphen2, SIFT, CADD, PROVEAN, VarSome, VARSEAK etc (Table 3.8). None of the identified variants was found in homozygous state in any human genome variations databases including gnomAD, ExAC, dbSNP, 1000 Genome Project and ethnically matched healthy controls.

Table 3.7: Clinical details, age and cast of twelve unrelated patients with Mucopolysaccharidosis type I subjected to Whole Genome Sequencing.

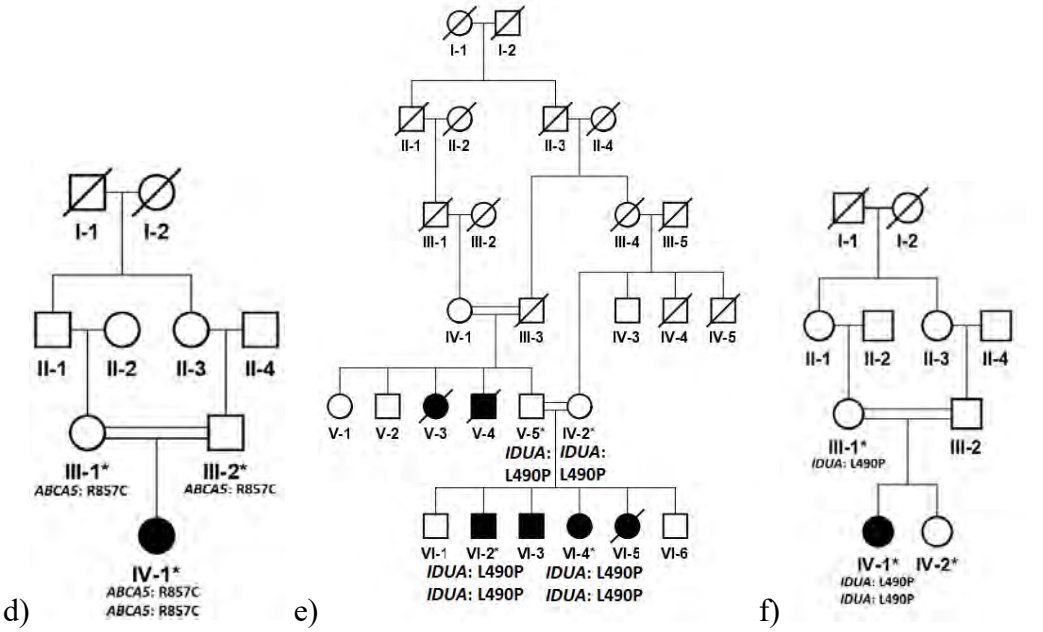
Pedigree no.	Patient ID	Age	Cast	Clinical profiles
A	V-4	1y	Rajpoot	Squint eyes, tip-toe walking, kyphosis/scoliosis, gibbus formation, frontal bossing, depressed nasal bridge, umbilical hernia, narrowing of spinal cord, dysostosis, throat problem, weak muscles, more crying episodes, two central incisors widely spaced, joint contractures, coarse facial features, Pectus carinatum, Haemoglobin = 8.6 g/dL (12-15), HCT = 32 (35-55), MCV = 63.9 fl (75-100), MCH = 17.2 pg (25-35), MCHC = 26.9 g/dL (31-38), Neutrophils = 40.4% (60-70), ESR = 28 mm/Hr (0-15)
B	V-1	1y	Malik	Delay in achieving milestone, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, delayed developmental Q, depressed nasal bridge, squinted eyes
C	V-1	4y	Pathan	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, short stature, respiratory/ear infections, coarse facial features, broad based gait, joint contractures, short trunk, scoliosis/kyphosis, hepatosplenomegaly, low IQ, intellectual disability, low haemoglobin
D	IV-1	1.5y	Sheikh	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, gibbus formation, short stature, upper airways obstruction, kyphosis, hazel corneas, coarse facial features, hearing loss, joint contractures, hepatosplenomegaly, intellectual disability, hypertrichosis
E	VI-4	8y	Rajpoot	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding
	VI-2	3y		Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding
F	IV-1	16.5y	Narman	Delay in achieving milestones, abdominal distention, hernia, short stature, coarse facial features, joint contractures, short trunk, hepatosplenomegaly

G	IV-4	6y	Rajpoot	Delay in achieving milestones, abdominal distention, aggressive behavior, short stature, respiratory/ear infection, joint contractures, short trunk, speech delay, coarse facial features, hepatosplenomegaly, intellectual disability
H	IV-2	14y	Suddhan	Delay in achieving milestones, abdominal distention, coarse facial features, aggressive behavior, short stature, respiratory/ear infections, hip dysplasia, joint stiffness, hepatosplenomegaly, difficulty in walking Craniofacial disproportion, J-shaped sella, L1 vertebral body is hypoplastic and slightly posteriorly displaced, exaggerated lumbar lordosis, RBG = 71 mg/dL (80-160), serum urea = 10 mg/dL (12-50), loss of biconcavity of metacarpals with proximal tapping
I	V-1	12y	Mughal	Joint contractures, corneal clouding, very less eyesight, abdominal distention, coarse facial features, short stature, underweight, umbilical hernia, cannot walk, developmental delay, cardiac issues, hearing issues, respiratory problems/loud snoring
J	V-2	2.5y	Mughal	Delay in achieving milestones, aggressive behavior with multiple crying episodes, respiratory/ear infection, no walk, intellectual disability
K	IV-4	8.5y	Abbasi	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, umbilical hernia, gibbus formation, short stature, constipation, coarse facial features, cannot walk (wheel-chair bound), joint contractures, scoliosis and kyphosis, hepatosplenomegaly, fish mouth vertebrae, oar shaped ribs, low IQ, widely spaced eyes, narrowing of eye opening (right), drooping eyelids, intellectual disability, developmental delay Red cell count = 2.68 million/ul (4-5.2), Haemoglobin = 7.2 g/dL (11.5-15.5), HCT = 21.8 (35-45), Platelet count = 141×1000/ul (170-450), Neutrophils = 53.4% (60-70)
L	IV-5	8y	Yousafzai	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, low IQ, intellectual disability, dysostosis multiplex

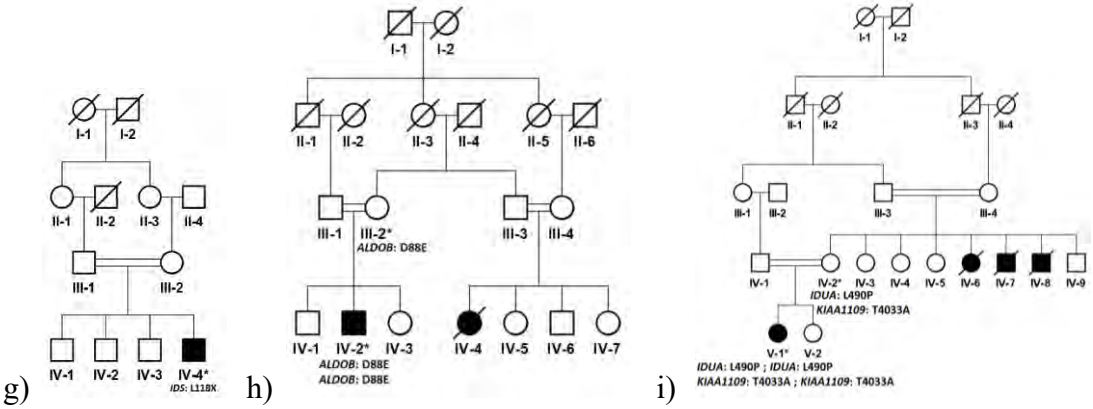
ESR: erythrocytes sedimentation rate, HCT: hematocrit, IQ: intelligent quotient, MCV: mean corpuscle volume, MCV: mean corpuscle volume, MCH: mean corpuscle hemoglobin, MCHC: mean corpuscle hemoglobin concentration, RBG: random blood glucose, y: years



a) Pedigree A b) Pedigree B c) Pedigree C



d) Pedigree D e) Pedigree E f) Pedigree F



g) Pedigree G h) Pedigree H i) Pedigree I

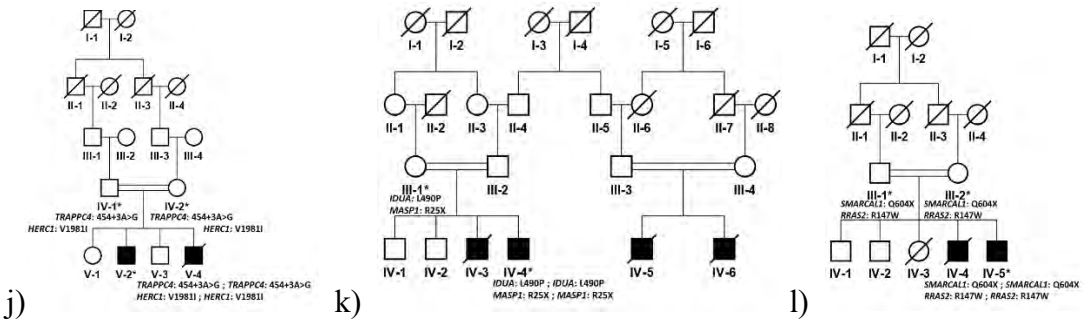


Figure 3.10: Pedigrees of twelve families (A-L) inheriting MPS I phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

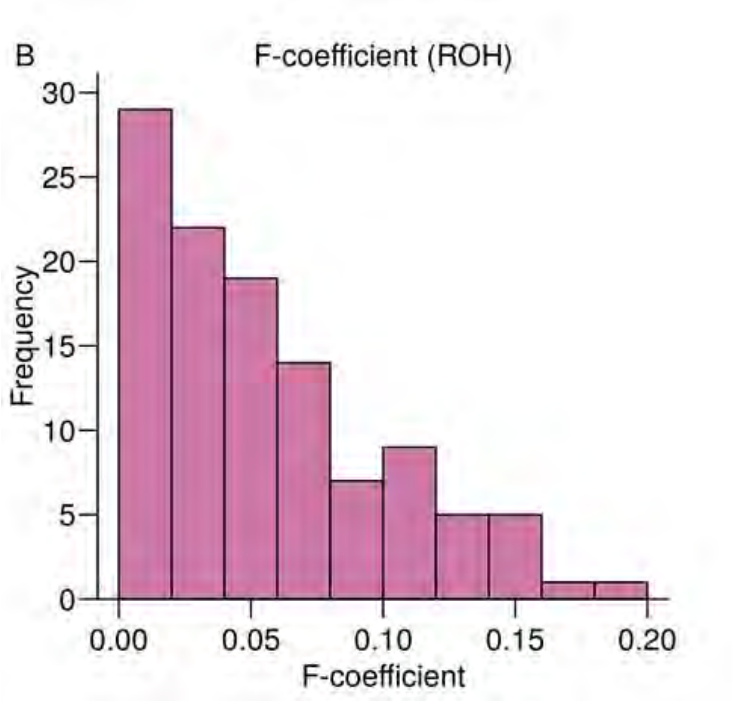


Figure 3.11: Distribution of runs of homozygosity-based F-coefficient of the MPS families.

Table 3.8: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in twelve MPS I recruited patients subjected to Whole Genome Sequencing.

Pedigree/Patient ID	Gene variant	Patients' Genotype	dbSNP	*In-Silico* tools with deleterious output	MAF gnomAD	Classification ACMG	VarSome	HGMD accession number
A/V-4	NM_000512.5 (*GALNS*); c.871G>A (p.Ala291Thr)	Homozygous	rs118204448	4/7	0.00001972	Pathogenic PP3, PP4, PP5, PM1, PM2	Pathogenic	CM950535
B/V-1	NM_000203.5 (*IDUA*); c.1172-1173insGCTGCTGGC (p.Gly391insLeuLeuAla)	Homozygous	rs1560548356	2/4	0.000007	Likely pathogenic PP4, PM2, PM4	Likely pathogenic	Current study
C/V-1	NM_000199.5 (*SGSH*); c.220G>A (p.Arg74Cys)	Homozygous	rs104894636	17/17	0.0002270	Pathogenic PP5, PS3, PM1, PM2, PM5, PP3	Pathogenic	CM971355
	NM_000152.5 (*GAA*); c.1930G>T (p.Ala644Ser)		rs752735168	16/18	0.00003306	Likely pathogenic PM1, PM2, PP3, PP4	Pathogenic	Current study
D/IV-1	NM_172232.4 (*ABCA5*); c.26569G>A (p.Arg857Cys)	Homozygous	rs747667004	7/7	0.0001090	Likely pathogenic PM1, PM2, PP3, PP4	Uncertain significance	Current study
E/IV-2, IV-4	NM_000203.5 (*IDUA*); c.1469T>C (p.Leu490Pro)	Homozygous	rs121965027	7/7	0.00004681	Pathogenic PS3, PM1, PM2, PP4, PP5, PS3	Pathogenic	CM950685
F/IV-1	NM_000203.5 (*IDUA*); c.1469T>C (p.Leu490Pro)	Homozygous	rs121965027	7/7	0.00004681	Pathogenic PS3, PM1, PM2, PP4, PP5, PS3	Pathogenic	CM950685
G/IV-4	NM_001166550.4 (*IDS*); c.353T>A (p.Leu118Ter)	Hemizygous	-	4/4	-	Pathogenic PVS1, PM2, PP3, PP4	Pathogenic	Current study
H/IV-2	NM_000035.4 (*ALDOB*); c.264G>T (p.Asp88Glu)	Homozygous	rs200585150	11/18	0.0004219	VUS PM1, PM2, PP3 BP1,	Uncertain significance	CM1618755

I/V-1	NM_000203.5 (*IDUA*); c.1469T>C (p.Leu490Pro)	Homozygous	rs121965027	7/7	0.00004681	Likely pathogenic PS3, PM1, PM2, PP4, PP5	Pathogenic	CM950685
	NM_001384125.1 (*KIAA1109*); c.1297A>G (p.Thr4033Ala)		rs773658128	6/7	0.00001213	VUS PM1, PM2	Uncertain significance	Current study
J/V-2	NM_016146.6 (*TRAPPC4*); c.454+3A>G	Homozygous	rs375776811	N/A	0.0002403	Pathogenic PP5, PM2, PP3	Likely pathogenic	Not available
	NM_003922.4 (*HERC1*); c.5941G>A (p.Val1981Ile)		-	4/7	-	VUS PM2, BP4	Uncertain significance	Current study
K/IV-4	NM_000203.5 (*IDUA*); c.1469T>C (p.Leu490Pro)	Homozygous	rs121965027	7/7	0.00004681	Likely pathogenic PS3, PM1, PM2, PP4, PP5	Pathogenic	CM950685
	NM_139125.4 (*MASP1*); c.C73T (p.Arg25Ter)		rs369042881	N/A	0.0001591	VUS PM2, BS1, BS2	Uncertain significance	Current study
L/IV-5	NM_014140.4 (*SMARCAL1*); c.1810C>T (p.Gln604Ter)	Homozygous	-	6/7	-	Pathogenic PVS1, PM2, PP3	Likely pathogenic	Current study
	NM_012250.6 (*RRAS2*); c.439C>T (p.Arg147Trp)		-	16/18	-	Likely pathogenic PM1, PM2, PP2, PP3	Likely pathogenic	Not available in HGMD

HGMD: human genome mutation database, MAF: minor allele frequency, N/A: not applicable, PM: moderate pathogenic, PP: possibly pathogenic, PS: strong pathogenic, PVS: very strong pathogenic, VUS: variant of uncertain significance

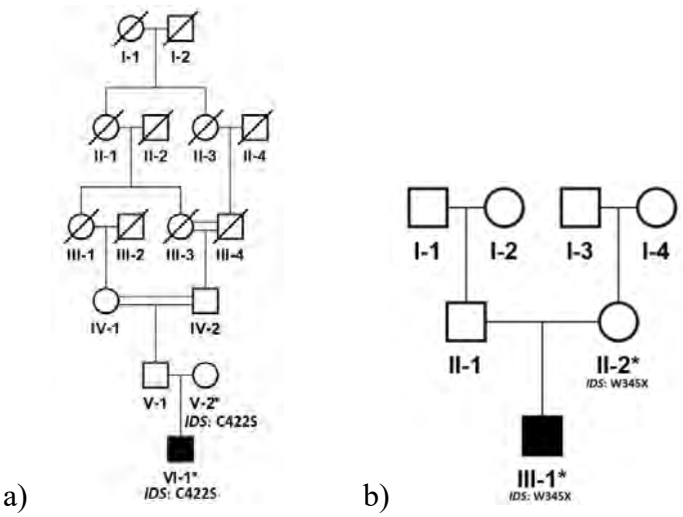
3.3.2. Study of Clinical Aspects and Genetic Screening of Two Mucopolysaccharidosis type II Families.

Two suspected families for MPS II or hunter syndrome have been recruited based on their presenting clinical profiles (Table 3.9). WGS data analysis revealed two pathogenic variations in gene *IDS* located on X-chromosome (Table 3.10). Parental consanguinity could not be linked in these families as it is X-linked disorder (Figure 3.12).

Table 3.9: Clinical details, age and cast of two unrelated patients with Mucopolysaccharidosis type II subjected to Whole Genome Sequencing.

Pedigree no.	Patient ID	Age	Cast	Clinical profiles
A	VI-1	4.5y	Rajpoot Bhatti/Khosa	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, constipation, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability, hernia
B	III-1	5y	Aryn	Delay in achieving milestones, hernia, short stature, coarse facial features, intellectual disability

y: years



a) b)

Pedigree A Pedigree B

Figure 3.12: Pedigrees of families inheriting MPS II phenotype consistent with X-linked mode of inheritance of disease.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Variants identified in the present families are shown below each participating individual

Table 3.10: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in two MPS II recruited patients subjected to Whole Genome Sequencing.

Pedigree / Patient ID	Gene variant	Patients' Genotype	*In-Silico* tools with deleterious output	Classification ACMG	VarSome	HGMD accession number
A/VI-1	NM_000202.8 (*IDS*); c.1264C>G (p.Cys422Ser)	Hemizygous	17/23	Likely pathogenic PM1, PM2, PP3, PP4, BP1	Likely pathogenic	Current study
B/III-1	NM_000202.8 (*IDS*); c.1035C>T (p.Trp345Ter)	Hemizygous	11/11	Pathogenic PVS1, PM2, PP3,	Pathogenic	Current study

ACMG: american college of medical genetics, BP: benign supporting, HGMD: human genome mutation database, PM: moderate pathogenic, PP: possibly pathogenic, PVS: very strong pathogenic

3.3.3. Study of Clinical Aspects and Genetic Screening of Six Mucopolysaccharidosis type III or Sanfilippo Syndrome Families.

There were six families recruited based on their presenting clinical profiles as MPS III or Sanfilippo syndrome affected. There were various overlapping features among all patients mentioned in detail in table 3.11. Detailed pedigrees showing non/consanguineous parents are mentioned in figure 3.13. After subjected to the WGS, five families were affected with MPS III including sub-type A and B: NM_000263.4 (*NAGLU*); c.2020C>T, (p.Arg674Cys) in pedigree A; NM_000199.5 (*SGSH*); c.221C>T, (p.Arg74His) in pedigree B; NM_000199.5 (*SGSH*); c.364C>T, (p.Gly122Arg) in pedigree D; NM_000199.5 (*SGSH*); c.258A>C, (p.Asn86Lys) in pedigree E and NM_000199.5 (*SGSH*); c.817G>A, (p.Asp273Asn) in pedigree F while one family had shown variant in the gene causing Spinocerebral Ataxia, NM_144992 (*VWA3B*); c.2951A>C, (p.Glu984Ala) in pedigree C. All variants were found in homozygous state in patients while in heterozygous state in parents. Phenotypically normal siblings in families were either homozygous wild type or heterozygous carriers for that particular variant.

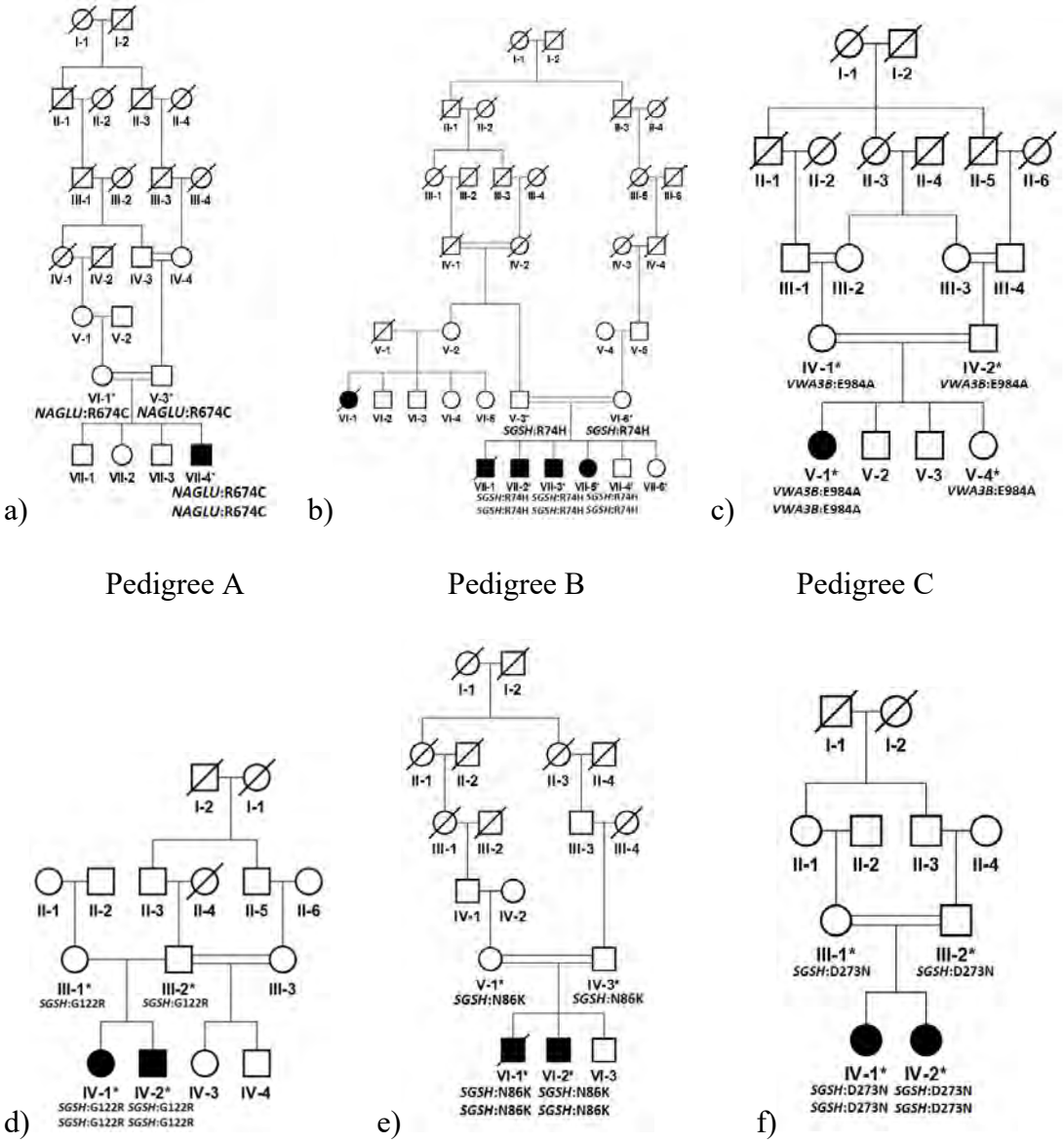
These variants were predicted to be deleterious/pathogenic by various mutation prediction tools like MutationTaster, SIFT, Polyphen-2, CADD, VarSome, VARSEAK and PROVEAN (Table 3.12). All these predicted variants were not found in any human genome variation databases including ExAC, gnomAD, 1000 Genome Project, dbSNP and healthy control samples from same ethnicity.

Table 3.11: Clinical details, age and cast of six unrelated patients with Mucopolysaccharidosis type III subjected to Whole Genome Sequencing.

Pedigree no.	Patient ID	Age	Cast	Clinical profiles
A	VII-4	5y	Sardar	Delay in achieving milestone, abdominal distention, aggressive behavior (sometime), speech delay, short stature, increased body hair, coarse facial features, broad based gate, joint contracture, short trunk, hepatosplenomegaly, self-biting, intellectual disability ALT = 49 U/L (up to 40), AST = 46 U/L (up to 40), Bilirubin total = 1.9 mg/dL (<1.2), Hb = 11.1 g/dL (12-1.4), RBC count = 4.08 mil/cmm (4.75-4.85), Hematocrit = 34% (36-42), Monocytes = 06% (4-5), Serum Sodium = 155 mmol/L (132-145), Serum Potassium = 5.5 mmol/L (3.1-5.1), Serum Ferritin = 27.84 ng/mL (30-400), beaking at L1 and L2 vertebrae, Glucose (R) = 76 mg/dL (80-130)
B	VII-2	14.5y	Khawaja	Delay in achieving milestone, chronic diarrhea, abdominal distention, aggressive behaviors, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, poor IQ
	VII-5	10.5y		Delay in achieving milestone, chronic diarrhea, abdominal distention, aggressive behaviors, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, poor IQ Hb = 9.9 g/dL (13.5-19.5), RBC = 3.92×10^{6}/ul (4.0-5.9), MCH = 25.2 pg (28-32), MCHC = 27.7 g/dL (29-37), Neutrophils = 76.1 % (50-75), Eosinophils = 0.1 % (1-6)
	VII-3	9y		Delay in achieving milestone, chronic diarrhea, abdominal distention, aggressive behaviors, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, poor IQ
C	V-1	5y	Aryn	Delay in achieving milestones, chronic diarrhea, aggressive behavior, speech delay, short stature, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability
	V-4	1.5y		Delay in achieving milestones, chronic diarrhea,, short stature, joint contractures, hepatosplenomegaly, intellectual disability
D	IV-1	2y	Qaisrani baloch	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, respiratory and ear infection, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability 3^{rd} and 4^{th} ventricle dilation, Mild periventricular ischemia, Wide cisterna magna, Mild cerebral atrophy, Urea = 12 mg/dL (15-45), Creatinine = 0.5 g/dL (0.6-1.1), ALT = 39 u/L (5-31), Hb = 10.1 g/dL (12-16), HCT = 35.7 % (37-47), MCV = 67.2 fL (80-97), MCH = 19 pg (26-32), MCHC = 28.3 g/dL (32-36), Microcytes = Positive, Hypochromia = Positive, Poikilocytes = Positive
	IV-2	1y		Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, respiratory and ear infection (otitis media), coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Mild cerebral atrophy, mild mucosal thickness of right ethmoid and right maxillary sinus, Urea = 10 mg/dL (15-45), Creatinine = 0.5 mg/dL (0.7-1.4), Uric acid = 2.8 md/dL (3.6-7.7), Serum Ferritin = 4.75 µg/L (30-400), Hb = 10.1g/dL (14-18), HCT = 36.5 % (41-51), MCV = 69.4 fL (80.97), MCH = 19.2 pg (26-

				32), MCHC = 27.7 g/dL (32-36), Microcytes = Positive, Hypochromia = Positive, Poikilocytes = Positive
E	VI-1	14y	Pathan	Delay in achieving milestone, aggressive behavior, speech delay, gibbus formation, short stature, respiratory/ear infection, coarse facial features, excessive hair growth, broad based/waddling, joint contractures, short stature, scoliosis/kyphosis, intellectual disability Diffuse cerebral cortical atrophy, AV and VA concordance, Dolichocephaly, Mild atrophy, Hb = 11.4 g/dL (14-18), Neutrophils = 44% (50-75), Lymphocytes = 49% (20-40), Monocytes = 01% (2-6)
	VI-2	8y		Delay in achieving milestone, aggressive behavior, speech delay, gibbus formation, short stature, respiratory/ear infection, coarse facial features, excessive hair growth, broad based/waddling, joint contractures, short stature, scoliosis/kyphosis, intellectual disability Small acquired porencephalic cyst left peri-sylvian mid parietal region, Hb = 11 g/dL (14-18), WBCs = 3900/cmm (4000-11000), Lymphocytes = 47% (20-45), Monocytes = 01% (2-6)
F	IV-1	10y	Kakakhel	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, respiratory/ear infection (decreased hearing), coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Spina bifida S1, Irregular acetabular margins, Pallor = Positive, Thick calvarium
	IV-2	7y		Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, respiratory/ear infection, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Hb = 11.7 g/dL (12-15), Neutrophils = 16.7% (60-70), Eosinophils = 25.2% (0.0-4.0), Irregular acetabular with small lytic areas with small femoral epiphysis, Spina bifida S1, Thick calvarium, J-shaped sella, Hip joint juvenile arthritis

ALT: alanine transaminase, AST: aspartate aminotransferase, AV/VA: atrioventricular/ventriculoarterial, Hb: hemoglobin, HCT: hematocrit, IQ: intelligent quotient, MCV: mean corpuscular volume, MCHC: mean corpuscular hemoglobin concentration, MCH: mean corpuscular hemoglobin, RBC: red blood cells, WBCs: white blood cells, y: years



Pedigree A Pedigree B Pedigree C

Pedigree D Pedigree E Pedigree F

Figure 3.13: Pedigrees of families inheriting MPS III phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

Table 3.12: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in six MPS III recruited patients subjected to Whole Genome Sequencing.

Pedigree/Patient ID	Gene variants	Patient genotype	dbSNP	*In-silico* tools with deleterious output	ACMG classification	Varsome	HGMD accession number
A/VII-4	NM_000263.4 (*NAGLU*); c.2020C>T (p.Arg674Cys)	Homozygous	rs763299645	9/9	Pathogenic PP5, PP3, PM5, PM2	Pathogenic	CM981369
B/VII-2, VII-3, VII-5	NM_000199.5 (*SGSH*); c.221C>T (p.Arg74His)	Homozygous	rs778336949	9/9	Pathogenic PP5, PS3, PM5, PP3, PM2	Pathogenic	CM971354
C/V-1, V-4	NM_144992 (*VWA3B*); c.2951A>C (p.Glu984Ala)	Homozygous	N/A	8/9	Likely benign BP1, BP4, PM2	Uncertain significance	Current study
D/IV-1, IV-2	NM_000199.5 (*SGSH*); c.364C>T (p.Gly122Arg)	Homozygous	rs761607612	9/9	Pathogenic PP5, PP3, PM2	Likely pathogenic	CM971358
E/VI-1, VI-2	NM_000199.5 (*SGSH*); c.258A>C (p.Asn86Lys)	Homozygous	N/A	8/9	Uncertain significance PM2, PP3	Uncertain significance	Current study
F/IV-1, IV-2	NM_000199.5 (*SGSH*); c.817G>A (p.Asp273Asn)	Homozygous	rs1046551417	7/9	Uncertain significance PM2, PP5	Uncertain significance	CM002397

ACMG: american college medical college, BP: benign supporting, HGMD: human genome mutation database, N/A: not applicable, PM: moderate pathogenic, PP: possibly pathogenic, PS: strong pathogenic

3.3.4. Study of Clinical Aspects and Genetic Screening of Eight Mucopolysaccharidosis type IV or Morquio Syndrome Families.

From clinical and phenotypic diagnosis, eight families affected with Morquio syndrome (MPS IV) were recruited to be the part of this study spanning on duration of 2017 to 2020. MPS IV patients showing overlapping features like short stature, coarse facial features, scoliosis/kyphosis, gibbus formation and many others (Table 3.13). All families of MPS IV show consanguinity except pedigree C (Figure 3.14).

Data analysis revealed six families affected with MPS IVA: NM_000512.5 (*GALNS*); c. 1365-2A>G in pedigree A; NM_000512.5 (*GALNS*); c.107A>C, (p.Leu36Arg) in pedigree B; NM_000512.5 (*GALNS*); c.1259G>C, (p.Pro420Arg) in pedigree D; NM_000512.5 (*GALNS*), c.423-1G>A in pedigree F; NM_000512.5 (*GALNS*); c.107A>C, (p.Leu36Arg) in pedigree G; NM_000512.5 (*GALNS*); c.107A>C, (p.Leu36Arg) in pedigree H and one family affected with Mucolipidosis III gamma i.e NM_032520.5 (*GNPTG*); c.179G>A, (p.Gly60Glu) in pedigree E; and one family with multigenic inheritance including MPS IVA and Biotinidase deficiency, NM_001281723.2 (*BTD*); c.1336G>C, (p.Asp446His) in pedigree H. All variants were found in homozygous state in patients while in heterozygous state in parents. Phenotypically normal siblings in families were either homozygous wild type or heterozygous carriers for that particular variant. In pedigree C (MPS IV) initial analysis of coding regions in patients did not reveal any pathogenic variant.

These variants were predicted to be deleterious/pathogenic by various mutation prediction tools like MutationTaster, SIFT, Polyphen-2, CADD, VarSome, VARSEAK and PROVEAN (Table 3.14). All these predicted variants were not found in any human genome variation databases including ExAC, gnomAD, 1000 Genome Project, dbSNP and healthy control samples from same ethnicity.

Table 3.13: Clinical details, age and cast of eight unrelated patients with Mucopolysaccharidosis type IV subjected to Whole Genome Sequencing.

Pedigree no	Patient ID	Age	Cast	Clinical profiles
A	IV-1	3y	Gujjar	Delay in achieving milestones, abdominal distention, gibbus formation, short stature, height 95 cm,12 kg weight, head circumference 48 cm, coarse facial features, broad based gate, joint contractures, kyphosis/scoliosis, bullet shaped metacarpals, beaking of vertebrae, oar shaped ribs, widening of medullary canal, coarse trabecular pattern Metaphyseal dysplasia, Blood sugar Random = 70 mg/dL (80-160), Hb = 11.7 g/dL (14-18), HCT = 34.3% (35-55), Neutrophils = 39.8% (60-70), Lymphocytes = 42.1% (30-40), Vitamin D3 = 18ng/mL (Deficient <20)
B	IV-5	7.5y	Awan	Speech delay, gibbus formation, short stature, (92 cm height, weight 15 kg, head circumference 49 cm {F}), coarse facial features, broad based gate, bone pain, joint contractures, short trunk, scoliosis/kyphosis, pectus carinatum
	IV-4	5y		Abdominal distention, aggressive behavior, short stature, constipation, (89 cm height, 14 kg weight, head circumference 51 cm {M})
C	III-1	12y	Pathan	Asthma (use inhaler), sleep apnea (suddenly wake up), no hearing, very loose joints, thin bones, underweight, deep voice, macrocephaly, carpel tunnel syndrome, short stature, spinal cord is bent, all body muscles and bones deform, weak muscles, aggressive, coarse facial features, jerky movement, widely spaced teeth, large tongue, mild ID, flat nasal bridge, skin problem
	III-2	9y		Sleep apnea (suddenly wake up), very loose joints, thin bones, underweight, deep voice, macrocephaly, short stature, spinal cord is bent, all body muscles and bones deform, weak muscles, aggressive, coarse facial features, jerky movement, widely spaced teeth, large tongue, mild ID, flat nasal bridge
D	IV-1	10.5y	Mohmand	Delay in achieving milestone, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/kyphosis, hernia, broad band
	IV-2	9y		Delay in achieving milestone, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/kyphosis, hernia, broad band
	IV-3	8y		Delay in achieving milestone, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/kyphosis, hernia, broad band
E	VI-1	10y		Aggressive behavior, short stature, gibbus formation, waddling gait, joint contractures, short trunk, scoliosis/kyphosis, liver problem, anemia, sleep apnea
	VI-2	8y	Qureshi	Aggressive behavior, short stature, gibbus formation, waddling gait, joint contractures, short trunk, scoliosis/kyphosis, liver problem, anemia, sleep apnea
F	IV-1	13y	Rana	Short stature, gibbus formation, joint contractures, short trunk, scoliosis/kyphosis, coarse facial features, spondyloepiphyseal dysplasia, Height = 80cm, Weight = 10kg

G	IV-3	4y	Awan	Speech delay, abdominal distention, short stature, gibbus formation, respiratory/ear infection, coarse facial features, short trunk, joint contractures, scoliosis/kyphosis, hernia, hepatosplenomegaly
	IV-5	3y		Speech delay, abdominal distention, short stature, gibbus formation, respiratory/ear infection, coarse facial features, short trunk, joint contractures, scoliosis/kyphosis, hernia, hepatosplenomegaly (maternal twins)
	IV-6	3y		
H	V-3	5y	Pakhtoon	Delay in achieving milestones, short stature, coarse facial features, scoliosis/kyphosis, abdominal distention, hepatosplenomegaly, joint contractures, beaking of chest Hb = 11.8 g/dl (12-14), Bilirubin = 16.4 mg/dl (0.1-1.0), Serum growth hormone = 1.62 ng/mL (2-5)

F: female, Hb: hemoglobin, HCT: hematocrit, ID: intellectual disability, Kg: kilogram, M: male, y: years

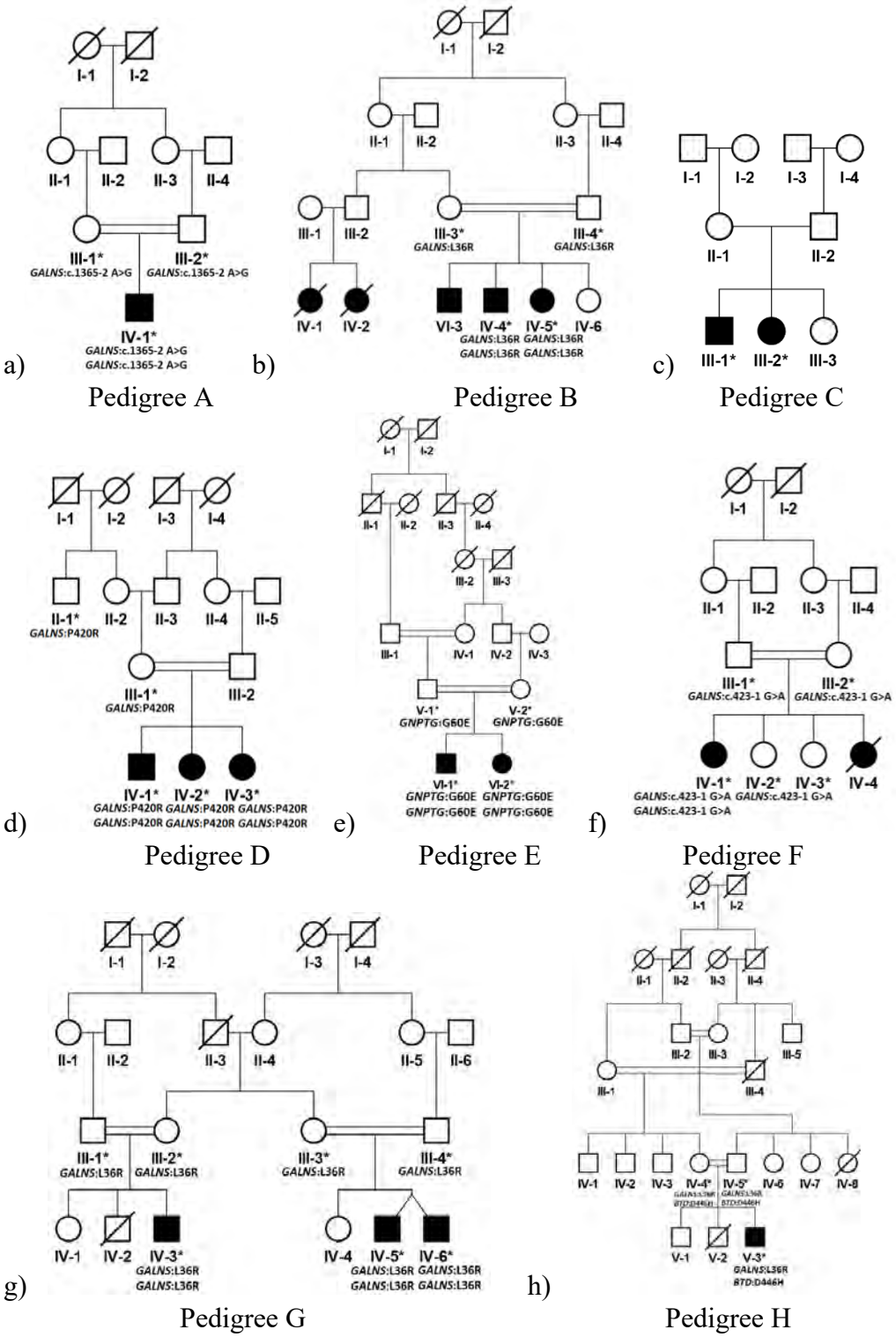


Figure 3.14: Pedigrees of families inheriting MPS IV phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

Table 3.14: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in eight MPS IV recruited patients subjected to Whole Genome Sequencing.

Pedigree/ Patient ID	Gene variants	Genotype	dbSNP	*In-Silico* prediction for pathogenesis	ACMG classification	Varsome	HGMD accession number
A/IV-1	NM_000512.5 (*GALNS*); c. 1365-2A>G	Homozygous	rs1281515662	N/A	Likely pathogenic PVS1, PM2	Likely pathogenic	Current study
B/VI-3,VI-4,VI-5	NM_000512.5 (*GALNS*); c.107A>C (p.Leu36Arg)	Homozygous	rs755832705	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571
C/III-1,III-2	Unsolved	N/A	N/A	N/A	N/A	N/A	N/A
D/IV-1,IV-2,IV-3	NM_000512.5 (*GALNS*); c.1259G>C (p.Pro420Arg)	Homozygous	N/A	9/9	Uncertain significance PP3, PM2, PP5	Uncertain significance	CM145597
E/VI-1,VI-2	NM_032520.5 (*GNPTG*); c.179G>A (p.Gly60Glu)	Homozygous	N/A	8/9	Uncertain significance PP3, PM2	Uncertain significance	Current study
F/IV-1	NM_000512.5 (*GALNS*); c.423-1G>A	Homozygous	N/A	N/A	Pathogenic PVS1, PP5, PM2	Pathogenic	Current study
G/IV-3/IV-5,IV-6	NM_000512.5 (*GALNS*); c.107A>C (p.Leu36Arg)	Homozygous	rs755832705	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571
H/V-3	NM_000512.5 (*GALNS*); c.107A>C (p.Leu36Arg)	Homozygous	rs755832705	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571
	NM_001281723.2 (*BTD*); c.1336G>C (p.Asp446His)	Homozygous	N/A	8/9	Pathogenic PS3, PS1, PP5, PM1, PM5	Pathogenic	Current study

ACMG: american college of medical genetics, HGMD: human genome mutation database, N/A: not applicable, PM: moderate pathogenic, PP: possibly pathogenic, PS: strong pathogenic, PVS: very strong pathogenic

3.3.5. Study of Clinical Aspects and Genetic Screening of Three Gaucher Disease Families.

Among six collected families based on their presenting clinical profiles as GD affected, three were subjected to Sanger Sequencing at Molecular Biology lab, Department of Zoology, Faculty of Biological Sciences, Quaid-I-Azam University, Islamabad, Pakistan as mentioned in section 3.3.2. While three families were subjected to Whole Genome Sequencing at Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark during six months stay funded by the Higher Education Commission, Pakistan, under International Research Support Initiative Program (Annexure 2.1). Their clinical profiles are quite resembling with each other (Table 3.15). All patients were born to first degree consanguineous couples (Figure 3.15).

Whole genome sequencing revealed two novel variants of *NPC1* gene, NM_000271.5 (*NPC1*); c.1610T>C, (p. Phe537Ser), and NM_000271.5 (*NPC1*); c.3293C>T, (p. Thr1098Ile), in pedigree A. All the variants showed heterozygous carrier status in parents and homozygous mutant in patients. Both variants in *NPC1* found in pedigree A are previously not reported in HGMD for causing any pathogenesis. The *In-Silico* analysis showed the pathogenesis of the variant by applying various tools (Table 3.16). Remaining two families, pedigree B and pedigree C remained unsolved and could not be identified with any genetic variant in coding regions.

Table 3.15: Clinical details, age and cast of three unrelated patients with Gaucher disease subjected to Whole Genome Sequencing.

Pedigree no.	Patient ID	Age	Cast	Clinical profiles
A	IV-1	9m	Suddhan	Delay in achieving milestones, respiratory/ear infections, abdominal distention, weak muscles, hepatosplenomegaly, hyper cellular fragments and trails, microcytosis, hypochromia, anisocytosis, poikilocytosis, target cells, fragmented cells WBC = 8.6/cmm (6000-14500), Hemoglobin = 9.8 g/dL (11-15), Hematocrit = 31.5% (40-68), MCV = 55.9 fl (62-98), MCH = 17.4 pg (25-32), Platelets = 177/cmm (140,000-450,000), Red cell count = 6.25×10^6/uL, MCV = 53.4 fL (78-96), MCHC = 16.8 pg (27-32), Reticulocyte count = 0.1% (1.0-6.0), ESR = 15 mm/hr (0-12), Neutrophils = 38% (40-80), Lymphocytes = 49% (20-40), ALP = 214 U/L (upto 135)
B	IV-3	1.5y	Chodhary	Delay in achieving milestones, abdominal distention, hepatosplenomegaly, anemia, weak muscles, microcytosis, hypochromia, anisocytosis Hemoglobin = 8.8 g/dL (12-15), HCT = 29.2 (35-55), MCV = 67.5 fL (75-100), MCH = 20.4 pg (25-35), MCHC = 30.2 g/dL (31-38), Platelet count = 2×1000/uL (100-400), RDW-CV = 20% (11-15), Neutrophils = 37% (60-70), Lymphocytes = 55% (30-40), Monocytes = 6% (0-5), Iron binding capacity = 419 ug/dL (100-400)
C	V-2	1.5y	Abbasi	Delay in achieving milestones, abdominal distention, hepatosplenomegaly, nasal bleeding, anemia, anisocytosis, fragmented cells, macrocytes, leucoerythroblastic picture, poikilocytosis, polychromasia Red cell count = 2.09 million/ul (3.5-5.5), Hemoglobin = 5.8 g/dL (12-15), HCT = 18.5 (35-55), Platelet count = 10×1000/uL (100-400), RDW-CV = 22.9% (11-15), Neutrophils = 25% (60-70), Lymphocytes = 60% (30-40), Monocytes = 8% (0-5), ALP = 216 U/L (upto 135), Serum urea = 10 mg/dL (12-50), Serum Sodium = 133 m.mol/l (135-145), Iron = 8.8 µg/dL (40-100), TIBC = 495 µg/dL (100-400), Ferritin = 224 ng/ml (5-67), Uric acid = 1.5 mg/dL (4.5-8.2)

ALP: alkaline phosphatase, ESR: erythrocyte sedimentation rate, HCT: hematocrit, m: months, MCV: mean corpuscular volume, MCHC: mean corpuscular hemoglobin concentration, MCH: mean corpuscular hemoglobin, RBC: red blood cells, RDW-CV: red cell distribution width, TIBC: total iron binding capacity, WBCs: white blood cells, y: years

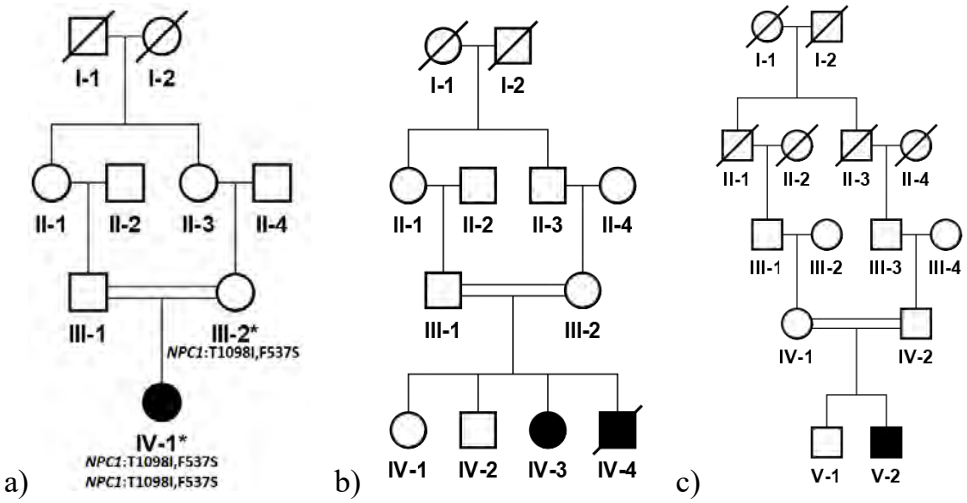


Figure 3.15: Pedigrees of families inheriting Gaucher disease phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual

Table 3.16: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in family recruited with Gaucher disease and subjected to Whole Genome Sequencing.

Pedigree/ Patient ID	Gene variant	Genotype	*In-Silico* predictions for pathogenesis	ACMG classification	Varsome	HGMD accession number
A/IV-I	NM_000271.5 (*NPC1*): c.3293C>T: p.(Thr1098Ile)	Homozygous	8/9	Likely pathogenic PM2, PP3	Likely pathogenic	Current study
	NM_000271.5 (*NPC1*): c.1610T>C: p.(Phe537Ser)	Homozygous	8/9	Uncertain significance PM5, PM2, PP3	Uncertain significance	Current study

ACMG: American college of medical genetics, HGMD: human genome mutation database, PM: moderate pathogenic, PP: possibly pathogenic

3.3.6. Study of Clinical Aspects and Genetic Screening of two Niemann-Pick Disease Families.

Two families were suspected as Niemann-Pick Disease affected based on their presenting clinical profiles and bone marrow diagnosis. Most of the phenotypes among patients were overlapping like abdominal distention hepatosplenomegaly, delay in achieving milestones (Table 3.17). All patients were born to first-degree consanguineous couples (Figure 3.16).

Whole genome sequencing revealed a homozygous sequence variant; NM_000543.5 (*SMPD1*); c.1268A>G, (p. His423Arg) in pedigree A and B. All the variants showed heterozygous carrier status in parents and homozygous mutant in patients. The missense variant of *SMPD1* found in pedigree A and B has been previously reported for causing NP disease in population of European/White and Ashkenazi Jewish origin (McGovern *et al.,* 2006). The *In-Silico* analysis showed the pathogenesis of the variant by applying various tools (Table 3.18).

Table 3.17: Clinical details, age and cast of two unrelated patients with Niemann Pick disease subjected to Whole Genome Sequencing.

Pedigree no.	Patient ID	Age	Cast	Clinical profiles
A	IV-1	4y	Pakhtoon	Delay in achieving milestones, cannot walk, ear problem, recurrent seizures, hypopigmented spots on trunk, facial angiofibroma, cherry red spot at macula, splenomegaly, exaggerated reflexes, no speech, very loose body cannot with stand
B	VI-1	5m	Pakhtoon	Delay in achieving milestones, cherry red spot, abdominal distention, respiratory/ear infections, hepatosplenomegaly TLC = 24.7×10^9/L (6-18), MCH = 23.6 pg (24-30), RDW-CV = 14.7% (11-14), Neutrophils = 30% (60-70), Lymphocytes = 57% (30-40), Polychromasia +/+, Reticulocyte's count = 0.7% (2.0-6.0), Cholesterol-S = 5.2 mmol/L (1.8-4.5), Hemoglobin = 11.5 g/dL (12-15), MCV = 72.3 fl (75-100), MCH = 23.2 pg (25-35), Neutrophils = 30.1% (60-70), Lymphocytes = 56.4% (30-40), Serum Urea = 11 mg/dL (12-50), Serum ALT = 328 U/L (Upto 42), Serum Alkaline Phosphatase = 500 U/L (Upto 135)

ALT: alanine transaminase, Hb: hemoglobin, HCT: hematocrit, m: months, MCV: mean corpuscular volume, MCH: mean corpuscular hemoglobin, RDW-CV: red cell distribution width, TLC: total leukocyte count, y: years

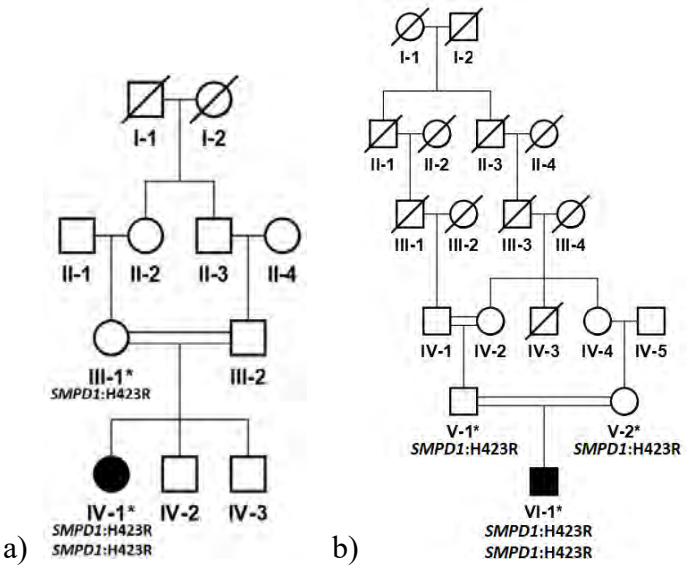


Figure 3.16: Pedigrees of families inheriting Niemann Pick disease phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

Table 3.18: Genotype, *In-Silico* analysis and ACMG classification status of the identified variants in two Niemann-Pick disease recruited patients subjected to Whole Genome Sequencing.

Pedigree/ Patient ID	Gene variant	Genotype	*In-Silico* predictions for pathogenesis	ACMG classification	Varsome	HGMD accession number
A/IV-1	NM_000543.5 (*SMPD1*); c.1268A>G: (p.His423Arg)	Homozygous	7/9	Pathogenic PM1, PM5, PM2	Pathogenic	CM068450
B/VI-1	NM_000543.5 (*SMPD1*); c.1268A>G: (p.His423Arg)	Homozygous	7/9	Pathogenic PM1, PM5, PM2	Pathogenic	CM068450

ACMG: American college of medical genetics, HGMD: human genome mutation database, PM: moderate pathogenic

3.3.7. Study of Clinical Aspects and Genetic Screening of Pycnodysostosis Family.

The patient is a 6 year old boy of a first degree cousin couple (Figure 3.17). The patient was anemic with low hemoglobin level (11.1g/dL). The height was 94 cm and weight were 12 kg. Nails were deshaped and had fungal infection. There was overcrowding of teeth, PCV/HCT was lower than the marginal value i.e 34.3 (35-55), platelets counts were higher i.e 537×1000/µl (100-400), RDW-CV value is also higher than the marginal value i.e 15.2% (11.6-15), neutrophils are lower than the marginal value i.e 52.4% (60-70), serum calcium is higher than the marginal value i.e 2.7 mmol/l (1.0-2.4) and RBCs are also higher in urine 5-6 HPF (0-1 HPF).

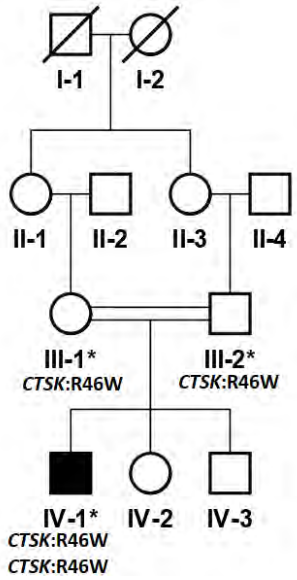


Figure 3.17: Pedigree of family inheriting Pycnodysostosis phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

Whole Genome sequencing revealed a previously reported variant NM_000396.3 (*CTSK*); c. 136C>T, (p. Arg46Trp) (Schilling *et al.,* 2007; Kausar *et al.,* 2017) in exon 3 of *CTSK* gene. The variant is in homozygous pattern in the patient while both parents were confirmed as heterozygous carriers. As per ACMG the variant

is pathogenic (PP5, PM5, PP3, PM2). According to HOPE, tryptophan is bigger in size, neutral, and hydrophobic in nature as compared to the wild type of residue i.e Arginine. Clustal-W has shown a highly conserved pattern for the wild type of Arginine at position 46 in different orthologues (Figure 3.18A). PMut predicted the variation as disease causing with a score of 0.58 (82%). PhD-SNP has a reliability index of 7 and called it as disease-related polymorphism. The results of PANTHER also confirmed this variant as probably damaging with a probability of deleterious effect with a score of 0.89 and preservation time as 1629 million years. The interaction of CTSK protein with other proteins is mentioned in figure 3.18B. I-Mutant has shown decrease in stability of protein with a DDG score of -0.98 due to substitution of Arginine with Tryptophan. MUpro also revealed the decrease in protein stability with delta g score of -1.07. SIFT explained this variation affects protein structure with a score of 0.00 (<0.05 deleterious). PROVEAN also confirms this variant a deleterious with score of -6.11 (cutoff value = -2.5). PolyPhen-2 explains it as probably damaging with a score of 1.00. Fathmm predicted this variant as damaging with a score of -2.90. And MutPred2 probability score 0.28 and P-value of 0.01 (<0.05) with relative loss of solvent accessibility.

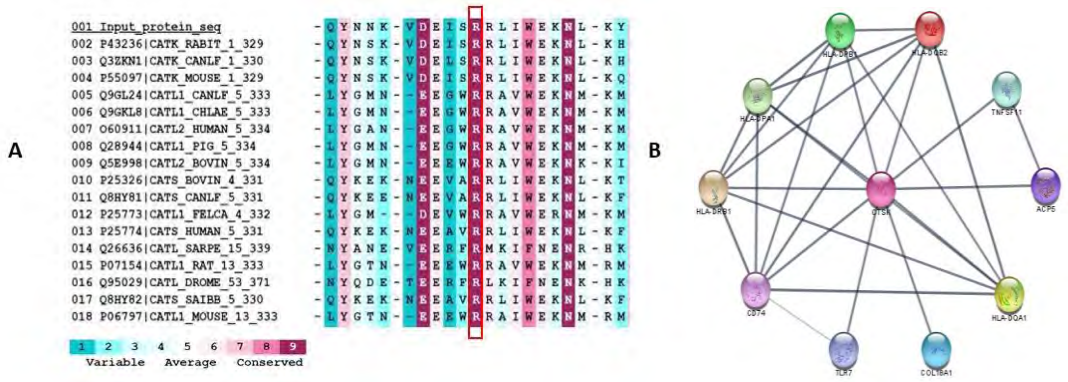


Figure 3.18: (A) Conservation pattern of Arginine at position 46 in different orthologues. (B) CTSK protein interactions with other functionally active proteins.

3.3.8. Study of Clinical Aspects and Genetic Screening of Griscelli Syndrome Family.

One of the recruited family affected with one of the silver hair syndromes termed as Griscelli syndrome is subjected to WGS and found to be affected with type 1 caused by the mutation in gene *RAB27A* (Table 3.19). The identified variant is NM_183235.3 (*RAB27A*); c.258_260dup, (p. Ala87dup). The identified variant is present in homozygous form in both affected siblings and also confirmed as heterozygous carrier form in mother. As per ACMG, variant is of uncertain significance (PM4, PM2). One of the affected siblings has been died previously from same condition as mentioned by the parents. Parents were having first degree consanguinity and also showing various cousin marriages in the previous successive generations (Figure 3.19). The STRING analysis showed MLPH and MYO5A as strong functional partners with a predicted score of 0.999 (Figure 3.20).

Table 3.19: Clinical details, age and cast of patients with Griscelli syndrome subjected to Whole Genome Sequencing.

Pedigree no./Patient ID	Age	Cast	Clinical profiles
A/VI-1,VI-2	6.5y	Rajpoot	Silvery grey hairs, delay in achieving milestone, chronic diarrhea, aggressive behavior, speech babbling, blond hair, abdominal distention, waddling gate, unable to walk, stand only with support for short time, few episodes of fits, fever, vomiting, failure to thrive, hypotonia/floppy, lethargy/drowsy, RDW-CV= 14.2 % (11-14), Neutrophils = 21.8 % (60-70), Lymphocytes = 65.7 % (30-40)
	5.5y		Delay in achieving milestone, chronic diarrhea, aggressive behavior, speech babbling, blond hair, abdominal distention, broad based gate TLC = 18.1 $\times 10^9$/L, (6-18), Haemoglobin = 10.9 g/dL (11.1-14.1), MCH = 24.8 pg (25-29), Neutrophils = 0% (60-70), Lymphocytes= 83.1% (30-40)

MCH: mean corpuscular hemoglobin, RDW-CV: red cell distribution width, TLC: total leukocyte count, y: years

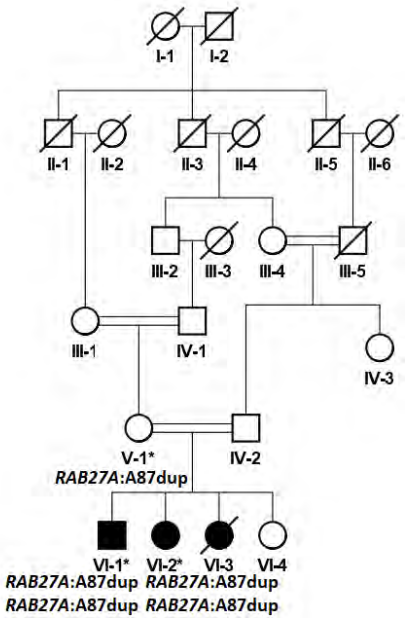


Figure 3.19: Pedigree of family inheriting Griscelli syndrome phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

Figure 3.20: RAB27A protein interactions with other functionally active proteins including MLPH and MYO5A.

3.3.9. Study of Clinical Aspects and Genetic Screening of Chediak-Higashi Syndrome Family.

The patient recruited on the phenotypic basis of Chediak-Higashi syndrome like grey-silver hairs, weak body, delay in achieving milestones, etc. remained unsolved after whole genome sequencing data analysis targeting the coding regions. The patient was born to a consanguineous couple with first degree of consanguinity and a family history in previous generation depicted with the same clinical presentations according to the information provided by one of the family member (Figure 3.21).

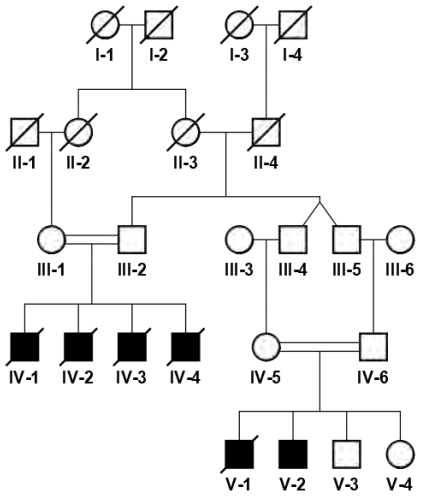


Figure 3.21: Pedigree of family inheriting Chediak-Higashi syndrome phenotype.

Squares and circles denote males and females respectively. Filled and unfilled symbols indicate affected and unaffected individuals respectively and consanguinity is represented by double marriage lines. Asterisk symbol above square or circle shows availability of the individual. Genotypes of participating individuals are shown below symbol of each individual.

4. DISCUSSION

During current study, 45 families were enrolled, and their initial recruitment was based upon their phenotypic presentation and basic clinical evaluations as prescribed by the physician like blood tests, radiological findings, urine sample analysis etc. Initial diagnosis was made by an expert pediatrician and all families diagnosed to be affected with different sub-types of LSDs like MPS I, MPS II, MPS III, MPS IV, Niemann Pick disease, Griscelli disease and Gaucher disease were enrolled. There are very few studies published from Pakistan regarding LSDs including MPS I, MPS II, MPS IV, Niemann Pick, Pycnodysostosis, and Gaucher disease (Kausar *et al.,* 2017; Ullah *et al.,* 2017; Zahoor *et al.,* 2019; Zubaida *et al.,* 2019; Cheema *et al.,* 2020). These studies were composed mainly of the patients from Punjab region, and some patients from KPK as well. As LSDs are rare genetic disorders and mainly prevalent in consanguineous populations like Pakistan. Majority of the LSDs are autosomal recessive, so their expected prevalence is also high in our population where every 8/10 marriages are consanguineous (Gul *et al.,* 2021). The diagnosis of LSD's is complicated and the gold standard method for diagnosis include mass spectrophotometry to measure enzyme levels in the leukocytes or fibroblasts which is one of the most accurate and quick technique. However, if any LSD is due to defect in a non-enzymatic protein, then mass spectrometry can lead us to false negative results. For this purpose, whole genome sequencing or whole exome sequencing are proved to be the best methods for identifying the underlying causes of pathophysiology and provide differential diagnosis for sub-types as well (Han and Lee, 2020). As per previous studies including current study published from Pakistan, the mortality and morbidity ratio of LSDs affected patients is high due to lack of early and advance diagnostic facilities and proper medical treatment like enzyme replacement therapy (ERT) (Gul *et al.,* 2021). There are multiple approved enzymes for ERTs to treat various LSDs involving enzymatic protein deficiencies that are currently being used in the developed countries which are lacking here is Pakistan except for Aldurazyme® for treating MPS I, Elaprase® for treating MPS II and Vpriv® for treating Gaucher disease on very limited access and availability throughout country. These enzymes are being used for treating patients only in one public sector hospital in Lahore with the help of a Non-Government Organization (NGO) from the

United States of America. For a country like Pakistan heavily populated with more than 220 million people and high ratios of consanguinity, this facility equalizes to nothing. Early and on time diagnosis is the key to getting some control over early morbidity and mortality in the patients. Across the globe, the treatment options in practice could only be effective if the diagnosis is early. For example, for bone marrow transplant (BMT) or hematopoietic stem cell transplant (HSCT), if the disease is identified in the patient during first year of life, then it will be possible to revert symptoms except the ones involving central nervous system. The same limitation is faced by ERT, as the exogenous enzyme is large enough to cross the blood brain barrier and is ineffective in neurological manifestation cases (Afroze and Brown, 2017; Scarpa *et al.,* 2017). About 60% of the LSDs involve CNS involvement (Gul *et al.,* 2021). However, there are a lot of treatment options in their final clinical trials that could be the ray of hope for treating such disorders e.g., pharmacological chaperon therapy in which chaperons are small enough to cross the blood brain barrier and proved to be helpful by stopping the progression of the disease in the patient with growing age and also by reverting the symptoms (Schiffmann *et al.,* 2020; Gul *et al.,* 2021). In a recently published study, there is a successful in-utero enzyme replacement therapy for treating infantile-onset Pompe disease, where after receiving in-utero ERT and then standard post-natal therapy the patient is showing normal development in milestones (Cohen *et al.,* 2022).

In current study due to diagnostic challenges e.g., unavailability of enzyme analysis, diagnosis based on presenting features and radiological findings, after performing Next Generation Sequencing (NGS) there are deviations of sub types of LSDs between initially diagnosed cases and confirmation after performing whole genome sequencing. This study includes cost benefits of WGS supported financially by the HEC, Pakistan and Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark. Here, MPS I cases have deviation ratio of 83% (7/12), MPS III has 16% (1/6), MPS IV has 25% (2/8) and GD has 100% (3/3) Figure 4.1 (A and B). There are also six families identified with multigenic inheritance, confirming the homozygous status of each variant in the affected patients and heterozygous carrier's parents. Due to complexity of phenotypes, late presentation due to lack of awareness, poor diagnostics, even no enzyme analysis for all suspected

cases and unavailability of molecular genetic testing, there is increased incidence of premature mortality caused by LSDs in our population.

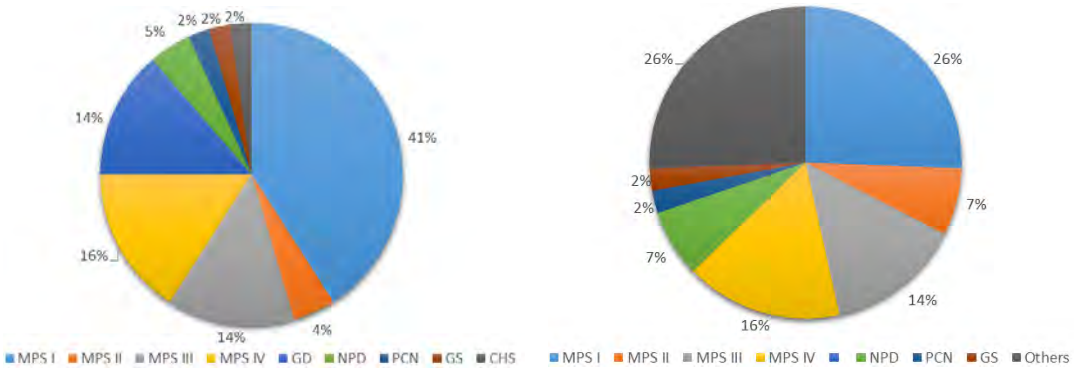


Figure 4.1: (A) Pie chart showing families recruited on the basis of their phenotypic and initial clinical diagnosis. (B) Pie chart showing families diagnosed with LSDs subtypes and other disorders after Whole Genome Sequencing.

4.1. LSDs' Families Subjected to Sanger's Sequencing

Thirty-five of the families were collected from Children hospital, Pakistan Institute of Medical Sciences, Islamabad on course of two years. After one year of collection, nine families including six MPS I and three GD families, were subjected to Sanger's sequencing.

4.1.1. Mucopolysaccharidosis type I

Among mucopolysaccharidosis, MPS I with different phenotypic variations is found to be more prevalent sub-type in the world including Pakistani population (Chkioua *et al.,* 2011; Gul *et al.,* 2021). The ratios of different phenotypes of MPS 1 were reported as Hurler 60.9%, Hurler/Scheie 23%, Sheie 12.9% and 3.2% undetermined by *Beck et al.,* (2014). Variants in a single gene are responsible for three MPS 1 phenotypes; thus, identification of causative variants for each patient is important to develop genotype-phenotype correlation (De R *et al.,* 2012). Consequently, molecular genetics diagnosis leads to therapeutic management as hematopoietic stem cell transplantation is used for patients presenting with severe phenotype i.e., Hurler and show better results with earlier age at treatment (Ghosh *et al.,* 2017). This data necessitates molecular genetic screening of MPS 1 cases

especially from consanguineous populations that will not only identify prevalent variants to establish molecular genetic diagnostic testing but will have clinical utility for disease management. For this purpose, the present study was performed to screen *IDUA* gene in the six MPS 1 segregating unrelated consanguineous Pakistani families through Sanger Sequencing (Table 3.2). Variant's analysis of *IDUA* gene confirmed one homozygous disease-causing variant in each of our screened families.

In pedigree A the novel nonsense mutation p. Glu486Ter, is responsible for causing the phenotype. The patient is of 16 months and has severe phenotype with presenting features like asthma, sleep apnea, joint contractures, kyphosis, ear/nose/throat infections, coarse facial features, hepatosplenomegaly, inguinal hernia, aggressive, heavy eyebrows, large tongue, thick lips and rough skin. The variant was neither found in ExAC nor in 1000 genome browser. On further analysis of the variant this is the deleterious mutation as it causes the truncation of protein which is otherwise 653 amino acid long. The trimmed protein is of 485 amino acid which is degraded through process of non-sense mediated decay (NMD). All transcripts in the mammals which generate pre-mature termination codon (PTC), 50-55 base pairs upstream of exon-exon junction complex (EJC) undergo the process of NMD. Thus, there is no α-L-iduronidase enzyme production for degradation of GAGs and results in their excessive accumulation inside the cells which causes the severe phenotype i.e. Hurler. Non-sense mutations are more prevalent in Hurler phenotype with ranging ratio of 60-80% (Maquat, 2005). The mutation we found in our patient is also non-sense, thus we can relate the severe phenotype of our patient with the genotype.

A previously reported missense variant i.e., p. Leu490Pro was detected in three out of six analyzed families, and it also provided the possibility of this variant to be a founder variant in Pakistani cohort but still more studies are required to confirm its status (Table 3.3) (Tieu *et al.,* 1995). The disease phenotype of all of p. Leu490Pro carrying patients was Hurler-Scheie similar to previously reported Pakistani patients carrying same variant (Arora *et al.,* 2007; Zahoor *et al.,* 2019; Gul *et al.,* 2020). The substitution of leucine at position 490 with proline may affect enzymatic activity of protein causing an intermediate phenotype (Ghosh *et al.,* 2017). The nature and position of substituted amino acid and also newly inserted amino acid describe the

impact of that very substitution on structure of protein (Firasat *et al.,* 2018). According to HOPE, leucine at position 490 is located in helix-loops-helix region (482-508 amino acids residues) of β-sandwich domain (482-508) which is important for interaction with other two domains of enzyme, and its substitution with proline may result in loss of this interaction (Maita *et al.,* 2013; *Gul et al.,* 2020).

Sequencing of *IDUA* coding sequence in remaining two families (pedigree B and D) revealed, a novel deletion of 14 nucleotides (c. 568_581delAACGTCTCCATGAC) in exon 5 and a previously reported single nucleotide deletion i.e., c.784delC, in exon 6 respectively. According to mutation taster both deletions lead to the frameshift of the coding sequence thus creating a pre-mature stop codon i.e., c. 568_581delAACGTCTCCATGAC leading to Asn190Hisfs*204 and c.784delC causing p. His262Thrfs*55 with truncated proteins of 392 and 315 amino acids respectively. Previously, c.784delC causing p. His262Thrfs*55 was reported by Al-Jasmi *et al.,* 2012 in an MPS 1 patient of Pakistani origin with central nervous system involvement indicating Hurler Phenotype (Al-Jasmi *et al.,* 2012). Here in our study Pedigree F patient carrying same variant also had by birth onset of the disease. According to Maquat, 2005, all those transcripts which generate pre-mature termination codon at 50-55 base pairs upstream of exon-exon junction complex undergo the process of non-sense mediated decay (NMD) in mammals (Maquat, 2005). Furthermore, *IDUA* variants which cause premature protein truncations lead to absence of enzyme and severe phenotype (Keeling *et al.,* 2013; Huang *et al.,* 2018). Thus, both variants lead to truncating proteins predicted to be degraded by NMD leading to absence of α-L-iduronidase enzyme and excessive accumulation of GAGs hence showing severe phenotype with by birth onset of the disease and neuropathic involvement in our patients. Furthermore, asparagine amino acid at 190 positions of IDUA protein is among one of the six potential sites i.e N110, N190, N336, N372, N415 and N451 which are subjected to post-translational glycosylation (Scott *et al.,* 1991), thus substitution with histidine at this position may impact protein function as described previously by Maita *et al.,* (2013).

Remaining 17 variants including 9 missense and 8 synonymous were benign variants according to Mutation taster (Table 3.4). Previously it was reported that benign variants may also contribute to variable enzyme activity in healthy subjects

(Scott *et al.,* 1993) thus may contribute to the phenotypic variability among patients, along with different pathogenic variants. Consequently, including one novel nonsense mutation and one novel deletion, we described four variants and their genotypic-phenotypic correlation in the Pakistani MPS 1 patients.

4.1.2. Gaucher disease

Previous reports suggested that among Gaucher disease (GD) cases, type 1 GD is the most common which is non-neurological form of the disease (Marques and Saftig, 2019; Sheth *et al.,* 2019). However, among our three GD affected families, affected members of all except pedigree 1 displayed neurological symptoms (Table 3.6). To unveil mutational spectrum of *GBA* in our enrolled GD families we performed Sanger's Sequencing of coding exons of *GBA* gene. Data analysis showed a previously reported missense mutation p. Leu483Pro (rs421016), segregating with disease phenotype in all three families (Table 3.6). Recently, Sheth *et al.,* 2019 identified p. Leu483Pro as the most common GD causing mutation in Indian population irrespective of ethnic groups which is consistent with our findings (Sheth *et al.,* 2019). This mutation was also reported to be detected with 41% Japanese, 60% Thai and 2.84% Jewish GD cases (Grabowski, 1997; Eto and Ida, 1999; Tammachote *et al.,* 2013). Although p. Leu483Pro is reported to be associated with all subtypes of GD (Koprivica *et al.,* 2000; Tammachote *et al.,* 2013; Sheth *et al.,* 2019) but Koprivica *et al.,* 2000 suggested that p. Leu483Pro is mainly associated with GD sub types showing neurological involvement which is also found in this study (Koprivica *et al.,* 2000). Screening of *GBA* gene in our enrolled families and identification of same mutation in all of them suggest high prevalence of p. Leu483Pro allele in our population. Affected individual of one of our GD family (pedigree A) segregating p. Leu483Pro has no neurological symptoms which is consistent with GD type 1, but the possibility of neuronal manifestations at the later stage of life could not be ruled out as reported by Alfonso *et al.,* 2007.

Leucine at 483 positions of GBA protein lye in β-barrel domain (residues 30–75 and 431–497) of Beta-glucosidase which resembles an immunoglobulin fold (Lieberman *et al.,* 2009). According to previous studies this substitution leads to alteration of secondary structure of protein also this variant lye close to the catalytic

domain of enzyme thus results in loss of enzyme activity. This is supported by our HOPE data as shown in figure 3.9D, the mutated residue is in a domain which is important for the activity of the protein and leucine 483 makes contact with residues in another domain (Figure 3.9D) further highlighting importance of this amino acid in wild type structure and function of GBA protein.

4.2. LSDs' Families Subjected to Whole Genome Sequencing

Remaining 36 families (excluding nine analyzed by Sanger's Sequencing) were subjected to the whole genome sequencing to identify their genetic variations at University of Copenhagen, Denmark during six months research stay funded by the Higher Education Commission, Islamabad, Pakistan (Annexure 2.2).

4.2.1. Mucopolysaccharidosis type I

In the present study, out of eighteen families, twelve (A-L) were clinically characterized as MPS I. The diagnosis was based on clinical phenotypes including facial features, stature, laboratory tests and inheritance of the phenotypes in pedigree. WGS revealed homozygous variants in previously MPS-associated genes (*IDUA, GALNS, SGSH* and *IDS*) and a novel LSDs-associated candidate gene (*ABCA5*) in affected individuals. In addition to these genes, variants were found in seven genes (*ALDOB, MASP1, KIAA1109, TRAPPC4, SMARCAL1, HERC1, RRAS2*) that were previously not reported in MPS but associated with other metabolic disorders.

Our findings in pedigree A (i.e., MPS I), reported a homozygous sequence variant (c.871G>A; p. Ala291Thr) in *GALNS*. The gene encodes enzyme N-acetylgalactosamine-6-sulfatase, a lysosomal exohydrolase which is required for the degradation of the glycosaminoglycans, chondroitin 6-sulfate and keratan sulfate. Previously, the identified variant (p. Ala291Thr) has been reported in compound heterozygous and homozygous forms in two different families of Pakistani origin segregating MPS IV (Tomatsu *et al.,* 1995; Ullah *et al.,* 2017). *In-Silico* analysis revealed that missense variant (p. Ala291Ser) leads to the formation of additional intramolecular interactions, most likely affecting the conformation of the active site (Asp288 and Asn289) and its substrate binding ability. Therefore, variation at amino

acid position 291 of GALNS would affect structure and enzymatic function of the protein.

Multiple variants in *IDUA* gene have been identified and associated with MPS I (Tieu *et al.,* 1995). A novel homozygous insertion variant (c.1172-1173insGCTGCTGGC; p. Gly391insLeuLeuAla) in *IDUA* was identified in an affected individual in pedigree B. Glycine at position 391 is involved in the formation of torsion angles due to its high degree of rotational ability which maybe lost after the insertion of LLA leading to abnormal function of the enzyme. Furthermore, affected individuals in pedigree E, F, I and K also revealed a previously reported variant (c.1469T>C; p. Leu490Pro) in *IDUA* (Gul *et al.,* 2020). In these families, a high interfamilial variability in the phenotypes of patients segregating same sequence variant (*IDUA*; c.1469T>C; p. Leu490Pro) was observed in our pedigrees (E,F,I,K). Deep analysis of the genome of individuals revealed additional pathogenic variants in genes (c.1297A>G; p. Thr4033Ala in *KIAA1109* in pedigree I, and c.C73T; p. Arg25Ter in *MASP1* in pedigree K) that may act as modifiers or dual molecular diagnosis. Previously sequence variants in *KIAA1109* (4q27) encoding a protein playing role in endosomal trafficking and endosome recycling of lipids have been reported to cause Alkuraya-Kucinskas syndrome characterized by brain abnormalities associated with the underdevelopment of cerebral parenchyma, arthrogryposis, clubfoot, and global developmental delay (Kumar *et al.,* 2020). Patients in family K showed phenotypes of MPS as well as Alkuraya syndrome. Similarly, sequence variants in *MASP1* (3q27.3) encoding a serine protease having a crucial role in the innate and adaptive immunity and have been associated with 3MC syndrome 1 characterized by widely spaced eyes (hypertelorism), a narrowing of the eye opening (blepharophimosis), droopy eyelids (ptosis), highly arched eyebrows (Durmaz and Altıner, 2021). Patients in the present pedigree K showed phenotypes of 3MC syndrome in addition to MPS I.

IDUA, KIAA1109 and *MAPS1* are located at different loci of human genome. Due to higher rate of consanguinity, the number of homozygous regions increases leading to a high risk of occurrence of multiple pathogenic variants. This might lead to complex phenotypes due to co-segregation of two syndromes in the same patient.

Therefore, analyzing genomes of populations with a high rate of consanguinity, dual molecular diagnosis should be considered.

In pedigree C, we can appreciate a multi-genic inheritance, segregating two damaging variants in known LSDs enzymes *SGSH* (c.220G>A; p. Arg74Cys) and *GAA* (c.1930G>T; p. Ala644Ser) in homozygous state, both of which lie within the mapped interval at 17q25.3. *SGSH* encodes a lysosomal enzyme sulfamidase involved in the lysosomal degradation of heparan sulfate. Pathogenic sequence variants in the gene lead to MPS IIIA (Weber *et al.,* 1997). Arginine at amino acid position 74 is likely to be involved in the formation of the active site of sulfamidase. Transition of arginine at the position to a non-conserved Cysteine may grossly affect enzyme activity. *GAA* encodes a lysosomal enzyme alpha-glucosidase. The lysosomal alpha-glucosidase is essential for the degradation of glycogen to glucose in lysosomes. Defects in the enzyme lead to glycogen storage disease. In addition to MPS phenotypes associated with variants in *SGSH*, the patient showed low hemoglobin concentration which is found in glycogen storage disease. These findings, suggest that the disease phenotypes in the present patient could be due to the contribution of variants in both genes, *SGSH* and *GAA*.

In pedigree D, we have identified a missense homozygous variant (c.26569G>A; p. Arg857Cys) in *ABCA5*. The ABCA5 protein is a member of the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins play an important role in transporting various molecules across extra- and intracellular membranes. Loss of function variants in human *ABCA5* have been associated with hair overgrowth in human (De Stefano *et al.,* 2014; Raza *et al.,* 2022). On the other side, abca5-/- knockout mice developed lysosomal disease-like symptoms (Kubo *et al.,* 2005). Interestingly, the affected individual in pedigree 4 carrying *ABCA5* missense variant showed phenotypes overlying with MPS associated with hair overgrowth. To the best of our knowledge, this is the first human case associating a variation in *ABCA5* with LSD associated with hair overgrowth.

ABCA5 has 1642 amino acids long lysosomal peptide which is composed of two domains including ABC transporter 1 (478-713) and ABC transporter 2 (1290-1533). The protein determines the fate of cholesterol derived from lipoprotein. The

functional loss of this protein impairs the integrity of lysosomes, disrupts the intracellular flow of free cholesterol, and causes cholesterol to accumulate intra-endo-lysosomally, adding to the organelles' malfunction (De Stefano *et al.,* 2014; Raza *et al.,* 2022). The substitution p. Arg857Cys, replaces a positively charged, less hydrophobic bigger amino acid (Arginine) to a neutral, more hydrophobic, and smaller amino acid (Cysteine). Change in charge, size, and nature of amino acids (Arginine and Cysteine) at position 857 may affect interactions of the protein with other molecules and residues leading to improper protein folding and intra-lysosomal accumulation of cholesterol.

In pedigree G, a hemizygous nonsense variant c.353T>A; p. Leu118Ter was identified in *IDS* gene which is involved in the X-linked inheritance of Hunter syndrome. *IDS* gene encodes a protein iduronate 2-sulfatase that catalyzes the degradation of heparan sulfate and dermatan sulfate (Gomes *et al.,* 2020).

In pedigree H, a homozygous missense variant (c.264G>T; p. Asp88Glu) was identified in *ALDOB* gene. *ALDOB* encodes fructose-1,6-bisphosphate aldolase, a tetrameric glycolytic enzyme that catalyzes the reversible conversion of fructose-1,6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. Pathogenic variants in the gene have been associated with fructose intolerance. The specific variant (c.264G>T) identified in pedigree H was previously reported by Retterer *et al.,* (Retterer *et al.,* 2016) in a family segregating abnormality of central nervous system. Interestingly, the affected individual in our pedigree H, showed delay in achieving milestones, abdominal distention, coarse facial features, aggressive behavior, short stature, respiratory/ear infections, hip dysplasia, joint stiffness, hepatosplenomegaly, difficulty in walking, craniofacial disproportion, J-shaped sella, L1 vertebral body is hypoplastic and slightly posteriorly displaced, exaggerated lumbar lordosis. In addition to these phenotypes, the patient showed hypoglycemia, which is an indication of fructosuria. The amino acid aspartic acid (D) at position 88 make salt bridge with R57, L92 and L321, the mutant residue glutamic acid (E) may not be able to interact. Wild type residue D is in its preferred secondary structure, a turn, while mutant residue E prefers to be in another secondary structure, which may therefore destabilize the protein conformation. Due to mutated fructose-1,6-

bisphosphate aldolase, accumulation of fructose-1,6-bisphosphate in different organs of the body may lead to multi-organ abnormality.

Pedigree J revealed two homozygous variants, one is a missense variant c.5941G>A; p. Val1981Ile in *HERC1* gene and other is splice site variant (c.454+3A>G) in *TRAPPC4*. HERC1 protein is involved in the membrane trafficking via guanine nucleotide exchange factors (GEF) and previously reported as a novel candidate gene for causing intellectual disability (Ortega-Recalde *et al.,* 2015). Recently, Van Bergen and colleagues (Van Bergen *et al.,* 2020) reported the same splice variant in *TRAPPC4* underlying early-onset seizures, developmental delay, microcephaly, sensorineural deafness, spastic quadriparesis and progressive cortical and cerebellar atrophy in families of Caucasian, Turkish and French-Canadian ethnicities. Other authors also identified the same splice site variant in a family of Indian origin having progressive encephalopathy and muscle involvement (Kaur *et al.,* 2020). The primary phenotypes of our patient in pedigree J include delay in achieving milestones, aggressive behavior with multiple crying episodes, respiratory/ear infection, unable to walk, and intellectual disability. Interestingly, phenotypes like muscles weakness, deafness, microcephaly, and seizures were not found in our case. The phenotypic variability might be due to different familial backgrounds of the families and effect of rare SNPs in modifier genes. The variant is predicted to cause frameshift of the protein due to mis-splicing of mRNA.

In affected individual of pedigree L, a novel homozygous missense variant c.439C>T; p. Arg147Trp in *RRAS2* gene and a novel nonsense variant (c.1810C>T; Gln604Ter) was found in *SMARCAL1*. The substitution in gene *RRAS2* of a positively charged arginine at position 147 with a neutral residue tryptophan may lead to the loss of ionic interactions with surrounding residues. It also leads to distortion of shape as mutant residue tryptophan is bigger in size than wild type residue arginine. RRAS2 protein is involved in the regulation of MAPK signaling pathway (Capri *et al.,* 2019). The involvement of *RRAS2* gene in some disease is still to be explored and linked with disease etiology in human genome mutation database. The encoded protein SMARCAL1 is a member of the SWI/SNF family of proteins. Members of this family have helicase and ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin structure around those genes. Pathogenic variants in

the gene have been reported in nephrotic syndrome and a metabolic syndrome Schimke immune-osseous dysplasia characterized by short stature, kidney disease, and a weakened immune system (Jin *et al.,* 2019). The patient in the present study carrying a nonsense variant in the gene showed a severe phenotype of delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability, and dysostosis multiplex (Table 3.7). Neurological phenotypes including intellectual disability, aggressive behavior, joint contractures, and hepatosplenomegaly showed by the present case were not described in previous cases. The variation in the phenotypes of previous and the present cases might be due to the nature and/or position of variant, different familial background, and ages of the affected individuals. The identified nonsense variant (p. Gln604Ter) in *SMARCAL1* is predicted to cause loss of function of the protein either through nonsense-mediated mRNA decay or production of truncated protein.

4.2.2. Mucopolysaccharidosis type II

Regarding families diagnosed as MPS II (pedigree A and B) we found pathogenic variants (c.1264C>G; p. Cys422Ser in pedigree 13 and c.1035C>T; p. Trp345Ter in pedigree B) in *IDS* segregating with the phenotypes. The novel missense variant p. Cys422Ser affects a non-conserved residue among human sulfatases. Cysteine is more hydrophobic in nature than serine, which causes loss of hydrophobic interactions. Cysteine at position 422 is involved in the sulphide bridge formation which will be lost after this change with serine thus it will affect the stability of protein structure and ultimately loss of function. The other nonsense variant (p. Trp345Ter) in *IDS* was previously identified in Japanese patient affected with sever phenotypes of MPS II (Sukegawa *et al.,* 1995). The authors found a smaller precursor protein in cells transfected with p. Trp345Ter. Therefore, it is predicted that the variant leads to the production of truncated enzyme. The affected individual in pedigree B showed mild phenotypes of short stature, coarse facial features and hernia as compared to those reported by Sukegawa *et al.,* (Sukegawa *et al.,* 1995).

4.2.3. Mucopolysaccharidosis type III

Sanfilippo syndrome (MPS III) affected six families were recruited from different regions of Punjab, Khyber Pakhtunkhwa, Balochistan and Azad Jammu and Kashmir. Initial recruitment and diagnosis were based mainly on their clinical profiles and phenotypes (Table 3.9). Whole genome sequencing revealed six different variants in each family. All found variants are missense. In pedigree A, the variant is present in *NAGLU* gene confirming the involvement of MPS III type B, while in pedigree B, D, E and F, the variations are present in *SGSH* gene, leading to the MPS III type A. In pedigree C the variant is identified in *VWA3B* gene causing cerebral ataxia.

In pedigree A, the missense variant is c. 2020C>T (p. Arg674Cys) (rs763299645) in exon 6 of *NAGLU* gene, reported firstly in 1998 (Zhao *et al.,* 1998). The patient is homozygous for the variation while parents are heterozygous with carrier status. Another missense variant was also reported in 1996 at the same codon causing variation c.2021G>A (p. Arg674His) (Zhao *et al.,* 1996). This locus is the mutational host spot for substitution of Arginine. Among all point mutations, 36% are on the locus R674, which also has CpG dinucleotide involvement (Zhao *et al.,* 1996). Cysteine is smaller in size than wild type residue, so it causes loss of interactions, and it also causes loss of ionic interactions and hydrogen bonding due to absence of charge and its hydrophobic nature. The genotype of the patient explained the severity of the phenotype (Table 3.9). Patient was born to a first-degree cousin once removed couple (Figure 3.14) with successive cousin marriages in the family. The *In-Silico* analysis of the variant is described in the table 3.10.

In pedigree B, all three patients are homozygous for the variation c. 221C>T (p. Arg74His) (rs778336949), and both parents are heterozygous for the variant. This pathogenic variant is in exon 2 in *SGSH* gene and reported firstly in 1997 (Bunge *et al.,* 1997) and the pathogenic effect of the variant is mentioned in the table 3.10. Same codon is altered for another substitution where Arginine is replaced by Cysteine (c. 222G>A, p. Arg74Cys) (Bunge *et al.,* 1997). This locus is likely to be the mutational hot spot for the disease. This residue is conserved among all mammalian sulfatases and stabilizes the active site of the enzyme (Bunge *et al.,* 1997). This substitution might cause loss of ionic interactions formed by positively charged arginine, also

difference in sizes causes loss of interactions with other residues and domains. There are four affected individuals, 1 female and 3 males (1 deceased male at the age of 14 years). One female patient in another loop in the same family, but she was also deceased at the age of ~12 years. Two healthy alive siblings of the patients, 1 female aged 21 years and male aged 15 years both are having the carrier status. Genetic counselling was done to the family about carriers and their future marriages. Parents showing first cousin once removed status (Figure 3.14).

In pedigree C, the identified novel variant (c.2951A>C; p. Asp984Ala, exon 22) is in a gene *VWA3B* which is previously reported for causing cerebral ataxia. In both patients the variant is in homozygous form while in parents it is heterozygous. This variation causes the loss of interactions and loss of hydrogen bonds thus effect the correct folding of protein. There are only 4 variants reported for causing cerebral ataxia in gene *VWA3B* as per HGMD (Accessed on 20-May-2022). The presenting phenotypes at the time of enrollment in elder sibling is quite resembling to MPS III phenotypes including weak muscles, intellectual disability, aggressiveness, and delayed growth, but in younger sibling the phenotypes were not so appealing due to the early age years. The phenotypes for MPS III are progressive and worsen with growing age. The patients were born to a first-degree cousin couple with successive cousin marriages in the generations (Figure 3.14).

In pedigree D, the missense variation c. 364C>T (p. Gly122Arg) in exon 4 in the gene *SGSH* causing the phenotype in both patients. Both affected individuals are homozygous while parents are heterozygous in state. This variant was firstly reported in 1997 in Netherland and Arabs (Bunge *et al.,* 1997). Wild type residue is smaller in size so cannot fit in the core. Insertion of a charged residue arginine will cause distortion in the structure and might cause disruption of correct folding of protein. Clinical details and pathogenicity of the variant is described in table 3.9 and table 3.10, respectively. There are two affected individuals, one sister and one brother born to a non-consanguineous couple, without any previous family history of the disease but all the successive marriages in previous generations are intra-cast marriages (Figure 3.14).

In pedigree E, the variant found in homozygous state in the patients is c. 258A>C (p. Asn86Lys) in exon 3 of *SGSH* gene. Both parents are heterozygous for the variation. This variant was not previously reported from any population to cause Sanfilippo syndrome. This family has two affected individuals, both are males and born to a first cousin once removed couple (Figure 3.14). One patient died after sample collection in a month at the age of 15 years. The detailed *In-Silico* analysis showing the pathogenicity of this variant is mentioned in table 3.10. The wild-type residue forms a hydrogen bond with methionine at position 88, the difference in size of wild type and mutated residues causes the loss of that interaction. The variation is in a domain that is important for the activity of the protein and interact with other domain which might affect the function of protein.

In pedigree F, the identified pathogenic variant is c. 817G>A (p. Asp273Asn) in exon 7 of *SGSH* gene, found in both affected patients in homozygous state and in parents in heterozygous state. This variant was reported firstly in 2000 (Beesley *et al.,* 2000). Both affected were females and born to a first-degree cousin couple (Figure 3.14). This variant is also located in the CpG dinucleotide site which is the mutational host spot of the gene (Beesley *et al.,* 2000). *In-Silico* studies are mentioned in the table 3.10. The wild-type residue is negative, so it is involved in interaction with metal ion "Ca++" which is lost with this substitution, thus the stability is reduced and disturb the domain. Wild type residue also forms hydrogen bonds with Asp31 and Asp32 and salt bridge with Arg74, Leu123 and Arg282, which might be lost due to substitution of Asp273 with a neutral Asn residue, disturbing the ionic interactions.

4.2.4. Mucopolysaccharidosis type IV

During this study eight families were recruited on the basis of their presenting phenotypes as Morquio syndromes (MPS IV) from Punjab, Azad Jammu and Kashmir and Khyber Pakhtunkhwa. Among suspected eight families, six families had genotype confirmation through Whole Genome Sequencing for Morquio syndrome type A. There are four identified variants including two splicing variants and two missense variants segregating along with the phenotype.

In pedigree A, single affected individual was born to a first-degree cousin couple (Figure 3.15). The details related to clinicals, and phenotypes are mentioned

in table 3.11. The identified variant in *GALNS* gene is a splice acceptor site variant g.38845A>G (chr16:88884534T>C); c.1365-2A>G (Intron 12) in homozygous mutant form in the patient and heterozygous in both parents. The identified variant was previously not reported for causing Morquio syndrome anywhere in the world.

In pedigree B, G and H the identified missense variant in *GALNS* gene is c.107A>C; p. Leu36Arg (rs755832705) in exon 1 was previously reported by Morrone *et al.,* 2014. This variant was described in Asian-multiethnic populations to cause Morquio syndrome type A. At same loci another amino acid substitution c. 107 T>C, p. Leu36Pro was reported by Tomatsu *et al.,* 2005. Parents were first degree cousin married (Figure 3.15). The variant is homozygous mutant in all patients while heterozygous with carrier status in parents. The clinical analysis and phenotypes are explained in table 3.11. The variation is in the catalytic domain which can disturb the function of protein. It can also hinder the correct protein folding due to charged residue i.e., Arginine. It can cause loss of hydrophobic interactions. Pedigree H also has another homozygous missense variant c.1336G>C; p. Asp446His (Exon 4), in *BTD* gene causing Biotinidase deficiency (EC 3.5.1.12). This enzyme is involved in the recycling of biotin bounded to protein by releasing biotin and lysine (Pispa, 1965; Wolf *et al.,* 1985). The wild-type residue forms hydrogen bonds with Gly423, Tyr456 and Gln458, and salt bridge with Arg544, but as a result of substitution these hydrogen bonding and ionic interaction will be lost. Wild type residue is buried inside the core while mutant residue is bigger in size and probably will not fit. Detailed *In-Silico* analysis is mentioned in table 3.12.

In pedigree D, there were three affected individuals, two females and a male born to a first-degree cousin parent (Figure 3.15). The identified variant is c.1259G>C; p. Pro420Arg (Exon 12) previously reported by Morrone *et al.,* 2014 in *GALNS* gene. All patients were homozygous recessive mutant while parents were heterozygous for the variant. Proline is rigid so provide a backbone for the structural conformation, that might be lost after this substitution. It also affects the function of protein. Arginine is positively charged so it causes the repulsion between neighboring residues. It also causes loss of hydrophobic interactions.

In pedigree E, the identified missense variant c.179G>A; p. Gly60Glu, is in the gene *GNPTG* which is responsible for causing Mucolipidoses type III. This gene encodes gamma subunit of enzyme N-acetylglucosamine-1-phosphotransferase (EC 2.7.8.17) (Braulke *et al.,* 2008; Qian *et al.,* 2010). The clinical profiles for ML-III also include phenotypes like short stature, scoliosis and joint contractures (Tiede *et al.,* 2005; Bargal *et al.,* 2006), which was also present in our patients (Table 3.11). Both patients are homozygous while parents were heterozygous for the variant. This residue is highly conserved, so its substitution is probably damaging for the protein. The mutant residue is negatively charged so its incorporation will lead to repulsion of ligands and other residues of same charge. The mutant residue is big so it might lead to bumps. Glycine is able to form torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and it disturb the local structure. The detailed *In-Silico* analysis is given in table 3.12.

In pedigree F, there were two affected individuals one is alive, and one was deceased at the age of 14 years. Another splice acceptor site variant c.423-1G>A in homozygous mutant state is present in the patient, while segregating in heterozygous state in both parents. Parents were first degree cousins (Figure 3.15). This variant is also previously not reported as the causative for Morquio syndrome.

4.2.5. Niemann Pick disease

In pedigree A, the identified variations in gene *NPC1* are (p. Phe537Ser) and (p. Thr1098Ile). The family belongs to Azad Jamu and Kashmir valley at the border of India and Pakistan with first-degree cousin marriage among parents (Figure 3.16). For both variants, parents showed heterozygosity while patient is homozygous mutant. This is the case of multiple mutations. Both variants were previously not linked with pathogenicity of NP disease as per HGMD. The protein structural analysis showed that mutant amino acid Isoleucine is bigger in size and is more hydrophobic than wild type residue threonine at position 1098. This difference in size and hydrophobicity may cause disturbance of hydrogen bonding with isoleucine at position 1094. This loss of hydrogen bonding may affect the correct folding of protein. The region for this substitution is in transmembrane domain which further affects the interactions with lipid membrane. In second variation i.e., (p. His537Arg), the wild

type of residue histidine at position 537 is smaller in size than mutant arginine residue. Wild type residue (Histidine) is more hydrophobic than mutant (Arginine). This variation is in the stretch of repeated amino acids called Lumenal (HOPE analysis). This substitution will lead to an empty space and loss of hydrophobic interactions in the core of protein. Most phenotypes of the patients are quite resembling with NP disease A/B, however some additional findings like reduced WBCs, reduced platelets, elevated ESR value and elevated ALP values can be observed in these patients (Table 3.13).

In pedigree B and C, the pathogenic variation is (p. His423Arg) in *SMPD1* which is involved in causing NP disease type A/B. Both families were unrelated but have same ethnicity i.e Pathan, the caste with very high rate of consanguinity in Pakistan due to their social norms (Figure 3.16). The variant has been previously reported in literature for causing NP disease type A in European and Ashkenazi Jewish population (McGovern *et al.,* 2006). The phenotypes presented in our patients are quite similar and overlapping with the ones mentioned by McGovern *et al.,* 2006 like cherry red spots, weak body muscles, hepatosplenomegaly, respiratory issues and low levels of hemoglobin (Table 3.13). According to bioinformatics structural analysis of amino acid, the mutant amino acid (Arginine) is bigger in size than wild type (Histidine). Mutant amino acid is charged positively while wild type residue is neutral that forms hydrogen bonds with Tyrosine 276 and Histidine 511 and because of this substitution these bonding were lost causing distortion in the structure of enzyme. The amino acid His423 is partially conserved. The wild type residue was buried in the core of protein whereas the mutant residue being bigger in size may leads to disturbance in protein folding.

4.2.6. Pycnodysostosis

In this research, a family from Khyber Pakhtunkhwa, a province of Pakistan, is presented with a patient suffering from severe bone pain. The clinical examination and diagnosis confirmed it to be a type of bone related genetic disorder called pycnodysostosis. Biochemical and radiological findings go along with the phenotype. The height and weight of the patient was less than third centile. Whole genome sequencing confirms it as an autosomal recessive bone related disease. The variant p.

Arg46Trp of *CTSK* gene was found in the homozygous state in patient while both parents were heterozygous for the same variant. This variant was firstly reported by Shilling *et al.,* 2007 in a Kashmiri origin family. Later, this same variant was reported by Kauser *et al.,* 2017 again in a Kashmiri based family with a proposed founder effect. But the family presented in this study has origin from Pakistan explaining a possibility of common ancestor of these different families.

In-Silico studies have shown the effect of this variation on structure and function of protein is seriously damaging. Difference in size of mutant residue effects the positioning of bonding with other amino acids and domains. Arginine interacts with glutamic acid at position 85 through hydrogen bonding. It also makes salt bridge with aspartic acid 42, aspartic acid 80, glutamic acid 84 and glutamic acid 85. Being neutral in charge tryptophan is unable to form ionic bonding thus affecting the structure and function of cathepsin K protein. Wild type residue is also really conserved in different orthologues thus making this variation probably damaging (Figure 3.18A). As per conservation pattern and conservation time, more the conservation time is, more are the chances for a variant to be pathogenic. Also, according to Richard *et al.,* 2015, when a variant is confirmed in heterozygous state in parents without any family history and also is predicted as pathogenic/deleterious by multiple *In-Silico* tools, it is pathogenic variant (Richard *et al.,* 2015).

4.2.7. Griscelli syndrome

In pedigree A, three affected individuals born to a first-degree cousin couple (Figure 3.19) two alive and one deceased. The details related to clinicals, and phenotypes are mentioned in table 3.15. The identified variant in *RAB27A* gene is a duplication variant *RAB27A* (NM_183235.3): c.258_260dup; (p. Ala87dup) in homozygous mutant form in both alive patients and heterozygous in mother. The identified variant was previously not reported for causing Griscelli syndrome anywhere in the world. *RAB27A* gene is responsible for causing GS type 2 with phenotypes like recurrent infections due to weak immunity which is also reported in our patients. Unlike reports of no neurological abnormalities based on literature, both our patients have shown neurological manifestations like delay in achieving milestones (no proper speech, need support to walk) (Table 3.15).

4.3. Conclusion and Future Perspectives

Out of forty-five families, forty-one families are solved during this study based on molecular genetics results. A total of 23 novel disease-causing mutations were found in 20 families while 14 reported mutations were found in 30 families. There are four unsolved families due to poor quality of the samples and re-sampling in not possible because of the death of the patients. Furthermore, in this study *ABCA5* gene which was previously reported to cause LSD like phenotype in mouse model, was identified for the first time in human which might play a significant role in understanding of complex pathogenesis of the disease (Kubo *et al.,* 2005). Functional studies for this gene are under progress and will be available soon. Findings of these novel/recurrent mutations may be targeted for future therapeutic research and could be used as key diagnostics biomarkers for LSDs. The enrolled families were highly inbred, to reduce incidence of disease in the future generations, genetic counselling was provided to each family. The findings of this study suggest a public awareness campaign about recessively inherited disorders in our population to lessen the burden of these diseases and their associated morbidities and mortalities in the future. The incidence of the LSDs in our population is high, current study necessitates future molecular genetic analysis of inborn errors of metabolism using cases belonging to our population. Findings of current study also signify the importance of Whole Genome Sequencing for unveiling genetic basis of inherited disorders especially complex genetic diseases. Based on current research the prevalent variants like p.Leu490Pro in *IDUA* gene can be used as routine marker for diagnosing MPS-I in affected patients. It will help in reducing the cost of genetic testing through WES/WGS. This study also emphasize the importance of molecular diagnosis for LSD patients. The ethical dilemma of the lack of treatment available after even diagnosing LSDs is complicated molecular mechanism involved in causing these disorders and most of them are still unknown as mentioned in introduction section.

5. REFERENCES

Abdelwahab, M., Blankenship, D., & Schiffmann, R. (2016). Long-term follow-up and sudden unexpected death in Gaucher disease type 3 in Egypt. Neurology Genetics, 2(2).

Abghari, F. Z., Bayat, F., Razipour, M., Karimipoor, M., Taghavi-Basmenj, M., Zeinali, S., & Davoudi-Dehaghani, E. (2019). Characterization of Niemann-Pick diseases genes mutation spectrum in Iran and identification of a novel mutation in SMPD1 gene. Medical Journal of the Islamic Republic of Iran, 33, 126.

Abumansour, I. S., Yuskiv, N., Paschke, E., & Stockler-Ipsiroglu, S. (2020). Morquio-B disease: Clinical and genetic characteristics of a distinct GLB1-related dysostosis multiplex. JIMD reports, 51(1), 30-44.

Pruneri, G., De Braud, F., Sapino, A., Aglietta, M., Vecchione, A., Giusti, R., ... & Jommi, C. (2021). Next-generation sequencing in clinical practice: is it a cost-saving alternative to a single-gene testing approach?. PharmacoEconomics-Open, 5, 285-298.

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. Human Immunology, 82(11), 801-811.

Schuler, B. A., Nelson, E. T., Koziura, M., Cogan, J. D., Hamid, R., & Phillips, J. A. (2022). Lessons learned: next-generation sequencing applied to undiagnosed genetic diseases. The Journal of Clinical Investigation, 132(7).

Torshizi, A. D., & Wang, K. (2018). Next-generation sequencing in drug development: target identification and genetically stratified clinical trials. Drug Discovery Today, 23(10), 1776-1783.

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., ... & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. Nature methods, 7(4), 248-249.

Afroze, B., & Brown, N. (2017). Ethical issues in managing Lysosomal storage disorders in children in low and middle income countries. Pakistan Journal of Medical Sciences, 33(4), 1036.

Akyol, M. U., Alden, T. D., Amartino, H., Ashworth, J., Belani, K., Berger, K. I., ... & Scarpa, M. (2019). Recommendations for the management of MPS IVA: systematic evidence-and consensus-based guidance. Orphanet journal of rare diseases, 14(1), 1-25.

Aldenhoven, M., Wynn, R. F., Orchard, P. J., O'Meara, A., Veys, P., Fischer, A., ... & Boelens, J. J. (2015). Long-term outcome of Hurler syndrome patients after hematopoietic cell transplantation: an international multicenter study. Blood, The Journal of the American Society of Hematology, 125(13), 2164-2172.

Alegra, T., Koppe, T., Acosta, A., Sarno, M., Burin, M., Kessler, R. G., … Schwartz, I. V. (2014). Pitfalls in the prenatal diagnosis of mucolipidosis II alpha/beta: A case report. Meta Gene, 1(2), 403–406.

Alfonso, P., Aznarez, S., Giralt, M., Pocovi, M., & Giraldo, P. (2007). Mutation analysis and genotype/phenotype relationships of Gaucher disease patients in Spain. Journal of human genetics, 52(5), 391-396.

Al-Jasmi, F. A., Tawfig, N., Berniah, A., Ali, B. R., Taleb, M., Hertecant, J. L., ... & Souid, A. K. (2012). Prevalence and novel mutations of lysosomal storage disorders in United Arab Emirates. In JIMD Reports-Volume 10 (pp. 1-9). Springer, Berlin, Heidelberg.

Andrade, F., Aldámiz-Echevarría, L., Llarena, M., & Couce, M. L. (2015). Sanfilippo syndrome: Overall review. Pediatrics International, 57(3), 331-338.

Arman, A., Bereket, A., Coker, A., Kiper, P. Ö. Ş., Güran, T., Özkan, B., ... & Turan, S. (2014). Cathepsin K analysis in a pycnodysostosis cohort: demographic, genotypic and phenotypic features. Orphanet journal of rare diseases, 9(1), 1-8.

Arora, R. S., Mercer, J., Thornley, M., Tylee, K., & Wraith, J. E. (2007). Enzyme replacement therapy in 12 patients with MPS I–H/S with homozygous p. Leu490Pro mutation. Journal of Inherited Metabolic Disease: Official Journal of the Society for the Study of Inborn Errors of Metabolism, 30(5), 821-821.

Baehner, F., Schmiedeskamp, C., Krummenauer, F., Miebach, E., Bajbouj, M., Whybra, C., ... & Beck, M. (2005). Cumulative incidence rates of the mucopolysaccharidoses in Germany. Journal of Inherited Metabolic Disease:

Official Journal of the Society for the Study of Inborn Errors of Metabolism, 28(6), 1011-1017.

Baglam, T., Binnetoglu, A., Topuz, M. F., Ikizoglu, N. B., Ersu, R., Turan, S., & Sarı, M. (2017). Pycnodysostosis at otorhinolaryngology. International journal of pediatric otorhinolaryngology, 95, 91-96.

Baguez-Cesar, A. B. (1943). Neutoropenia cronica maligna familiar con granulaciones atipicas de los leucocitos. Bol. Soc. Cubbana Pediatr., 15, 900-922.

Bahadoran, P., Aberdam, E., Mantoux, F., Buscà, R., Bille, K., Yalman, N., ... & Ballotti, R. (2001). Rab27a: a key to melanosome transport in human melanocytes. The Journal of Cell Biology, 152(4), 843-850.

Bajaj, L., Lotfi, P., Pal, R., Ronza, A. D., Sharma, J., & Sardiello, M. (2019). Lysosome biogenesis in health and disease. Journal of neurochemistry, 148(5), 573-589.

Bame, K. J., & Rome, L. H. (1986) a. Acetyl-coenzyme A: alpha-glucosaminide N-acetyltransferase. Evidence for an active site histidine residue. Journal of Biological Chemistry, 261(22), 10127-10132.

Bame, K. J., & Rome, L. H. (1986) b. Genetic evidence for transmembrane acetylation by lysosomes. Science, 233(4768), 1087-1089.

Barbosa, M. D., Nguyen, Q. A., Tchernev, V. T., Ashley, J. A., Detter, J. C., Blaydes, S. M., ... & Kingsmore, S. F. (1996). Identification of the homologous beige and Chediak–Higashi syndrome genes. Nature, 382(6588), 262-265.

Bargal, R., Zeigler, M., Abu-Libdeh, B., Zuri, V., Mandel, H., Neriah, Z. B., ... & Raas-Rothschild, A. (2006). When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients. Molecular genetics and metabolism, 88(4), 359-363.

Barton, N. W., Brady, R. O., Dambrosia, J. M., Di Bisceglie, A. M., Doppelt, S. H., Hill, S. C., ... & Yu, K. T. (1991). Replacement therapy for inherited enzyme deficiency—macrophage-targeted glucocerebrosidase for Gaucher's disease. New England Journal of Medicine, 324(21), 1464-1470.

Bava, K. A., Gromiha, M. M., Uedaira, H., Kitajima, K., & Sarai, A. (2004). ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic acids research, 32(suppl_1), D120-D121.

BdO, S., Quaio, C. R., Spolador, G. M., de Paula, A. C., Curiati, M. A., Martins, A. M., ... & Kim, C. A. (2021). Impact of ERT and Follow up of 17 Patients from the Same Family with Mild form of MPS II.

Beck, M., Arn, P., Giugliani, R., Muenzer, J., Okuyama, T., Taylor, J. and Fallet, S (2014). The natural history of MPS I: global perspectives from the MPS I Registry. Genet Med 16(10):759-765.

Beesley, C. E., Burke, D., Jackson, M., Vellodi, A., Winchester, B. G., & Young, E. P. (2003). Sanfilippo syndrome type D: identification of the first mutation in the N-acetylglucosamine-6-sulphatase gene. Journal of medical genetics, 40(3), 192-194.

Beesley, C. E., Meaney, C. A., Greenland, G., Adams, V., Vellodi, A., Young, E. P., & Winchester, B. G. (2001). Mutational analysis of 85 mucopolysaccharidosis type I families: frequency of known mutations, identification of 17 novel mutations and in vitro expression of missense mutations. Human genetics, 109(5), 503-511.

Beesley, C. E., Young, E. P., Vellodi, A., & Winchester, B. G. (2000). Mutational analysis of Sanfilippo syndrome type A (MPS IIIA): identification of 13 novel mutations. Journal of Medical Genetics, 37(9), 704-707.

Benetó, N., Vilageliu, L., Grinberg, D., & Canals, I. (2020). Sanfilippo syndrome: Molecular basis, disease models and therapeutic approaches. International Journal of Molecular Sciences, 21(21), 7819.

Bernhardt, R. R., & Schachner, M. (2000). Chondroitin sulfates affect the formation of the segmental motor nerves in zebrafish embryos. Developmental biology, 221(1), 206-219.

Bertola, D., Amaral, C., Kim, C., Albano, L., Aguena, M., & Passos-Bueno, M. R. (2010). Craniosynostosis in pycnodysostosis: broadening the spectrum of the

cranial flat bone abnormalities. American Journal of Medical Genetics Part A, 152(10), 2599-2603.

Bertola, F., Filocamo, M., Casati, G., Mort, M., Rosano, C., Tylki-Szymanska, A., ... & Parini, R. (2011). IDUA mutational profiling of a cohort of 102 European patients with mucopolysaccharidosis type I: identification and characterization of 35 novel α-L-iduronidase (IDUA) alleles. Human mutation, 32(6), E2189-E2210.

Besley, G. T. N., & Wraith, J. E. (1997). Lysosomal disorders. Current Paediatrics, 7(2), 128-134.

Bie, H., Yin, J., He, X., Kermode, A. R., Goddard-Borger, E. D., Withers, S. G., & James, M. N. (2013). Insights into mucopolysaccharidosis I from the structure and action of α-L-iduronidase. Nature chemical biology, 9(11), 739-745.

Bink, R. J., Habuchi, H., Lele, Z., Dolk, E., Joore, J., Rauch, G. J., ... & Zivkovic, D. (2003). Heparan sulfate 6-O-sulfotransferase is essential for muscle development in zebrafish. Journal of Biological Chemistry, 278(33), 31118-31127.

Bizaoui, V., Michot, C., Baujat, G., Amouroux, C., Baron, S., Capri, Y., ... & Cormier-Daire, V. (2019). Pycnodysostosis: Natural history and management guidelines from 27 French cases and a literature review. Clinical genetics, 96(4), 309-316.

Bondeson, M. L., Dahl, N., Malmgren, H., Kleijer, W. J., Tönnesen, T., Carlberg, B. M., & Pettersson, U. (1995a). Inversion of the IDS gene resulting from recombination with IDS-related sequences in a common cause of the Hunter syndrome. Human molecular genetics, 4(4), 615-621.

Bondeson, M. L., Malmgren, H., Dahl, N., Carlberg, B. M., & Pettersson, U. (1995b). Presence of an IDS-related locus (IDS2) in Xq28 complicates the mutational analysis of Hunter syndrome. European journal of human genetics, 3(4), 219-227.

Braulke, T., Pohl, S., & Storch, S. (2008). Molecular analysis of the GlcNac-1-phosphotransferase. Journal of inherited metabolic disease, 31(2), 253-257.

Brömme, D., & Okamoto, K. (1995). Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution.

Brunetti-Pierri, N., & Scaglia, F. (2008). GM1 gangliosidosis: review of clinical, molecular, and therapeutic aspects. Molecular genetics and metabolism, 94(4), 391-396.

Bultron, G., Kacena, K., Pearson, D., Boxer, M., Yang, R., Sathe, S., ... & Mistry, P. K. (2010). The risk of Parkinson's disease in type 1 Gaucher disease. Journal of inherited metabolic disease, 33(2), 167-173.

Bunge, S., Ince, H., Steglich, C., Kleijer, W. J., Beck, M., Zaremba, J., ... & Gal, A. (1997). Identification of 16 sulfamidase gene mutations including the common R74C in patients with mucopolysaccharidosis type IIIA (Sanfilippo A). Human mutation, 10(6), 479-485.

Burton, B. K., Kronn, D. F., Hwu, W. L., Kishnani, P. S., & Pompe Disease Newborn Screening Working Group. (2017). The initial evaluation of patients after positive newborn screening: recommended algorithms leading to a confirmed diagnosis of Pompe disease. Pediatrics, 140(Supplement_1), S14-S23.

Çağdaş, D., Özgür, T. T., Asal, G. T., Tezcan, İ., Metin, A., Lambert, N., ... & Sanal, Ö. (2012). Griscelli syndrome types 1 and 3: analysis of four new cases and long-term evaluation of previously diagnosed patients. European journal of pediatrics, 171(10), 1527-1531.

Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L., & Casadio, R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. Human mutation, 30(8), 1237-1244.

Capri, Y., Flex, E., Krumbach, O. H., Carpentieri, G., Cecchetti, S., Lißewski, C., ... & Zenker, M. (2019). Activating mutations of RRAS2 are a rare cause of Noonan syndrome. The American Journal of Human Genetics, 104(6), 1223-1232.

Capriotti, E., Calabrese, R., & Casadio, R. (2006). Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics, 22(22), 2729-2734.

Castaño-Jaramillo, L. M., Lugo-Reyes, S. O., Cruz Munoz, M. E., Scheffler-Mendoza, S. C., Duran McKinster, C., Yamazaki-Nakashimada, M. A., ... & Saez-de-Ocariz Gutierrez, M. D. M. (2021). Diagnostic and therapeutic caveats in Griscelli syndrome. Scandinavian Journal of Immunology, 93(6), e13034.

Cathey, S. S., Leroy, J. G., Wood, T., Eaves, K., Simensen, R. J., Kudo, M., ... & Friez, M. J. (2010). Phenotype and genotype in mucolipidoses II and III alpha/beta: a study of 61 probands. Journal of medical genetics, 47(1), 38-48.

Celik, B., Tomatsu, S. C., Tomatsu, S., & Khan, S. A. (2021). Epidemiology of mucopolysaccharidoses update. Diagnostics, 11(2), 273.

Chediak, M. (1952). Nouvelle anomalie leucocytaire de caractère constitutionnel et familial. Rev. Hematol, 7, 362-367.

Cheema, H. A., Malik, H. S., Parkash, A., & Fayyaz, Z. (2016). Spectrum of inherited metabolic disorders in Pakistani children presenting at a tertiary care centre. J Coll Physicians Surg Pak, 26(6), 498-502.

Cheema, H. A., Rasool, I. G., Anjum, M. N., & Zahoor, M. Y. (2020). Mutational spectrum of SMPD1 gene in Pakistani Niemann-Pick disease patients. Pakistan journal of medical sciences, 36(3), 479.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34(17), i884-i890.

Chen, Z., Pang, Y., Liu, X., Wang, X., Deng, Z., & Sun, X. (2001). The Metabolic and Molecular. In Bases of Inherited Disease; McGraw-Hill, Inc., Health Professions Division.

Chien, Y. H., Lee, N. C., Chen, P. W., Yeh, H. Y., Gelb, M. H., Chiu, P. C., ... & Hwu, W. L. (2020). Newborn screening for Morquio disease and other lysosomal storage diseases: results from the 8-plex assay for 70,000 newborns. Orphanet journal of rare diseases, 15(1), 1-7.

Chih-Kuang, C., Shuan-Pei, L., Shyue-Jye, L., & Tuen-Jen, W. (2002). MPS screening methods, the berry spot and acid turbidity tests, cause a high incidence of false-negative results in sanfilippo and morquio syndromes. Journal of clinical laboratory analysis, 16(5), 253-258.

Chkioua, L., Khedhiri, S., Turkia, H.B., Chahed, H., Ferchichi, S., Dridi, M.F.B., Laradi, S. and Miled, A (2011). Hurler disease (mucopolysaccharidosis type IH): clinical features and consanguinity in Tunisian population. Diagn. Pathol 6(1):1-7.

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics, 31(16), 2745-2747.

Choy, F. Y., Zhang, W., Shi, H. P., Zay, A., Campbell, T., Tang, N., & Ferreira, P. (2007). Gaucher disease among Chinese patients: review on genotype/phenotype correlation from 29 patients and identification of novel and rare alleles. Blood Cells, Molecules, and Diseases, 38(3), 287-293.

Chudley, A. E., & Chakravorty, C. (2002). Genetic landmarks through philately: Luís Morquio 1867-1935. Clinical genetics, 62(6), 438-439.

Cohen, J. L., Chakraborty, P., Fung-Kee-Fung, K., Schwab, M. E., Bali, D., Young, S. P., ... & MacKenzie, T. C. (2022). In Utero Enzyme-Replacement Therapy for Infantile-Onset Pompe's Disease. New England Journal of Medicine, 387(23), 2150-2158.

Coutinho, M. F., Prata, M. J., & Alves, S. (2012). Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. Molecular genetics and metabolism, 105(4), 542-550.

Cox, T., Lachmann, R., Hollak, C., Aerts, J., Van Weely, S., Hrebícek, M., ... & Zimran, A. (2000). Novel oral treatment of Gaucher's disease with N-butyldeoxynojirimycin (OGT 918) to decrease substrate biosynthesis. The Lancet, 355(9214), 1481-1485.

Crocker, A. C. (1961). The cerebral defect in Tay-Sachs disease and Niemann-Pick disease. Journal of neurochemistry, 7(1), 69-80.

Dahms, N. M., Lobel, P., & Kornfeld, S. (1989). Mannose 6-phosphate receptors and lysosomal enzyme targeting. Journal of Biological Chemistry, 264(21), 12115-12118.

Dai, R., Wu, Z., Chu, H. Y., Lu, J., Lyu, A., Liu, J., & Zhang, G. (2020). Cathepsin K: The action in and beyond bone. Frontiers in Cell and Developmental Biology, 8, 433.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience, 10(2), giab008.

Dardis, A., Zampieri, S., Gellera, C., Carrozzo, R., Cattarossi, S., Peruzzo, P., ... & Bembi, B. (2020). Molecular genetics of Niemann–Pick type C disease in Italy: an update on 105 patients and description of 18 NPC1 novel variants. Journal of clinical medicine, 9(3), 679.

Daykin, E. C., Ryan, E., & Sidransky, E. (2021). Diagnosing neuronopathic Gaucher disease: New considerations and challenges in assigning Gaucher phenotypes. Molecular Genetics and Metabolism, 132(2), 49-58.

De Duve, C. (2005). The lysosome turns fifty. Nature cell biology, 7(9), 847-849.

De Pasquale, V., & Pavone, L. M. (2019). Heparan sulfate proteoglycans: The sweet side of development turns sour in mucopolysaccharidoses. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1865(11), 165539.

De Ru, M.H., Teunissen, Q.G., Van Der Lee, J.H., Beck, M., Bodamer, O.A., Clarke, L.A., Hollak, C.E., Lin, S.P., Rojas, M.V.M., Pastores, G.M. and Raiman, J.A (2012). Capturing phenotypic heterogeneity in MPS I: results of an international consensus procedure. Orphanet J Rare Dis 7(1):1-9.

Del Grosso, A., Parlanti, G., Mezzena, R., & Cecchini, M. (2022). Current treatment options and novel nanotechnology-driven enzyme replacement strategies for lysosomal storage disorders. Advanced Drug Delivery Reviews, 114464.

Demydchuk, M., Hill, C. H., Zhou, A., Bunkóczi, G., Stein, P. E., Marchesan, D., ... & Read, R. J. (2017). Insights into Hunter syndrome from the structure of iduronate-2-sulfatase. Nature communications, 8(1), 1-9.

DeStefano, G. M., Kurban, M., Anyane-Yeboa, K., Dall'Armi, C., Di Paolo, G., Feenstra, H., ... & Christiano, A. M. (2014). Mutations in the cholesterol

transporter gene ABCA5 are associated with excessive hair overgrowth. PLoS genetics, 10(5), e1004333.

Di Lorenzo, G., Velho, R. V., Winter, D., Thelen, M., Ahmadi, S., Schweizer, M., ... & Pohl, S. (2018). Lysosomal proteome and secretome analysis identifies missorted enzymes and their nondegraded substrates in mucolipidosis III mouse cells. Molecular & Cellular Proteomics, 17(8), 1612-1626.

Dierks, T., Lecca, M. R., Schmidt, B., & von Figura, K. (1998). Conversion of cysteine to formylglycine in eukaryotic sulfatases occurs by a common mechanism in the endoplasmic reticulum. FEBS letters, 423(1), 61-65.

Douillard-Guilloux, G., Raben, N., Takikita, S., Batista, L., Caillaud, C., & Richard, E. (2008). Modulation of glycogen synthesis by RNA interference: towards a new therapeutic approach for glycogenosis type II. Human molecular genetics, 17(24), 3876-3886.

Douillard-Guilloux, G., Raben, N., Takikita, S., Ferry, A., Vignaud, A., Guillet-Deniau, I., ... & Richard, E. (2010). Restoration of muscle functionality by genetic suppression of glycogen synthesis in a murine model of Pompe disease. Human molecular genetics, 19(4), 684-696.

Durchfort, N., Verhoef, S., Vaughn, M. B., Shrestha, R., Adam, D., Kaplan, J., & Ward, D. M. (2012). The enlarged lysosomes in beigej cells result from decreased lysosome fission and not increased lysosome fusion. Traffic, 13(1), 108-119.

Durmaz, C. D., & Altıner, Ş. (2021). MASP1-related 3MC syndrome in a patient from Turkey. American Journal of Medical Genetics Part A, 185(7), 2267-2270.

Eapen, M., DeLaat, C. A., Baker, K. S., Cairo, M. S., Cowan, M. J., Kurtzberg, J., ... & Filipovich, A. H. (2007). Hematopoietic cell transplantation for Chediak–Higashi syndrome. Bone marrow transplantation, 39(7), 411-415.

El-Hattab, A. W. (2015). Inborn errors of metabolism. Clinics in perinatology, 42(2), 413-439.

El-Helou, S. M., Biegner, A. K., Bode, S., Ehl, S. R., Heeg, M., Maccari, M. E., ... & Grimbacher, B. (2019). The German national registry of primary immunodeficiencies (2012–2017). Frontiers in immunology, 10, 1272.

Eto, Y., & Ida, H. (1999). Clinical and molecular characteristics of Japanese Gaucher disease. Neurochemical research, 24(2), 207-211.

Everts, V., Aronson, D. C., & Beertsen, W. (1985). Phagocytosis of bone collagen by osteoclasts in two cases of pycnodysostosis. Calcified tissue international, 37(1), 25-31.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 32(19), 3047-3048.

Fan X, Zhang H, Zhang S, Bagshaw RD, Tropak MB, Callahan JW, Mahuran DJ (2006) Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). Am J Hum Genet 79: 738–744.

Fedele, A. O. (2015). Sanfilippo syndrome: causes, consequences, and treatments. The application of clinical genetics, 8, 269.

Feldhammer, M., Durand, S., Mrázová, L., Boucher, R. M., Laframboise, R., Steinfeld, R., ... & Pshezhetsky, A. V. (2009) b. Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (HGSNAT) gene. Human mutation, 30(6), 918-925.

Feldhammer, M., Durand, S., Mrazova, L., van Diggelen, O. P., Hrebicek, M., Kmoch, S., & Pshezhetsky, A. V. (2009) a. 54. Sanfilippo syndrome type C: Novel mutations in the HGSNAT gene. Molecular Genetics and Metabolism, 96(2), S24-S24.

Ferraz, M. J., Marques, A. R., Gaspar, P., Mirzaian, M., van Roomen, C., Ottenhoff, R., ... & Aerts, J. M. (2016). Lyso-glycosphingolipid abnormalities in different murine models of lysosomal storage disorders. Molecular genetics and metabolism, 117(2), 186-193.

Ferreira, C. R., van Karnebeek, C. D., Vockley, J., & Blau, N. (2019). A proposed nosology of inborn errors of metabolism. Genetics in Medicine, 21(1), 102-106.

Fields, T., Patterson, M., Brémovà-Ertl, T., Belcher, G., Billington, I., Churchill, G. C., ... & Strupp, M. (2021). A master protocol to investigate a novel therapy acetyl-

L-leucine for three ultra-rare neurodegenerative diseases: Niemann-Pick type C, the GM2 gangliosidoses, and ataxia telangiectasia. Trials, 22(1), 1-15.

Firasat, S., Kaul, H., Ashfaq, U.A. and Idrees, S (2018). In silico analysis of five missense mutations in CYP1B1 gene in Pakistani families affected with primary congenital glaucoma. Int Opthamol 38(2):807-814.

Fonteles, C. S. R., Chaves Jr, C. M., Da Silveira, A., Soares, E. C. S., Couto, J. L. P., & de Azevedo, M. D. F. V. (2007). Cephalometric characteristics and dentofacial abnormalities of pycnodysostosis: report of four cases from Brazil. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, 104(4), e83-e90.

Forgac, M. (2007). Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. Nature reviews Molecular cell biology, 8(11), 917-929.

Fuller, M., Meikle, P. J., & Hopwood, J. J. (2006). Epidemiology of lysosomal storage diseases: an overview. Fabry disease: perspectives from 5 years of FOS.

Garnero, P., Borel, O., Byrjalsen, I., Ferreras, M., Drake, F. H., McQueney, M. S., ... & Delaissé, J. M. (1998). The collagenolytic activity of cathepsin K is unique among mammalian proteinases. Journal of Biological Chemistry, 273(48), 32347-32352.

Gaucher, P. C. E. (1882). On primary epithelioma of the spleen: idiopathic hypertrophy of the spleen without leukemia. Paris Fac Med Paris.

Gelb, B. D., Edelson, J. G., & Desnick, R. J. (1995). Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping. Nature genetics, 10(2), 235-237.

Ghosh, A., Mercer, J., Mackinnon, S., Yue, W.W., Church, H., Beesley, C.E., Broomfield, A., Jones, S.A. and Tylee, K (2017). IDUA mutational profile and genotype–phenotype relationships in UK patients with Mucopolysaccharidosis Type I. Hum mutat 38(11):1555-1568.

Gil-Krzewska, A., Wood, S. M., Murakami, Y., Nguyen, V., Chiang, S. C., Cullinane, A. R., ... & Krzewski, K. (2016). Chediak-Higashi syndrome: Lysosomal trafficking regulator domains regulate exocytosis of lytic granules but not

cytokine secretion by natural killer cells. Journal of Allergy and Clinical Immunology, 137(4), 1165-1177.

Gironi, L. C., Zottarelli, F., Savoldi, G., Notarangelo, L. D., Basso, M. E., Ferrero, I., ... & Savoia, P. (2019). Congenital hypopigmentary disorders with multiorgan impairment: a case report and an overview on Gray Hair syndromes. Medicina, 55(3), 78.

Gleitz, H. F., Liao, A. Y., Cook, J. R., Rowlston, S. F., Forte, G. M., D'souza, Z., ... & Bigger, B. W. (2018). Brain-targeted stem cell gene therapy corrects mucopolysaccharidosis type II via multiple mechanisms. EMBO molecular medicine, 10(7), e8730.

Goasdoué, K., Miller, S. M., Colditz, P. B., & Björkman, S. T. (2017). The blood-brain barrier; protecting the developing fetal brain. Placenta, 54, 111-116.

Goker-Alpan, O. (2011). Therapeutic approaches to bone pathology in Gaucher disease: past, present and future. Molecular genetics and metabolism, 104(4), 438-447.

Goker-Alpan, O., Schiffmann, R., Park, J. K., Stubblefield, B. K., Tayebi, N., & Sidransky, E. (2003). Phenotypic continuum in neuronopathic Gaucher disease: an intermediate phenotype between type 2 and type 3. The Journal of pediatrics, 143(2), 273-276.

Gomes, C. P., Marins, M. M., Motta, F. L., Kyosen, S. O., Curiati, M. A., D'Almeida, V., ... & Pesquero, J. B. (2020). A new mutation in IDS gene causing hunter syndrome: a case report. Frontiers in genetics, 10, 1383.

Grabowski, G. A. (1997). Gaucher disease: gene frequencies and genotype/phenotype correlations. Genetic testing, 1(1), 5-12.

Greer, W. L., Riddell, D. C., Byers, D. M., Welch, J. P., Girouard, G. S., Sparrow, S. M., ... & Neumann, P. E. (1997). Linkage of Niemann-Pick disease type D to the same region of human chromosome 18 as Niemann-Pick disease type C. The American Journal of Human Genetics, 61(1), 139-142.

Grewal, S., Kilic, Ö., Savci-Heijink, C. D., & Kloen, P. (2019). Disturbed remodeling and delayed fracture healing in pediatric pycnodysostosis patients. Journal of orthopaedics, 16(5), 373-377.

Griscelli, C., Durandy, A., Guy-Grand, D., Daguillard, F., Herzog, C., & Prunieras, M. (1978). A syndrome associating partial albinism and immunodeficiency. The American journal of medicine, 65(4), 691-702.

Gueneau, L., Fish, R. J., Shamseldin, H. E., Voisin, N., Mau-Them, F. T., Preiksaitiene, E., ... & Reymond, A. (2018). KIAA1109 variants are associated with a severe disorder of brain development and arthrogryposis. The American Journal of Human Genetics, 102(1), 116-132.

Guffon, N., Bertrand, Y., Forest, I., Fouilhoux, A., & Froissart, R. (2009). Bone marrow transplantation in children with Hunter syndrome: outcome after 7 to 17 years. The Journal of pediatrics, 154(5), 733-737.

Gul, R., Firasat, S., Hussain, M., Afshan, K., & Nawaz, D. (2020). IDUA gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families. Congenital anomalies, 60(4), 126-127.

Gul, R., Firasat, S., Hussain, M., Tufail, M., Ahmad, W., & Afshan, K. (2021). Neurological manifestations in Pakistani lysosomal storage disorders patients and molecular characterization of Gaucher disease. Genetika, 53(3), 1017-1029.

Hagege, E., Grey, R. J., Lopez, G., Roshan Lal, T., Sidransky, E., & Tayebi, N. (2017). Type 2 Gaucher disease in an infant despite a normal maternal glucocerebrosidase gene. American Journal of Medical Genetics Part A, 173(12), 3211-3215.

Hahn, A., Praetorius, S., Karabul, N., Dießel, J., Schmidt, D., Motz, R., ... & Mengel, E. (2014). Outcome of patients with classical infantile pompe disease receiving enzyme replacement therapy in Germany. In JIMD Reports, Volume 20 (pp. 65-75). Springer, Berlin, Heidelberg.

Han, J. Y., & Lee, I. G. (2020). Genetic tests by next-generation sequencing in children with developmental delay and/or intellectual disability. Clinical and experimental pediatrics, 63(6), 195.

Han, T. U., Sam, R., & Sidransky, E. (2020). Small molecule chaperones for the treatment of Gaucher disease and GBA1-associated Parkinson disease. Frontiers in Cell and Developmental Biology, 8, 271.

Hendriksz, C. J., Burton, B., Fleming, T. R., Harmatz, P., Hughes, D., Jones, S. A., ... & Dummer, W. (2014). Efficacy and safety of enzyme replacement therapy with BMN 110 (elosulfase alfa) for Morquio A syndrome (mucopolysaccharidosis IVA): a phase 3 randomised placebo-controlled study. Journal of inherited metabolic disease, 37(6), 979-990.

Hendriksz, C. J., Harmatz, P., Beck, M., Jones, S., Wood, T., Lachman, R., ... & Tomatsu, S. (2013). Review of clinical presentation and diagnosis of mucopolysaccharidosis IVA. Molecular genetics and metabolism, 110(1-2), 54-64.

Henry, B., Ziobro, R., Becker, K. A., Kolesnick, R., & Gulbins, E. (2013). Acid sphingomyelinase. Sphingolipids: Basic Science and Drug Development, 77-88.

Henter, J. I., von Bahr Greenwood, T., & Bergsten, E. (2020). Emapalumab in primary hemophagocytic lymphohistiocytosis. The New England journal of medicine, 383(6), 596-598.

Hermans, M. M., Wisselaar, H. A., Kroos, M. A., Oostra, B. A., & Reuser, A. J. J. (1993). Human lysosomal α-glucosidase: functional characterization of the glycosylation sites. Biochemical Journal, 289(3), 681-686.

Hers, H. G. (1963). α-Glucosidase deficiency in generalized glycogen-storage disease (Pompe's disease). Biochemical Journal, 86(1), 11.

Higashi, O. (1954). Congenital Gigantism of Peroxidase Granules The First Case ever Reported of Qualitative Abnormity of Peroxidase. The Tohoku journal of experimental medicine, 59(3), 315-332.

Ho, M. C., & Hsieh, Y. T. (2013). Mixed hyperpigmentation and hypopigmentation of iris and choroid in Chediak-Higashi syndrome. Journal of American Association for Pediatric Ophthalmology and Strabismus, 17(5), 558-560.

Hoefsloot, L. H., Hoogeveen-Westerveld, M., Reuser, A. J., & Oostra, B. A. (1990). Characterization of the human lysosomal α-glucosidase gene. Biochemical journal, 272(2), 493-497.

Hollak, C. E. M., De Sonnaville, E. S. V., Cassiman, D., Linthorst, G. E., Groener, J. E., Morava, E., ... & Poorthuis, B. J. H. M. (2012). Acid sphingomyelinase (Asm) deficiency patients in The Netherlands and Belgium: disease spectrum and natural course in attenuated patients. Molecular genetics and metabolism, 107(3), 526-533.

Hoops, S. L., Kolter, T., & Sandhoff, K. (2007). Saposin C and other sphingolipid activator proteins. GAUCHER.

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., ... & Flicek, P. (2021). Ensembl 2021. Nucleic acids research, 49(D1), D884-D891.

Hrebicek M, Mrazova L, Seyrantepe V, et al (2006) Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Am J Hum Genet 79: 807–819.

Huang, L., Low, A., Damle, S.S., Keenan, M.M., Kuntz, S., Murray, S.F., Monia, B.P. and Guo, S (2018). Antisense suppression of the nonsense mediated decay factor Upf3b as a potential treatment for diseases caused by nonsense mutations. Genom biol 19(1):1-16.

Huang, X., Wu, D., Zhu, L., Wang, W., Yang, R., Yang, J., ... & Zhao, Z. (2022). Application of a next-generation sequencing (NGS) panel in newborn screening efficiently identifies inborn disorders of neonates. Orphanet Journal of Rare Diseases, 17(1), 66.

Hughes, D., Mikosch, P., Belmatoug, N., Carubbi, F., Cox, T., Goker-Alpan, O., ... & Deegan, P. (2019). Gaucher disease in bone: from pathophysiology to practice. Journal of Bone and Mineral Research, 34(6), 996-1013.

Huizing, M., Helip-Wooley, A., Westbroek, W., Gunay-Aygun, M., & Gahl, W. A. (2008). Disorders of lysosome-related organelle biogenesis: clinical and molecular genetics. Annual review of genomics and human genetics, 9, 359.

Hume, A. N., Collinson, L. M., Hopkins, C. R., Strom, M., Barral, D. C., Bossi, G., ... & Seabra, M. C. (2002). The leaden gene product is required with Rab27a to recruit myosin Va to melanosomes in melanocytes. Traffic, 3(3), 193-202.

Hussain, R., & Bittles, A. H. (1998). The prevalence and demographic characteristics of consanguineous marriages in Pakistan. Journal of biosocial science, 30(2), 261-275.

Iaselli, F., Rea, G., Cappabianca, S., Fabozzi, G., Montemarano, M., Vitale, C., ... & Rotondo, A. (2011). Adult-onset pulmonary involvement in Niemann-Pick disease type B. Monaldi Archives for Chest Disease, 75(4).

International HapMap 3 Consortium. (2010). Integrating common and rare genetic variation in diverse human populations. Nature, 467(7311), 52.

Introne, W. J., Westbroek, W., Cullinane, A. R., Groden, C. A., Bhambhani, V., Golas, G. A., ... & Toro, C. (2016). Neurologic involvement in patients with atypical Chediak-Higashi disease. Neurology, 86(14), 1320-1328.

Jackson, J., Titman, P., Butler, S., Bond, K., Rao, A., Veys, P., ... & Rao, K. (2013). Cognitive and psychosocial function post hematopoietic stem cell transplantation in children with hemophagocytic lymphohistiocytosis. Journal of allergy and clinical immunology, 132(4), 889-895.

Jain, E., Flanagan, M., Sheth, S., Patel, S., Gan, Q., Patel, B., ... & Zustiak, S. P. (2020). Biodegradable polyethylene glycol hydrogels for sustained release and enhanced stability of rhGALNS enzyme. Drug delivery and translational research, 10(5), 1341-1352.

Jin, J., Wu, K., Liu, Z., Chen, X., Jiang, S., Wang, Z., & Li, W. (2019). Whole exome sequencing identified a novel biallelic SMARCAL1 mutation in the extremely rare disease SIOD. Frontiers in genetics, 10, 565.

Jonsson, J. J., Aronovich, E. L., Braun, S. E., & Whitley, C. B. (1995). Molecular diagnosis of mucopolysaccharidosis type II (Hunter syndrome) by automated sequencing and computer-assisted interpretation: toward mutation mapping of the iduronate-2-sulfatase gene. American journal of human genetics, 56(3), 597.

Kaplan, J., De Domenico, I., & Ward, D. M. (2008). Chediak-higashi syndrome. Current opinion in hematology, 15(1), 22-29.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., ... & MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 581(7809), 434-443.

Kaur, P., Kadavigere, R., Girisha, K. M., & Shukla, A. (2020). Recurrent bi-allelic splicing variant c. 454+ 3A> G in TRAPPC4 is associated with progressive encephalopathy and muscle involvement. Brain, 143(4), e29-e29.

Kausar, M., Ashraf, N., Hayat, F., Hashmi, A. H., Siddiqi, S., & Anees, M. (2017). A Recurrent Mutation in CTSK Gene is Responsible for Autosomal Recessive Pycnodysostosis in Consanguineous Pakistani Families. Pakistan Journal of Zoology, 49(5).

Kaya, Z., Ehl, S., Albayrak, M., Maul-Pavicic, A., Schwarz, K., Kocak, U., ... & Gursel, T. (2011). A novel single point mutation of the LYST gene in two siblings with different phenotypic features of Chediak Higashi syndrome. Pediatric blood & cancer, 56(7), 1136-1139.

Keeling, K.M., Wang, D., Dai, Y., Murugesan, S., Chenna, B., Clark, J., Belakhov, V., Kandasamy, J., Velu, S.E., Baasov, T. and Bedwell, D.M (2013). Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression. PloS one 8(4):60478.

Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. Genome research, 12(4), 656-664.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. Genome research, 12(6), 996-1006.

Khan, B., Ahmed, Z., & Ahmad, W. (2010). A novel missense mutation in cathepsin K (CTSK) gene in a consanguineous Pakistani family with pycnodysostosis. Journal of Investigative Medicine, 58(5), 720-724.

Khan, M. A., Hussain, A., Sher, G., Zubaida, B., & Naeem, M. (2018). Genetic testing of a large consanguineous Pakistani family affected with mucolipidosis III gamma through next-generation sequencing. Genetic Testing and Molecular Biomarkers, 22(9), 541-545.

Khan, M. A., Zubaida, B., Karim, N., Cheema, H. A., & Naeem, M. (2020). Identification of two novel variants in GNPTAB underlying mucolipidosis II in a Pakistani family. Journal of Pediatric Endocrinology and Metabolism, 33(5), 647-651.

Khan, S. A., Peracha, H., Ballhausen, D., Wiesbauer, A., Rohrbach, M., Gautschi, M., ... & Tomatsu, S. (2017). Epidemiology of mucopolysaccharidoses. Molecular genetics and metabolism, 121(3), 227-240.

Kiely, B. T., Kohler, J. L., Coletti, H. Y., Poe, M. D., & Escolar, M. L. (2017). Early disease progression of Hurler syndrome. Orphanet journal of rare diseases, 12(1), 32.

Kim, H. M., Jeon, S., Chung, O., Jun, J. H., Kim, H. S., Blazyte, A., ... & Bhak, J. (2020). Comparative analysis of seven short-reads sequencing platforms using the Korean Reference Genome: MGI and Illumina sequencing benchmark for whole-genome sequencing. bioRxiv.

Kim, S. R., Kissoon-Larkin, T., Horn, B., & Elder, M. (2019). Anakinra as an agent to control hemophagocytic lymphohistiocytosis in Griscelli type 2. Pediatric blood & cancer, 66(12), e27997.

Kingma, S. D., & Jonckheere, A. I. (2021). MPS I: early diagnosis, bone disease and treatment, where are we now?. Journal of Inherited Metabolic Disease.

Kingma, S. D., Bodamer, O. A., & Wijburg, F. A. (2015). Epidemiology and diagnosis of lysosomal storage disorders; challenges of screening. Best Practice & Research Clinical Endocrinology & Metabolism, 29(2), 145-157.

Kishnani, P. S., Hwu, W. L., Mandel, H., Nicolino, M., Yong, F., Corzo, D., & Infantile-Onset Pompe Disease Natural History Study Group. (2006). A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease. The Journal of pediatrics, 148(5), 671-676.

Kishnani, P. S., Steiner, R. D., Bali, D., Berger, K., Byrne, B. J., Case, L. E., ... & Watson, M. S. (2006). Pompe disease diagnosis and management guideline. Genetics in Medicine, 8(5), 267-288.

Kohler, L., Puertollano, R., & Raben, N. (2018). Pompe disease: from basic science to therapy. Neurotherapeutics, 15(4), 928-942.

Koprivica, V., Stone, D. L., Park, J. K., Callahan, M., Frisch, A., Cohen, I. J., ... & Sidransky, E. (2000). Analysis and classification of 304 mutant alleles in patients with type 1 and type 3 Gaucher disease. The American Journal of Human Genetics, 66(6), 1777-1786.

Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R., & Remm, M. (2018). Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics, 34(11), 1937-1938.

Krivit, W., Sung, J. H., Shapiro, E. G., & Lockman, L. A. (1995). Microglia: the effector cell for reconstitution of the central nervous system following bone marrow transplantation for lysosomal and peroxisomal storage diseases. Cell transplantation, 4(4), 385-392.

Krzewski, K., & Cullinane, A. R. (2013). Evidence for defective Rab GTPase-dependent cargo traffic in immune disorders. Experimental cell research, 319(15), 2360-2367.

Kubo, Y., Sekiya, S., Ohigashi, M., Takenaka, C., Tamura, K., Nada, S., ... & Yamaguchi, A. (2005). ABCA5 resides in lysosomes, and ABCA5 knockout mice develop lysosomal disease-like symptoms. Molecular and cellular biology, 25(10), 4138-4149.

Kumar, K., Bellad, A., Prasad, P., Girimaji, S. C., & Muthusamy, B. (2020). KIAA1109 gene mutation in surviving patients with Alkuraya-Kučinskas syndrome: a review of literature. BMC Medical Genetics, 21(1), 1-11.

Kwon, H. J., Abi-Mosleh, L., Wang, M. L., Deisenhofer, J., Goldstein, J. L., Brown, M. S., & Infante, R. E. (2009). Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. Cell, 137(7), 1213-1224.

Lachmann, R. H. (2020). Treating lysosomal storage disorders: what have we learnt?. Journal of inherited metabolic disease, 43(1), 125-132.

Lange, M., Linden, T., Müller, H. L., Flasskuehler, M. A., Koester, H., Lehmberg, K., ... & Beske, F. (2021). Primary haemophagocytic lymphohistiocytosis (Chédiak-Higashi Syndrome) triggered by acute SARS-CoV-2 infection in a six-week-old infant. British Journal of Haematology.

Langford, G. M., & Molyneaux, B. J. (1998). Myosin V in the brain: mutations lead to neurological defects. Brain research reviews, 28(1-2), 1-8.

Lawrence, R., Van Vleet, J. L., Mangini, L., Harris, A., Martin, N., Clark, W., ... & Crawford, B. E. (2019). Characterization of glycan substrates accumulating in GM1 Gangliosidosis. Molecular genetics and metabolism reports, 21, 100524.

LeBlanc, S., & Savarirayan, R. (2020). Pycnodysostosis.

Leonard, J. V., & Morris, A. A. (2006). Diagnosis and early management of inborn errors of metabolism presenting around the time of birth. Acta paediatrica, 95(1), 6-14.

Leroy, J. G., & Martin, J. J. (1975). Mucolipidosis II (I-cell disease): present status of knowledge. Birth defects original article series, 11(6), 283-293.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, H. Y., Ma, H. W., Wang, H. Q., & Ma, W. H. (2009). Molecular analysis of a novel cathepsin K gene mutation in a Chinese child with pycnodysostosis. Journal of International Medical Research, 37(1), 264-269.

Li, P., Bellows, A. B., & Thompson, J. N. (1999). Molecular basis of iduronate-2-sulphatase gene mutations in patients with mucopolysaccharidosis type II (Hunter syndrome). Journal of medical genetics, 36(1), 21-27.

Lieberman, R. L., D'aquino, J. A., Ringe, D., & Petsko, G. A. (2009). Effects of pH and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability. Biochemistry, 48(22), 4816-4827.

Lin, H. Y., Lee, C. L., Chang, C. Y., Chiu, P. C., Chien, Y. H., Niu, D. M., ... & Lin, S. P. (2020). Survival and diagnostic age of 175 Taiwanese patients with mucopolysaccharidoses (1985–2019). Orphanet Journal of Rare Diseases, 15(1), 1-11.

Lin, H. Y., Lin, S. P., Chuang, C. K., Niu, D. M., Chen, M. R., Tsai, F. J., ... & Lin, J. L. (2009). Incidence of the mucopolysaccharidoses in Taiwan, 1984–2004. American journal of medical genetics Part A, 149(5), 960-964.

Lin, S. P., Lin, H. Y., Wang, T. J., Chang, C. Y., Lin, C. H., Huang, S. F., ... & Chuang, C. K. (2013). A pilot newborn screening program for Mucopolysaccharidosis type I in Taiwan. Orphanet journal of rare diseases, 8(1), 147.

Linhart, A., & Elliott, P. M. (2007). The heart in Anderson-Fabry disease and other lysosomal storage disorders. Heart, 93(4), 528-535.

Liu, S., Zhang, W., Shi, H., Meng, Y., & Qiu, Z. (2014). Three novel homozygous mutations in the GNPTG gene that cause mucolipidosis type III gamma. Gene, 535(2), 294-298.

Long, G. L., Winfield, S., Adolph, K. W., Ginns, E. I., & Bornstein, P. (1996). Structure and organization of the human metaxin gene (MTX) and pseudogene. Genomics, 33(2), 177-184.

Lotfi, P., Tse, D. Y., Di Ronza, A., Seymour, M. L., Martano, G., Cooper, J. D., ... & Sardiello, M. (2018). Trehalose reduces retinal degeneration, neuroinflammation and storage burden caused by a lysosomal hydrolase deficiency. Autophagy, 14(8), 1419-1434.

Lozano, M. L., Rivera, J., Sánchez-Guiu, I., & Vicente, V. (2014). Towards the targeted management of Chediak-Higashi syndrome. Orphanet journal of rare diseases, 9(1), 1-11.

Lücke, T., Billing, H., Sloan, E. A., Boerkoel, C. F., Franke, D., Zimmering, M., ... & Das, A. M. (2005). Schimke-immuno-osseous dysplasia: New mutation with weak genotype–phenotype correlation in siblings. American Journal of Medical Genetics Part A, 135(2), 202-205.

Luzio, J. P., Hackmann, Y., Dieckmann, N. M., & Griffiths, G. M. (2014). The biogenesis of lysosomes and lysosome-related organelles. Cold Spring Harbor perspectives in biology, 6(9), a016840.

Mabe, P., Valiente, A., Soto, V., Cornejo, V., & Raimann, E. (2004). Evaluation of reliability for urine mucopolysaccharidosis screening by dimethylmethylene blue and Berry spot tests. Clinica chimica acta, 345(1-2), 135-140.

Maita, N., Tsukimura, T., Taniguchi, T., Saito, S., Ohno, K., Taniguchi, H. and Sakuraba, H (2013). Human α-L-iduronidase uses its own N-glycan as a substrate-binding and catalytic module. Proc Natl Acad Sci 110(36):14628-14633.

Maquat, L. E. (2005). Nonsense-mediated mRNA decay in mammals. Journal of cell science, 118(9), 1773-1776.

Maroteaux, P., & Lamy, M. (1962) b. 2 cases of a condensing osseous disease: pynodysostosis. Archives francaises de pediatrie, 19, 267-274.

Maroteaux, P., & Lamy, M. (1966). Hurler's pseudo-polydystrophy. La Presse Medicale, 74(55), 2889-2892.

Maroteaux, P., & Lamy, M. P. (1962) a. in French. Presse Med, 70, 999-1002.

Marques, A. R., & Saftig, P. (2019). Lysosomal storage disorders–challenges, concepts and avenues for therapy: beyond rare diseases. Journal of cell science, 132(2), jcs221739.

Marsh, R. A., & Haddad, E. (2018). How i treat primary haemophagocytic lymphohistiocytosis. British journal of haematology, 182(2), 185-199.

Martell, L., Lau, K., Mei, M., Burnett, V., Decker, C., & Foehr, E. D. (2011). Biomarker analysis of Morquio syndrome: identification of disease state and drug responsive markers. Orphanet Journal of Rare Diseases, 6(1), 1-10.

Martínez-Arias, R., Calafell, F., Mateu, E., Comas, D., Andrés, A., and Bertranpetit, J. (2001). Sequence variability of a human pseudogene. Genome research, 11(6), 1071-1085.

Massaro, G., Hughes, M. P., Whaler, S. M., Wallom, K. L., Priestman, D. A., Platt, F. M., ... & Rahim, A. A. (2020). Systemic AAV9 gene therapy using the synapsin I promoter rescues a mouse model of neuronopathic Gaucher disease but with limited cross-correction potential to astrocytes. Human molecular genetics, 29(12), 1933-1949.

Masue, M., Sukegawa, K., Orii, T., & Hashimoto, T. (1991). N-acetylgalactosamine-6-sulfate sulfatase in human placenta: purification and characteristics. The Journal of Biochemistry, 110(6), 965-970.

Masuno, M., Tomatsu, S., Nakashima, Y., Hori, T., Fukuda, S., Masue, M., ... & Orii, T. (1993). Mucopolysaccharidosis IV A: assignment of the human N-acetylgalactosamine-6-sulfate sulfatase (GALNS) gene to chromosome 16q24. Genomics, 16(3), 777-778.

Matalonga, L., Arias, A., Coll, M. J., Garcia-Villoria, J., Gort, L., & Ribes, A. (2014). Treatment effect of coenzyme Q 10 and an antioxidant cocktail in fibroblasts of patients with Sanfilippo disease. Journal of inherited metabolic disease, 37(3), 439-446.

Matesic, L. E., Yip, R., Reuss, A. E., Swing, D. A., O'Sullivan, T. N., Fletcher, C. F., ... & Jenkins, N. A. (2001). Mutations in Mlph, encoding a member of the Rab effector family, cause the melanosome transport defects observed in leaden mice. Proceedings of the National Academy of Sciences, 98(18), 10238-10243.

Matos, L., Canals, I., Dridi, L., Choi, Y., Prata, M. J., Jordan, P., ... & Vilageliu, L. (2014). Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations. Orphanet journal of rare diseases, 9(1), 1-12.

McGovern, M. M., Aron, A., Brodie, S. E., Desnick, R. J., & Wasserstein, M. P. (2006). Natural history of Type A Niemann-Pick disease: possible endpoints for therapeutic trials. Neurology, 66(2), 228-232.

McGovern, M. M., Avetisyan, R., Sanson, B. J., & Lidove, O. (2017). Disease manifestations and burden of illness in patients with acid sphingomyelinase deficiency (ASMD). Orphanet journal of rare diseases, 12(1), 1-13.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research, 20(9), 1297-1303.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., ... & Cunningham, F. (2016). The ensembl variant effect predictor. Genome biology, 17(1), 1-14.

McNally, K. E., & Cullen, P. J. (2018). Endosomal retrieval of cargo: retromer is not alone. Trends in cell biology, 28(10), 807-822.

Mehta, A. B., & Winchester, B. (Eds.). (2022). Lysosomal storage disorders: a practical guide. John Wiley & Sons.

Meikle, P. J., Fietz, M. J., & Hopwood, J. J. (2004). Diagnosis of lysosomal storage disorders: current techniques and future directions. Expert review of molecular diagnostics, 4(5), 677-691.

Ménasché, G., Ho, C. H., Sanal, O., Feldmann, J., Tezcan, I., Ersoy, F., ... & de Saint Basile, G. (2003). Griscelli syndrome restricted to hypopigmentation results from a melanophilin defect (GS3) or a MYO5A F-exon deletion (GS1). The Journal of clinical investigation, 112(3), 450-456.

Ménasché, G., Pastural, E., Feldmann, J., Certain, S., Ersoy, F., Dupuis, S., ... & de Saint Basile, G. (2000). Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome. Nature genetics, 25(2), 173-176.

Metcalf, J. A., Linders, B., Wu, S., Bigg, P., O'Donnell, P., Sleeper, M. M., ... & Ponder, K. P. (2010). Upregulation of elastase activity in aorta in mucopolysaccharidosis I and VII dogs may be due to increased cytokine expression. Molecular genetics and metabolism, 99(4), 396-407.

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L. P., Mushayamaha, T., & Thomas, P. D. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. Nucleic acids research, 49(D1), D394-D403.

Mohammadzadeh Shanehsaz, S., Rezazadeh, A., & Dandashli, A. (2015). Elejalde syndrome (ES). Dermatology Online Journal, 21(3).

Montano, A. M., Tomatsu, S., Brusius, A., Smith, M., & Orii, T. (2008). Growth charts for patients affected with Morquio A disease. American Journal of Medical Genetics Part A, 146(10), 1286-1295.

Moreland, R. J., Jin, X., Zhang, X. K., Decker, R. W., Albee, K. L., Lee, K. L., ... & Canfield, W. M. (2005). Lysosomal acid α-glucosidase consists of four different peptides processed from a single chain precursor. Journal of Biological Chemistry, 280(8), 6780-6791.

Morquio, L. (1929). Sur une forme de dystrophie osseuse familiale. Bull Soc Pediatr Paris, 27, 145-152.

Morrone, A., Tylee, K. L., Al-Sayed, M., Brusius-Facchin, A. C., Caciotti, A., Church, H. J., ... & Miller, N. (2014). Molecular testing of 163 patients with Morquio A (Mucopolysaccharidosis IVA) identifies 39 novel GALNS mutations. Molecular genetics and metabolism, 112(2), 160-170.

Motta, I., Consonni, D., Stroppiano, M., Benedetto, C., Cassinerio, E., Tappino, B., ... & Cappellini, M. D. (2021). Predicting the probability of Gaucher disease in subjects with splenomegaly and thrombocytopenia. Scientific Reports, 11(1), 1-6.

Muenzer, J. (2004). The mucopolysaccharidoses: a heterogeneous group of disorders with variable pediatric presentations. The Journal of pediatrics, 144(5), S27-S34.

Muenzer, J., Wraith, J. E., & Clarke, L. A. (2009). Mucopolysaccharidosis I: management and treatment guidelines. Pediatrics, 123(1), 19-29.

Mujawar, Q., Naganoor, R., Patil, H., Thobbi, A. N., Ukkali, S., & Malagi, N. (2009). Pycnodysostosis with unusual findings: a case report. Cases Journal, 2(1), 1-4.

MWer, S., Dykes, D., & Polesky, H. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids res, 16(3), 1215.

Naeem, M., Sheikh, S., & Ahmad, W. (2009). A mutation in CTSK gene in an autosomal recessive pycnodysostosis family of Pakistani origin. BMC medical genetics, 10(1), 1-5.

Nagle, D. L., Karim, M. A., Woolf, E. A., Holmgren, L., Bork, P., Misumi, D. J., ... & Moore, K. J. (1996). Identification and mutation analysis of the complete gene for Chediak–Higashi syndrome. Nature genetics, 14(3), 307-311.

Nair, V., Belanger, E. C., & Veinot, J. P. (2019). Lysosomal storage disorders affecting the heart: a review. Cardiovascular Pathology, 39, 12-24.

Nakanishi, M., & Otsu, M. (2012). Development of Sendai virus vectors and their potential applications in gene therapy and regenerative medicine. Current gene therapy, 12(5), 410-416.

Naz, A., Abedin, Q., Ahmed, S., Siddiqui, S., & Shamsi, T. (2018). Identification of GBA gene mutations in 19 Pakistani unrelated patients of Gaucher disease. Blood, 132, 4952.

Nelson, J., Crowhurst, J., Carey, B., & Greed, L. (2003). Incidence of the mucopolysaccharidoses in Western Australia. American journal of medical genetics Part A, 123(3), 310-313.

Neufeld, E. F. (2001). The mucopolysaccharidoses. The metabolic and molecular bases of inherited disease, 3421-3452.

Newton, J., Milstien, S., & Spiegel, S. (2018). Niemann-Pick type C disease: The atypical sphingolipidosis. Advances in biological regulation, 70, 82-88.

Ngiwsara, L., Ketudat-Cairns, J. R., Sawangareetrakul, P., Charoenwattanasatien, R., Champattanachai, V., Kuptanon, C., ... & Svasti, J. (2018). p. X654R IDUA variant among Thai individuals with intermediate mucopolysaccharidosis type I and its residual activity as demonstrated in COS-7 cells. Annals of Human Genetics, 82(3), 150-157.

Niemann, C., & Link, K. P. (1933). Synthesis of the Hexuronic Acids: II. The Synthesis of D-Mannuronic Acid from D-Mannosaccharic Acid. Journal of Biological Chemistry, 100(2), 407-413.

Norton, W. H., Ledin, J., Grandel, H., & Neumann, C. J. (2005). HSPG synthesis by zebrafish Ext2 and Extl3 is required for Fgf10 signalling during limb development.

Ohira, M., Kikuchi, E., Mizuta, S., Yoshida, N., Onodera, M., Nakanishi, M., ... & Mashima, R. (2021). Production of therapeutic iduronate-2-sulfatase enzyme with a novel single-stranded RNA virus vector. Genes to Cells.

Okada, S., & O'Brien, J. S. (1968). Generalized gangliosidosis: beta-galactosidase deficiency. Science, 160(3831), 1002-1004.

Ortega-Recalde, O., Beltrán, O. I., Gálvez, J. M., Palma-Montero, A., Restrepo, C. M., Mateus, H. E., & Laissue, P. (2015). Biallelic HERC1 mutations in a syndromic form of overgrowth and intellectual disability. Clinical genetics, 88(4), e1-e3.

Orvisky, E., Park, J. K., LaMarca, M. E., Ginns, E. I., Martin, B. M., Tayebi, N., & Sidransky, E. (2002). Glucosylsphingosine accumulation in tissues from patients with Gaucher disease: correlation with phenotype and genotype. Molecular genetics and metabolism, 76(4), 262-270.

Otaify, G. A., Abdel-Hamid, M. S., Mehrez, M. I., Aboul-Ezz, E., Zaki, M. S., Aglan, M. S., & Temtamy, S. A. (2018). Genetic study of eight Egyptian patients with pycnodysostosis: identification of novel CTSK mutations and founder effect. Osteoporosis International, 29(8), 1833-1841.

Ou, L., DeKelver, R. C., Rohde, M., Tom, S., Radeke, R., Martin, S. J. S., ... & Whitley, C. B. (2019). ZFN-mediated in vivo genome editing corrects murine hurler syndrome. Molecular Therapy, 27(1), 178-187.

Oussoren, E., van Eerd, D., Murphy, E., Lachmann, R., van der Meijden, J. C., Hoefsloot, L. H., … Langeveld, M. (2018). Mucolipidosis type III, a series of adult patients. Journal of Inherited Metabolic Disease, 41, 839–848.

Ozkinay, F., Emecen, D. A., Kose, M., Isik, E., Bozaci, A. E., Canda, E., ... & Onay, H. (2021). Clinical and genetic features of 13 patients with mucopolysaccarhidosis type IIIB: Description of two novel NAGLU gene mutations. Molecular Genetics and Metabolism Reports, 27, 100732.

Pagel, K. A., Pejaver, V., Lin, G. N., Nam, H. J., Mort, M., Cooper, D. N., ... & Radivojac, P. (2017). When loss-of-function is loss of function: assessing mutational signatures and impact of loss-of-function genetic variants. Bioinformatics, 33(14), i389-i398.

Pastural, E., Barrat, F. J., Dufourcq-Lagelouse, R., Certain, S., Sanal, O., Jabado, N., ... & Basile, G. D. S. (1997). Griscelli disease maps to chromosome 15q21 and is

associated with mutations in the myosin-Va gene. Nature genetics, 16(3), 289-292.

Pastural, E., Ersoy, F., Yalman, N., Wulffraat, N., Grillo, E., Ozkinay, F., ... & de Saint Basile, G. (2000). Two genes are responsible for Griscelli syndrome at the same 15q21 locus. Genomics, 63(3), 299-306.

Peracha, H., Sawamoto, K., Averill, L., Kecskemethy, H., Theroux, M., Thacker, M., ... & Tomatsu, S. (2018). Molecular genetics and metabolism, special edition: Diagnosis, diagnosis and prognosis of Mucopolysaccharidosis IVA. Molecular genetics and metabolism, 125(1-2), 18-37.

Perou, C. M., Moore, K. J., Nagle, D. L., Misumi, D. J., Woolf, E. A., McGrail, S. H., ... & Kaplan, J. (1996). Identification of the murine beige gene by YAC complementation and positional cloning. Nature genetics, 13(3), 303-308.

Pick, L. (1933). Niemann–Pick‘s disease and other forms of so called xanthomatosis. Am J Med Sci, 185, 601-616.

Piraud, M., Boyer, S., Mathieu, M., & Maire, I. (1993). Diagnosis of mucopolysaccharidoses in a clinically selected population by urinary glycosaminoglycan analysis: a study of 2,000 urine samples. Clinica chimica acta, 221(1-2), 171-181.

Pispa, J. (1965). Animal biotinidase. In Annales medicinaeexperimentalis et biologiaeFenniae (Vol. 43, pp. 1-39).

Platt, F. M., Boland, B., & van der Spoel, A. C. (2012). Lysosomal storage disorders: The cellular impact of lysosomal dysfunction. Journal of Cell Biology, 199(5), 723-734.

Pohl, S., Encarnacão, M., Castrichini, M., Müller-Loennies, S., Muschol, N., & Braulke, T. (2010). Loss of N-acetylglucosamine-1-phosphotransferase gamma subunit due to intronic mutation in GNPTG causes mucolipidosis type III gamma: Implications for molecular and cellular diagnostics. American Journal of Medical Genetics, Part A, 152A(1), 124–132.

Polymeropoulos, M. H., Ortiz De Luna, R. I., Ide, S. E., Torres, R., Rubenstein, J., & Francomano, C. A. (1995). The gene for pycnodysostosis maps to human chromosome 1cen–q21. Nature genetics, 10(2), 238-239.

Pompe, J. C. (1932). Over idiopathische hypertrophie van het hart. Ned Tijdschr Geneeskd, 76, 304-311.

Qian, Y., Lee, I., Lee, W. S., Qian, M., Kudo, M., Canfield, W. M., ... & Kornfeld, S. (2010). Functions of the α, β, and γ subunits of UDP-GlcNAc: lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. Journal of Biological Chemistry, 285(5), 3360-3370.

Raas-Rothschild, A., Cormier-Daire, V., Bao, M., Genin, E., Salomon, R., Brewer, K., ... & Canfield, W. M. (2000). Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). The Journal of clinical investigation, 105(5), 673-681.

Raas-Rothschild, A., Pohl, S., & Braulke, T. (2012). A. B. Mehta, & B. Winchester (Eds.), Lysosomal storage diseases: A practical guide: Multiple enzyme deficiencies: Defects in transport: Mucolipidosis II; mucolipidosis III alpha/beta and mucolipidosis III gamma (pp. 121–126). Oxford, UK: Wiley-Blackwell (Oxford).

Rampanelli, E., Ochodnicky, P., Vissers, J. P., Butter, L. M., Claessen, N., Calcagni, A., ... & Leemans, J. C. (2018). Excessive dietary lipid intake provokes an acquired form of lysosomal lipid storage disease in the kidney. The Journal of pathology, 246(4), 470-484.

Raza, R., Ullah, A., Haider, N., Krishin, J., Shah, M., Khan, F. U., ... & Basit, S. (2022). Exome sequencing reveals the first intragenic deletion in ABCA5 underlying autosomal recessive hypertrichosis. Clinical and Experimental Dermatology.

Reitman, M. L., & Kornfeld, S. (1981). UDP-N-acetylglucosamine: glycoprotein N-acetylglucosamine-1-phosphotransferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes. Journal of Biological Chemistry, 256(9), 4275-4281.

Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., ... & Bale, S. (2016). Clinical application of whole-exome sequencing across clinical indications. Genetics in Medicine, 18(7), 696-704.

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic acids research, 39(17), e118-e118.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine, 17(5), 405-423.

Ridaura-Sanz, C., Durán-McKinster, C., & Ruiz-Maldonado, R. (2018). Usefulness of the skin biopsy as a tool in the diagnosis of silvery hair syndrome. Pediatric dermatology, 35(6), 780-783.

Rivera-Colón, Y., Schutsky, E. K., Kita, A. Z., & Garman, S. C. (2012). The structure of human GALNS reveals the molecular basis for mucopolysaccharidosis IV A. Journal of molecular biology, 423(5), 736-751.

Robertson DA, Callen DF, Baker EG, Morris CP, Hopwood JJ (1988) Chromosomal localization of the gene for human glucosamine-6-sulphatase to 12q14. Hum Genet 79: 175–178.

Rothenbühler, A., Piquard, C., Gueorguieva, I., Lahlou, N., Linglart, A., & Bougneres, P. (2010). Near normalization of adult height and body proportions by growth hormone in pycnodysostosis. The Journal of Clinical Endocrinology & Metabolism, 95(6), 2827-2831.

Ruijter, G. J. G., Valstar, M. J., van de Kamp, J. M., Van der Helm, R. M., Durand, S., Van Diggelen, O. P., ... & Wijburg, F. A. (2008). Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. Molecular genetics and metabolism, 93(2), 104-111.

Runz, H., Dolle, D., Schlitter, A. M., & Zschocke, J. (2008). NPC-db, a Niemann-Pick type C disease gene variation database. Human mutation, 29(3), 345-350.

Sachdev, M., Bansal, M., Chakraborty, S., Hamal, S., Bhargava, R., & Dua, V. (2021). Haploidentical Stem Cell Transplant With Post-transplant Cyclophosphamide for Chediak-Higashi Syndrome: A Very Rare Case Report. Journal of pediatric hematology/oncology, 43(7), e1030-e1032.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol: chloroform. Cold Spring Harbor Protocols, 2006(1), pdb-prot4455.

Sánchez-Guiu, I., Antón, A. I., García-Barberá, N., Navarro-Fernández, J., Martínez, C., Fuster, J. L., ... & Lozano, M. L. (2014). Chediak–H igashi syndrome: description of two novel homozygous missense mutations causing divergent clinical phenotype. European journal of haematology, 92(1), 49-58.

Sanfilippo, S. J. (1963). Mental retardation associated with acid mucopoly-saccharidurie (heparan sulfate type). J Pediatr, 63, 837-838.

Sano, M., Nakasu, A., Ohtaka, M., & Nakanishi, M. (2019). A Sendai Virus-Based Cytoplasmic RNA Vector as a Novel Platform for Long-Term Expression of MicroRNAs. Molecular Therapy-Methods & Clinical Development, 15, 371-382.

Saraiva, J. M., Dinis, A., Resende, C., Faria, E., Gomes, C., Correia, A. J., ... & da Fonseca, N. (1999). Schimke immuno-osseous dysplasia: case report and review of 25 patients. Journal of medical genetics, 36(10), 786-789.

Sattar, F., Shan, T., Virk, R., Hassan, S. W., & Gul, B. (2021). Molecular and Genetic Analysis of Selected Exons of GBA gene in Pakistani Families affected with Gaucher's Disease. Ilkogretim Online, 20(6), 501-508.

Sayed Amr, K., El-Bassyouni, H. T., Abdel Hady, S., Mostafa, M. I., Mehrez, M. I., Coviello, D., & El-Kamah, G. Y. (2021). Genetic and Molecular Evaluation: Reporting Three Novel Mutations and Creating Awareness of Pycnodysostosis Disease. Genes, 12(10), 1552.

Scarpa, M., Orchard, P. J., Schulz, A., Dickson, P. I., Haskins, M. E., Escolar, M. L., & Giugliani, R. (2017). Treatment of brain disease in the mucopolysaccharidoses. Molecular genetics and metabolism, 122, 25-34.

Scerra, G., De Pasquale, V., Scarcella, M., Caporaso, M. G., Pavone, L. M., & D'Agostino, M. (2022). Lysosomal positioning diseases: beyond substrate storage. Open Biology, 12(10), 220155.

Schapira, A. H. (2015). Glucocerebrosidase and Parkinson disease: recent advances. Molecular and Cellular Neuroscience, 66, 37-42.

Schiffmann, R., Sevigny, J., Rolfs, A., Davies, E. H., Goker-Alpan, O., Abdelwahab, M., ... & Zimran, A. (2020). The definition of neuronopathic Gaucher disease. Journal of inherited metabolic disease, 43(5), 1056-1059.

Schilling, A. F., Mülhausen, C., Lehmann, W., Santer, R., Schinke, T., Rueger, J. M., & Amling, M. (2007). High bone mineral density in pycnodysostotic patients with a novel mutation in the propeptide of cathepsin K. Osteoporosis international, 18(5), 659-669.

Schmidt, B., Selmer, T., Ingendoh, A., & von Figurat, K. (1995). A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell, 82(2), 271-278.

Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., ... & Orlando, L. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nature protocols, 9(5), 1056-1082.

Schuchman, E. H. (2001). Niemann-Pick disease types A and B: acid sphingomyelinase deficiencies. The metabolic and molecular bases of inherited disease, 3, 3589-3610.

Schuchman, E. H., Levran, O., Pereira, L. V., & Desnick, R. J. (1992). Structural organization and complete nucleotide sequence of the gene encoding human acid sphingomyelinase (SMPD1). Genomics, 12(2), 197-205.

Schuh, R. S., Gonzalez, E. A., Tavares, A. M. V., Seolin, B. G., de Souza Elias, L., Vera, L. N. P., ... & Baldo, G. (2020). Neonatal nonviral gene editing with the CRISPR/Cas9 system improves some cardiovascular, respiratory, and bone disease features of the mucopolysaccharidosis I phenotype in mice. Gene therapy, 27(1), 74-84.

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. Nature methods, 11(4), 361-362.

Scott, C. R., Elliott, S., Buroker, N., Thomas, L. I., Keutzer, J., Glass, M., ... & Turecek, F. (2013). Identification of infants at risk for developing Fabry, Pompe, or mucopolysaccharidosis-I from newborn blood spots by tandem mass spectrometry. The Journal of pediatrics, 163(2), 498-503.

Scott, C., & Ioannou, Y. A. (2004). The NPC1 protein: structure implies function. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1685(1-3), 8-13.

Scott, H. S., Anson, D. S., Orsborn, A. M., Nelson, P. V., Clements, P. R., Morris, C. P., & Hopwood, J. J. (1991). Human alpha-L-iduronidase: cDNA isolation and expression. Proceedings of the National Academy of Sciences, 88(21), 9695-9699.

Scott, H. S., Blanch, L., Guo, X. H., Freeman, C., Orsborn, A., Baker, E., ... & Hopwood, J. J. (1995). Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. Nature genetics, 11(4), 465-467.

Scott, H. S., Bunge, S., Gal, A., Clarke, L. A., Morris, C. P., & Hopwood, J. J. (1995). Molecular genetics of muccpolysaccharidosis type I: Diagnostic, clinical, and biological implications. Human mutation, 6(4), 288-302.

Scott, H. S., Guo, X. H., Hopwood, J. J., & Morris, C. P. (1992c). Structure and sequence of the human α-L-iduronidase gene. Genomics, 13(4), 1311-1313.

Scott, H., V. Nelson, P., Litjens, T., J. Hopwood, J. and Morris, C.P (1993). Multiple polymorphisms within the α-L-iduronidase gene (IDUA): implications for a role in modification of MPS-I disease phenotype. Hum Mol Genet 2(9):1471-1473.

Seker Yilmaz, B., Davison, J., Jones, S. A., & Baruteau, J. (2021). Novel therapies for mucopolysaccharidosis type III. Journal of Inherited Metabolic Disease, 44(1), 129-147.

Settembre, C., & Ballabio, A. (2014). Lysosomal adaptation: how the lysosome responds to external cues. Cold Spring Harbor perspectives in biology, 6(6), a016907.

Settembre, C., Fraldi, A., Medina, D. L., & Ballabio, A. (2013). Signals from the lysosome: a control centre for cellular clearance and energy metabolism. Nature reviews Molecular cell biology, 14(5), 283-296.

Shah, B. J., Jagati, A. K., Katrodiya, N. K., & Patel, S. M. (2016). Griscelli syndrome type-3. Indian dermatology online journal, 7(6), 506.

Shapiro, E. G., Jones, S. A., & Escolar, M. L. (2017). Developmental and behavioral aspects of mucopolysaccharidoses with brain manifestations—neurological signs and symptoms. Molecular genetics and metabolism, 122, 1-7.

Shapiro, E. G., Nestrasil, I., Delaney, K. A., Rudser, K., Kovac, V., Nair, N., ... & Whitley, C. B. (2016). A prospective natural history study of mucopolysaccharidosis type IIIA. The Journal of pediatrics, 170, 278-287.

Shapiro, E. G., Nestrasil, I., Rudser, K., Delaney, K., Kovac, V., Ahmed, A., ... & Whitley, C. B. (2015). Neurocognition across the spectrum of mucopolysaccharidosis type I: age, severity, and treatment. Molecular genetics and metabolism, 116(1-2), 61-68.

Sharma, P., Nicoli, E. R., Serra-Vinardell, J., Morimoto, M., Toro, C., Malicdan, M. C. V., & Introne, W. J. (2020). Chediak-Higashi syndrome: a review of the past, present, and future. Drug Discovery Today: Disease Models, 31, 31-36.

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic acids research, 29(1), 308-311.

Sheth, J., Bhavsar, R., Mistri, M., Pancholi, D., Bavdekar, A., Dalal, A., ... & Sheth, F. (2019). Gaucher disease: single gene molecular characterization of one-hundred Indian patients reveals novel variants and the most prevalent mutation. BMC medical genetics, 20(1), 1-11.

Sidransky, E., Orvisky, E., Lwin, A., Stubblefield, B., Schiffmann, R., & Goker-Alpan, O. (2003). Role of genetic modifiers in determining the phenotype in Gaucher

disease: A closer-look at the genotype L444P/L444P. In AMERICAN JOURNAL OF HUMAN GENETICS (Vol. 73, No. 5, pp. 451-451). 1427 E 60TH ST, CHICAGO, IL 60637-2954 USA: UNIV CHICAGO PRESS.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research, 40(W1), W452-W457.

Simonaro, C. M., D'Angelo, M., Haskins, M. E., & Schuchman, E. H. (2005). Joint and bone disease in mucopolysaccharidoses VI and VII: identification of new therapeutic targets and biomarkers using animal models. Pediatric research, 57(5), 701-707.

Simonaro, C. M., Park, J. H., Eliyahu, E., Shtraizent, N., McGovern, M. M., & Schuchman, E. H. (2006). Imprinting at the SMPD1 Locus: Implications for Acid Sphingomyelinase–Deficient Niemann-Pick Disease. The American Journal of Human Genetics, 78(5), 865-870.

Spada, M., Pagliardini, V., Ricci, F., Biamino, E., Mongini, T., & Porta, F. (2018). Early higher dosage of alglucosidase alpha in classic Pompe disease. Journal of Pediatric Endocrinology and Metabolism, 31(12), 1343-1347.

Stark, Z., & Savarirayan, R. (2009). Osteopetrosis. Orphanet journal of rare diseases, 4(1), 1-12.

Stinchcombe, J. C., Barral, D. C., Mules, E. H., Booth, S., Hume, A. N., Machesky, L. M., ... & Griffiths, G. M. (2001). Rab27a is required for regulated secretion in cytotoxic T lymphocytes. The Journal of cell biology, 152(4), 825-834.

Stirnemann, J., Belmatoug, N., Camou, F., Serratrice, C., Froissart, R., Caillaud, C., ... & Berger, M. G. (2017). A review of Gaucher disease pathophysiology, clinical presentation and treatments. International journal of molecular sciences, 18(2), 441.

Stockler-Ipsiroglu, S., Yazdanpanah, N., Yazdanpanah, M., Moisa Popurs, M., Yuskiv, N., Schmitz Ferreira Santos, M. L., ... & Giugliani, R. (2021). Morquio-like dysostosis multiplex presenting with neuronopathic features is a distinct GLB1-related phenotype. JIMD reports, 60(1), 23-31.

Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques, 28(6), 1102-1104.

Strom, M., Hume, A. N., Tarafder, A. K., Barkagianni, E., & Seabra, M. C. (2002). A family of Rab27-binding proteins: melanophilin links Rab27a and myosin Va function in melanosome transport. Journal of Biological Chemistry, 277(28), 25423-25430.

Sugahara, K., & Mikami, T. (2007). Chondroitin/dermatan sulfate in the central nervous system. Current opinion in structural biology, 17(5), 536-545.

Sugahara, K., Mikami, T., Uyama, T., Mizuguchi, S., Nomura, K., & Kitagawa, H. (2003). Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. Current opinion in structural biology, 13(5), 612-620.

Sukegawa, K., Tomatsu, S., Fukao, T., Iwata, H., Song, X. Q., Yamada, Y., ... & Orii, T. (1995). Mucopolysaccharidosis type II (Hunter disease): identification and characterization of eight point mutations in the iduronate-2-sulfatase gene in Japanese patients. Human mutation, 6(2), 136-143.

Sun, A. (2018). Lysosomal storage disease overview. Annals of translational medicine, 6(24).

Tammachote, R., Tongkobpetch, S., Srichomthong, C., Phipatthanananti, K., Pungkanon, S., Wattanasirichaigoon, D., ... & Shotelersuk, V. (2013). A common and two novel GBA mutations in Thai patients with Gaucher disease. Journal of human genetics, 58(9), 594-599.

Thiele, H., & Nürnberg, P. (2005). HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics, 21(8), 1730-1732.

Thomas, A. S., Mehta, A. B., & Hughes, D. A. (2013). Diagnosing Gaucher disease: an on-going need for increased awareness amongst haematologists. Blood Cells, Molecules, and Diseases, 50(3), 212-217.

Thurberg, B. L., Diaz, G. A., Lachmann, R. H., Schiano, T., Wasserstein, M. P., Ji, A. J., ... & Peterschmitt, M. J. (2020). Long-term efficacy of olipudase alfa in adults with acid sphingomyelinase deficiency (ASMD): Further clearance of

hepatic sphingomyelin is associated with additional improvements in pro-and anti-atherogenic lipid profiles after 42 months of treatment. Molecular Genetics and Metabolism, 131(1-2), 245-252.

Tiede, S., Muschol, N., Reutter, G., Cantz, M., Ullrich, K., &Braulke, T. (2005). Missense mutations in N-acetylglucosamine-1-phosphotransferase α/β subunit gene in a patient with mucolipidosis III and a mild clinical phenotype. American Journal of Medical Genetics Part A, 137(3), 235-240.

Tieu, P. T., Bach, G., Matynia, A., Hwang, M., & Neufeld, E. F. (1995). Four novel mutations underlying mild or intermediate forms of α-L-iduronidase deficiency (MPS IS and MPS IH/S). Human mutation, 6(1), 55-59.

Tomatsu, S., Fukuda, S., Cooper, A., Wraith, J. E., Rezvi, G. M. M., Yamagishi, A., ... & Orii, T. (1995). Mucopolysaccharidosis type IVA: identification of six novel mutations among non-Japanese patients. Human molecular genetics, 4(4), 741-743.

Tomatsu, S., Montaño, A. M., Nishioka, T., Gutierrez, M. A., Peña, O. M., Trandafirescu, G. G., ... &Orii, T. (2005). Mutation and polymorphism spectrum of the GALNS gene in mucopolysaccharidosis IVA (Morquio A). Human mutation, 26(6), 500-512.

Tomatsu, S., Nishioka, T., Montaño, A. M., Gutierrez, M. A., Pena, O. S., Orii, K. O., ... & Noguchi, A. (2004). Mucopolysaccharidosis IVA: identification of mutations and methylation study in GALNS gene. Journal of medical genetics, 41(7), e98-e98.

Trinh, M. N., Brown, M. S., Seemann, J., Goldstein, J. L., & Lu, F. (2018). Lysosomal cholesterol export reconstituted from fragments of Niemann-Pick C1. Elife, 7, e38564.

Tsuji, S., Choudary, P. V., Martin, B. M., Stubblefield, B. K., Mayor, J. A., Barranger, J. A., and Ginns, E. I. (1987). A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. New England Journal of Medicine, 316(10), 570-575.

Turan, S. (2014). Current research on pycnodysostosis. Intractable & rare diseases research, 3(3), 91-93.

Turk, V., Turk, B., & Turk, D. (2001). Lysosomal cysteine proteases: facts and opportunities. The EMBO journal, 20(17), 4629-4633.

Tüysüz, B., Kasapçopur, Ö., Alkaya, D. U., Şahin, S., Sözeri, B., & Yeşil, G. (2018). Mucolipidosis type III gamma: three novel mutation and genotype-phenotype study in eleven patients. Gene, 642, 398-407.

Tylki-Szymańska, A., Groener, J. E., Kamiński, M. L., Ługowska, A., Jurkiewicz, E., & Czartoryska, B. (2011). Gaucher disease due to saposin C deficiency, previously described as non-neuronopathic form—No positive effects after 2-years of miglustat therapy. Molecular genetics and metabolism, 104(4), 627-630.

Ullah, A., Zubaida, B., Cheema, H. A., & Naeem, M. (2020). Identification of two novel variants in GAA underlying infantile-onset Pompe disease in two Pakistani families. Journal of Pediatric Endocrinology and Metabolism, 33(4), 553-556.

Ullah, I., Nasir, A., Mehmood, S., Ahmed, S., Ullah, M. I., Ullah, A., ... & Ahmad, W. (2017). Identification and in silico analysis of GALNS mutations causing Morquio A syndrome in eight consanguineous families. Turkish Journal of Biology, 41(3), 458-468.

Uttarilli, A., Ranganath, P., Matta, D., Md Nurul Jain, J., Prasad, K., Babu, A. S., ... & Puri, R. D. (2016). Identification and characterization of 20 novel pathogenic variants in 60 unrelated Indian patients with mucopolysaccharidoses type I and type II. Clinical genetics, 90(6), 496-508.

Valente, N. Y. S., & Doche, I. (2019). 12.8 Silvery Hair Syndromes. Fundamentals of Hair and Scalp Dermoscopy, 180.

Valstar, M. J., Marchal, J. P., Grootenhuis, M., Colland, V., & Wijburg, F. A. (2011). Cognitive development in patients with Mucopolysaccharidosis type III (Sanfilippo syndrome). Orphanet journal of rare diseases, 6(1), 1-6.

Valstar, M. J., Neijs, S., Bruggenwirth, H. T., Olmer, R., Ruijter, G. J., Wevers, R. A., ... & Wijburg, F. A. (2010). Mucopolysaccharidosis type IIIA: clinical spectrum and genotype-phenotype correlations. Annals of neurology, 68(6), 876-887.

Valstar, M. J., Ruijter, G. J. G., Van Diggelen, O. P., Poorthuis, B. J., & Wijburg, F. A. (2008). Sanfilippo syndrome: a mini-review. *Journal of inherited metabolic disease*, *31*(2), 240-252.

Van Bergen, N. J., Guo, Y., Al-Deri, N., Lipatova, Z., Stanga, D., Zhao, S., ... & Christodoulou, J. (2020). Deficiencies in vesicular transport mediated by TRAPPC4 are associated with severe syndromic intellectual disability. Brain, 143(1), 112-130.

Van den Hout, H. M., Hop, W., van Diggelen, O. P., Smeitink, J. A., Smit, G. P. A., Bakker, H. D., ... & van der Ploeg, A. T. (2003). The natural course of infantile Pompe's disease: 20 original cases compared with 133 cases from the literature. Pediatrics, 112(2), 332-340.

Van der Auwera, G. A., & O'Connor, B. D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media.

Vanier, M. T. (2010). Niemann-Pick disease type C. Orphanet journal of rare diseases, 5(1), 1-18.

Vanier, M. T., & Millat, G. (2003). Niemann–Pick disease type C. Clinical genetics, 64(4), 269-281.

Vasiljeva, O., Reinheckel, T., Peters, C., Turk, D., Turk, V., & Turk, B. (2007). Emerging roles of cysteine cathepsins in disease and their potential as drug targets. Current pharmaceutical design, 13(4), 387-403.

Velho, R. V., Alegra, T., Sperb, F., Ludwig, N. F., Saraiva-Pereira, M. L., … Schwartz, I. V. (2014). A de novo or germline mutation in a family with mucolipidosis III gamma: Implications for molecular diagnosis and genetic counseling. Molecular Genetics and Metabolism Reports, 1, 98–102.

Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-

Science approach with life scientist friendly interfaces. BMC bioinformatics, 11(1), 1-10.

Verma, V., & Singh, R. K. (2020). IGFA Case Report of Pycnodysostosis Associated with Multiple Pituitary Hormone Deficiencies and Response to Treatment. Journal of clinical research in pediatric endocrinology, 12(4), 444.

Vijay, S., & Ed Wraith, J. (2005). Clinical presentation and follow-up of patients with the attenuated phenotype of mucopolysaccharidosis type I. Acta Pædiatrica, 94(7), 872-877.

Voorink-Moret, M., Goorden, S. M. I., van Kuilenburg, A. B. P., Wijburg, F. A., Ghauharali-Van Der Vlugt, J. M. M., Beers-Stet, F. S., ... & Vaz, F. M. (2018). Rapid screening for lipid storage disorders using biochemical markers. Expert center data and review of the literature. Molecular genetics and metabolism, 123(2), 76-84.

Wada, M., Shimada, Y., Iizuka, S., Ishii, N., Hiraki, H., Tachibana, T., ... & Kobayashi, H. (2020). Ex Vivo Gene Therapy Treats Bone Complications of Mucopolysaccharidosis Type II Mouse Models through Bone Remodeling Reactivation. Molecular Therapy-Methods & Clinical Development, 19, 261-274.

Walsh, E. C., & Stainier, D. Y. (2001). UDP-glucose dehydrogenase required for cardiac valve formation in zebrafish. Science, 293(5535), 1670-1673.

Wasserstein, M. P., Diaz, G. A., Lachmann, R. H., Jouvin, M. H., Nandy, I., Ji, A. J., & Puga, A. C. (2018). Olipudase alfa for treatment of acid sphingomyelinase deficiency (ASMD): safety and efficacy in adults treated for 30 months. Journal of inherited metabolic disease, 41(5), 829-838.

Weber, B., Guo, X. H., Wraith, J. E., Cooper, A., Kleijer, W. J., Bunge, S., & Hopwood, J. J. (1997). Novel mutations in Sanfilippo A syndrome: implications for enzyme function. Human molecular genetics, 6(9), 1573-1579.

Wilson, S. M., Yip, R., Swing, D. A., O'Sullivan, T. N., Zhang, Y., Novak, E. K., ... & Jenkins, N. A. (2000). A mutation in Rab27a causes the vesicle transport defects

observed in ashen mice. Proceedings of the National Academy of Sciences, 97(14), 7933-7938.

Winder-Rhodes, S. E., Garcia-Reitböck, P., Ban, M., Evans, J. R., Jacques, T. S., Kemppinen, A., ... & Spillantini, M. G. (2012). Genetic and pathological links between Parkinson's disease and the lysosomal disorder Sanfilippo syndrome. Movement disorders, 27(2), 312-315.

Winfield, S. L., Tayebi, N., Martin, B. M., Ginns, E. I., and Sidransky, E. (1997). Identification of three additional genes contiguous to the glucocerebrosidase locus on chromosome 1q21: implications for Gaucher disease. Genome research, 7(10), 1020-1026.

Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. F1000Research, 7.

Winkler, M. B., Kidmose, R. T., Szomek, M., Thaysen, K., Rawson, S., Muench, S. P., ... & Pedersen, B. P. (2019). Structural insight into eukaryotic sterol transport through Niemann-Pick type C proteins. Cell, 179(2), 485-497.

Winsor, E. J., & Welch, J. P. (1978). Genetic and demographic aspects of Nova Scotia Niemann-Pick disease (type D). American journal of human genetics, 30(5), 530.

Wisselaar, H. A., Kroos, M. A., Hermans, M. M., Van Beeumen, J., & Reuser, A. J. (1993). Structural and functional changes of lysosomal acid alpha-glucosidase during intracellular transport and maturation. Journal of Biological Chemistry, 268(3), 2223-2231.

Wolf, B., Grier, R. E., SecorMcVoy, J. R., & Heard, G. S. (1985). Biotinidase deficiency: a novel vitamin recycling defect. Journal of inherited metabolic disease, 8(1), 53-58.

Wolf, B., Heard, G. S., Weissbecker, K. A., McVoy, J. R. S., Grier, R. E., & Leshner, R. T. (1985). Biotinidase deficiency: initial clinical features and rapid diagnosis. Annals of neurology, 18(5), 614-617.

Wolfe, L. S., Senior, R. G., & Kin, N. N. Y. (1974). The structures of oligosaccharides accumulating in the liver of GM1-gangliosidosis, type I. Journal of Biological Chemistry, 249(6), 1828-1838.

World Medical Association (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. Jama 310(20):2191-2194.

Wraith, J. E. (1996). Mucopolysaccharidoses. Current Paediatrics, 6(2), 74-79.

Yang, C., & Wang, X. (2021). Lysosome biogenesis: Regulation and functions. Journal of Cell Biology, 220(6), e202102001.

Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. Nucleic acids research, 43(W1), W174-W181.

Yang, M., Cho, S. Y., Park, H. D., Choi, R., Kim, Y. E., Kim, J., ... & Jin, D. K. (2017). Clinical, biochemical and molecular characterization of Korean patients with mucolipidosis II/III and successful prenatal diagnosis. Orphanet Journal of Rare Diseases, 12(1), 1-9.

Young, I. D., & Harper, P. S. (1983). The natural history of the severe form of Hunter's syndrome: a study based on 52 cases. Developmental Medicine & Child Neurology, 25(4), 481-489.

Zahoor, M.Y., Cheema, H.A., Ijaz, S., Anjum, M.N., Ramzan, K. and Bhinder, M.A (2019). Mapping of IDUA gene variants in Pakistani patients with mucopolysaccharidosis type 1. Ann Pediatr Endocrinol Metab 32(11):1221-1227.

Zelei, T., Csetneki, K., Voko, Z., & Siffel, C. (2018). Epidemiology of Sanfilippo syndrome: results of a systematic literature review. Orphanet journal of rare diseases, 13(1), 1-11.

Zelickson, A. S., Windhorst, D. B., White, J. G., & Good, R. A. (1967). The Chediak-Higashi syndrome: formation of giant melanosomes and the basis of hypopigmentation. J Invest Dermatol, 49(6), 575-581.

Zhao HG, Li HH, Bach G, Schmidtchen A, Neufeld EF (1996) The molecular basis of Sanfilippo syndrome type B. Proc Natl Acad Sci U S A 93: 6101–6105.

Zhao, H. G., Aronovich, E. L., & Whitley, C. B. (1998). Genotype-phenotype correspondence in Sanfilippo syndrome type B. The American Journal of Human Genetics, 62(1), 53-63.

Zhao, J., Benlekbir, S., & Rubinstein, J. L. (2015). Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. Nature, 521(7551), 241-245.

Zimran, A., Szer, J., & Revel-Vilk, S. (2020). Impact of Gaucher disease on COVID-19. Internal Medicine Journal, 50(7), 894.

Zubaida, B., Batool, H., Cheema, H. A., Waheed, N., & Naeem, M. (2019). Novel IDS Variants Identified in Three Unrelated Pakistani Patients Affected with Mucopolysaccharidosis Type II (Hunter Syndrome). Human Heredity, 84(6), 279-286.

Zubaida, B., Hashmi, M. A., Cheema, H. A., & Naeem, M. (2018). Identification of a novel GLB1 mutation in a consanguineous Pakistani family affected by rare infantile GM1 gangliosidosis. Journal of genetics, 97(5), 1445-1449.

ANNEXURE 1

1.1. Consent form:

مائی عزویہ کے موروثی امراض کی تحقیق میں شمولیت کی رضامندی

ادارہ: جامعہ قائداعظم

تعارف

مائی عزویہ کے بہت سے پیچیدہ امراض موروثی ھیں جن کے بارے میں معلومات بہت کم ھیں اور ان امراض پر مکمل طور پر قابو نہیں پایا جاسکتا۔ ہم قائداعظم یونیورسٹی میں ان امراض پر تحقیق کر رھے ھیں۔ جس میں آپ کی شمولیت انتہائی اھمیت کی حامل ھے۔

زیادہ تر یہ امراض پیدائشی ھوتے ھیں۔ ان کی تشخیص کا طریقہ کار پیچیدہ ہے۔ بروقت تشخیص سے مرض کو شدت اختیار کرنے سے روکا جاسکتا ہے۔ تحقیق سے تشخیص کے آسان، کم خرچ اور معیاری طریقہ کار وضع کیے جاسکتے ہیں۔ علاوہ ازیں آئندہ آنے والی نسلوں کو ان مسائل سے بچایا جاسکتا ہے۔

آپ کی رضامندی سے تحقیقاتی عملہ آپ کے اور آپ کے خاندان کے دوسرے ارکان کے طبی مسائل کے بارے میں معلومات حاصل کرے گا اور بوجہ ضرورت آپ کی طبی تفصیلات / رپورٹ بھی طلب کے جائیں گی اور تقریباً ۵ سے ۱۰ ملی لیٹر خون بطور نمونہ بھی لیا جائے گا۔ عملہ آپ کی اور آپ کے خاندان کی اہم معلومات صیغہ راز میں رکھنے کا ذمہ دار ہے۔

اجازت نامہ

میں یہ تعارف پڑھنے اور تحقیق کار سے سوالات پوچھنے کے بعد اپنی مرضی سے اس تحقیق کا حصہ بننا چاہتا / چاہتی ہوں۔

تحقیق میں شامل نابالغ کے والدین / سرپرست کے دستخط

تحقیق میں شامل بالغ مریض کے دستخط

تحقیق کار کے دستخط

1.2. Identification form:

جامعہ قائداعظم، اسلام آباد

بیماری:_____

نام: _____ والد/شوہر کا نام :_____

عمر:_____ قوم /ذات : _____

پیشہ:_____بیماری کی تشخیص کے وقت مریض کی عمر:_____

مریض کا ہسپتال کا رجسٹر نمبر:_____

مکمل پتہ:_____

گھر کے قریب کوئی مشہور شخصیت / نشانی:_____

گھر کے سربراہ کے فون نمبرز: موبائل نمبر:_____پی ٹی سی ایل نمبر:_____

کیا مریض کے والدین کی شادی خاندان میں ہوئی ہے؟ ہاں نہیں

خاندان میں متاثرہ افراد:_____

_____

1.3. Questionnaire

Family ID: **Hospital:**

Quaid-I-Azam University, Islamabad

Patients hospital ID: **Type:**

Patient name: **S/D/o:**

Blood group: **Age:**

Parental consanguinity: Yes or No

Other diseases:

PRESENTING COMPLAINTS:

1. Delay in achieving milestone
2. Abdominal distention
3. Aggressive behavior
4. Speech delay
5. Hernia
6. Gibbus formation
7. Short stature
8. Respiratory/ear infection
9. Eye dysfunction/corneal clouding

EXAMINATION:

1. Height
2. Weight
3. Head circumference
4. Facial features

LOCOMOTOR EXAM:

1. Gait
2. Joint contractures
3. Trunk
4. Spine (scoliosis/kyphosis)

GIT

1. Hepatosplenomegaly

NEUROLOGIC EXAM:

1. I.Q
2. Intellectual disability

LABORATORIES:

1. CBC ē PF
2. LFT's
3. RFT's
4. ECHO
5. MRI
6. Radiological findings
7. Enzyme analysis

Signature

1.4. Bioethical committee approval (Quaid-I-Azam University, Islamabad, Pakistan)

قائداعظم یونیورسٹی
QUAID-I-AZAM UNIVERSITY
Faculty of Biological Sciences
Bioethics Committee

No. #BEC-FBS-QAU2019-198 Dated: 06-12-19

Dr. Sabika Firasat,
Department of Animal Sciences,
Faculty of Biological Sciences,
Quaid-i-Azam University, Islamabad
45320, Pakistan

Subject:- **"Molecular analysis of lysosomal storage disorders in Pakistani population."**

Dear Dr. Sabika Firasat,

We wish to inform you that your subject research study has been reviewed and is hereby granted approval for implementation by Bio-Ethical Committee (BEC) of Quaid-i-Azam University, Your study has been assigned protocol **#BEC-FBS-QAU2019-198.**

While the study is in progress, please inform us of any adverse events or new, relevant information about risks associated with the research. In case changes have to be made to the study procedure, the informed consent from and or informed consent process, the BEC must review and approve any of these changes prior to implementation.

Sincerely,

Sarwat Jahan

Prof. Dr. Sarwat Jahan
Department of Animal Sciences

cc:
Dean, F.B.S

ANNEXURE 2

2.1. Award letter International Research Support Initiative Program (IRSIP), Higher Education Commission (HEC), Pakistan

Higher Education Commission
H-9, Islamabad (Pakistan). Phone: (0092-51)90408036, Fax: (0092-51)90808038, E-mail: jkhan@hec.gov.pk
Project Director

No: 1-8/HEC/HRD/2020/10579
PIN: IRSIP 45 BMS 44
Dated: July 14th, 2020

Ms. Rutaba Gul
House no. 5, Khyber Block, Model Town, Haripur

Subject: **Award of scholarship under "International Research Support Initiative Program"**

Dear, Ms. Gul,

I am delighted to inform you upon your selection as a recipient of a grant under the International Research Support Initiative Program (IRSIP) for University of Copenhagen, Denmark. Details of the award, along with terms and conditions are as follows:

a. Return air travel (Economy Class) at actual and up to US$ 1300/- (Thirteen Hundred US Dollar) whichever is less.

b. Stipend/living allowance EUR€ 900/- per month for a maximum period of six months.

c. EUR€ 2550/- as bench fee to be paid to foreign university. (Payment is subject to the provision of fee invoice from Foreign University).

2. All other expenses are to be covered by you.

3 You will be required to submit legal bonds, undertaking and guarantee to HEC to return and complete PhD studies at parent University/Institution.

4. You are required to initiate visa process immediately and the offer will stand withdrawn without further notice in case you do not join foreign university by December 31st, 2020.

5. Before return to Pakistan, you will be required to submit us a progress report on your visit on prescribed form duly endorsed by your foreign supervisor.

6. Any other terms/conditions and TORs of the program, m issued by HEC from time to time will be binding for all purposes.

Yours Sincerely,

Jehanzeb Khan
Project Director

Cc:

1. Prof. Dr. Torben Hansen, Maersk Tower, Building: 07-8-26, Blegdamsvej 3B, 2200 Kobenhavn N, University of Copenhagen, Denmark.
2. Dr. Sabika Firasat, Dept. of Animal Sciences, Quaid-i-Azam University, Islamabad.
3. Head of Department. Dept. of Animal Sciences, Quaid-i-Azam University, Islamabad.
4. The Honorable Ambassador,Embassy of Pakistan, Valeursvej 17, 2900 Hellerup, Denmark.
5. Personal File

2.2. Acceptance letter Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark

UNIVERSITY OF COPENHAGEN
FACULTY OF HEALTH AND MEDICAL SCIENCES

To whom it may concern

AUGUGST 13, 2019

THE NOVO NORDISK FOUNDATION
CENTER FOR BASIC METABOLIC
RESEARCH

MAERSK TOWER
BLEGDAMSVEJ 3B
DK-2200 COPENHAGEN N

Research placement Ms. Rutaba Gul

To whom it may concern,

I hereby invite Ms. Rutaba Gul to carry out a six month research placement within our laboratory to conduct research on "**Molecular Analysis of Lysosomal Storage Disorders in Pakistani Population**".

During her visit Ms. Rutaba Gul will perform the whole genome sequencing of all collected samples from Pakistani families. Also the data will be analyzed and in-silico studies will be performed with collaboration of bioinformatics department of the University of Copenhagen.

Our laboratory is equipped with all the facilities required for this project. The bench fee of 2550 euro will be a requirement for Ms. Rutaba visit. If any further information is required, please do not hesitate to contact me.

Yours sincerely

Torben Hansen, MD, PhD
Professor of Metabolic Genetics
Novo Nordisk Foundation Center for Basic Metabolic Research
Human Genomics and Metagenomics in Metabolism
Faculty of Health and Medical Sciences
University of Copenhagen

Received: 12 June 2019 | Revised: 25 July 2019 | Accepted: 31 July 2019

DOI: 10.1111/cga.12354

NOTE

Congenital Anomalies | WILEY

IDUA gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families

Rutaba Gul[1] | Sabika Firasat[1] | Mulazim Hussain[2] | Kiran Afshan[1] | Dil Nawaz[1]

[1]Department of Animal Sciences, Faculty of Biological Sciences, Quaid-i-Azam University, Islambad, Pakistan

[2]The Children Hospital, Pakistan Institute of Medical Sciences, Islamabad, Pakistan

Correspondence
Sabika Firasat, Department of Animal Sciences, Faculty of Biological Sciences, Quaid-i-Azam University, University road, Islamabad 45320, Pakistan.
Email: sabika.firasat@qau.edu.pk

Funding information
Scientific Instrumentation Program, Higher Education Commission, Islamabad, Pakistan

Placeholder Text

Mucopolysaccharidosis type-I (MPS-1) is an autosomal recessive lysosomal storage0020disorder caused by mutations in the α-L-iduronidase (*IDUA*) gene (OMIM 252800).[1] IDUA (E.C.3.2.1.76) catalyzes degradation of glycosaminoglycans heparin sulphate and dermatan sulphate which otherwise accumulate in the lumen of lysosomes leading to MPS-I phenotypes. Based on disease severity, there are three clinical subtypes of MSP-1 named as Hurler syndrome (MPS I-H) (OMIM 607014) the most severe form,

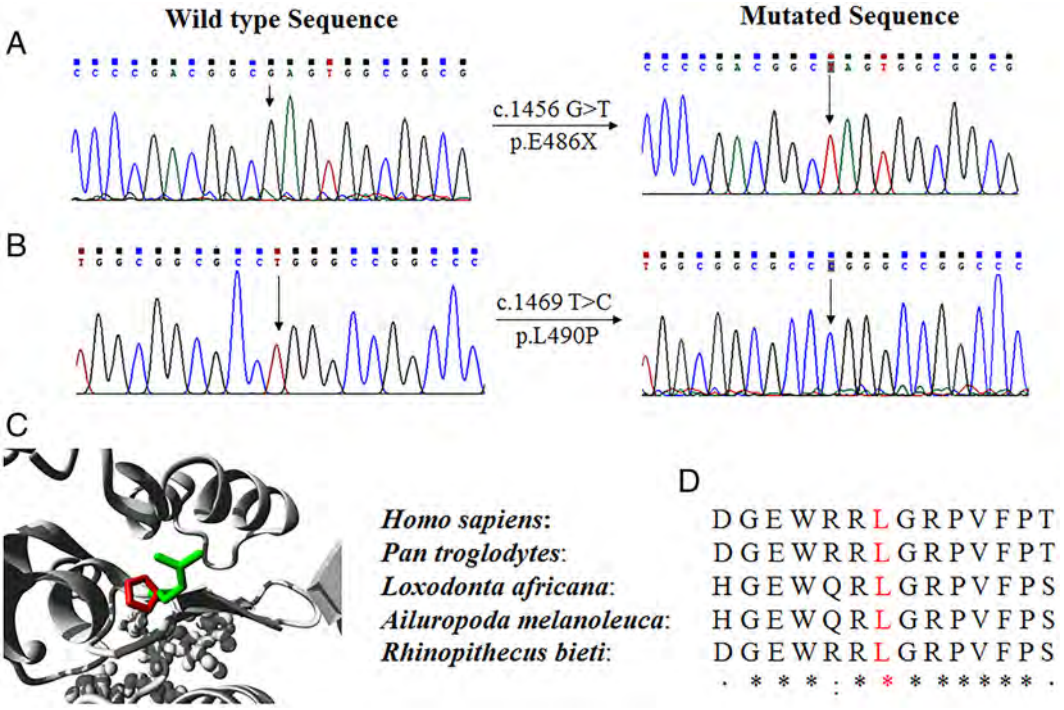


FIGURE 1 A, DNA sequence chromatogram showing novel variant c.1456G>T causing a p.(gLU486*) in affected individual (right) and unaffected control (left). B, DNA sequence chromatogram showing previously reported missense variant c.1469T>C causing a p.(Leu490Pro) in affected individual (right) and unaffected control (left). C, In-silico prediction of p.(Leu490Pro) as generated by HOPE. Wild type amino acid residue is colored as green while mutant is colored as red. D, Clustal omega multiple sequence alignment results for amino acid leucine at position 490 of IDUA protein showing its 100% conservation among mammals (indicated by an asterisk and shown in red)

 © 2019 Japanese Teratology Society

Hurler-Scheie syndrome (MPS I H/S) (OMIM 607015) intermediate form and Scheie syndrome (MPS I-S) (OMIM 607016) being milder of all.[2] Thus identification of the causative mutation for each patient is important to develop genotype-phenotype correlation. Here we identified a novel truncating *IDUA* mutation and a previously reported missense mutation each in a Hurler and a Hurler-Scheie patient, respectively.

We recruited two inbred Pakistani families that is, Family-A and Family-B (Figure S1, Supporting Information), each having one alive MPS-1 patient. Dried blood spot tandem mass spectrometry confirmed MPS-1 diagnosis with enzyme values lower than the cut-off values (>0.7). *IDUA* gene was screened to find out the causative mutation underlying to enzyme defect. All research was conducted according to principles of world medical association of Helsinki[3] and was approved from the Bioethical Committee of institution. Blood samples were collected after signed informed consents from patient's families. DNA was extracted from the blood samples as detailed previously.[4] AMplification of *IDUA* exons was done through polymerase chain reaction as described by Afzal et al, 2018.[5] Sanger sequencing was performed to sequence amplified products and data was analyzed by Sequencher software (version 5.4.6).

In family A, we found a novel variant NM_000203.5 (IDUA_v001): c.1456G>T in exon 10 of *IDUA* gene, leading to a premature termination codon that is, p. (Glu486*) resulting in a truncated protein of 485 AMino acids rather than 653 AMino acids (NP_000194.2) (Figure 1A). Patient of family A aged 16 months and had severe phenotype with presenting features including asthma, sleep apnea, joint contractures, kyphosis, ear/nose/throat infections, coarse facial features, hepatosplenomegaly, inguinal hernia, aggressive, heavy eye brows, large tongue, thick lips, and rough skin. In family B, a previously reported missense mutation that is, NM_000203.5 (IDUA_v001):c.1469T>C (rs121965027) linked to Hurler-Scheie phenotype was detected in exon 10 causing substitution of leucine with proline at position 490 (NP_000194.2) (Figure 1B).[6] In-silico predictions of NM_000203.5 (IDUA_v001):c.1469T>C mutation on protein structure through HOPE, is shown in Figure 1C. The p. (Leu490Pro) substitution was predicted to be pathogenic by Mutation taster[7] and Sorting Intolerant from Tolerant (SIFT).[8] Clustal omega results showed conservation of leucine residue at position 490 (Figure 1D).

In conclusion, we found genotype-phenotype correlation in two analyzed cases of MPS-1 which will help affected families in clinical management and future reproductive planning.

ACKNOWLEDGMENTS

We thank patients and their families for their participation. This study was financially supported (in part) by access to Scientific Instrumentation Program, Higher Education Commission, Islamabad, Pakistan.

DISCLOSURE OF INTEREST

None.

ORCID

Sabika Firasat https://orcid.org/0000-0003-4959-7832

REFERENCES

1. Neufeld EFMJ. The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic & Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:3421-3452.
2. De RMH, Teunissen QG, der Lee V, et al. Capturing phenotypic heterogeneity in MPS I: results of an international consensus procedure. *Orphanet J Rare Dis*. 2012;7:22.
3. Association WM. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194.
4. Grimberg J, Nawoschik S, Belluscio L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Res*. 1989;17(20):8390.
5. Afzal R, Firasat S, Kaul H, et al. Mutational analysis of the CYP1B1 gene in Pakistani primary congenital glaucoma patients: identification of four known and a novel causative variant at the 3′ splice acceptor site of intron 2. *Congenit Anom*. 2018;1-10. https://doi.org/10.1111/cga.12312.
6. Tieu PT, Bach G, Matynia A, Hwang M, Neufeld EF. Four novel mutations underlying mild or intermediate forms of α-L-iduronidase deficiency (MPS IS and MPS IH/S). *Hum Mutat*. 1995;6(1):55-59.
7. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7(8):575-576.
8. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812-3814.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Gul R, Firasat S, Hussain M, Afshan K, Nawaz D. *IDUA* gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families. *Congenit Anom*. 2019;1–2. https://doi.org/10.1111/cga.12354

UDC 575.
https://doi.org/10.2298/GENSR2103017G
Original scientific article

NEUROLOGICAL MANIFESTATIONS IN PAKISTANI LYSOSOMAL STORAGE DISORDERS PATIENTS AND MOLECULAR CHARACTERIZATION OF GAUCHER DISEASE

Rutaba GUL[1], Sabika FIRASAT[1]*, Mulazim HUSSAIN[2], Muhammad TUFAIL[1], Waheed AHMAD[1], Kiran AFSHAN[1]

[1]Department of Zoology, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan
[2] The Children Hospital, Pakistan Institute of Medical Sciences, Islamabad, Pakistan

Gul R., S. Firasat, M. Hussain, M. Tufail, W. Ahmad, K. Afshan (2021). *Neurological manifestations in Pakistani lysosomal storage disorders patients and molecular characterization of Gaucher disease*. - Genetika, Vol 53, No.3, 1017-1029.

Lysosomal storage disorders (LSDs) are a large group of inborn errors of metabolism each caused by genetic mutations of a particular lysosomal protein encoding gene. These inherited conditions are characterized by lysosomal dysfunction with wide variety of organ impact sometimes organ failure with growing age. Neurological complications in LSD cases range from severe neurodegenerations in 70% cases to mild symptoms or absence of neuropathy in others. Each LSD is monogenic but heterogeneous from a molecular standpoint with a large number of mutations described in the respective gene. Some mutations are particular to specific populations, reflecting consequences of founder effect. Present study aimed to access the demographic and clinical profiles of forty-five LSD affected families enrolled during January 2018 to December 2019 at local hospitals to find out neurological symptoms in Pakistani LSD cases. Furthermore, molecular genetic analysis of Gaucher's disease affected families was performed to unveil underlying disease causing mutation/s. Neurological manifestations were present in twenty-eight families including eleven Mucopolysaccharidosis-1 (MPS-I), four Gaucher's disease (GD) and all MPS-II, MPS-III, Niemann-Pick, Griscelli and Chediak-Higashi cases. Neurological involvement was not found in eight MPS-I, one GD, all MPS-IV and Pycnodysostosis affected families. Screening of *GBA* gene in GD families revealed a

Corresponding author: Sabika Firasat, Department of Zoology, Faculty of Biological Sciences, Quaid-i-Azam University, University road, 45320, Islamabad Pakistan, Telephone: 00925190643209, E-mail: sabika.firasat@qau.edu.pk.

.

reported missense mutation p.L483P in all analyzed families. Clinical heterogeneity of MPS-1 and GD is evident from literature however mutational analysis of all enrolled GD families depicted segregation of a reported missense variant p.L483P of *GBA* gene with disease phenotype in all families. Our findings highlight importance of homeostatic role of lysosomes in neuronal development as twenty eight out of forty families had neurological manifestations. Furthermore, identification of same mutation in GD patients with or without neuronal involvement may be related to some unknown differences in the expression of genetic modifiers or exposure to environmental triggers.

Key words: Lysosomal Storage Disorders, Gaucher Disease, Neurological manifestation, *GBA* gene, Pakistan

INTRODUCTION

All functional cells of body have central organelle for degradation i.e lysosome. Functional lysosomes are important for degradation of cellular autophagic contents to regulate the quality of cytoplasm by eliminating cellular macromolecular aggregates (SERANOVA *et al.,* 2017). The substrate for the lysosome is obtained through multiple pathways like phagocytosis, autophagy and endocytosis. These substrates are degraded in lysosomes via lysosomal hydrolases thus causing disposal and recycling of cellular waste (BALLABIO *et al.*, 2020). If this homeostatic role of lysosomes is disturbed due to compromised or absent lysosomal enzyme/s it may lead to inappropriate storage of material interfering with normal cell function and causing multisystem disorders including lysosomal storage disorders (LSDs), neurodegenerative diseases and cancer (PLATT *et al.*, 2012). LSDs are a large group of inherited disorders causing multiple organ failure. Furthermore, about two third of LSD's have central nervous system involvement (PARÁ *et al.*, 2020). Most of these disorders are caused by recessively inherited Mendelian mutations in genes encoding lysosomal proteins (MARQUES *et al.*, 2019; PARÁ *et al.*, 2020). Frequency of all LSDs in different population is 1:7500 in live births (MEIKLE *et al.*, 1999; POUPETOVA *et al.*, 2010). However, exact prevalence in some populations may even be higher because many cases remain undiagnosed (PLOTEGHER *et al.*, 2017). In addition, in societies like Pakistan, due to high ratio of consanguinity, recessively inherited disorders are more common than in any randomly mating populations (PRADHAN *et al.*, 2011; SHAHID *et al.*, 2020).

LSDs related damages are progressive therefore patient management requires early precise diagnosis of disease subtype (ZHOU *et al.*, 2011) for therapeutic management, which is a serious problem. Widely used therapeutic option is enzyme replacement therapy (ERT) (AFROZE *et al.*, 2017). ERT limitations include high cost, lifelong treatment requirement, inability to stop progression of neuronal symptoms and possibility of immune response against an exogenous enzyme (PARENTI *et al.*, 2015; AFROZE *et al.*, 2017). Recently, pharmacological chaperon (PC) therapies, are being developed to use chaperones which increase the stability of mutated protein inside the cell by avoiding endoplasmic reticulum (ER) associated degradation. Interestingly, PCs are capable of crossing the blood-brain-barrier, thus showing promising results in stabilizing the mutant enzymes which otherwise get degraded by ER. However, PCs also showed some limitations due to type of mutations affecting the enzyme structure (PARENTI *et al.*, 2015; MOHAMED *et al.*, 2017). Other treatment options including bone marrow transplant or human

stem cell transplant are only possible after early precise diagnosis within the first few months of life. Thus, molecular genetics diagnosis of LSDs along with enzyme analysis for accurate early detection and patient management is required (MITTAL, 2015; MOKHTARIYE *et al.*, 2019).

LSDs are considered a major reason of neurodegeneration in childhood (VERITY *et al.*, 2010). LSDs including some forms of Gaucher disease, Niemann-Pick disease C, neuronal lipofuscinosis and mucopolysaccharidosis manifest mild to severe neurodegenerative symptoms (MEIKLE *et al.*, 1999; SHETH *et al.*, 2004; MARQUES *et al.*, 2019). Due to continuous spectrum of phenotypes among LSDs severe cases are presented at infancy or early childhood while mild cases are being observed at adult age (VERITY *et al.*, 2010; MARQUES *et al.*, 2019). This study is the first report aimed to access neurological manifestations in Pakistani LSDs affected cases.

MATERIALS AND METHODS

Approval of the study was taken from bioethical committee of Quaid-i-Azam University Islamabad, Pakistan. A total of forty five unrelated LSD affected Pakistani families were enrolled through local hospitals during January 2018 to December 2019 from local hospitals. All diagnosis were made on the basis of clinical and biochemical analysis by expert pediatricians. For diagnosis of each LSD subtype enzyme analysis by dry blood spot, radiological findings and blood complete picture was performed. For each enrolled family, detailed family/medical history, pedigree drawing and blood samples of available affected and unaffected family members were collected after informed consent in accordance with the tenets of the Declaration of Helsinki (ASSOCIATION, 2013). Demographic and clinical profile of proband of each Gaucher disease family is listed in table 1 (Table 1). Five Gaucher disease (GD) affected families (Family A-E) (Figure1; Table 2) were selected for mutational analysis of *GBA* gene based on definitive diagnosis and availability of blood samples of affected and unaffected family members. All GD patients were born to consanguineous couples (Table 1). The age of diagnosis was between 2 to 14 months (Table 1). Probands of family A and C had one affected deceased sibling each at the age of 1.5 years and 7 years, respectively.

Table 1. Demographic and clinical details of proband of each enrolled Gaucher Disease family

Family ID	Ethnicity	Family history/ No. of affected	Parental Consanguinity	Neuronal manifestations	Other clinical features
Family A	Punjabi	Yes/1	Yes	No	Chronic diarrhea, AD, HSM, anemia
Family B	KPK	Nil	Yes	Yes	Breathing problem, AD, HSM, weak muscles
Family C	Punjabi	Yes/1	Yes	Yes	AD, HSM, DD
Family D	Punjabi	Nil	Yes	Yes	DD, AD, HSM
Family E	KPK	Nil	Yes	Yes	AD, HSM, weak muscles

KPK = Khyber Pakhtunkhwa, Abdominal Distension=AD, hepatosplenomegaly=HSM, Developmental Delay=DD

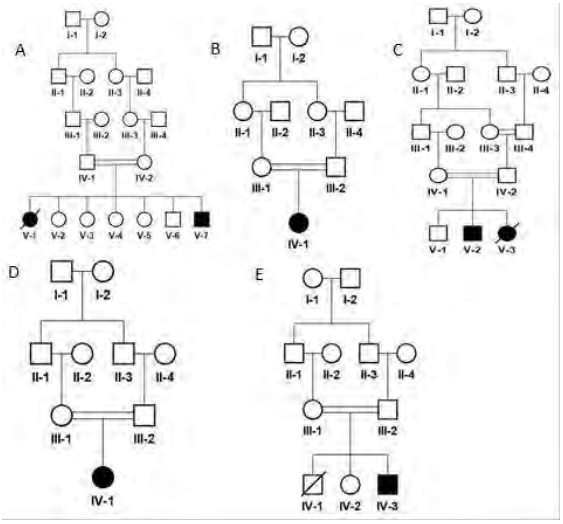


Figure 1. Pedigrees of Gaucher disease (GD) families (Family A-E) consistent with autosomal recessive inheritance pattern

Squares indicate males while circles indicate females. Filled and hollow symbols denote GD patients and normal individuals respectively. A double marriage line represents consanguinity.

Table 2. Clinical, biochemical and molecular genetics diagnosis details of proband of each Gaucher disease family.

Clinical data		Family A	Family B	Family C	Family D	Family E
Gender		M	F	M	M	M
Age at diagnosis/ age at examination		± 1 year/ 6.5 years	3 months/ 13 months	4 years/ 12 years	6 months/ 1 year	6 months/ 2 years
Treatment		ERT	N/A	ERT	ERT	ERT
Splenectomy		Yes	N/A	N/A	N/A	N/A
GCase activity		0.63 nmol/ml/h	N/A	0.2 µmol/L/h	0.5 µmol/L/h	0.1 µmol/L/h
RBCs		4.55 million/µL	3.22×10^{-6}/µL	5.62 10^{6}/uL	4.33 10^{6}/uL	3.87 million/µL
ALT		22 U/L	52U/L	20 U/L	16 U/L	N/A
WBCs		N/A	16.6×10^{-3}/µL	9.1 10^{3}/uL	4.8 10^{3}/uL	7000/mm^{3}
Age at diagnosis		5 months	3 months	6 months	14 months	3 months
Hemoglobin		11.9 g/dL	7.4g/dL	10.5 g/dL	13.4 g/dL	8.2g/dL
Platelets		220×10^{-3}/µL	16×10^{-3}/µL	531 10^{3}/uL	105 10^{3}/uL	154×10^{-3}/µL
Genotype	DNA mutation	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C	c.1448T>C/ c.1448T>C
	Protein change	p.L483P/ p.L483P	p.L483P/ p.L483P	p.L483P/ p.L483P	p.L483P/ p.L483P	p.L483P/ p.L483P

M: Male, F: Female, N/A: Not available, ERT: Enzyme Replacement Therapy

Mutational analysis of GD families

Blood samples of enrolled affected and unaffected members of selected GD families were stored in 10 mL vacutainer tubes (BD vacutainer K2 EDTA 18 mg). Genomic DNA was extracted by a modified method as described by KAUL *et al.* (2010). Quantity and purity of extracted DNA was accessed using a μDrop Plate reader (Multiskan^{TM}, Thermo Fisher Scientific, and Waltham, MA, USA).The extracted DNA was stored at 4°C. Primers of *GBA* gene were designed by using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/) and are listed in table 3 (Table 3). Primers were further confirmed for binding and self ligation through BLAT and Primer stat respectively (https://www.bioinformatics.org/sms2/pcr_primer_stats.html) (https://genome.ucsc.edu/cgi-bin/hgBlat?command=start). Samples were amplified by polymerase chain reaction (PCR) using T100 thermal cycler (Bio-Rad, CA, USA) with a cycling program of 95°C for 5 min, followed by 10 cycles of 95°C for 45s, 69°C- 64°C (according to melting temperature of each primer pair) for 45s with an increment of -1°C in each subsequent cycle, 72°C for 45s again followed by 30 cycles of 95°C for 45s, 59°C- 54°C (according to melting temperature of each primer pair) for 45s, 72°C for 45s and a final extension at 72°C for 10 min followed by a final hold at 25°C. The amplified PCR products were loaded on the 1.5% agarose gel along with 1kb size ladder to evaluate product size and purified by using DNA purification Kit (*Wiz Bio* Solutions, Seongnam, Korea). The sequencing reaction was performed using Big Dye Terminator Ready reaction mix (Applied Biosystems) following manufacturer instructions. Sanger's sequencing was done on an automated ABI 3730 genetic analyzer by Macrogen, Korea.

The sequenced data was analyzed by using BioEdit (version 7.2). The variants were confirmed by using Mutation taster software (http://www.mutationtaster.org/). Further the effect of mutation on the structure and function of protein was studied by using a bioinformatics tool HOPE (https://www3.cmbi.umcn.nl/hope/input/).

Table 3. Primers used for PCR amplification of GBA gene in Gaucher disease affected families

S. No.	Primer ID	3'-5' sequence	GC contents	Melting temperature	Product size
1	1,2 F	GGGAAGCCGGAATTACTTG	52.63	59.51	850 bp
2	1,2 R	GAGCCAAAATTGCACCACT	47.37	58.71	
3	3,4 F	ACCGTGTTCAGTCTCTCCTAGC	54.5	62.9	652 bp
4	3,4 R	GACAGAATGGGCAGAGTGAG	55	59.9	
5	5,6 F	AGGAGCCCAAGTTCCCTTT	52.63	60	806 bp
6	5,6 R	CTGATGGAGTGGGCAAGATT	50	60	
7	7 F	AGGCTGTTCTCGAACTCCTG	55	59.6	496 bp
8	7 R	GGGAATGGTGCTCTAGGAATC	52.38	59.92	
9	8 F	GTT GCA TTC TTC CCG TCA CC	55	62.8	367 bp
10	8 R	CTG GAC AGG AAG GGC TTC TG	60	62.3	
11	9 F	CTCTCCCACATGTGACCCTTA	52.3	59.9	399 bp
12	9 R	GCCTCCATGGTGCAAAAGGGG	52.9	59.1	
13	10,11 F	GCAGAAAAGCAGGGTCAGTG	55	60.98	575 bp
14	10,11 R	TGCTGTGCCCTCTTTAGTCA	50	59.59	

RESULTS

A total of forty-five families suffering from Lysosomal Storage Disorders (LSDs) were recruited during January, 2018 to December, 2019 through expert pediatrics from local hospitals. Among enrolled families, affected individuals of twenty eight 62.2% (28/45) families had neurological involvement. Majority of the patients i.e., 68.8% (31/45) were from Punjab province of Pakistan whereas 22.2% (10/45) and 8.8% (4/45) belonged to Khyber Pakhtunkhwa and Azad Jammu and Kashmir (AJK) respectively. All patients belonged to inbred families and in forty three families affected cases were products of marriage between first cousins. Average age of proband of each enrolled family was 11.04 ± 8.10 years. Majority of families were affected with Mucopolysaccharidosis (35/45) followed by Gaucher disease (5/45), Niemann-Pick disease (2/45), Chédiak-Higashi syndrome (1/45), Pycnodysostosis (1/45) and Griscelli disease (1/45). All collected patients of same subtype had comparable phenotypes e.g., all Gaucher disease cases had abdominal distension and hepatosplenomegalies (Table 1 and 2).

Gaucher's disease (GD) cases were selected for mutational analysis of *GBA* gene due to: a) limited financial resources for sequencing, b) clinical confirmation of disease for selection of gene and c) availability of blood samples of affected and unaffected family members. Out of five GD probands, four were males and one was female. In family A, there is a five generations pedigree with consanguinity and two affected individuals (V-1, V-7), among them one is deceased at the age of one and half year. While second affected is of 7 years and is still alive. The neurological degeneration is not present in this subject. The dry blood spot (DBS) measurement for enzyme concentration in index case i.e., V-1 was found to be lower than cut off value (Table 2) i.e 2.3-14.1 nmol/mL/h. Also chitotriosidase level is elevated (285.99 nmol/h/mL, normal: <150.0) which is in alignment with the Gaucher Disease. The bone marrow aspiration report revealed hyperplastic erythropoiesis and myelopoiesis. Increased atypical histiocytes with fibrillary cytoplasm and eccentric nucleous were observed. Splenectomy explains abundant eosinophilic cytoplasm having a wrinkled tissue paper appearance. Patient also has chronic diarrhea in common, abdominal distention and hepatospleomegaly. In family B, there is one affected individual with abdominal distention, severe breathing problem and require continuous oxygen supply, jaundice, weak muscles, swelling in feet, inguinal hernia, repeat episodes of shocks, macrocephaly, hepatosplenomegaly with gross ascites, gum hypertrophy, respiratory dyspnea, wheezing, localized paralysis, seizures and neurological features. The patient also reported severe cough attacks and intermittent high fever. In family C and D both patients were having hepatosplenomegaly, abdominal distention, delay in achieving milestones, neuronal dysfunction, respiratory distress and muscle weakness. In family E, enzyme analysis in DBS is below the cut off value (> 1.5 μmol/L/h), indicating the GD. The patient has abdominal distention, neuronal dysfunction, hepatosplenomegaly, inguinal hernia and delayed milestones.

Mutational analysis of Gaucher disease families

Analysis of sequencing results revealed a previously reported missense variant p.L483P (c.1448T>C) in exon 10 (rs421016) (TSUJI *et al.*, 1987) of *GBA* gene in all five GD affected families segregating with disease phenotype. All the affected cases were homozygous for mutated allele whereas their parents were heterozygous carriers of mutation. In addition, we

performed a control sample sequencing of exon 10 of *GBA* gene in order to confirm the normal homozygous state (Figure 2B).

Here we used in-silico analysis tool i.e., Have Our Protein Explained (HOPE) for the first time to check impact of p.L483P substitution, which highlighted that the mutated residue is located in a domain that is important for the activity of the protein and in contact with residues in another domain. It is possible that this interaction is important for the correct function of the protein. The mutation can affect this interaction and subsequently normal protein function. The wild-type and mutant amino acids differ in size. The mutant residue is smaller than the wild-type residue. The mutation was predicted to cause an empty space in the core of the protein (Figure 2D).

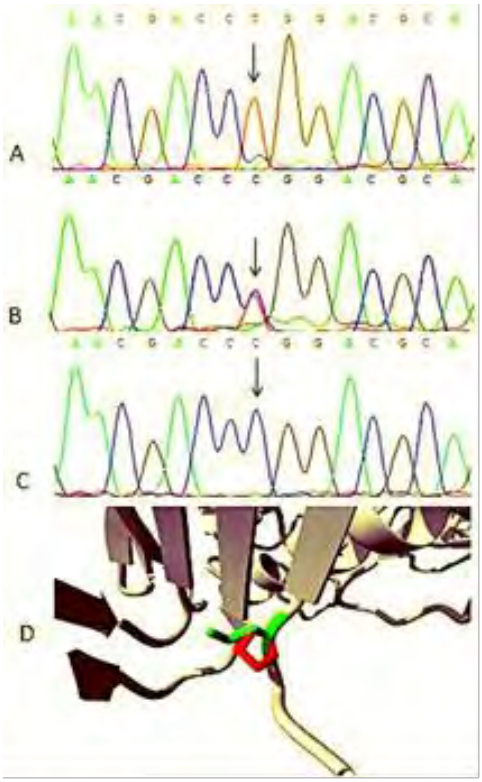


Figure 2. A: Sequence chromatogram of normal allelic sequence. B: Sequence chromatogram of heterozygous carrier for c.1448T>C. C: Sequence chromatogram of homozygous mutant for c.1448T>C in affected individuals of Gaucher affected families. D: Structural effect of the amino acid substitution as shown by HOPE analysis. Mutated amino acid is shown in red color while wild type amino acid is shown in green color.

DISCUSSION

Genetic defect/s leading to LSDs predisposes to progressive accumulation of material with growing age and show many devastating phenotypes including neurodegenerative diseases (PLATT *et al.*, 2012; SERANOVA *et al.*, 2017). Even most of the clinical phenotypes appear in later stages of life, but by then those damages are irreversible (COX *et al.*, 2012). Approximately 45% cases of neurodegenerations in the world are due to LSDs (VERITY *et al.*, 2010). However, central nervous system defects are reported in more than 70% cases of LSDs and rest ±30% cases may have no nervous involvement or have mild symptoms (JARDIM *et al.*, 2010; MARQUES *et al.*, 2019; SHETH *et al.*, 2019). Hence, understanding the pathophysiology of LSDs can be useful to unveil neurodegenerative mechanisms. Etiology of most LSDs predominantly involves recessively inherited Mendelian mutations. In societies like Pakistan, recessively inherited disorders are more common due to consanguinity (AFROZE *et al.*, 2017; GUL *et al.*, 2019; MARQUES *et al.*, 2019). Therefore present study was designed to study two aspects of enrolled LSD affected families from local population. Firstly, to identify neuronal manifestation in enrolled LSD subtypes and secondly to screen *GBA* gene in Gaucher's disease (GD) affected families.

Out of forty-five LSDs affected families enrolled during January, 2018 to December, 2019 twenty eight (62.2%) families had neurological involvement. All recruited cases of LSDs are novel and was not previously studied by any other group. Our data revealed that mucopolysaccharidosis (MPS) is the predominant LSD in our study with thirty five affected families segregating MPS-I, MPS-II, MPS-III and MPS IV phenotype in 19, 3, 6 and 7 enrolled families respectively . Our findings are consistent with Hutchesson et al., 1998 and Afroze et al., 2016 reports of increased frequency of MPS in Pakistani children (HUTCHESSON *et al.*, 1998; AFROZE *et al.*, 2016; AFROZE *et al.*, 2017); however data presented by CHEEMA *et al*. (2016) showed Gaucher disease as predominant LSD in cases presented at a tertiary care hospital Lahore, Pakistan (CHEEMA *et al.*, 2016). MPS-1 has further three subtypes including the most severe form Hurler (MPS I-H), Hurler-Scheie (MPS I-H/S) and the less severe Scheie (MPS I-S) (DE RU *et al.*, 2012). Previously, we reported a non-sense mutation (p.E486X) of IDUA gene in MPS-1H affected family displaying neuronopathy whereas another MPS1-H/S affected family was segregating a missense variant (p.L490P) with no neuronal involvement at early age (GUL *et al.*, 2019). Similarly, remaining seventeen unscreened MPS 1 families have different clinical manifestations but the underlying genetic mutations are yet to be clarified for exact molecular genetic subtyping of disease.

Unlike, MPS-1 all enrolled affected cases of MPS-II (Hunter syndrome) and MPS-III (Sanfillipo syndrome) displayed neuronal involvement. Multiple interviews of Sanfillipo affected families revealed that the neuronopathies got severe with age. Rest of the cases of Niemann-Pick, Griscelli and Chediak-Higashi disease also had severe neuronopathic symptoms. Previous reports suggest that among Gaucher disease (GD) cases, type 1 GD is most common which is non-neurological form of disease (MARQUES *et al.*, 2019; SHETH *et al.*, 2019). However, among our five GD affected families, affected members of all except Family A displayed neurological symptoms (Table 2). To unveil mutational spectrum of *GBA* in our enrolled GD families we performed Sanger's sequencing of coding exons of *GBA* gene. Data analysis showed a previously reported missense mutation p.L483P (rs421016), segregating with disease

phenotype in all five families. Recently, SHETH *et al*. (2019) identified p.L483P as the most common GD causing mutation in Indian population irrespective of ethnic groups (SHETH *et al.*, 2019) which is consistent with our findings. This mutation was also reported to be detected with 41% Japanese, 60% Thai and 2.84% Jewish GD cases (GRABOWSKI, 1997; ETO *et al.*, 1999; TAMMACHOTE *et al.*, 2013) but in this study we detected this mutation in 100% analyzed cases. Although p.L483P is reported to be associated with all subtypes of GD (KOPRIVICA *et al.*, 2000; TAMMACHOTE *et al.*, 2013; SHETH *et al.*, 2019) but KOPRIVICA *et al*. (2000) suggested that p.L483P is mainly associated with GD sub types showing neurological involvement which is also found in this study (KOPRIVICA *et al.*, 2000). Screening of *GBA* gene in our enrolled families and identification of same mutation in all of them suggest high prevalence of p.L483P allele in our population. Affected individual of one of our GD family (family A) segregating p.L483P has no neurological symptoms which is consistent with GD type 1, but the possibility of neuronal manifestations at the later stage of life could not be ruled out as reported by ALFONSO *et al*. (2007) Leucine at 483 position of GBA protein lies in β-barrel domain (residues 30–75 and 431–497) of Beta-Glucosidase which resembles an immunoglobulin fold (LIEBERMAN *et al.*, 2009). According to previous studies this substitution leads to alteration of secondary structure of protein also this variant lie close to the catalytic domain of enzyme thus results in loss of enzyme activity. This is supported by our HOPE data as shown in figure 2D, the mutated residue is located in a domain which is important for the activity of the protein and leucine 483 makes contact with residues in another domain (Figure 2D) further highlighting importance of this amino acid in wild type structure and function of GBA protein.

CONCLUSION

We concluded that presence of neuropathies in 62% of our enrolled LSD cases, involvement of lysosomal defects in age related neurodegenerative conditions, including Parkinson's and Alzheimer's (MARQUES *et al.*, 2019) highlights importance of studies on molecular basis of LSDs to unveil molecular networks and links between different autophagy associated neurodegenerative conditions. Additionally, identification of p.Leu483Pro in all of GD cases in our study, a high reported carrier frequency of this variant in Indian population (SHETH *et al.*, 2018) and reported association of carriers of this allele with an increased risk of Parkinson disease (PD) development (WANG *et al.*, 2012) necessitates screening of this mutation in all PD and GD cases from our population. As the recessive genes causing LSDs are predominantly common in consanguineous populations like Pakistan, and the Pakistani population is less explored for the genetic basis of these conditions, then further genetic studies are indispensible which may reveal novel variants for diagnostics and treatment regimes.

ACKNOWLEDGMENT

Permission has been obtained from patients and patients' families. We thank them for their participation in this study. We sincerely thank the staff of The Children Hospital, Pakistan Institute of Medical Sciences (PIMS), Islamabad, Pakistan and Lady Reading Hospital (LRH) Peshawar, Pakistan.

Received, September 20th, 2020
Accepted May 18th, 2021

REFERENCES

AFROZE, B. and N., BROWN (2017): Ethical issues in managing Lysosomal storage disorders in children in low and middle income countries. Pak. J. Med. Sci., *33*: 1036-1041.

AFROZE, B., L., LAKHANI, F., NAZ, S., SOMANI, Z.M., YUNUS, N., BROWN (2016): Challenges identified in the management of patients with inherited metabolic disorders–A five year experience from Pakistan. Egyptian J. Med. Human Gen., *17*: 259–264.

ALFONSO, P., S., AZNAREZ, M., GIRALT, M., POCOVI, P., GIRALDO, R. (2007): Mutation analysis and genotype/phenotype relationships of Gaucher disease patients in Spain. J. Human Gen., *52*: 391-396.

ASSOCIATION, W.M. (2013): World Medical association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA, *20*: 2191-2194.

BALLABIO, A. and J.S., BONIFACINO (2020): Lysosomes as dynamic regulators of cell and organismal homeostasis. Nat. Rev. Mol. Cell Biol., *21*: 101-118.

CHEEMA, H.A., H.S., MALIK, A., PARKASH, Z., FAYYAZ (2016): Spectrum of Inherited Metabolic Disorders in Pakistani Children Presenting at a Tertiary Care Centre. J. Coll Physicians Surg. Pak., *26*: 498-502.

COX, T.M. and M.B., CACHON-GONZALEZ (2012): The cellular pathology of lysosomal diseases. J. Pathol., *226*: 241-254.

DE RU, M.H., Q.G., TEUNISSEN, J.H., VAN DER LEE, M., BECK, O.A., BODAMER, L.A., CLARKE, C.E., HOLLAK, S.P., LIN, M.V., ROJAS, G.M., PASTORES, J.A., RAIMAN, M., SCARPA, E. P., TREACY, A., TYLKI-SZYMANSKA, J. E., WRAITH, J., ZEMAN, F.A., WIJBURG (2012): Capturing phenotypic heterogeneity in MPS I: results of an international consensus procedure. Orphanet J. Rare Diseases, *7*: 22.

ETO, Y. and H., IDA (1999): Clinical and molecular characteristics of Japanese Gaucher disease. Neurochemical Res., *24*: 207-211.

GRABOWSKI, G.A. (1997): Gaucher disease: gene frequencies and genotype/phenotype correlations. Genetic Testing, *1*: 5-12.

GUL, R., S., FIRASAT, M., HUSSAIN, K., AFSHAN, D., NAWAZ (2019): IDUA gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families. Congenital anomalies.

HUTCHESSON, A.C., S., BUNDEY, M.A., PREECE, S.K., HALL, A., GREEN (1998): A comparison of disease and gene frequencies of inborn errors of metabolism among different ethnic groups in the West Midlands, UK. J. Med. Gen., *35*: 366-370.

JARDIM, L.B., M.M., VILLANUEVA, C.F., DE SOUZA, C.B., NETTO (2010): Clinical aspects of neuropathic lysosomal storage disorders. J. Inherit. Metab. Dis., *33*: 315-329.

KAUL, H., S.A., RIAZUDDIN, M., SHAHID, S., KOUSAR, N.H., BUTT, A.U., ZAFAR, S.N., KHAN, T., HUSNAIN, J., AKRAM, J.F., HEJTMANCIK, S., RIAZUDDIN (2010): Autosomal recessive congenital cataract linked to EPHA2 in a consanguineous Pakistani family. Molecular Vision, *16*: 511-517.

KOPRIVICA, V., D.L., STONE, J.K., PARK, M., CALLAHAN, A., FRISCH, I.J., COHEN, N., TAYEBI, E., SIDRANSKY (2000): Analysis and classification of 304 mutant alleles in patients with type 1 and type 3 Gaucher disease. Am. J. Human Gen., *66*: 1777-1786.

LIEBERMAN, R.L., A., D'AQUINO J.D., RINGE, G.A., PETSKO (2009): Effects of pH and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability. Biochemistry, *48*: 4816-4827.

MARQUES, A.R.A. and P., SAFTIG (2019): Lysosomal storage disorders - challenges, concepts and avenues for therapy: beyond rare diseases. J. Cell Sci., *132*.

MEIKLE, P.J., J.J., HOPWOOD, A.E., CLAGUE, W.F., CAREY (1999): Prevalence of lysosomal storage disorders. JAMA, *281*: 249-254.

MITTAL, R.D. (2015): Tandem mass spectroscopy in diagnosis and clinical research. Indian J. Clin. Biochem., *30*: 121-123.

MOHAMED, F.E., L., AL-GAZALI, F., AL-JASMI, B.R., ALI (2017): Pharmaceutical Chaperones and Proteostasis Regulators in the Therapy of Lysosomal Storage Disorders: Current Perspective and Future Promises. Front. Pharmacol., *8*: 448.

MOKHTARIYE, A., L., HAGH-NAZARI, A.R., VARASTEH, F., KEYFI (2019): Diagnostic methods for Lysosomal Storage Disease. Rep. Bioch. Mol. Biol., *7*: 119-128.

PARÁ, C., P., BOSE, A.V., PSHEZHETSKY (2020): Neuropathophysiology of lysosomal storage diseases: synaptic dysfunction as a starting point for disease progression. J. Clin. Med., *9*: 616.

PARENTI, G., G., ANDRIA, K.J., VALENZANO (2015): Pharmacological Chaperone Therapy: Preclinical Development, Clinical Translation, and Prospects for the Treatment of Lysosomal Storage Disorders. Mol. Ther., *23*: 1138-1148.

PLATT, F.M., B., BOLAND, A.C., VAN DER SPOEL (2012): The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. J. Cell Biol., *199*: 723-734.

PLOTEGHER, N. and M.R., DUCHEN (2017): Mitochondrial Dysfunction and Neurodegeneration in Lysosomal Storage Disorders. Trends Mol. Med., *23*: 116-134.

POUPETOVA, H., J., LEDVINOVA, L., BERNA, L., DVORAKOVA, V., KOZICH, M., ELLEDER (2010): The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. J. Inherit. Metab. Dis., *33*: 387-396.

PRADHAN, S., M., SENGUPTA, A., DUTTA, K., BHATTACHARYYA, S.K., BAG, C., DUTTA, K., RAY (2011): Indian genetic disease database. Nuc. Ac. Res., *39*: D933-938.

SERANOVA, E., K. J., CONNOLLY, M., ZATYKA, T.R., ROSENSTOCK, T., BARRETT, R., I., TUXWORTH, S., SARKAR (2017): Dysregulation of autophagy as a common mechanism in lysosomal storage diseases. Essays in Bioch., *61*: 733-749.

SHAHID, M., S., FIRASAT, H.S., SATTI, T.M., SATTI, T., GHAFOOR, I., SHARIF, K., AFSHAN (2020): Screening of the FANCA gene mutational hotspots in the Pakistani fanconi anemia patients revealed 19 sequence variations. Congenital Anomalies, *60*: 32-39.

SHETH, J., R., BHAVSAR, M., MISTRI, D., PANCHOLI, A., BAVDEKAR, A., DALAL, P., RANGANATH, K.M., GIRISHA, A., SHUKLA, S., PHADKE, R., PURI, I., PANIGRAHI, A., KAUR, M., MURANJAN, M., GOYAL, R., RAMADEVI, R., SHAH, S., NAMPOOTHIRI, S., DANDA, C., DATAR, S., KAPOOR, S., BHATWADEKAR, F., SHETH (2019): Gaucher disease: single gene molecular characterization of one-hundred Indian patients reveals novel variants and the most prevalent mutation. BMC Med. Gen., *20*: 31.

SHETH, J., D., PANCHOLI, M., MISTRI, P., NATH, C., ANKLESHWARIA, R., BHAVSAR, R., PURI, S., PHADKE, F., SHETH (2018): Biochemical and molecular characterization of adult patients with type I Gaucher disease and carrier frequency analysis of Leu444Pro - a common Gaucher disease mutation in India. BMC Med. Gen., *19*: 178.

SHETH, J., P., PATEL, F., SHETH, R., SHAH (2004): Lysosomal storage disorders. Indian Pediatr., *41*: 260-265.

TAMMACHOTE, R., S., TONGKOBPETCH, C., SRICHOMTHONG, K., PHIPATTHANANANTI, S., PUNGKANON, D., WATTANASIRICHAIGOON, K., SUPHAPEETIPORN, V., SHOTELERSUK (2013): A common and two novel GBA mutations in Thai patients with Gaucher disease. J. Human Gen., *58*: 594-599.

TSUJI, S., P.V., CHOUDARY, B.M., MARTIN, B.K., STUBBLEFIELD, J.A., MAYOR, J.A., BARRANGER, E.I., GINNS (1987): A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. The New England J. Med., *316*: 570-575.

VERITY, C., A.M., WINSTONE, L., STELLITANO, R., WILL, A., NICOLL (2010): The epidemiology of progressive intellectual and neurological deterioration in childhood. Arch. Dis. Child., *95*: 361-364.

WANG, Y., L., LIU, J., XIONG, X., ZHANG, Z., CHEN, L., YU, C., CHEN, J., HUANG, Z., ZHANG, A.A., MOHMED, Z., LIN, N., XIONG, T., WANG (2012): Glucocerebrosidase L444P mutation confers genetic risk for Parkinson's disease in central China. Behavioral and Brain Func.: BBF, *8*: 57.

ZHOU, H., P., FERNHOFF, R.F., VOGT (2011): Newborn bloodspot screening for lysosomal storage disorders. J. Pediatr., *159*: 7-13 e11.

NEUROLOŠKE MANIFESTACIJE POREME.AJ A SKLADIŠTENJA LIZOZOMA I MOLEKULARNA KARAKTERIZACIJA GAUCHER -ove BOLESTI U PAKISTANU

Rutaba GUL[1], Sabika FIRASAT[1]*, Mulazim HUSSAIN[2], Muhammad TUFAIL[1], Waheed AHMAD[1], Kiran AFSHAN[1]

[1]Departman za zoologiju, Fakultet za biološke nauke, Quaid-i-Azam Univerzitet, Islamabad, Pakistan

[2] Dečija bolnica, Pakistanski institut za medincinske nauke, Islamabad, Pakistan

Izvod

Poreme. aji skladištenja lizozoma (LSD) su velika grupa urođenih grešaka u metabolizmu, od kojih je svaka uzrokovana genetskim mutacijama određenog gena koji kodira lizozomalni protein. Ova nasledna stanja karakteriše lizozomalna disfunkcija sa velikim brojem uticaja na organe, a sa godinama ponekad i otkazivanje organa. Neurološke komplikacije u slučajevima LSD -a kreću se od teških neurodegeneracija u 70% slučajeva do blagih simptoma ili odsustva neuropatije u drugih. Svaki LSD je monogen, ali heterogen sa molekularnog stanovišta sa velikim brojem mutacija opisanih u odgovarajućem genu. Neke mutacije su specifične za određene populacije, odražavajući posledice efekta osnivača. Ova studija imala je za cilj da pristupi demografskim i kliničkim profilima četrdeset pet porodica pogođenih LSD-om upisanih u periodu od januara 2018. do decembra 2019. godine u lokalne bolnice kako bi se otkrili neurološki simptomi u pakistanskim slučajevima LSD-a. Štaviše, izvršena je molekularno - genetska analiza porodica pogođenih Gaucherovom bolešću kako bi se otkrile osnovne bolesti koje izazivaju mutacije. Neurološke manifestacije bile su prisutne u dvadeset osam porodica, uključujući jedanaest mukopolisaharidoza-1 (MPS-I), četiri Gaucherove bolesti (GD) i sve slučajeve MPS-II, MPS-III, Niemann-Pick, Griscelli i Chediak-Higashi. Neurološka uključenost nije pronađena u osam porodica pogođenih MPS-I, jednim GD-om, svim MPS-IV i piknodizostozom. Skrining GBA gena u GD porodicama otkrio je prijavljenu mutaciju p.L483P u svim analiziranim porodicama. Klinička heterogenost MPS-1 i GD evidentna je iz literature, međutim mutaciona analiza svih upisanih GD porodica pokazala je segregaciju prijavljene varijante p.L483P GBA gena sa fenotipom bolesti u svim porodicama. Klinička heterogenost MPS-1 i GD evidentna je iz literature, međutim mutaciona analiza svih upisanih GD porodica pokazala je segregaciju prijavljene varijante p.L483P GBA gena sa fenotipom bolesti u svim porodicama. Naši nalazi ukazuju na važnost homeostatske uloge lizozoma u razvoju neurona jer je dvadeset osam od četrdeset porodica imalo neurološke manifestacije. Štaviše, identifikacija iste mutacije kod pacijenata sa GD sa ili bez neuronske uključenosti može biti povezana sa nekim nepoznatim razlikama u ekspresiji genetskih modifikatora ili izloženosti okidačima iz okruženja.

Primljeno 20.IX.2020.
Odobreno 18. V. 2021.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 April 2023
DOI 10.3389/fgene.2023.1128850

Check for updates

OPEN ACCESS

EDITED BY
Elsayed Abdelkreem,
Sohag University, Egypt

REVIEWED BY
Sheng Wang,
University of California, San Francisco, United States
Daniela Melis,
University of Salerno, Italy

*CORRESPONDENCE
Sabika Firasat,
sabika.firasat@qau.edu.pk
Torben Hansen,
torben.hansen@sund.ku.dk

SPECIALTY SECTION
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

RECEIVED 21 December 2022
ACCEPTED 13 March 2023
PUBLISHED 05 April 2023

CITATION
Gul R, Firasat S, Schubert M, Ullah A, Peña E, Thuesen ACB, Hussain M, Staeger FF, Gjesing AP, Albrechtsen A and Hansen T (2023), Identifying the genetic causes of phenotypically diagnosed Pakistani mucopolysaccharidoses patients by whole genome sequencing. *Front. Genet.* 14:1128850. doi: 10.3389/fgene.2023.1128850

COPYRIGHT
© 2023 Gul, Firasat, Schubert, Ullah, Peña, Thuesen, Hussain, Staeger, Gjesing, Albrechtsen and Hansen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identifying the genetic causes of phenotypically diagnosed Pakistani mucopolysaccharidoses patients by whole genome sequencing

Rutaba Gul[1,2], Sabika Firasat[1]*, Mikkel Schubert[2], Asmat Ullah[2], Elionora Peña[2], Anne C. B. Thuesen[2], Mulazim Hussain[3], Frederik F. Staeger[4], Anette P. Gjesing[2], Anders Albrechtsen[4] and Torben Hansen[2]*

[1]Department of Zoology, Faculty of Biological Sciences, Quaid-I-Azam University, Islamabad, Pakistan, [2]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [3]The Children Hospital, Pakistan Institute of Medical Sciences (PIMS), Islamabad, Pakistan, [4]Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Background: Lysosomal storage disorders (LSDs) are a group of inherited metabolic diseases, which encompass more than 50 different subtypes of pathologies. These disorders are caused by defects in lysosomal enzymes, transporters, and other non-lysosomal proteins. Mucopolysaccharidosis (MPS) is the most common subgroup of lysosomal storage disorders in which the body is unable to properly breakdown mucopolysaccharides. The aim of the present study was to identify novel genes and pathogenic variants in families from diverse regions of Pakistan with clinically diagnosed mucopolysaccharidosis type I and mucopolysaccharidosis type II.

Methods: Clinical diagnosis identified 12 with mucopolysaccharidosis I and 2 with mucopolysaccharidosis II in 14 families and whole genome sequencing (WGS) was performed to identify the causative variations in 15 affected individuals. Twenty-two unaffected individuals including parents or normal siblings of patients were also sequenced. Putative causal variants were identified by co-segregation and functional annotation.

Results: Analysis of whole genome sequencing data revealed ten novel and six previously reported variants in lysosomal storage disorders-associated genes (*IDUA, GALNS, SGSH, GAA, IDS, ALDOB, TRAPPC4, MASP1, SMARCAL, KIAA1109, HERC1, RRAS2*) and a novel candidate gene (*ABCA5*) for lysosomal storage disorder-like phenotypes, which has previously been associated with symptoms strongly related with lysosomal storage disorder in animal models.

Conclusion: Multigenic inheritance was found in several families highlighting the importance of searching for homozygous pathogenic variants in several genes also in families with a high degree of consanguinity.

KEYWORDS

lysosomal storage disorder, mucopolysaccharidosis, Pakistani families, whole genome sequencing, ABCA5

1 Introduction

Lysosomal storage disorders (LSDs) are a diverse group of inherited conditions primarily caused by loss of function of lysosomal proteins leading to the dysregulation of many lysosomes-linked activities including autophagy, lipid homeostasis, membrane repair, cell viability, exocytosis and signaling cascades (Parenti et al., 2015; Marques and Saftig, 2019). Common clinical phenotypes of LSDs include delay in achieving milestones, abdominal distention, hepatosplenomegaly, short stature, coarse facial features, and joint contractures (Gul et al., 2021). LSDs are classified into subtypes based on accumulation of undegraded substrates like glycosaminoglycans (GAGs), glycogens and sphingolipids accumulated as a result of compromised activities of enzymes in lysosomes.

Autosomal recessive disorders are more prevalent in populations like the Pakistani where 60% of all marriages are consanguineous and among them more than 80% are first cousin unions (Hussain and Bittles, 1998). Globally, the prevalence for LSDs ranges from 7.5/100,000 to 23.5/100,000 live births, with sphingolipidoses as the highest prevalent disorder followed by mucopolysaccharidoses (MPSs) (Kingma et al., 2015). Mucopolysaccharidoses (MPS) (OMIM # 252700) are carbohydrate metabolism disorders caused by the deficiency of lysosomal enzymes required for hydrolyzing glycosaminoglycans (GAGs; negatively charged polysaccharides). MPS are characterized by gradual accumulation of various types of GAGs within cells of multiple organs leading to somatic manifestations such as coarse facial features, cardiac issues, skeletal deformities, hepatosplenomegaly, respiratory problems, hematologic, neurologic, and ocular phenotypes. MPS include a spectrum of clinical phenotypes from severe to attenuated forms including all MPS sub-types like MPS I, MPS II, MPS III, MPS IV, MPS VI, MPS VII with overlapping physical features (Parenti and Giugliani, 2022). Most MPSs are inherited as autosomal recessive disorders except MPS II which follows an X-linked inheritance pattern (Linhart and Elliott, 2007).

Genetic defects in genes encoding lysosomal enzymes including *IDUA, IDS, SGSH, NAGLU, HGSNAT, GNS, GALNS, GLB1* and *GAA* lead to MPS. Additionally, pathogenic variants in other lysosomal proteins also lead to hereditary metabolic disorders with phenotypes that overlap with MPS (Kubo et al., 2005), indicating that other candidate genes may also cause or contribute to the disease. For example, ATP binding cassette (ABC) are group of transporters involved in transportation of substrates across membranes. *ABCA5* is a member of this group of transporters (Brooks-Wilson et al., 1999). ABCA5 is located in the lysosomes and late endosomes and its homozygous knockout results in lysosomal disease like symptoms in mice (Kubo et al., 2005).

Therapeutic management is a serious problem in LSDs as enzyme replacement therapy (ERT) has limitations including high cost, life-long treatment, inability to stop progression of neuronal symptoms and possibility of immune response against an exogenous enzyme (Arora et al., 2007). Globally, combinational therapies including ERT, and pharmacological chaperone therapy (PCT) are proving fruitful. For PCT, knowledge of the enzyme mutations is required for selection of a specific chaperone, reflecting the importance of identifying a molecular diagnosis. Other emerging treatment options including gene therapy, genome editing, and anti-inflammatory therapy also rely on identifying the genetic cause of the disease (Kingma and Jonckheere, 2021; Schuh et al., 2029). Early diagnosis for LSDs represents a big challenge and it is crucial to prevent multi-organ failure (Mokhtariye et al., 2019). The use of comprehensive genetic sequencing methodologies such as whole genome sequencing (WGS) are particularly suitable to elucidate the genetic background in rare diseases such as MPS I and MPS II, by enabling the detection of rare variants in known as well as novel genes. The aim of this study was to identify disease-causing variants using WGS data of 14 Pakistani families initially diagnosed as MPS I and MPS II based on clinical presentation.

2 Materials and methodology

2.1 Ethical approvals

Ethical approval was granted by the Bioethical Review board, Quaid-i-Azam University, Islamabad, Pakistan (BEC-FBS-QAU2019-198) and Shaheed Zulfiqar Ali Bhutto Medical University, Islamabad, Pakistan (F.1-1/2105/ERB/SZABMU/179). Before collection of blood samples, all participants or their guardians signed an informed consent form. The study was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013).

2.2 Participants

In the present study, we recruited 14 families in total (A–N). Twelve of them (A–L) were diagnosed as MPS I with 26 affected individuals (Figure 1) and the other two (M, N) as MPS II with two affected males (Figure 2). In total, 37 individuals including 15 affected and 22 unaffected participated in the present study. Families were recruited from diverse regions of Pakistan (i.e., Punjab, Islamabad capital territory, Khyber Pakhtunkhwa, and Azad Jammu and Kashmir) (Figure 3). Moreover, 77 additional individuals with Pakistani origin that are part of an ongoing LSDs study (unpublished) were also sequenced.

Inclusion criteria were based on their physical phenotypes, radiological findings, and blood testing (Table 1) at outpatient department (OPD), Children Hospital, Pakistan Institute of Medical Sciences, Islamabad, Pakistan. Pedigrees were drawn by using the information provided by the accompanying elder attendant of the patients.

2.3 Blood collection and DNA extraction

Peripheral blood samples of 3–5 ml were taken in an EDTA vacutainer (BD vacutainer K2 EDTA 18 mg) and stored as whole blood at −20°C before DNA extraction. DNA extraction was performed using standard Phenol-Chloroform method (Grimberg et al., 1989). DNA was quantified by using the μDrop Plate reader (MultiskanTM, Thermo Fisher Scientific, and Waltham, MA, United States).

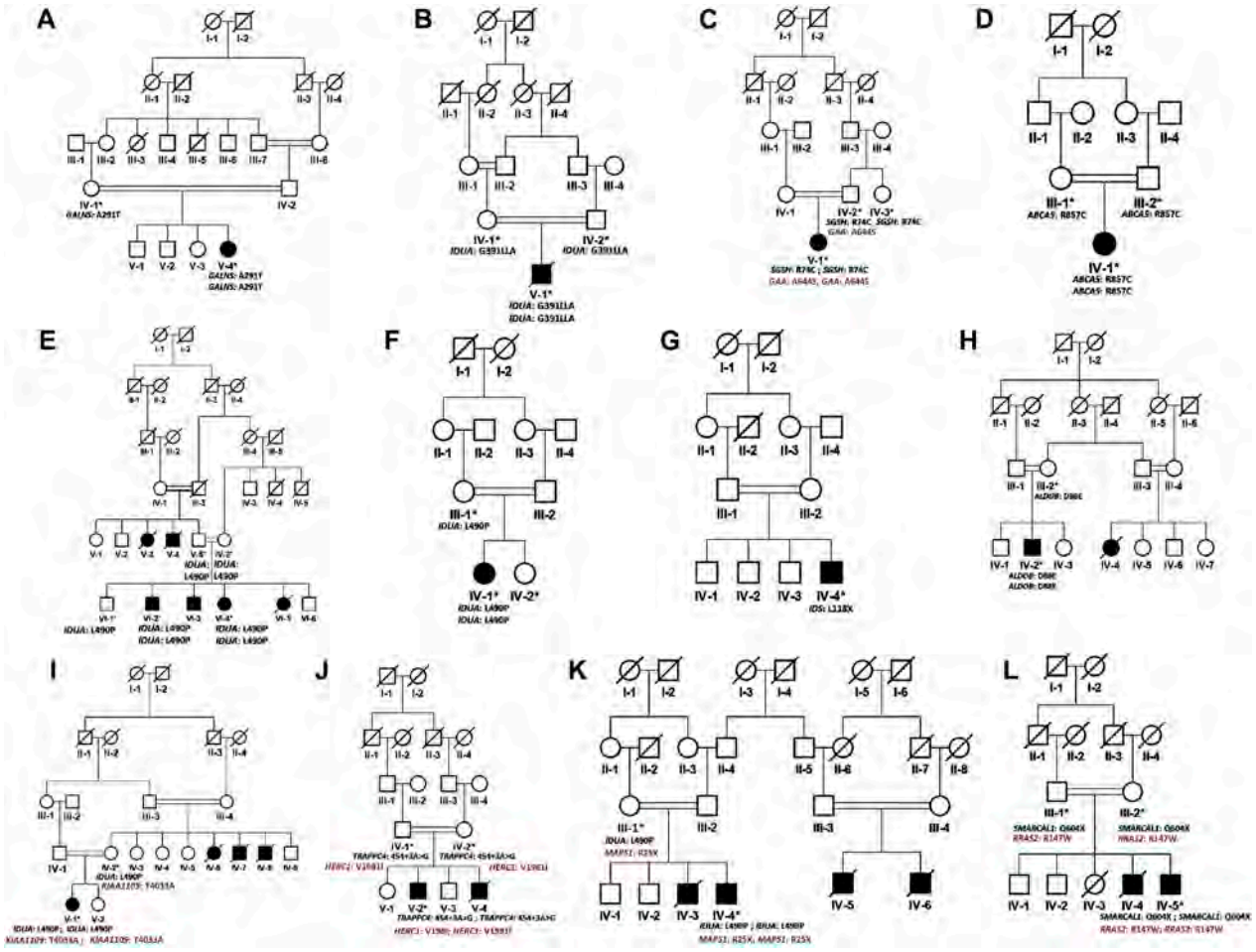


FIGURE 1
Pedigrees of families **(A–L)** inheriting MPS I phenotype consistent with autosomal recessive mode of inheritance of disease. Square and circles denote males and females respectively, filled symbols indicate affected individuals and consanguinity is represented by double marriage lines. Genotypes of participating individuals are shown beneath each individual.

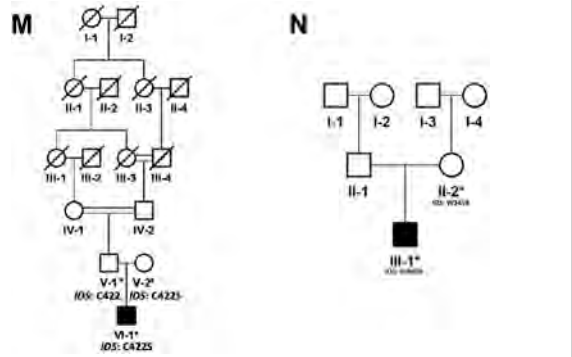


FIGURE 2
Pedigrees of families **(M,N)** inheriting MPS II phenotype consistent with X-linked mode of inheritance of disease. Square and circles denote males and females respectively, filled symbols indicate affected individuals and consanguinity is represented by double marriage lines. Genotypes of participating individuals are shown beneath each individual.

2.4 Whole genome sequencing and variant calling

Read mapping and genotype calling was performed using a modified version of the PALEOMIX pipeline (Schubert et al., 2014). The modified pipeline and detailed instructions are available at https://github.com/Hansen-Group/Gul2022.

Briefly, pre-analyses quality assurance was performed using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC v1.10 (Ewels et al., 2016). Read pairs were subsequently processed using fastp v0.20.1 (Chen et al., 2018) to trim adapter sequences, remove reads with low complexity, and merge and correct overlapping read pairs (with a minimum overlap of 11 bp). Reads were mapped using BWA mem v0.7.17-r1188 (Li, 2013). Against the hg38 human reference genome distributed as part of the GATK resource bundle (https://gatk.broadinstitute.org/hc/en-us/articles/360035890811-Resource-bundle), (McKenna et al., 2010) including alternative and decoy contigs. Mapped reads were post-processed using samtools v1.11 (Danecek et al., 2021),

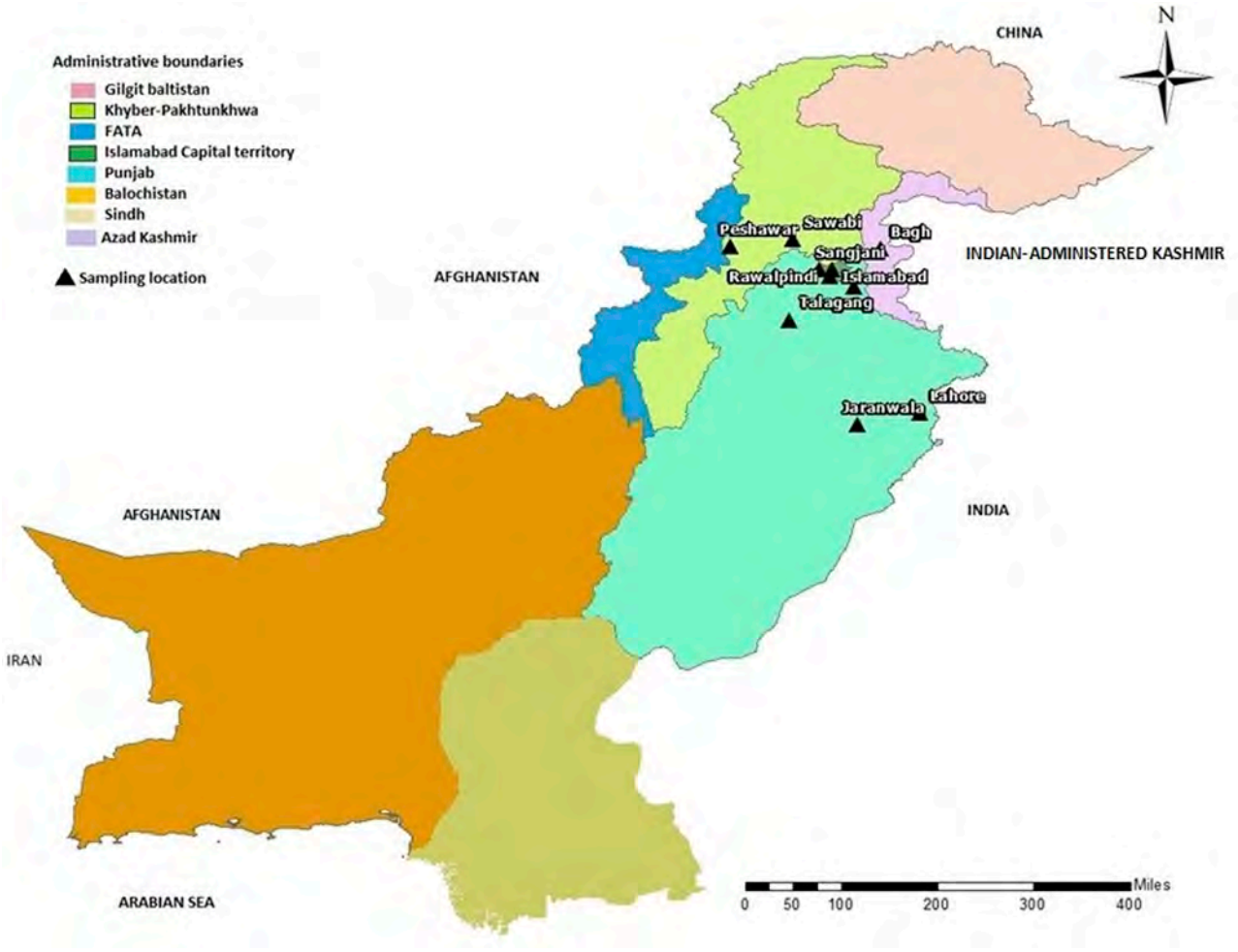


FIGURE 3
Map of Pakistan showing the locations (black triangles) of cities/towns from where families included in this study were collected.

and the "bwa-postalt.js" script included in the BWA-kit which post-processes alignments mapping to alternative contigs such as HLA variants. Duplicate, unmapped, and orphaned reads, and secondary and supplementary alignments were filtered. The resulting BAMs were recalibrated using GATK ApplyBQSR trained using dbSNP release 151 (Sherry et al., 2001) for known variable sites.

Genotypes (GVCFs) were called for each sample using GATK HaplotypeCaller, merged, and called using GATK GenotypeGVCFs. Variant recalibration was carried out separately for SNPs and indels using GATK ApplyVQSRNode, trained using positive lists from the GATK resource bundle. Features used in the calibration were ExcessHet, DP, MQ, QD, SOR, FS, ReadPosRankSum, MQRankSum, and BaseQRankSum (only for SNPs). We kept only variables within the upper 98.0 tranche for indel and 99.6 tranche for SNPs. The resulting VCFs were annotated using VEP v104 (McLaren et al., 2016) using the Ancestral Allele, ExACpLI, GERP Conservation Scores, and LOFTEE v1.0.3 (https://github.com/konradjk/loftee) (Karczewski et al., 2020) plugins. Additional custom annotation was included as described in the github repository.

2.4.1 Data analysis and variant prioritization

All variants were evaluated as per ACMG guidelines (Richards et al., 2020). Variants that were classified to be pathogenic, likely pathogenic or uncertain significance according to ACMG guidelines were prioritized.

Runs of homozygosity (ROH)-based inbreeding coefficient (F) and principal component analysis was calculated using PLINK(v1.90b6) on a dataset including both affected and unaffected individuals as well as 77 additional individuals with Pakistani origin that are part of an ongoing LSDs study (unpublished) for better allele frequency estimates. Results are only shown for the individuals in present study.

For ROH calculation, only variants with minor allele frequency MAF > 5%, and no missingness were used. ROHs were defined as having more than 100 variants, a total length larger than 1 MB, on average at least 1 variant per 50 kB, consecutive variants less than 1 MB apart, and at most 1 heterozygous call per window of 50 variants. The total length of ROHs was then divided by the total length of the genome (3 GB).

Principal component analysis was performed on by first merging with the 1000 genomes project (https://www.nature.com/articles/nature15393). Non-overlapping sites and sites with a MAF difference

10.3389/fgene.2023.1128850

TABLE 1 Showing the details of each patient affected with Mucopolysaccharidosis type I and II.

Family ID	Individual no.	No. of affected individuals	Age/ Sex	Cast/Region/ Consanguinity	Clinical profile
A	V-4	1	9 m/F	Rajput/ICT/Yes	Squint eyes, tip-toe walking, kyphosis, frontal bossing, depressed nasal bridge, umbilical hernia, narrowing of spinal cord, dysostosis, throat problem, weak muscles, more crying episodes, two central incisors widely spaced, joint contractures, coarse facial features
					Pectus carinatum, Hemoglobin = 8.6 g/dL (12–15), HCT = 32 (35–55), MCV = 63.9 fl (75–100), MCH = 17.2 pg (25–35), MCHC = 26.9 g/dL (31–38), Neutrophils = 40.4% (60–70), ESR = 28 mm/Hr (0–15)
B	V-1	1	1 y/M	Malik/Punjab/Yes	Delay in achieving milestone, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, developmental delay, depressed nasal bridge, squinted eyes
C	V-1	1	4 y/F	Pathan/ICT/Yes	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, short stature, respiratory/ear infections, coarse facial features, broad based gait, joint contractures, short trunk, scoliosis/kyphosis, hepatosplenomegaly, low IQ, intellectual disability, low hemoglobin
D	IV-1	1	1.5 y/F	Sheikh/Punjab/Yes	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, gibbus formation, short stature, upper airways obstruction, kyphosis, hazel corneas, coarse facial features, hearing loss, joint contractures, hepatosplenomegaly, Intellectual disability, hypertrichosis
E	VI-4	6	8 y/F	Rajput/AJK/Yes	Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding
	VI-2		3 y/M		Delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, hernia, gibbus formation, short stature, respiratory/ear infection, coarse facial features, waddling in gait, joint contractures, scoliosis, hepatosplenomegaly, corneal clouding
F	IV-1	1	16.5 y/F	Narman/AJK/Yes	Delay in achieving milestones, abdominal distention, hernia, short stature, coarse facial features, joint contractures, short trunk, hepatosplenomegaly
G	IV-4	1	6 y/M	Rajput/Punjab/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior, short stature, respiratory/ear infection, joint contractures, short trunk, speech delay, coarse facial features, hepatosplenomegaly, intellectual disability
H	IV-2	2	14 y/M	Suddhan/ICT/Yes	Delay in achieving milestones, abdominal distention, coarse facial features, aggressive behavior, short stature, respiratory/ear infections, hip dysplasia, joint stiffness, hepatosplenomegaly, difficulty in walking
					Craniofacial disproportion, J-shaped sella, L1 vertebral body is hypoplastic and slightly posteriorly displaced, exaggerated lumbar lordosis, RBG = 71 mg/dL (80–160), serum urea = 10 md/dL (12–50), loss of biconcavity of metacarpels with proximal tapping

(Continued on following page)

TABLE 1 (*Continued*) Showing the details of each patient affected with Mucopolysaccharidosis type I and II.

Family ID	Individual no.	No. of affected individuals	Age/ Sex	Cast/Region/ Consanguinity	Clinical profile
I	V-1	4	12 y/F	Mughal/Punjab/Yes	Joint contractures, corneal clouding, very less eyesight, abdominal distention, coarse facial features, short stature, underweight, umbilical hernia, cannot walk, developmental delay, cardiac issues, hearing issues, respiratory problems/loud snoring
J	V-2	2	2.5 y/M	Mughal/KPK/Yes	Delay in achieving milestones, aggressive behavior with multiple crying episodes, respiratory/ear infection, no walk, intellectual disability
K	IV-4	4	8.5 y/M	Abbasi/Punjab/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, umbilical hernia, gibbus formation, short stature, constipation, coarse facial features, cannot walk (Wheel chair bound), joint contractures, scoliosis and kyphosis, hepatosplenomegaly, fish mouth vertebrae, oar shaped ribs, low IQ, widely spaced eyes, narrowing of eye opening (right), drooping eyelids, intellectual disability, developmental delay
					Red cell count = 2.68 million/ul (4–5.2), Hemoglobin = 7.2 g/dL (11.5–15.5), HCT = 21.8 (35–45), Platelet count = 141 × 1000/ul (170–450), Neutrophils = 53.4% (60–70)
L	IV-5	2	8 y/M	Yousafzai/KPK/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, low IQ, intellectual disability, dysostosis multiplex
M	VI-1	1	4.5 y/M	Rajput Bhatti, Khosa Baloch/ Punjab/No	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, constipation, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability, hernia
N	II-1	1	5 y/M	Aryn/Punjab/No	Hernia, short stature, coarse facial features

AJK, Azad Jammu and Kashmir; ESR, erythrocytes sedimentation rate; F, female; HCT, hematocrit; ICT, Islamabad capital territory; IQ, Intelligent quotient; KPK, Khyber Pakhtunkhwa; M, male; MCV, mean corpuscle volume; MCH, mean corpuscle hemoglobin; MCHC, mean corpuscle hemoglobin concentration; m, months; RBG, random blood glucose; y, years.

between LSD-samples and the 1000 genomes South Asian super population of more than 25% were removed. The PCA was calculated using PLINK--pca function on remaining overlapping sites with overall MAF > 0.05, no missingness, and LD-pruned (window size of 1 MB, step size of 1, R2-threshold of 0.8).

Non-synonymous variants including missense, non-sense, frameshift and indels in coding and splice sites changes in non-coding regions with a MAF of less than 0.01 in large, outbred populations (gnomAD exomes of South Asian population) were considered as putative causal candidates. Further prioritization of these rare variants was based on segregation with disease phenotypes within respective family. As a first phase for WGS data analysis, already reported genes causing MPS were screened for pathogenic variants and during second phase, unsolved families were checked for variants in other LSDs-related genes. In 12 families (A–L) diagnosed with MPS I, variants homozygous in affected individuals and heterozygous in their parents were prioritized based on recessive inheritance pattern of the disease and consanguinity among parents revealed by their pedigrees. The lists of heterozygous variants in each family were searched for compound heterozygous or digenic inheritance. In two families (M, N) diagnosed with MPS II, pathogenic variants on X-chromosome segregated from the mother were extracted. In families where genotypes of parents were available, exomes of each family were analyzed for dual molecular diagnosis (Supplementary Figure S1). As a reference population, more than a 100 healthy control samples from the general Pakistani population were analyzed for the prevalence of potentially causal variant as per ACMG guidelines.

The identified variants were subjected to *in silico* analysis using the following tools: Have (y) Our Protein Explained (HOPE) (https://www3.cmbi.umcn.nl/hope/), Predictor of human deleterious single nucleotide polymorphisms (PhD-SNP) (https://snps.biofold.org/phd-snp/phd-snp.html), Protein analysis through evolutionary relationships (PANTHER) (http://www.pantherdb.org/tools/csnpScoreForm.jsp), I-Mutant (https://folding.biofold.

TABLE 2 List of variants identified in the current study.

Pedigree/ Patient ID	Gene variant	Patients' genotype	dbSNP	*In-silico* tools with deleterious output	MAF gnomAD	Classification ACMG	VarSome	HGMD accession number
A/V-4	*GALNS*	Homozygous	rs118204448	4/7	0.00001972	Pathogenic	Pathogenic	CM950535
	NM_000512.5:c.871G>A (p.Ala291Thr)					PP3, PP4, PP5, PM1, PM2		
B/V-1	*IDUA*	Homozygous	rs1560548356	2/4	0.000007	Likely pathogenic	Likely pathogenic	Current study
	NM_000203.5:c.1172-1173insGCTGCTGGC (p.Gua391insLeuLeuAla)					PP4, PM2, PM4		
C/V-1	*SGSH*	Homozygous	rs104894636	17/17		Pathogenic	Pathogenic	CM971355
	NM_000199.5:c.220G>A (p.Arg74Cys)				0.0002270	PP5, PS3, PM1, PM2, PM5, PP3		
	GAA		rs752735168	16/18	0.00003306	Likely Pathogenic	Pathogenic	Current study
	NM_000152.5:c.1930G>T (p.Ala644Ser)					PM1, PM2, PP3, PP4		
D/IV-1	*ABCA5*	Homozygous	rs747667004	7/7	0.0001090	Likely Pathogenic	Uncertain significance	Current study
	NM_172232.4: c.26569G>A (p.Arg857Cys)					PM1, PM2, PP3, PP4		
E/IV-2, IV-4	*IDUA*	Homozygous	rs121965027	7/7	0.00004681	Pathogenic	Pathogenic	CM950685
	NM_000203.5:c.1469T>C (p.Leu490Pro)					PS3, PM1, PM2, PP4, PP5, PS3		
F/IV-1	*IDUA*	Homozygous	rs121965027	7/7	0.00004681	Pathogenic	Pathogenic	CM950685
	NM_000203.5:c.1469T>C (p.Leu490Pro)					PS3, PM1, PM2, PP4, PP5, PS3		
G/IV-4	*IDS*	Hemizygous	—	4/4	—	Pathogenic	Pathogenic	Current study
	NM_001166550.4: c.353T>A (p.Leu118Ter)					PVS1, PM2, PP3, PP4		
H/IV-2	*ALDOB*	Homozygous	rs200585150	11/18	0.0004219	VUS	Uncertain Significance	CM1618755
	NM_000035.4:c.264G>T (p.Asp88Glu)					PM1, PM2, PP3 BP1		
I/V-1	*IDUA*	Homozygous	rs121965027	7/7		Likely Pathogenic	Pathogenic	CM950685
	NM_000203.5:c.1469T>C (p.Leu490Pro)				0.00004681	PS3, PM1, PM2, PP4, PP5		
	KIAA1109		rs773658128	6/7	0.00001213	VUS	Uncertain significance	Current study
	NM_001384125.1: c.1297A>G (p.Thr4033Ala)					PM1, PM2		
J/V-2	*TRAPPC4*	Homozygous	rs375776811	N/A	0.0002403	Pathogenic	Likely pathogenic	Not available
	NM_016146.6:c.454 + 3A>G				-	PP5, PM2, PP3		
	HERC1		—	4/7		VUS	Uncertain significance	Current study
	NM_003922.4:c.5941G>A (p.Val1981Ile)					PM2, BP4		
K/IV-4	*IDUA*	Homozygous	rs121965027	7/7	0.00004681	Likely Pathogenic	Pathogenic	CM950685

(Continued on following page)

TABLE 2 (*Continued*) List of variants identified in the current study.

Pedigree/ Patient ID	Gene variant	Patients' genotype	dbSNP	*In-silico* tools with deleterious output	MAF gnomAD	Classification ACMG	VarSome	HGMD accession number
	NM_000203.5:c.1469T>C (p.Leu490Pro)				0.0001591	PS3, PM1, PM2, PP4, PP5		
	MASP1		rs369042881	N/A		VUS	Uncertain significance	Current study
	NM_139125.4:c.C73T (p.Arg25Ter)					PM2, BS1, BS2		
L/IV-5	*SMARCAL1*	Homozygous	—	6/7	-	Pathogenic	Likely pathogenic	Current study
	NM_014140.4:c.1810C>T (p.Gln604Ter)				-	PVS1, PM2, PP3		
	RRAS2		—	16/18		Likely Pathogenic	Likely Pathogenic	Not available in HGMD
	NM_012250.6:c.439C>T (p.Arg147Trp)					PM1, PM2, PP2, PP3		
M/VI-1	*IDS*	Hemizygous	—	17/23	—	Likely Pathogenic	Likely pathogenic	Current study
	NM_000202.8:c.1264C>G (p.Cys422Ser)					PM1, PM2, PP3, PP4, BP1		
N/III-1	*IDS*	Hemizygous	—	11/11	—	Pathogenic	Pathogenic	Current study
	NM_000202.8:c.1035C>T (p.Trp345Ter)					PVS1, PM2, PP3		

N/A, not applicable; VUS, variant of uncertain significance; -, not found.

org/i-mutant/i-mutant2.0.html), MUpro (http://mupro.proteomics.ics.uci.edu/), VarSome (https://varsome.com/), and ACMG classification (http://wintervar.wglab.org/).

3 Results

3.1 Clinical characterization

In total, fourteen families (A-N) of Pakistani origin with phenotypes associated with MPS I and MPS II were included in the present study. Patients showed a wide variety of characteristic symptoms from both MPS I and MPS II including coarse facial features, joint contractures, abdominal distention, hepatosplenomegaly, hernia, and mental abnormalities among others (Table 1). Twelve of fourteen (A-L) families showed consanguinity and clinical phenotypes consistent with MPS I and were consequently labeled as suspected cases of MPS I by the clinicians (Table 1; Figure 1). The other two families (M and N) did not show consanguinity and phenotypes matched clinically with MPS II (Table 1; Figure 2).

3.2 Genetic characterization

Principal component analysis of the LSDs families merged with the 1000 genomes projects showed the LSDs families clustering with the South Asian (SAS) superpopulation (Figure 4A). The SAS population includes a Punjabi population (PJL) verifying the genetic ancestry of the LSDs families. Runs of homozygosity-based F-coefficient of the affected individuals showed that the high rate of consanguinity in MPS I families was also reflected in the genetic data (Figure 4B) and close to 0 for the non-consanguineous MPS II affected individuals.

Analysis of WGS data of available affected and unaffected individuals in each family revealed ten novel and six previously reported homozygous or hemizygous variants in the genes involved in MPS I, MPS II or other lysosomal storage disorders. Five families (C, I, K, J, L) showed multigenic inheritance (Table 2). All the variants were found in ROH in homozygous/hemizygous state in affected individuals while heterozygous or wild type in their parents. In family D, no homozygous or compound heterozygous variants were found in MPS-related genes. Therefore, screening for other LSD-associated or novel candidate genes was performed. The analysis revealed a homozygous missense variant (c.26569G>A; R857C) in *ABCA5* located in the ROH. Based on the gene expression, function and pathways, *ABCA5* was predicted to be the strongest candidate. Phenotypically normal siblings were either wild type or heterozygous carrier of the respective variants in each family (Figures 1, 2). Genes were annotated according to the following transcripts: NM_000512.5 (GALNS); NM_000203.5 (IDUA); NM_000199.5 (SGSH); NM_000152.5 (GAA); NM_172232.4 (ABCA5); NM_001166550.4 (IDS) NM_000035.4 (ALDOB); NM_001384125.1 (KIAA1109), NM_016146.6 (TRAPPC4); NM_003922.4 (HERC1); NM_139125.4 (MASP1); NM_014140.4; NM_012250.6 (RRAS2).

These variants were predicted pathogenic, likely pathogenic or variant of uncertain significance according to ACMG classification (Table 2). None of the identified variants was found in homozygous state in any human genome variations databases including

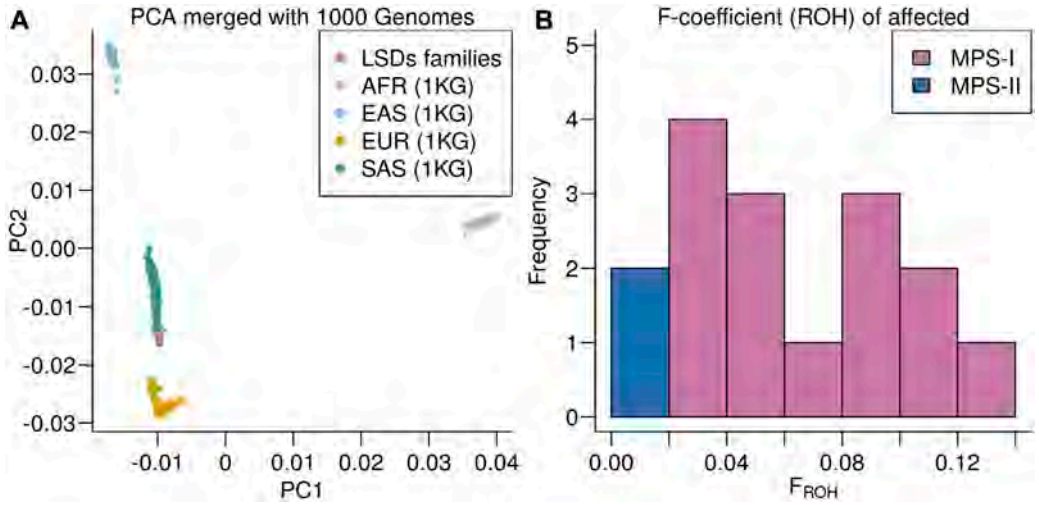


FIGURE 4
(A) Principal component analysis of the LSDs families merged with the 1000 Genomes project colored according to super population: African ancestry (AFR), East Asian ancestry (EAS), European ancestry (EUR), and South Asian ancestry (SAS). **(B)** Distribution of runs of homozygosity-based F-coefficient of the LSDs families.

gnomAD, ExAC, 1000 Genome Project and ethnically matched healthy controls.

4 Discussion

In the present study, fourteen families (A-N) initially diagnosed with MPS underwent WGS to identify potentially pathogenic variants. Among these, twelve families (A-L) were clinically characterized as MPS I while two (M-N) were clinically diagnosed as MPS II. The diagnosis was based on clinical phenotypes including facial features, stature, laboratory tests and inheritance of the phenotypes in pedigree. WGS revealed homozygous variants in previously MPS-associated genes (*IDUA, GALNS, SGSH, GAA* and *IDS*) and a novel LSDs-associated candidate gene (*ABCA5*) in affected individuals. In addition to these genes, variants were found in seven genes (*ALDOB, MASP1, KIAA1109, TRAPPC4, SMARCAL1, HERC1, RRAS2*) that were previously not reported in MPS but associated with other metabolic disorders.

Our findings in family A (i.e., MPS I), reported a homozygous sequence variant (c.871G>A; A291T) in *GALNS*. The gene encodes N-acetylgalactosamine-6-sulfatase which is a lysosomal exohydrolase required for the degradation of the glycosaminoglycans, keratan sulfate, and chondroitin 6-sulfate. Previously, the identified variant (A291T) has been reported in compound heterozygous and homozygous forms in two different families of Pakistani origin segregating MPS IV (Tomatsu et al., 1995; Ullah et al., 2017). *In-silico* analysis revealed that missense mutation (p.A291S) resulted in the formation of additional intramolecular interactions, most likely affecting the conformation of the active site (Asp288 and Asn289) and its substrate binding ability. Therefore, variation at amino acid position 291 of GALNS would affect structure and enzymatic function of the protein.

Multiple mutations in *IDUA* gene have been identified and associated with MPSI (Tieu et al., 1995). The *IDUA* gene is located at 4p16.3 and comprises 14 exons. IDUA codes for the enzyme α-L-iduronidase, which is required for the degradation of the GAGs dermatan and heparan sulfate. Deficiency of the enzyme leads to the accumulation of these GAGs in tissues throughout the body, playing a central role in the pathogenesis of MPS I. A novel homozygous insertion mutation (c.1172-1173insGCTGCTGGC; G391insLLA) in *IDUA* was identified in an affected individual in family B. Glycine at position 391 is involved in the formation of torsion angles due to its high degree of rotational ability which maybe lost after the insertion of LLA leading to abnormal function of the enzyme. Furthermore, affected individuals in families E, F, I and K also revealed a previously reported variant (c.1469T>C; L490P) in *IDUA* (Gul et al., 2020). In these cases, a high interfamilial variability in the phenotypes of patients segregating same sequence variant (*IDUA*; c.1469T>C; L490P) was observed in our families (E, F, I, K). Deep analysis of the genome of individuals revealed additional homozygous pathogenic variants in other genes (c.1297A>G; T4033A in *KIAA1109* in family I, and c.C73T; R25* in *MASP1* in family K) that may act as modifiers or dual molecular diagnosis. Previously sequence variants in *KIAA1109* (4q27) encoding a protein playing role in endosomal trafficking and endosome recycling of lipids have been reported to cause Alkuraya-Kucinskas syndrome characterized by brain abnormalities associated with cerebral parenchymal underdevelopment, arthrogryposis, clubfoot, and global developmental delay. Patients in family I showed phenotypes of MPS as well as Alkuraya syndrome. Similarly, sequence variants in *MASP1* (3q27.3) encoding a serine protease having an essential role in the innate and adaptive immune response have been associated with 3MC syndrome 1 characterized by widely spaced eyes (hypertelorism), a narrowing of the eye opening (blepharophimosis), droopy eyelids (ptosis) and highly arched eyebrows. Patients in the present family K showed phenotypes of 3MC syndrome in addition to MPS.

IDUA, *KIAA1109* and *MAPS1* are located at different loci of human genome. Due to higher rate of consanguinity, the number of homozygous regions increases leading to a high risk of occurrence of multiple pathogenic variants. This might lead to complex phenotypes due to co-segregation of two syndromes in the same patient. Therefore, analyzing genomes of populations with a high rate of consanguinity, dual molecular diagnosis should be considered.

In family C, we can appreciate a multi-genic inheritance, segregating two damaging variants in known LSDs enzymes SGSH (c.220G>A; R74C) and GAA (c.1930G>T; A644S) in homozygous state, both of which lie within the mapped interval at 17q25.3. *SGSH* encodes a lysosomal enzyme sulfamidase involved in the lysosomal degradation of heparan sulfate. Pathogenic sequence variants in the gene lead to MPS IIIA (Weber et al., 1997). Arginine at amino acid position 74 is likely to be involved in the formation of the active site of sulfamidase. Transition of arginine at the position to a non-conserved Cysteine may grossly affect enzyme activity. *GAA* encodes a lysosomal enzyme alpha-glucosidase. The lysosomal alpha-glucosidase is essential for the degradation of glycogen to glucose in lysosomes. Defects in the enzyme lead to glycogen storage disease. In addition to MPS phenotypes associated with variants in *SGSH*, the patient showed low hemoglobin concentration which is found in glycogen storage disease caused by pathogenic variant in *GAA*. These findings, suggest that the disease phenotypes in the present patient could be due to the contribution of variants in both genes, *SGSH* and *GAA*.

In family D, we have identified a missense homozygous variant (c.26569G>A; R857C) in *ABCA5*. The ABCA5 protein is a member of the superfamily of ATP binding cassette (ABC) transporters. ABC proteins play an important role in transporting various molecules across extra- and intracellular membranes. Loss of function variants in human *ABCA5* have been associated with hair overgrowth in human (DeStefano et al., 2014; Raza et al., 2020). On the other side, *abca5−/−* knockout mice developed lysosomal disease-like symptoms (Kubo et al., 2005). Interestingly, the affected individual in family D carrying *ABCA5* missense variant showed phenotypes overlying with MPS associated with hair overgrowth. To the best of our knowledge, this is the first human case associating a variation in *ABCA5* with LSD associated with hair overgrowth.

ABCA5 has 1642 amino acids long lysosomal peptide which is composed of two domains including ABC transporter 1 (478-713) and ABC transporter 2 (1290-1533). The protein determines the fate of cholesterol derived from lipoprotein. The functional loss of this protein impairs the integrity of lysosomes, disrupts the intracellular flow of free cholesterol, and causes cholesterol to accumulate intra-endo-lysosomally, adding to the organelles' malfunction (Raza et al., 2020). The substitution R857C, replaces a positively charged, less hydrophobic bigger amino acid (Arginine) to a neutral, more hydrophobic and smaller amino acid (Cysteine). Change in charge, size and nature of amino acids (Arginine and Cysteine) at position 857 may affect interactions of the protein with other molecules and residues leading to improper protein folding and intra-lysosomal accumulation of cholesterol.

In family G, a homozygous non-sense variant c.353T>A; p.L118* was identified in *IDS* gene which is involved in the X-linked inheritance of MPS II. IDS gene encodes a protein iduronate 2-sulfatase that catalyzes the degradation of heparan sulfate and dermatan sulfate (Gul et al., 2020).

In family H, a homozygous missense variant (c.264G>T; p.D88E) was identified in *ALDOB*. *ALDOB* encodes fructose-1,6-bisphosphate aldolase, a tetrameric glycolytic enzyme that catalyzes the reversible conversion of fructose-1,6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. Pathogenic variants in the gene have been associated with fructose intolerance. The specific variant (c.264G>T) identified in family H was previously reported by Retterer et al. (2016) in a family segregating abnormality of central nervous system. Interestingly, the affected individual in our family H, showed delay in achieving milestones, abdominal distention, coarse facial features, aggressive behavior, short stature, respiratory/ear infections, hip dysplasia, joint stiffness, hepatosplenomegaly, difficulty in walking, craniofacial disproportion, J-shaped sella, L1 vertebral body is hypoplastic and slightly posteriorly displaced and exaggerated lumbar lordosis. In addition to these phenotypes, the patient showed hypoglycemia, which is an indication of fructosuria. The amino acid aspartic acid D at position 88 make salt bridge with R57, L92 and L321, the mutant residue glutamic acid E may not be able to interact. Wild type residue D is in its preferred secondary structure, a turn, while mutant residue E prefers to be in another secondary structure, which may therefore destabilize the protein conformation. Due to mutated fructose-1,6-bisphosphate aldolase, accumulation of fructose-1,6-bisphosphate in different organs of the body may lead to multi-organ abnormality.

Family J revealed two homozygous variants, one is a missense variant c.5941G>A; p.V1981I in *HERC1* gene and other is splice site variant (c.454 + 3A>G) in *TRAPPC4*. HERC1 protein is involved in the membrane trafficking *via* guanine nucleotide exchange factors (GEF) and previously reported as a novel candidate gene for causing intellectual disability (Ortega-Recalde et al., 2015). Recently, Van Bergen and colleagues (Van Bergen et al., 2020) reported the same splice variant underlying early-onset seizures, developmental delay, microcephaly, sensorineural deafness, spastic quadriparesis and progressive cortical and cerebellar atrophy in families of Caucasian, Turkish and French-Canadian ethnicities. Other authors also identified the same splice site variant in a family of Indian origin having progressive encephalopathy and muscle involvement (Kaur et al., 2020). The primary phenotypes of our patient in family J include delay in achieving milestones, aggressive behavior with multiple crying episodes, respiratory/ear infection, unable to walk, and intellectual disability. Interestingly, phenotypes like muscles weakness, deafness, microcephaly and seizures were not found in our case. The phenotypic variability might be due to different familial backgrounds of the families and effect of rare SNPs in modifier genes. The variant is predicted to cause frameshift of the protein due to mis-splicing of mRNA.

In affected individual of family L, a novel homozygous missense variant c.439C>T; p.R147W in *RRAS2* gene and a novel non-sense variant (c.1810C>T; Q604*) was found in *SMARCAL1*. The substitution in gene *RRAS2* of a positively charged arginine at position 147 with a neutral residue tryptophan may lead to the loss of ionic interactions with surrounding residues. It also leads to distortion of shape as mutant residue tryptophan is bigger in size than wild type residue arginine. RRAS2 protein is involved in the regulation of MAPK signaling pathway (Capri et al., 2019). The involvement of *RRAS2* gene in some disease is still to be explored and linked with disease etiology in human genome mutation

database. The encoded protein SMARCAL1 is a member of the SWI/SNF family of proteins. Members of this family have helicase and ATPase activities and are thought to regulate transcription of certain genes by altering the chromatin structure around those genes. Pathogenic variants in the gene have been reported in nephrotic syndrome and a metabolic syndrome Schimke immuno-osseous dysplasia characterized by short stature, kidney disease, and a weakened immune system. The patient in the present study carrying a non-sense variant in the gene showed a severe phenotype of delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, hernia, short stature, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability and dysostosis multiplex (Table 1). Neurological phenotypes including intellectual disability, aggressive behavior, joint contractures, and hepatosplenomegaly showed by the present case were not described in previous cases. The variation in the phenotypes of previous and the present case might be due to the nature and/or position of mutation, different familial background and ages of the affected individuals. The identified non-sense variant (Q604*) in SMARCAL1 is predicted to cause loss of function of the protein either through non-sense-mediated mRNA decay or production of truncated protein.

Regarding families diagnosed as MPS II (M and N) we found pathogenic variants (c.1264C>G; C422S in family M and c.1035C>T; W345* in family N) in *IDS* segregating with the phenotypes. The novel missense variant C422S affects a non-conserved residue among human sulfatases. Cysteine is more hydrophobic in nature than serine, which causes loss of hydrophobic interactions. Cysteine at position 422 is involved in the sulphide bridge formation which will be lost after this change with serine thus it will affect the stability of protein structure and ultimately loss of function. The other non-sense variant (W345*) in IDS was previously identified in Japanese patient affected with sever phenotypes of MPS II (Sukegawa et al., 1995). The authors found a smaller precursor protein in cells transfected with W345*. Therefore, it is predicted that the variant leads to the production of truncated enzyme. The affected individual in family N showed mild phenotypes of short stature, coarse facial features and hernia as compared to those reported by Sukegawa et al. (1995).

Our study increases the knowledge of genetic factors involved in rare diseases. For this purpose, the Pakistani population is unique due to the high number of families with consanguineous marriages and the high frequency of large pedigrees. Furthermore, we have used WGS instead of phenotype-based diagnosis to diagnose subtypes of disorders having overlapping phenotypes. Thus, WGS is a fruitful strategy to identify novel causal homozygous variants in inbred populations.

The findings of this study have to be taken with some limitations. For example, limited clinical details were gathered from the recruited families due to limited resources like testing facilities at tertiary care hospital at the time of recruitment of the patients. The overlapping phenotypes among the recruited cases presented in this study is one of the main reasons for pseudo-identification of various pathologies diagnosed initially as MPS. After getting results from sequencing data, the families were recontacted for more detailed evaluation but most of the recruited patients have died and families flatly declined to provide more information. Thus, the present study could not include functional enzymatic studies and comprehensive analysis of 3D protein structure. Performing *in-vitro* and *in-vivo* functional studies would help in examining detailed biological effects of the identified variants. Further studies including more consanguineous families with similar phenotypes and combined with functional analysis should be performed to clarify the effect of the identified variants.

Genetic findings in the affected families showed that the diagnosis based on phenotypic presentation is not an optimal method. For this reason, the incorporation of next-generation sequencing (NGS) in consanguineous families represents one of the best ways to elucidate the underlying causes in rare diseases such as MPS I and MPS II and establish a genotype-phenotype relationship (Hoffman et al., 2013; Umair et al., 2018). NGS will not only be helpful in the correct diagnosis of the disease but it will also help in prenatal testing, family planning, carrier testing and genetic counseling in families affected with LSDs (Komlosi et al., 2016). Similarly, dual molecular diagnosis has been reported in several cases of consanguineous unions (Umair et al., 2017), which cannot be detected through conventional sequencing or enzymatic testing. Therefore, NGS is required to identify all the pathogenic and modifier variants in heterogeneous disorders like LSDs. Sequencing highly homologous regions (for example, *GBA1* locus) using Sanger or short-read NGS methods in suspected cases of LSDs can miss other mutant alleles. Because, the shorter reads cannot be mapped uniquely to the reference genome, especially in cases where there are recombinant alleles aligning to the homologous region. Therefore, long-read NGS is recommended in such cases to discriminate the functional genes from their pseudogenes (Woo et al., 2021).

Accurate and early diagnosis of LSDs in children represents a helpful step for designing therapeutic strategies to save different organs from permanent degeneration. Pre-natal screening and identification of carriers' status in an affected family for LSDs will be helpful for genetic counselling of the family. The identification of a novel MPS I gene (*ABCA5*) based on WGS proves to be helpful for expanding the limits of targeted therapies to treat LSDs.

Data availability statement

The data presented in the study are deposited in the ClinVar repository, accession numbers: SCV003842213 (*IDS*:c.1035G>T); SCV003842212 (*IDS*:c.1264T>G); SCV003842211 (*MASP1*:c.73A>T); SCV003842210 (*KIAA1109*:c.1297C>G); SCV003842209 (*ABCA5*:c.2569C>T); SCV003842208 (*SMARCAL1*:c.1810C>T); SCV003842207 (*HERC1*:c.5941G>A); SCV003842206 (*IDS*:c.353T>A); SCV003842205 (*GAA*:c.1930G>T); SCV003844061 (*RRAS2*:c.439C>T).

Ethics statement

The studies involving human participants were reviewed and approved by Bioethical Review board, Quaid-i-Azam University, Islamabad, Pakistan. Written informed consent to participate in this

study was provided by the participants' legal guardian/next of kin. The animal study was reviewed and approved by Bioethical Review board, Quaid-i-Azam University, Islamabad, Pakistan.

Author contributions

SF and TH designed and supervised the study experiments. RG and MH collected data and conducted the clinical assessments. RG performed initial lab work. MS, AU, EP, AT, FS, AG, and AA performed WGS and data analysis. RG, AU, and EP wrote the manuscript. All authors revised and approved the article.

Funding

The study was supported by the Novo Nordisk Foundation through the Challenge Program (NNF18OC0033950), and Higher Education Commission, Islamabad, Pakistan under International Research Support Initiative Program (IRSIP).

Acknowledgments

We would like to thank all participating families, staff and colleagues who helped us during this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1128850/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Schematic diagram of the study showing recruitment and clinical diagnosis of affected families followed by whole genome sequencing and variants prioritization.

References

Arora, R. S., Mercer, J., Thornley, M., Tylee, K., and Wraith, J. E. (2007). Enzyme replacement therapy in 12 patients with MPS I–H/S with homozygous p. Leu490Pro mutation. *J. Inherit. Metab. Dis.* 30, 821. doi:10.1007/s10545-007-0551-9

Brooks-Wilson, A., Marcil, M., Clee, S. M., Zhang, L. H., Roomp, K., van Dam, M., et al. (1999). Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat. Genet.* 22, 336–345. doi:10.1038/11905

Capri, Y., Flex, E., Krumbach, O. H., Carpentieri, G., Cecchetti, S., Lisewski, C., et al. (2019). Activating mutations of RRAS2 are a rare cause of Noonan syndrome. *Am. J. Hum. Genet.* 104, 1223–1232. doi:10.1016/j.ajhg.2019.04.013

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, 884–890. doi:10.1093/bioinformatics/bty560

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. doi:10.1093/gigascience/giab008

DeStefano, G. M., Kurban, M., Anyane-Yeboa, K., Dall'Armi, C., Di Paolo, G., Feenstra, H., et al. (2014). Mutations in the cholesterol transporter gene ABCA5 are associated with excessive hair overgrowth. *PLoS Genet.* 10, e1004333. doi:10.1371/journal.pgen.1004333

Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi:10.1093/bioinformatics/btw354

Grimberg, J., Nawoschik, S., Belluscio, L., McKee, R., Turck, A., and Eisenberg, A. (1989). A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Res.* 17, 8390. doi:10.1093/nar/17.20.8390

Gul, R., Firasat, S., Hussain, M., Afshan, K., and Nawaz, D. (2020). IDUA gene mutations in mucopolysaccharidosis type-1 patients from two Pakistani inbred families. *Congenit. Anom.* 60, 126–127. doi:10.1111/cga.12354

Gul, R., Firasat, S., Hussain, M., Tufail, M., Ahmad, W., and Afshan, K. (2021). Neurological manifestations in Pakistani lysosomal storage disorders patients and molecular characterization of Gaucher disease. *Genetika* 53, 1017–1029. doi:10.2298/gensr2103017g

Hoffman, J. D., Greger, V., Strovel, E. T., Blitzer, M. G., Umbarger, M. A., Kennedy, C., et al. (2013). Next-generation DNA sequencing of hexa: A step in the right direction for carrier screening. *Mol. Genet. Genomic* 1, 260–268. doi:10.1002/mgg3.37

Hussain, R., and Bittles, A. H. (1998). The prevalence and demographic characteristics of consanguineous marriages in Pakistan. *J. Biosoc. Sci.* 30, 261–275. doi:10.1017/s0021932098002612

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581, 434–443. doi:10.1038/s41586-020-2308-7

Kaur, P., Kadavigere, R., Girisha, K. M., and Shukla, A. (2020). Recurrent bi-allelic splicing variant c.454+ 3A>G in TRAPPC4 is associated with progressive encephalopathy and muscle involvement. *Brain* 143, e29. doi:10.1093/brain/awaa046

Kingma, S. D., Bodamer, O. A., and Wijburg, F. A. (2015). Epidemiology and diagnosis of lysosomal storage disorders; challenges of screening. *Best. Pract. Res. Clin. Endocrinol. Metab.* 29, 145–157. doi:10.1016/j.beem.2014.08.004

Kingma, S. D., and Jonckheere, A. I. (2021). MPS I: Early diagnosis, bone disease and treatment, where are we now? *J. Inherit. Metab. Dis.* 44, 1289–1310. doi:10.1002/jimd.12431

Komlosi, K., Solyom, A., and Beck, M. (2016). The role of Next-Generation Sequencing in the diagnosis of lysosomal storage disorders. *J. inborn errors Metab. screen* 4, 232640981666937–6. doi:10.1177/2326409816669376

Kubo, Y., Sekiya, S., Ohigashi, M., Takenaka, C., Tamura, K., Nada, S., et al. (2005). ABCA5 resides in lysosomes, and ABCA5 knockout mice develop lysosomal disease-like symptoms. *Mol. Cell Biol.* 25, 4138–4149. doi:10.1128/MCB.25.10.4138-4149.2005

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* [Preprint]. Available at: https://arxiv.org/abs/1303.3997.

Linhart, A., and Elliott, P. M. (2007). The heart in Anderson-Fabry disease and other lysosomal storage disorders. *Heart* 93, 528–535. doi:10.1136/hrt.2005.063818

Marques, A. R., and Saftig, P. (2019). Lysosomal storage disorders–challenges, concepts and avenues for therapy: Beyond rare diseases. *J. Cell Sci.* 132, jcs221739. doi:10.1242/jcs.221739

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi:10.1101/gr.107524.110

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The ensembl variant effect predictor. *Genome Biol.* 17, 122–14. doi:10.1186/s13059-016-0974-4

Mokhtariye, A., Hagh-Nazari, L., Varasteh, A. R., and Keyfi, F. (2019). Diagnostic methods for lysosomal storage disease. *Rep. Biochem. Mol. Biol.* 7, 119–128.

Ortega-Recalde, O., Beltrán, O. I., Gálvez, J. M., Palma-Montero, A., Restrepo, C., Mateus, H., et al. (2015). Biallelic HERC1 mutations in a syndromic form of overgrowth and intellectual disability. *Clin. Genet.* 88, e1–e3. doi:10.1111/cge.12634

Parenti, G., Andria, G., and Ballabio, A. (2015). Lysosomal storage diseases: From pathophysiology to therapy. *Annu. Rev. Med.* 66, 471–486. doi:10.1146/annurev-med-122313-085916

Parenti, G., and Giugliani, R. (2022). "The mucopolysaccharidoses," in *Physician's guide to the diagnosis, treatment, and follow-up of inherited metabolic diseases* (Cham: Springer), 1267–1286.

Raza, R., Ullah, A., Haider, N., Krishin, J., Shah, M., Khan, F. U., et al. (2020). Exome sequencing reveals the first intragenic deletion in ABCA5 underlying autosomal recessive hypertrichosis. *Clin. Exp. Dermatol* 47, 1137–1143. doi:10.1111/ced.15128

Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* 18, 696–704. doi:10.1038/gim.2015.148

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2020). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17, 405–424. doi:10.1038/gim.2015.30

Schubert, M., Ermini, L., Der Sarkissian, C., Jonsson, H., Ginolhac, A., Schaefer, R., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat. Protoc.* 9, 1056–1082. doi:10.1038/nprot.2014.063

Schuh, R. S., Gonzalez, E. A., Tavares, A. M. V., Seolin, B. G., Elias, L. d. S., Vera, L. N. P., et al. (2029). Neonatal nonviral gene editing with the CRISPR/Cas9 system improves some cardiovascular, respiratory, and bone disease features of the mucopolysaccharidosis I phenotype in mice. *Gene Ther.* 27, 74–84. doi:10.1038/s41434-019-0113-4

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29, 308–311. doi:10.1093/nar/29.1.308

Sukegawa, K., Tomatsu, S., Fukao, T., Iwata, H., Song, X. Q., Yamada, Y., et al. (1995). Mucopolysaccharidosis type II (hunter disease): Identification and characterization of eight point mutations in the iduronate-2-sulfatase gene in Japanese patients. *Hum. Mutat.* 6, 136–143. doi:10.1002/humu.1380060206

Tieu, P. T., Bach, G., Matynia, A., Hwang, M., and Neufeld, E. F. (1995). Four novel mutations underlying mild or intermediate forms of α-L-iduronidase deficiency (MPS IS and MPS IH/S). *Hum. Mutat.* 6, 55–59. doi:10.1002/humu.1380060111

Tomatsu, S., Fukuda, S., Cooper, A., Wraith, J. E., Rezvi, G. M., Yamagishi, A., et al. (1995). Mucopolysaccharidosis type IVA: Identification of six novel mutations among non-Japanese patients. *Hum. Mol. Genet.* 4, 741–743. doi:10.1093/hmg/4.4.741

Ullah, I., Nasir, A., Mehmood, S., Ahmed, S., Ullah, M. I., Ullah, A., et al. (2017). Identification and *in silico* analysis of GALNS mutations causing Morquio A syndrome in eight consanguineous families. *Turk J. Bio* 41, 458–468. doi:10.3906/biy-1607-81

Umair, M., Ahmad, F., and Ullah, A. (2018). Whole exome sequencing as a diagnostic tool for genetic disorders in Pakistan. *Pak J. Med. Res.* 57, 90–91.

Umair, M., Seidel, H., Ahmed, I., Ullah, A., Haack, T. B., Alhaddad, B., et al. (2017). Ellis-van Creveld syndrome and profound deafness resulted by sequence variants in the EVC/EVC2 and TMC1 genes. *J. Genet.* 96, 1005–1014. doi:10.1007/s12041-017-0868-6

Van Bergen, N. J., Guo, Y., Al-Deri, N., Lipatova, Z., Stanga, D., Zhao, S., et al. (2020). Deficiencies in vesicular transport mediated by TRAPPC4 are associated with severe syndromic intellectual disability. *Brain* 143, 112–130. doi:10.1093/brain/awz374

Weber, B., Guo, X. H., Wraith, J. E., Cooper, A., Kleijer, W. J., Bunge, S., et al. (1997). Novel mutations in sanfilippo A syndrome: Implications for enzyme function. *Hum. Mol. Genet.* 6, 1573–1579. doi:10.1093/hmg/6.9.1573

Woo, E. G., Tayebi, N., and Sidransky, E. (2021). Next-Generation Sequencing analysis of GBA1: The challenge of detecting complex recombinant alleles. *Front. Genet.* 12, 684067. doi:10.3389/fgene.2021.684067

World Medical Association (2013). World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *Jama* 310, 2191–2194.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 September 2023
DOI 10.3389/fgene.2023.1254909

Check for updates

OPEN ACCESS

EDITED BY
Ammar Husami,
Cincinnati Children's Hospital Medical Center, United States

REVIEWED BY
Abdul Nasir,
Second Affiliated Hospital of Zhengzhou University, China
Elsayed Abdelkreem,
Sohag University, Egypt

*CORRESPONDENCE
Sabika Firasat,
sabika.firasat@qau.edu.pk
Torben Hansen,
torben.hansen@Sund.ku.dk

RECEIVED 12 July 2023
ACCEPTED 17 August 2023
PUBLISHED 11 September 2023

CITATION
Gul R, Firasat S, Schubert M, Ullah A, Peña E, Thuesen ACB, Gjesing AP, Hussain M, Tufail M, Saqib M, Afshan K and Hansen T (2023), Identification of genetic variants associated with a wide spectrum of phenotypes clinically diagnosed as Sanfilippo and Morquio syndromes using whole genome sequencing.
Front. Genet. 14:1254909.
doi: 10.3389/fgene.2023.1254909

COPYRIGHT
© 2023 Gul, Firasat, Schubert, Ullah, Peña, Thuesen, Gjesing, Hussain, Tufail, Saqib, Afshan and Hansen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of genetic variants associated with a wide spectrum of phenotypes clinically diagnosed as Sanfilippo and Morquio syndromes using whole genome sequencing

Rutaba Gul[1,2], Sabika Firasat[1]*, Mikkel Schubert[2], Asmat Ullah[2], Elionora Peña[2], Anne C. B. Thuesen[2], Annete P. Gjesing[2], Mulazim Hussain[3], Muhammad Tufail[1], Muhammad Saqib[1,4], Kiran Afshan[1] and Torben Hansen[2]*

[1]Department of Zoology, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan, [2]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [3]The Children Hospital, Pakistan Institute of Medical Sciences (PIMS), Islamabad, Pakistan, [4]Department of Zoology, University of Lakki Marwat, Lakki Marwat, Khyber Pakhtunkhwa, Pakistan

Mucopolysaccharidoses (MPSs) are inherited lysosomal storage disorders (LSDs). MPSs are caused by excessive accumulation of mucopolysaccharides due to missing or deficiency of enzymes required for the degradation of specific macromolecules. MPS I-IV, MPS VI, MPS VII, and MPS IX are sub-types of mucopolysaccharidoses. Among these, MPS III (also known as Sanfilippo) and MPS IV (Morquio) syndromes are lethal and prevalent sub-types. This study aimed to identify causal genetic variants in cases of MPS III and MPS IV and characterize genotype-phenotype relations in Pakistan. We performed clinical, biochemical and genetic analysis using Whole Genome Sequencing (WGS) in 14 Pakistani families affected with MPS III or MPS IV. Patients were classified into MPS III by history of aggressive behaviors, dementia, clear cornea and into MPS IV by short trunk, short stature, reversed ratio of upper segment to lower segment with a short upper segment. Data analysis and variant selections were made based on segregation analysis, examination of known MPS III and MPS IV genes, gene function, gene expression, the pathogenicity of variants based on ACMG guidelines and *in silico* analysis. In total, 58 individuals from 14 families were included in the present study. Six families were clinically diagnosed with MPS III and eight families with MPS IV. WGS revealed variants in MPS-associated genes including *NAGLU, SGSH, GALNS, GNPTG* as well as the genes *VWA3B*, *BTD*, and *GNPTG* which have not previously associated with MPS. One family had causal variants in both *GALNS* and *BTD*. Accurate and early diagnosis of MPS in children represents a helpful step for designing therapeutic strategies to protect different organs from permanent damage. In addition, pre-natal screening and identification of genetic etiology will facilitate genetic counselling of the affected families. Identification of novel causal MPS genes might help identifying new targeted therapies to treat LSDs.

KEYWORDS

Sanfilippo syndrome, Morquio syndrome, Pakistani families, whole genome sequencing, *VWA3B*

Introduction

Background

Mucopolysaccharidoses (MPS) is a group of disorders belonging to lysosomal storage disorders. MPS are sub categorized into seven types (MPS I, II, III, IV, VI, VII, and IX) mostly inherited as autosomal recessive except for MPS II. Each type is caused by the deficiency of mucopolysaccharide-degrading enzymes. Due to the deficiency of these enzymes, the mucopolysaccharides (glycosaminoglycans) accumulate in different organs of the body including the arteries, skeleton, eyes, joints, ears, skin, teeth, respiratory system, liver, spleen, central nervous system, blood, and bone marrow. This excessive accumulation of mucopolysaccharides result in phenotypes of MPS overlapping with one another, making it difficult to diagnose without use of specific diagnostic approaches like the measurement of enzyme activity in leukocytes or molecular diagnosis (Besley and Wraith, 1997; Chih-Kuang et al., 2002).

Mucopolysaccharidosis type III (MPS III) or Sanfilippo syndrome is a subtype of MPS discovered 50 years ago with an autosomal recessive mode of inheritance. The disease affects 0.3–4.1 cases per 100,000 births (Sanfilippo, 1963; Valstar et al., 2008; Fedele, 2015; Khan et al., 2017). It has been further classified into four categories (MPS type IIIA—D) caused by variants in *SGSH*, *NAGLU, HGSNAT*, and *GNS* respectively. Enzymes encoded by these genes are involved in the degradation of a linear polysaccharide known as heparan sulfate (HS). HS has a crucial role in the development of central nervous system (CNS) which explains why patients suffering from Sanfilippo syndrome develop dementia from early childhood (De Pasquale and Pavone, 2019) and have neurological manifestations such as delayed behavioral and developmental milestones starting from early years of life (Shapiro et al., 2017; Ozkinay et al., 2021).

Mucopolysaccharidosis type IV, also known as Morquio syndrome, was first discovered by a pediatrician Luis Morquio in 1929. It has an autosomal recessive mode of inheritance (Morquio, 1929). It is a rare disorder with prevalence of 1 in 40,000 live births (Nelson et al., 2003). There are two types of Morquio syndrome (type A and type B) caused by pathogenic variants in *GALNS* and *GLB1,* respectively. The gene *GALNS* encodes the GALNS enzyme which has a role in the degradation of glycosaminoglycan including chondroitin-6-sulfate (C6S) and keratin sulfate (KS), while GLB1 encoded by *GLB1* is involved in the degradation of keratin sulfate only (Celik et al., 2021). The GLB1 enzyme is also involved in the hydrolyzation of terminal beta galactosyl residues of GM1 gangliosides and glycoproteins (Okada and John, 1968).

Diagnosis of MPS III involves either urinary HS excretion measurement or enzyme measurement from dry blood spot through tandem mass spectrometry or fluorimetry (Seker Yilmaz et al., 2021). The diagnosis of MPS IV involves radiological findings and urinary excretion or liquid chromatography–mass spectrometry (LC-MS) or MS-based method for KS measurements in plasma or urine (Piraud et al., 1993; Martell et al., 2011; Peracha et al., 2018; Chien et al., 2020; Lin et al., 2020). Due to enzymatic pseudo-deficiency, molecular genetic screening of the patients affected with MPS III or MPS IV is the most accurate method to diagnose any subtype of these disorders.

To the best of our knowledge, no prior molecular genetics research has been conducted in Pakistan regarding Sanfilippo cases. However, various international research groups have published studies highlighting different genetic variations found in cases of Pakistani descent across the globe. Notably, the literature includes instances of novel or previously reported variants of Sanfilippo syndrome originating from Pakistani patients. These include a unique deletion that triggers Sanfilippo syndrome type D, specifically the c.1169delA mutation within exon 10, resulting in a premature termination codon (Beesley et al., 2003). Additionally, a new nonsense variant associated with Sanfilippo syndrome type C, identified as p.Ser296Ter in exon 10, as well as a splice site variant designated c.744–2A>G in intron 7, have been documented (Feldhammer et al., 2009a; Feldhammer et al., 2009b). Several studies have been conducted in Pakistan related to cases of Morquio syndrome, revealing a total of seven new variants—namely, p. Phe216Ser, p. Met38Arg, p. Ala291Ser, p. Glu121Argfs*37, p. Tyr294Terfs—alongside two previously reported variants, p. Pro420Arg and p. Arg386Cys (Tomatsu et al., 2004; Morrone et al., 2014; Ullah et al., 2017; Zubaida et al., 2018).

Treatment options for MPS III and IV include enzyme replacement therapy (ERT) (MPS IV A), molecular chaperon therapy (MCT), substrate reduction therapy (SRT), hematopoietic stem cell transplant (HSCT), *in vivo* adeno-associated viral gene therapy, antioxidant therapy, inhibiting protein aggregation, stop-codon readthrough therapy and anti-inflammatory therapy (Hendriksz et al., 2013; Akyol et al., 2019; Beneto et al., 2020; Seker Yilmaz et al., 2021). In a recent study, recombinant human GALNS enzyme (rhGALNS) was infused with hydrogel like polyethylene glycol to deliver the exogenous enzyme successfully into the fibroblasts for sustainable and longer release (Jain et al., 2020). Molecular genetic testing, prenatal screening, carrier testing, and genetic counselling can help in the prevention of the disease and/or saving organs from permanent degeneration by allowing early diagnosis and appropriate treatment options (where available). This study was designed to evaluate the clinical profiles of Sanfilippo and Morquio syndromes patients from Pakistan, to identify causal gene variants using WGS and to correlate the genotypes with phenotypes.

Materials and methodology

Ethical approval and patient recruitment

The study was approved by the bioethical committee of Quaid-i-Azam University (BEC-FBS-QAU2019-198), Islamabad, Pakistan.

Patients were recruited and diagnosed based on their presenting clinical profiles (Table 1; Table 3) at outpatient department (OPD), Children Hospital, Pakistan Institute of Medical Sciences (PIMS), Islamabad, Pakistan. Before collecting blood samples from patients and accompanying parents/siblings, informed written consent forms were signed by the participants in accordance with the Declaration of Helsinki (World Medical Association, 2013). The detailed pedigrees were drawn using the information provided by well-informed elders of the family (Figure 1; Figure 2). Localities of families from different regions of Pakistan is mentioned in Figure 3.

10.3389/fgene.2023.1254909

TABLE 1 Clinical details of the patients suspected as Sanfilippo syndrome affected.

Pedigree no.	Individual no.	No. of affected individuals	Age/ Sex	Cast/Region/ Consanguinity	Phenotypes/Clinical profile
A	VII-4	1	5 y/M	Sardar/AJK/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior (sometimes), speech delay, short stature, increased body hair, coarse facial features, broad based gait, joint contracture, short trunk, hepatosplenomegaly, self-biting, intellectual disability ALT = 49 U/L (up to 40), AST = 46 U/L (up to 40), Bilirubin total = 1.9 mg/dL (<1.2), Hb = 11.1 g/dL (12–1.4), RBC count = 4.08 mil/cmm (4.75–4.85), Hematocrit = 34% (36–42), Monocytes = 06% (4–5), Serum Sodium = 155 mmol/L (132–145), Serum Potassium = 5.5 mmol/L (3.1–5.1), Serum Ferritin = 27.84 ng/mL (30–400), beaking at L1 and L2 vertebrae, Glucose (R) = 76 mg/dL (80–130)
B	VII-2	5	14.5 y/M	Khawaja/Punjab/Yes	Delay in achieving milestones, chronic diarrhea, abdominal distention, aggressive behavior, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, low IQ
	VII-3		9 y/M		Delay in achieving milestones, chronic diarrhea, abdominal distention, aggressive behavior, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, low IQ
	VII-4		10.5 y/F		Delay in achieving milestone, chronic diarrhea, abdominal distention, aggressive behaviors, few words speech only, hernia (surgery), short stature, constipation, coarse facial features, hepatosplenomegaly, low IQ Hb = 9.9 g/dL (13.5–19.5), RBC = 3.92×10^6/ul (4.0–5.9), MCH = 25.2 pg (28–32), MCHC = 27.7 g/dL (29–37), Neutrophils = 76.1% (50–75), Eosinophils = 0.1% (1–6)
C	V-1	2	5 y/F	Aryn/Punjab/Yes	Delay in achieving milestones, chronic diarrhea, aggressive behavior, speech delay, short stature, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability
	V-4		1.5 y/F		Delay in achieving milestones, chronic diarrhea, short stature, joint contractures, hepatosplenomegaly, intellectual disability
D	IV-2	2	1 y/M	Qaisrani Baloch/ICT/No	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, respiratory and ear infection (Otitis media), coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Mild cerebral atrophy, mild mucosal thickness of right ethmoid and right maxillary sinus, Urea = 10 mg/dL (15–45), Creatinine = 0.5 mg/dL (0.7–1.4), Uric acid = 2.8 md/dL (3.6–7.7), Serum Ferritin = 4.75 µg/L (30–400), Hb = 10.1 g/dL (14–18), HCT = 36.5% (41–51), MCV = 69.4 fL (80.97), MCH = 19.2 pg (26–32), MCHC = 27.7 g/dL (32–36), Microcytes = Positive, Hypochromia = Positive, Poikilocytes = Positive
	IV-1		2 y/F		Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, short stature, respiratory and ear infection, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability 3rd and 4th ventricle dilation, Mild periventricular

(Continued on following page)

TABLE 1 (*Continued*) Clinical details of the patients suspected as Sanfilippo syndrome affected.

Pedigree no.	Individual no.	No. of affected individuals	Age/ Sex	Cast/Region/ Consanguinity	Phenotypes/Clinical profile
					ischemia, Wide cisterna magna, Mild cerebral atrophy, Urea = 12 mg/dL (15–45), Creatinine = 0.5 g/dL (0.6–1.1), ALT = 39 u/L (5–31), Hb = 10.1 g/dL (12–16), HCT = 35.7% (37–47), MCV = 67.2 fL (80–97), MCH = 19 pg (26–32), MCHC = 28.3 g/dL (32–36), Microcytes = Positive, Hypochromia = Positive, Poikilocytes = Positive
E	VI-1	2	14 y/M	Pathan/KPK/Yes	Delay in achieving milestone, aggressive behavior, speech delay, gibbus formation, short stature, respiratory/ear infection, coarse facial features, excessive hair growth, broad based gait/waddling, joint contractures, short stature, scoliosis/kyphosis, intellectual disability Diffuse cerebral cortical atrophy, AV and VA concordance, Dolichocephaly, Mild atrophy, Hb = 11.4 g/dL (14–18), Neutrophils = 44% (50–75), Lymphocytes = 49% (20–40), Monocytes = 01% (2–6)
	VI-2		8 y/M		Delay in achieving milestone, aggressive behavior, speech delay, gibbus formation, short stature, respiratory/ear infection, coarse facial features, excessive hair growth, broad based gait/waddling, joint contractures, short stature, scoliosis/kyphosis, intellectual disability Small acquired porencephalic cyst left peri-sylvian mid parietal region, Hb = 11 g/dL (14–18), WBCs = 3900/cmm (4000–11000), Lymphocytes = 47% (20–45), Monocytes = 01% (2–6)
F	IV-1	2	10 y/F	Kaka khel/Punjab/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, respiratory/ear infection (decreased hearing), coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Spina bifida S1, Irregular acetabular margins, Pallor = Positive, Thick calvarium
	IV-2		7 y/F		Delay in achieving milestones, abdominal distention, aggressive behavior, speech delay, respiratory/ear infection, coarse facial features, joint contractures, hepatosplenomegaly, intellectual disability Hb = 11.7 g/dL (12–15), Neutrophils = 16.7% (60–70), Eosinophils = 25.2% (0.0–4.0), Irregular acetabular with small lytic areas with small femoral epiphysis, Spina bifida S1, Thick calvarium, J-shaped sella, Hip joint juvenile arthritis

AJK, azad jammu and kashmir; ALT, alanine transaminase; AST, aspartate aminotransferase; AV/VA, Atrioventricular/Ventriculoarterial; F*, father; F, female; Hb, Hemoglobin; HCT, hematocrit; ICT, islamabad capital territory; IQ, intelligent quotient; KPK, khyber pakhtunkhwa; M, male; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; MCH, mean corpuscular hemoglobin; M*, mother; P, patient; RBC, red blood cells; S, sibling; U, uncle; WBCs, White blood cells; y, years.

In presenting clinical profiles, height of each MPS IV patient was measured and compared to the healthy individual of same age to confirm their stature. Height is plotted on Center for Disease Control and Prevention (CDC) charts, used globally and a height that is two standard deviations below the mean height for age and sex (less than the 3rd percentile) is considered as short stature (Barstow and Rerucha, 2015). Complete skeletal survey through x-ray was performed which showed dysostosis multiplex, fish mouth vertebrae and bullet shaped metacarpals in Morquio syndrome patients.

Sampling and DNA extraction

Peripheral blood samples of 3–5 mL were collected from 58 individuals from 14 families, covering 26 affected and 32 unaffected individuals. The blood samples were collected in 10 mL EDTA vacutainers (BD vacutainer K2 EDTA 18 mg). DNA extraction was performed using the Phenol-Chloroform method and DNA was quantified using a µDrop Plate reader (Multiskan™, Thermo Fisher Scientific, and Waltham, MA, United States).

10.3389/fgene.2023.1254909

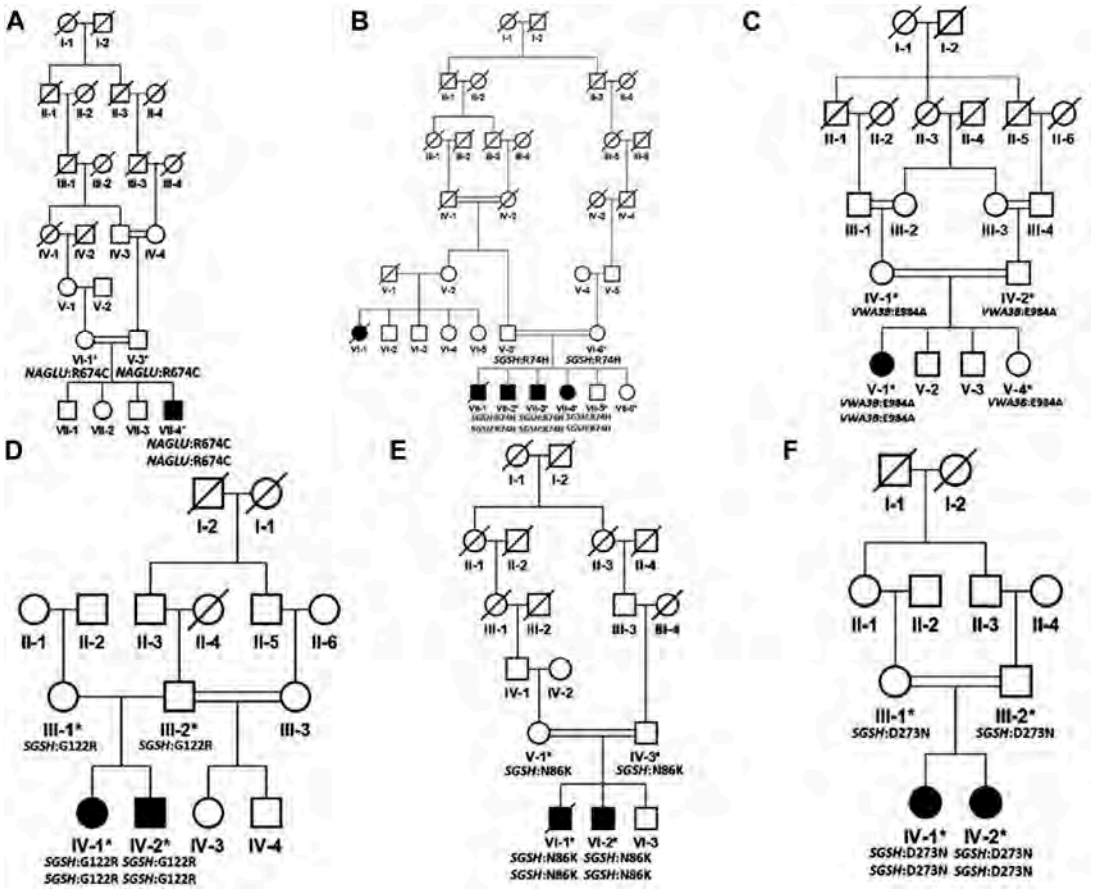


FIGURE 1
Pedigrees of families inheriting MPS III phenotypes **(A–F)** consistent with autosomal recessive mode of inheritance. Square and circles denote males and females respectively; filled symbols indicate affected individuals and consanguinity is represented by double marriage lines.

Whole genome sequencing and variant prioritization

Whole genome sequencing, reads mapping, genotype calling, pre-analysis quality assurance, and annotation were performed as describe previously by Gul et al. (2023).

Briefly, mapping and genotype calling was performed using a modified version of the PALEOMIX pipeline (https://pubmed.ncbi.nlm.nih.gov/24722405/), with initial quality assurance performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC (https://pubmed.ncbi.nlm.nih.gov/27312411/). Reads were processed using fastp (https://pubmed.ncbi.nlm.nih.gov/30423086/) and mapped against the hg38 human reference using BWA (https://arxiv.org/abs/1303.3997). Alignments were post-processed using samtools (https://pubmed.ncbi.nlm.nih.gov/33590861/) and 'bwa-postalt.js' from the BWA-kit. BAMs were calibrated, and genotypes were called and calibrated using GATK (https://pubmed.ncbi.nlm.nih.gov/20644199/). The resulting VCF was annotated using VEP (https://pubmed.ncbi.nlm.nih.gov/27268795/).

Initially, analysis of variants identification was based on MPS-associated genes. Variants in the exomes were extracted from whole genome sequencing data of all individuals. The variants including non-synonymous and splice sites having minor allele frequency (MAF) greater than 0.01 in larger outbred populations were excluded by using gnomAD. In pedigrees with more than one affected individual, variants shared by multiple patients were prioritized. Based on consanguinity and inheritance pattern, homozygous variants were selected and prioritized. During re-analysis, heterozygous variants were also screened for compound heterozygosity. The prioritization was based on segregation, gene function, association with disease and pathogenicity prediction according to ACMG guidelines and *in silico* analysis (Richards et al., 2015). Genes were annotated according to the following transcripts: NM_000263.4 (*NAGLU*); NM_144992 (*VWA3B*); NM_000199.5 (*SGSH*); NM_000512.5 (*GALNS*); NM_001281723.2 (*BTD*); NM_032520.5 (*GNPTG*).

All variants were found in homozygous state in patients while in heterozygous state in parents. Phenotypically normal siblings in the present families were either homozygous wild type or heterozygous carriers for that particular variant. All identified variants are listed in Table 2 and Table 4. The variants were evaluated using ACMG guidelines and were predicted as pathogenic, likely pathogenic or uncertain significance. To exclude non-disease associated variants, the genomes of more

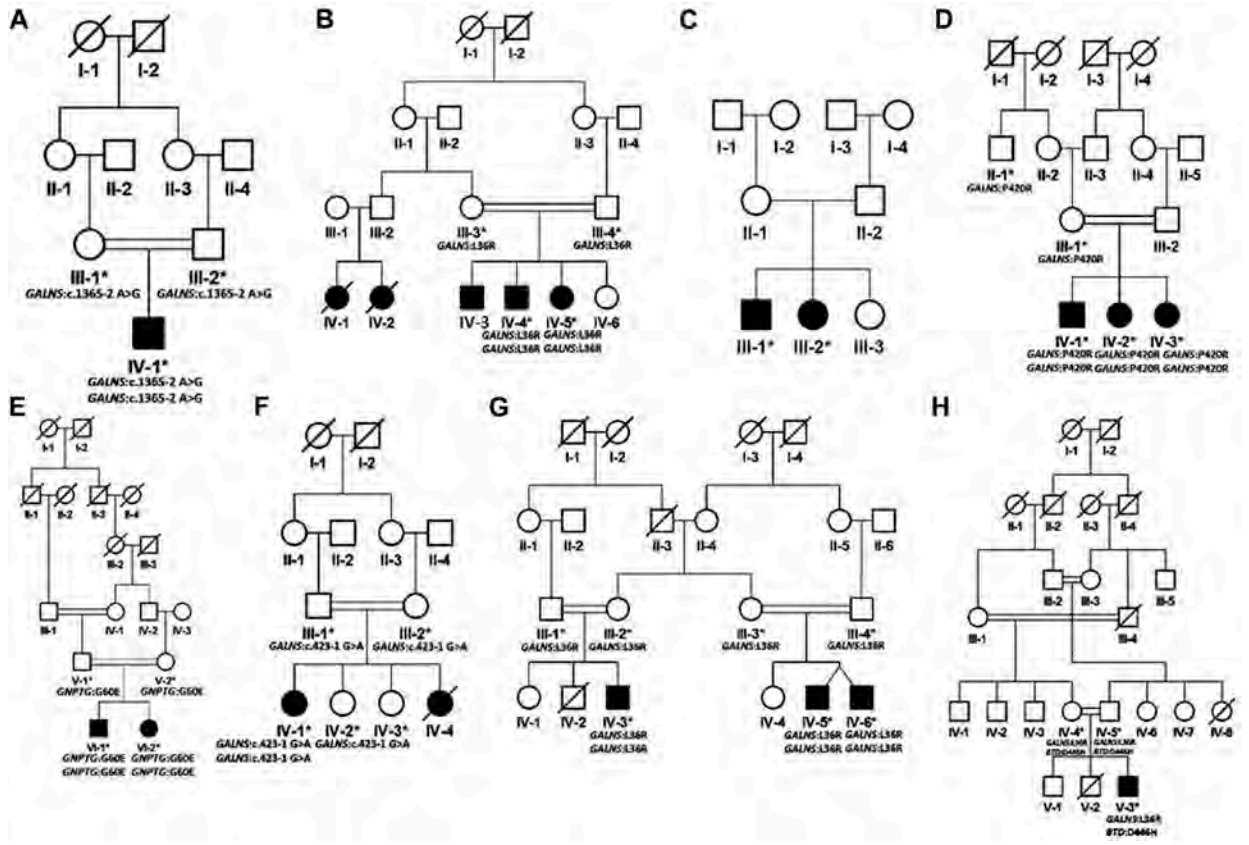


FIGURE 2
Pedigrees of families inheriting MPS IV phenotypes **(A–H)** consistent with autosomal recessive mode of inheritance. Square and circles denote males and females respectively; filled symbols indicate affected individuals and consanguinity is represented by double marriage lines.

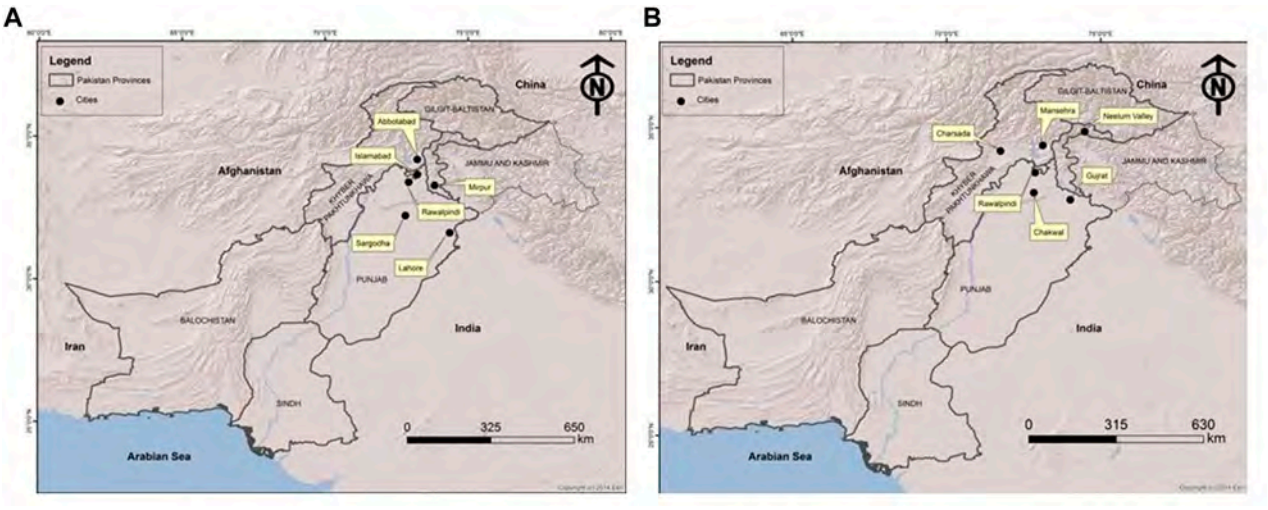


FIGURE 3
Map of Pakistan showing the locations (circles) of cities/towns from where families included in this study were collected. **(A)** All MPS III collected families. **(B)** All MPS IV collected families.

TABLE 2 *In-silico* **analysis and ACMG classification of the identified variants in suspected Sanfilippo syndrome cases.**

Pedigree/ Patient ID	Variants	Patient genotype	dbSNP	In-silico tools with deleterious output	ACMG classification	VarSome	HGMD accession number	ClinVar accession numbers
A/VII-4	NM_000263.4 (NAGLU), c.2020C>T (p.Arg674Cys)	Homozygous	rs763299645	9/9	Pathogenic PP5, PP3, PM5, PM2	Pathogenic	CM981369	SCV004012870
B/VII-2, VII-3, VII-5	NM_000199.5 (SGSH), c.221C>T (p.Arg74His)	Homozygous	rs778336949	9/9	Pathogenic PP5, PS3, PM5, PP3, PM2	Pathogenic	CM971354	SCV004012871
C/V-1, V-4	NM_144992 (VWA3B), c.2951A>C (p.Glu984Ala)	Homozygous	N/A	8/9	Likely benign BP1, BP4, PM2	Uncertain significance	Current study	SCV004012872
D/IV-1, IV-2	NM_000199.5 (SGSH), c.364C>T (p.Gly122Arg)	Homozygous	rs761607612	9/9	Pathogenic PP5, PP3, PM2	Pathogenic	CM971358	SCV004012873
E/VI-1, VI-2	NM_000199.5 (SGSH), c.258A>C (p.Asn86Lys)	Homozygous	N/A	8/9	Uncertain significance PM2	Uncertain significance	Current study	SCV004012874
F/IV-1, IV-2	NM_000199.5 (SGSH), c.817G>A (p.Asp273Asn)	Homozygous	rs1046551417	7/9	Likely pathogenic PP3, PM2, PP5	Likely pathogenic	CM002397	SCV004012875

BP, possibly benign; N/A, not available; PP, possibly pathogenic; PM, moderate pathogenic; PS, strong pathogenic.

than 100 ethnically matched *in-house* control samples were screened. The analysis plan for this study was same as mentioned in previously (Gul et al., 2023).

Protein modeling and protein-protein interaction

To generate the complete three-dimensional (3D) structures of *BTD, SGSH, VWA3B, GNTGP, GALNS, and NAGLU* proteins I-TASSER (Iterative Threading Assembly Refinement) (https://zhanggroup.org/I-TASSER/) was utilized. This is a bioinformatics tool that generates 3D protein structures using hierarchical techniques, first searching the PDB library for templates and then building the models using iterative template fragment assembly simulation. The degree of appropriateness of all structures were verified through VERIFY3D and ERRAT. UCSF Chimera software 1.14 was used to inspect all the structures graphically (Supplementary Figure S1).

Evolutionary conservation analysis

The degree of evolutionary conservation of the amino acids was calculated using the ConSurf (https://consurf.tau.ac.il/consurf_index.php) server to maintain the structural integrity and function of protein, yielding a conservation rating from 1-9, classified as variable, intermediate, or highly conserved protein site. Furthermore, properties of residues were categorized as either exposed (e), buried (b), highly conserved and exposed (functional, f), or highly conserved and buried (s) (Supplementary Figure S2). As a result, any change in the highly conserved area has a significant impact on the protein's integral structure and function.

Results

Clinical characterization

Based on clinical and phenotypic features, 11 patients (6 males and 5 females) with average age of ±7.2 years from six families were diagnosed with Sanfilippo syndrome (MPS III). The other 15 patients (9 males and 6 females) from 8 families were labeled as suspected cases of MPS IV with average age of ±7.3 years. MPS III patients showed a wide range of overlapping features like delay in achieving milestones, aggressive behavior, abdominal distention, increased body hair growth and intellectual disability (Table 1). MPS IV patients showing overlapping features like short stature, coarse facial features, scoliosis/kyphosis, gibbus formation and many others (Table 3). All families of MPS III show parental consanguinity except family D having intra-cast marriage without any direct blood relation among parents (Figure 1D). For MPS IV all families show consanguinity except family C (Figure 2C).

10.3389/fgene.2023.1254909

TABLE 3 Clinical details of the patients suspected as Morquio syndrome affected.

Pedigree no.	Individual no.	Affected individuals	Age/Sex	Cast/region/ Consanguinity	Phenotypes
A	IV-1	1	3 y/M	Gujjar/Punjab/Yes	Delay in achieving milestones, abdominal distention, gibbus formation, short stature, height 95 cm, 12 kg weight, head circumference 48 cm, coarse facial features, broad based gait, joint contractures, kyphosis/scoliosis, bullet shaped metacarpals, beaking of vertebrae, oar shaped ribs, widening of medullary canal, coarse trabecular pattern Metaphyseal dysplasia, Blood sugar Random = 70 mg/dL (80–160), Hb = 11.7 g/dL (14–18), HCT = 34.3% (35–55), Neutrophils = 39.8% (60–70), Lymphocytes = 42.1% (30–40), Vitamin D3 = 18 ng/mL (Deficient <20)
B	IV-4	5	5 y/M	Awan/Punjab/Yes	Abdominal distention, aggressive behavior, short stature, constipation, (89 cm height, 14 kg weight, head circumference 51 cm)
	IV-5		7.5 y/F		Speech delay, gibbus formation, short stature (F = 92 cm height, weight 15 kg, head circumference 49 cm, coarse facial features, broad based gait, bone pain, joint contractures, short trunk, scoliosis/kyphosis, pectus carinatum
C	III-1	2	12 y/M	Pakhtoon/KPK/No	Asthma (uses inhaler), sleep apnea (wakes up suddenly), no hearing, very loose joints, thin bones, underweight, deep voice, macrocephaly, carpel tunnel syndrome, short stature, spinal cord is bent, all body muscles and bones deform, weak muscles, aggressive, coarse facial features, jerky movement, widely spaced teeth, large tongue, mild ID, flat nasal bridge, skin problem
	III-2		9 y/F		Sleep apnea (wakes up suddenly), very loose joints, thin bones, underweight, deep voice, macrocephaly, short stature, spinal cord is bent, all body muscles and bones deform, weak muscles, aggressive, coarse facial features, jerky movement, widely spaced teeth, large tongue, mild ID, flat nasal bridge
D	IV-1	3	10.5 y/M	Mohmand/KPK/Yes	Delay in achieving milestones, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/ kyphosis, hernia, broad based gait
	IV-2		9 y/F		Delay in achieving milestones, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/ kyphosis, hernia, broad based gait
	IV-3		8 y/F		Delay in achieving milestones, abdominal distention, aggressive behavior, gibbus, short stature, respiratory/ear infection, coarse facial features, joint contractures, short trunk, scoliosis/ kyphosis, hernia, broad based gait
E	VI-1	2	10 y/M	Qureshi/Punjab/Yes	Aggressive behavior, short stature, gibbus formation, waddling gait, joint contractures, short trunk, scoliosis/kyphosis, liver problem, anemia, sleep apnea
	VI-2		8 y/F		Aggressive behavior, short stature, gibbus formation, waddling gait, joint contractures, short trunk, scoliosis/kyphosis, liver problem, anemia, sleep apnea
F	IV-1	1	13 y/F	*Rana*/AJK/Yes	Short stature, gibbus formation, joint contractures, short trunk, scoliosis/kyphosis, coarse facial features, spondyloepiphyseal dysplasia, height = 80cm, weight = 10 kg

(Continued on following page)

TABLE 3 (*Continued*) **Clinical details of the patients suspected as Morquio syndrome affected.**

Pedigree no.	Individual no.	Affected individuals	Age/Sex	Cast/region/ Consanguinity	Phenotypes
G	IV-3	3	4 y/M	Awan/KPK/Yes	Speech delay, abdominal distention, short stature, gibbus formation, respiratory/ear infection, coarse facial features, short trunk, joint contractures, scoliosis/kyphosis, hernia, hepatosplenomegaly
	IV-5 IV-6		3 y/M 3 y/M		Speech delay, abdominal distention, short stature, gibbus formation, respiratory/ear infection, coarse facial features, short trunk, joint contractures, scoliosis/kyphosis, hernia, hepatosplenomegaly (twins)
H	V-3	1	5 y/M	Pakhtun/KPK/Yes	Delay in achieving milestones, short stature, coarse facial features, scoliosis/kyphosis, abdominal distention, hepatosplenomegaly, joint contractures, beaking of chest Hb = 11.8 g/dL (12–14), Bilirubin = 16.4 mg/dL (0.1–1.0), Serum growth hormone = 1.62 ng/mL (2–5)

AJK, azad jammu and kashmir; cm, Centimeter; F, female; Hb, Hemoglobin; HCT, hematocrit; kg, Kilogram; KPK, Khyber Pakhtunkhwa; M, male; y, years.

Genetic characterization

Among 14 families diagnosed with MPS, 6 novel and 6 reported variants were identified, where 11 families carried variants in MPS-associated genes (*NAGLU, SGSH, GALNS*), two families carried variants in two genes (*VWA3B, GNPTG*) that were not previously associated with MPS phenotypes, while one family remained genetically undiagnosed. In family H we found dual molecular diagnosis (*GALNS* and *BTD* genes involvement). In family C (MPS IV), initial analysis of coding regions in patients did not reveal any homozygous or compound heterozygous pathogenic variant(s).

In silico analysis

Pathogenicity prediction

The identified variants were predicted to be deleterious/ pathogenic by various mutation prediction tools like MutationTaster, I-TASSER, SIFT, Polyphen-2, CADD, VarSome, VARSEAK and PROVEAN (Table 2; Table 4). Generated structures were selected based on their high C-score and Tm-score. Wild-type structure of proteins (*BTD*, *SGSH*, *VWA3B*, *GNPTG*, *NAGLU*, and *GALNS*) were retrieved from Alpha-fold. Comparison between the wild-type and mutant type was performed and the mutated structures showed significant structural change. The altered structure may result in a loss of function, altered interactions with ligands or neighboring proteins, or altered stability of the proteins. None of the identified variants were found in homozygous state in any human genome variation databases including ExAC, gnomAD, 1000 Genome Project, dbSNP and healthy control samples of same ethnicity.

Evolutionary conservation analysis

The evolutionary conservation analysis of candidate genes showed that, in BTD protein, p.Asp446His is an average (partially conserved) and exposed residue according to the NACSES. In the protein SGSH, the Asn86Lys is highly conserved and buried residue. In protein VWA3B, the Glu984Ala is highly conserved, exposed and functional residue. In protein GNPTG, the Gly60Glu is highly conserved, buried and structural residue according to NACSES. In the protein *NAGLU,* Arg674Cys is highly conserved, exposed and functional residue according to NACSES.

Discussion

In 14 families with phenotypically diagnosed MPS III (A-F) or MPS IV (A-H) we report that 11 families carried pathogenic variants in known MPS-associated genes. Two families had novel mutations in genes previously not associated with MPS and one family remains genetically undiagnosed. Additionally, one family (family H) had pathogenic variants in both *GALNS* and *BTD* segregating with the phenotype.

Mucopolysaccharidoses III

The initial recruitment and diagnosis of six families with Sanfilippo syndrome (MPS III) were based on their clinical profiles and phenotypes (Table 1).

We report two likely causal variations: One identified in *NAGLU* (family A) co-segregating with MPS III type B and the other in *SGSH* (families B, D, E and F) co-segregating with MPS III type A and *VWA3B* (Family C).

In family A, a homozygous missense variant c. 2020C>T (p. Arg674Cys) was found in exon 6 of *NAGLU* gene. This variant was first reported in compound heterozygous state by Zhao et al. (1998) in two different families affected with end-stage neurological debilitation and moderate mental retardation. Another missense variant at the same codon (c. 2021G>A (p. Arg674His)) in compound heterozygous state was reported by Zhao et al. (1996) in an Arab family. This locus is the mutational hot spot for substitution of Arginine. The locus has CpG dinucleotide

TABLE 4 In-silico analysis and ACMG classification of the identified variants in suspected Morquio syndrome cases.

Pedigree/ Patient ID	Variants	Patient genotype	dbSNP	In-silico tools with deleterious output	ACMG classification	VarSome	HGMD accession number	ClinVar accession numbers
A/IV-1	NM_000512.5 (*GALNS*), c. 1365-2A>G	Homozygous	rs1281515662	N/A	Likely pathogenic PVS1, PM2	Likely pathogenic	Current study	SCV004012876
B/IV-4,IV-5	NM_000512.5 (*GALNS*), c.107A>C (p.Leu36Arg)	Homozygous	rs755832705	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571	SCV004012877
C/III-1,III-2	Unsolved	Homozygous	N/A	N/A	N/A	N/A	N/A	N/A
D/IV-1,IV-2,IV-3	NM_000512.5 (*GALNS*), c.1259G>C (p.Pro420Arg)	Homozygous	N/A	9/9	Uncertain significance PP3, PM2, PP5	Uncertain significance	CM145597	SCV004012878
E/VI-1,VI-2	NM_032520.5 (*GNPTG*), c.179G>A (p.Gly60Glu)	Homozygous	N/A	8/9	Uncertain significance PP3, PM2	Uncertain significance	Current study	SCV004012879
F/IV-1	NM_000512.5 (*GALNS*), c.423-1G>A	Homozygous	N/A	N/A	Pathogenic PVS1, PP5, PM2	Pathogenic	Current study	SCV004012880
G/IV-3/IV-5,IV-6	NM_000512.5 (*GALNS*), c.107A>C (p.Leu36Arg)	Homozygous	N/A	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571	SCV004012877
H/V-3	NM_000512.5 (*GALNS*), c.107A>C (p.Leu36Arg)	Homozygous	N/A	8/9	Pathogenic PP5, PM1, PM5, PP3, PM2	Pathogenic	CM145571	SCV004012877
	NM_001281723.2 (*BTD*), c.1336G>C (p.Asp446His)	Homozygous	N/A	8/9	Pathogenic PS3, PS1, PP5, PM1, PM5	Pathogenic	Current study	SCV004012883

N/A, not available; PP, possibly pathogenic; PM, moderate pathogenic; PS, strong pathogenic, PVS, very strong pathogenic.

involvement (Zhao et al., 1996). Currently studied patient showed more severe phenotypes including delay in achieving milestone, abdominal distention, aggressive behavior, speech delay, short stature, increased body hair, coarse facial features, broad-based gait, joint contracture, short trunk, hepatosplenomegaly, self-biting and intellectual disability. The interfamilial variability in the severity of phenotypes might be due to homozygous substitution of Arginine at amino acid position 674 in the present patient and/or different ethnicities of the families reported by Zhao et al. (1996), Zhao et al. (1998). Furthermore, such variability in clinical presentations amongst patients harboring the same genetic variant have also been reported for other LSD subtypes, indicating the contribution of epigenetic modifications, environmental factors and/or genetic modifiers that is yet to be elucidated (Davidson et al., 2018; Sheth et al., 2022).

In family B, five individuals were affected including three males and two females. One female (VI-1) and one male (VII-1) patient had died at the age of 12 and 14 years respectively. All available patients were homozygous for the *SGSH* c. 221C>T (p. Arg74His) variant. The parents were heterozygous, while two healthy live siblings including a female (VII-5; 21 years) and a male (VII-4; 15 years) were wild type carriers. This pathogenic variant was first identified in compound heterozygous form in the Polish population (Bunge et al., 1997). The same codon is altered for another substitution where Arginine is replaced by Cysteine (c. 222G>A, p.Arg74Cys) (Bunge et al., 1997) resulting in a similar phenotype. This locus is likely to be the mutational hot spot for the disease. This residue is conserved among all mammalian sulfatases and stabilizes the active site of the enzyme (Bunge et al., 1997). The substitution might cause loss of ionic interactions formed by positively charged arginine; also, difference in sizes causes loss of interactions with other residues and domains.

In family C, a novel *VWA3B* variant (c.2951A>C; p.Glu984Ala, exon 22) was identified. The variant causes loss of interactions and loss of hydrogen bonds thus affecting the correct folding of protein (HOPE analysis https://www3.cmbi.umcn.nl/hope/).

Previously, four variants including three nonsense and one missense have been reported in *VWA3B* underlying neurological phenotypes including intellectual disability, cerebellar ataxia and autism spectrum disorder as per HGMD (Accessed on 20-May-2022). Affected individuals in the present study showed MPS III overlapping phenotypes including delay in achieving milestones, chronic diarrhea, aggressive behavior, speech delay, short stature, coarse facial features, joint contractures, hepatosplenomegaly, and intellectual disability. Variability in the phenotypes of previous reported cases and the patients in the present study might be due to different position of the variant, familial background, and/or the impact of SNPs in modifier genes.

In family D, the homozygous missense variant *SGSH* c. 364C>T (p. Gly122Arg) in exon 4 was associated with MPS III phenotype in both patients. This variant was first reported in 1997 in Dutch and Arab patients (Bunge et al., 1997). Insertion of a charged residue Arginine at amino acid position 122 of SGSH will cause distortion in the structure and might cause disruption of correct folding of protein. Clinical details and pathogenicity of the variant is described in Table 1 and Table 2, respectively. There were two affected individuals in the family, one sister and one brother born to a non-consanguineous couple, without any previous family history for the disease but all the successive marriages in previous generations are intra-cast marriages (Figure 1D).

In family E, a homozygous missense *SGSH* c. 258A>C (p. Asn86Lys) variant was identified. This variant is novel and has not previously been reported to cause Sanfilippo syndrome. The family had two affected males, born to a first-degree cousin couple (Figure 1E). One patient VI-1 died a month after sample collection at the age of 15 years. Asparagine at amino acid 86 is highly conserved in SGSH in different orthologues. The wild-type residue forms a hydrogen bond with methionine at position 88, the difference in size of wild type and mutated residues causes loss of interaction. The variation is in the catalytic domain that is important for the activity of the enzyme and interact with another domain which might affect the function of protein. As per ACMG guidelines, the variant was classified as a variant of uncertain significance (Table 2).

In family F, the identified pathogenic variant was *SGSH* c. 817G>A (p. Asp273Asn) in exon 7. This variant was reported in UK in 2000 (Beesley et al., 2000). The variant was found in homozygous state in both patients (IV-1, IV-2) and heterozygous in their parents. Both affected individuals were females and born to a first-degree cousin couple (Figure 1F). This variant is located in a CpG dinucleotide site which is the mutational hot spot of the gene (Beesley et al., 2000). As per HOPE analysis, the wild-type residue is negatively charged, and therefore, interaction with calcium ions is lost with this substitution, reducing the stability of the enzyme. The wild type residue Asp273 also forms hydrogen bonds with Asp31 and Asp32 and salt bridges with Arg74, Leu123, and Arg282, which might be lost due to substitution of Asp273 with a neutral Asparagine residue, disturbing the ionic interactions. As per ACMG guidelines the variant is classified as likely pathogenic (Table 2).

Functional lysosomes are important for autophagy to regulate the quality of cytoplasm by eliminating cellular macromolecular aggregates. If autophagy is disturbed due to compromised or absent lysosomal enzyme it may lead to inappropriate *storage* of material in different cells. This storage interferes with normal cell function and affects many organ systems including brain, viscera, bone and cartilage causing various diseases including lysosomal storage disorders (LSDs), neurodegenerative diseases and cancers (Saha et al., 2018; Scerra et al., 2022). Interestingly, in present study we identified missense disease-causing variants in all MPS-III diagnosed families, such variants could have little or no impact on the enzymatic activity of the mutant protein but may cause folding or tertiary structure alterations. Misfolded protein/s are retained in the endoplasmic reticulum for subsequent degradation thus causing enzyme deficiency and compromise in autophagy (Mohamed et al., 2017; Valencia et al., 2021). Therefore, neurological symptoms are one of the most common phenotypes observed in MPS III because neurons being post-mitotic cells could not dilute damaged cell organelles and protein aggregates (Valencia et al., 2021; Scerra et al., 2022). No therapeutic approach has been successful in reverting the symptoms of MPS III and neurological manifestations caused by it. Some of the common therapies tested include enzyme replacement therapy (ERT), substrate reduction therapy (SRT), pharmacological chaperon therapy (PCT) and hematopoietic stem cell transplant (HSCT) (Valstar et al., 2010; Beneto et al., 2020; Penon-Portmann et al., 2023). Some of the therapies applying on cell lines and animal models for treating MPS III include over expression of TFEB (Lotfi et al., 2018), master regulator of lysosomal biogenesis (Bajaj et al., 2019) and coenzyme Q_{10} (Matalonga et al., 2014).

Mucopolysaccharidoses IV

Eight families (A-H) were recruited on the basis of their MPS IV (Morquio syndrome type A) phenotypes. In six families (A, B, D, F, G, H) identification of *GALNS* mutations confirmed the diagnosis. In family E, a variant in the mucolipidosis-associated gene *GNPTG* segregated with the phenotype. No disease-causing variant could be identified in family C.

In family A, the patient had a novel homozygous *GALNS* splice site variant g.38845A>G (chr16:88884534T>C); c.1365–2A>G (intron 12). A splice site prediction tool varSEAK (https://varseak.bio/) predicted the identified splice site variant with class 5 (splicing effect). The identified variant is predicted to activate a cryptic site 34 nucleotide upstream of 3′splice site leading to the production of an abnormal protein.

Three families B, G and H shared a highly prevalent *GALNS* c.107T>G; p.Leu36Arg missense variant. Previously, two different variations (c. 107T>G; p. Leu36Arg; c. 107T>C, p.Leu36Pro) were identified on the same codon to cause Morquio syndrome type A in Asian-multiethnic and Mexican populations respectively (Tomatsu et al., 2005; Morrone et al., 2014). The variation (p.Leu36Arg) is in the catalytic domain, which can disturb the function of the enzyme. It may also hinder the correct protein folding due to the charged arginine residue leading to the loss of hydrophobic interactions. All the three families in the present study showed consanguinity and homozygosity of the variant in patients, while parents were heterozygous carriers (Figures 2B,G,H). Clinical examination demonstrated intra and interfamilial variability in the phenotypes of patients carrying the same variant (Table 3). The variability of phenotypes in this study might be due to different ethnicities of affected families. Additionally, family H showed dual molecular diagnosis as another homozygous missense novel variant in *BTD* gene, c. 1336G>C; p. Asp446His causing biotinidase deficiency (EC 3.5.1.12) segregated with the phenotype. This enzyme is involved in the recycling of biotin bound to protein by releasing biotin and lysine (Pispa, 1965; Wolf et al., 1985). The wild-type residue Asp at 446 forms hydrogen bonds with Gly423, Tyr456 and Gln458, and a salt bridge with Arg544, but because of the substitution, hydrogen bonding and ionic interaction will be lost.

In family D, there were three affected individuals, two females and a male born to first-degree cousin parents (Figure 2D). A homozygous *GALNS* c. 1259G>C; p. Pro420Arg missense variant segregated with disease. Previously, Morrone et al. (2014) identified the same variant to cause MPS IV in Asian-multiethnic population. Ullah et al. (2017) reported a consanguineous Pakistani family

affected with MPS IV segregating same disease-causing variant, i.e., p. Pro420Arg where substitution of proline residue with arginine distorts the protein conformation (Ullah et al., 2017).

In family E, a novel homozygous *GNPTG* c.179G>A; p.Gly60Glu missense variant was identified. Variants in *GNPTG* cause mucolipidoses type III (ML III). This gene encodes the gamma subunit of enzyme N-acetylglucosamine-1-phosphotransferase (EC 2.7.8.17) responsible for catalyzing the first step in synthesis of a mannose 6-phosphate lysosomal recognition marker (Braulke et al., 2008; Qian et al., 2010). The clinical profile for ML III includes phenotypes like short stature, scoliosis and joint contractures (Tiede et al., 2005; Bargal et al., 2006), which were also present in the patients in the present study (Table 3). As the initial diagnosis in the present study was based on clinical phenotypes, and the two disorders (MPS and ML) have overlapping phenotypes, we could not rule out the possibility of dual molecular diagnosis. Next-generation sequencing techniques should be used for the correct diagnosis of these complex overlapping phenotypes. Glycine at amino acid 60 in the GNPTG protein is highly conserved, so its substitution is probably damaging to the protein. The mutant residue is negatively charged so its incorporation will lead ligands and other residues of same charge to be repelled.

In family F, there were two affected individuals, one living while the other died at the age 14. A novel homozygous splice acceptor site *GALNS* c.423-1G>A variant was found in the patient. The parents were first degree cousins (Figure 2F). The identified variant is predicted to activate a cryptic site 1 nucleotide downstream of 3′splice site leading to a frameshift mutation (varSEAK Online - Splice Site Prediction Version 2.1).

Therapies for MPS IV include enzyme replacement therapy which is currently working only for type A not for type B. Some therapies in experimental phases include substrate reduction therapy, hematopoietic stem cell transplant, gene therapy, anti-oxidant therapy, inhibiting protein aggregation, stope codon read through therapy and anti-inflammatory therapy (Akyol et al., 2019; Seker Yilmaz et al., 2021; Penon-Portmann et al., 2023). In a recent therapy, the delivery of recombinant human GALNS enzyme (rhGALNS) into fibroblasts was achieved by infusing it with a hydrogel, such as polyethylene glycol. This innovative approach ensures sustainable and prolonged release of the exogenous enzyme. However, it is important to consider that this treatment might be costly and require frequent administration (Jain et al., 2020).

From Pakistan, very little knowledge about the genetic basis of MPS IV has revealed a common variant p.Leu36Arg in 3/8 families sequenced during current study, and another variant p.Pro420Arg reported formerly by Ullah et al., 2017, also found in family D with three affected siblings.

In Pakistan, the burden of recessively inherited lethal genetic disorders like Sanfilippo and Morquio syndromes is higher due to custom of intra-familial marriages, but the precise diagnosis is lacking due to unavailability of state-of-the-art diagnosis facilities (Hussain and Bittles, 1998; Shahid et al., 2020; Shahzadi et al., 2020; Gul et al., 2023). However, to ensure the health of children suffering from such life-threatening disorders, timely diagnosis and treatment is essential. Patients without proper medical care are affected by multiple medical conditions and progressive tissue degeneration. Therefore, early, and correct diagnosis of such disorders should be performed using targeted gene sequencing and/or WGS. Furthermore, molecular genetic diagnosis will help to provide premarital carrier diagnosis, genetic counselling and to devise prenatal screening tests to address this deficit in the foreseeable future.

Strength of Study: This study details extensive work from a highly consanguineous population of Pakistan, showing the need of genetic exploration for more detailed analysis related to lethal genetic disorders like Sanfilippo and Morquio syndromes.

Limitation of Study: The initial diagnosis of the patients included in this study, was performed based on clinical and radiographic examinations; thus, misdiagnosis of the patients could not be ruled out due to overlapping phenotypes of MPS, *VWA3B*-related phenotypes and ML. Furthermore, the effect of the identified variants on protein structure and function is based on *in silico* tools, and future functional studies are needed to validate the prediction/s.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving humans were approved by the Bioethical Review board, Quaid-i-Azam University, Islamabad, Pakistan (BEC-FBS-QAU2019-198) and Shaheed Zulfiqar Ali Bhutto Medical University, Islamabad, Pakistan. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

RG: Conceptualization, Formal Analysis, Methodology, Writing–original draft. SF: Conceptualization, Supervision, Writing–review and editing. MS: Data curation, Writing–review and editing. AU: Conceptualization, Formal Analysis, Methodology, Writing–original draft, Writing–review and editing. EP: Formal Analysis, Writing–review and editing. AT: Formal Analysis, Project administration, Writing–review and editing. AG: Conceptualization, Writing–original draft, Writing–review and editing. MH: Conceptualization, Investigation, Supervision, Writing–review and editing. MT: Investigation, Software, Visualization, Writing–review and editing. MS: Investigation, Methodology, Writing–review and editing. KA: Supervision, Writing–review and editing. TH: Conceptualization, Funding acquisition, Supervision, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by Center for Basic Metabolic Research (CBMR) Novo Nordisk Foundation (NNF) through the Challenge Program

(NNF18OC0033950) and Higher Education Commission, Islamabad, Pakistan under International Research Support Initiative Program (IRSIP) (1-8/HEC/HRD/2020/10579).

Acknowledgments

We would like to thank all participating families, staff and colleagues who helped us during this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1254909/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Three-dimensional (3D) of wild-type and mutants of *BTD, SGSH, VWA3B, GNTGP*. The zoom-in part represents the mutated residue of each protein. Mutated proteins show a significant change in their structure.

SUPPLEMENTARY FIGURE S2
Evolutionary conservation of mutated residue of *BTD, SGSH, VWA3B,* and *GNTGP*. The red boxes represent wild-type amino acids that might be altered by pathogenic mutations of high risk. The conservation scale has a score range of 1–9, with conserved sites receiving ratings in the range of 7–9. The tool illustrated the properties of residue that is either exposed (e), buried (b), highly conserved and exposed (functional, ,**(F)** or highly conserved and buried (b).

References

Akyol, M. U., Alden, T. D., Amartino, H., Ashworth, J., Belani, K., Berger, K. I., et al. (2019). Recommendations for the management of MPS IVA: systematic evidence-and consensus-based guidance. *Orphanet J. rare Dis.* 14 (1), 137–225. doi:10.1186/s13023-019-1074-9

Bajaj, L., Lotfi, P., Pal, R., Ronza, A. D., Sharma, J., and Sardiello, M. (2019). Lysosome biogenesis in health and disease. *J. Neurochem.* 148 (5), 573–589. doi:10.1111/jnc.14564

Bargal, R., Zeigler, M., Abu-Libdeh, B., Zuri, V., Mandel, H., Neriah, Z. B., et al. (2006). When mucolipidosis III meets mucolipidosis II: GNPTA gene mutations in 24 patients. *Mol. Genet. metabolism* 88 (4), 359–363. doi:10.1016/j.ymgme.2006.03.003

Barstow, C., and Rerucha, C. (2015). Evaluation of short and tall stature in children. *Am. Fam. Physician* 92 (1), 43–50.

Beesley, C. E., Burke, D., Jackson, M., Vellodi, A., Winchester, B. G., and Young, E. P. (2003). Sanfilippo syndrome type D: identification of the first mutation in the nacetylglucosamine-6-sulphatase gene. *J. Med. Genet.* 40 (3), 192–194. doi:10.1136/jmg.40.3.192

Beesley, C. E., Young, E. P., Vellodi, A., and Winchester, B. G. (2000). Mutational analysis of sanfilippo syndrome type A (MPS IIIA): identification of 13 novel mutations. *J. Med. Genet.* 37 (9), 704–707. doi:10.1136/jmg.37.9.704

Benetó, N., Vilageliu, L., Grinberg, D., and Canals, I. (2020). Sanfilippo syndrome: molecular basis, disease models and therapeutic approaches. *Int. J. Mol. Sci.* 21 (21), 7819. doi:10.3390/ijms21217819

Besley, G. T. N., and Wraith, J. E. (1997). Lysosomal disorders. *Curr. Paediatr.* 7 (2), 128–134. doi:10.1016/s0957-5839(97)80195-9

Braulke, T., Pohl, S., and Storch, S. (2008). Molecular analysis of the GlcNac-1-phosphotransferase. *J. Inherit. metabolic Dis.* 31 (2), 253–257. doi:10.1007/s10545-008-0862-5

Bunge, S., Ince, H., Steglich, C., Kleijer, W. J., Beck, M., Zaremba, J., et al. (1997). Identification of 16 sulfamidase gene mutations including the common R74C in patients with mucopolysaccharidosis type IIIA (Sanfilippo A). *Hum. Mutat.* 10 (6), 479–485. doi:10.1002/(SICI)1098-1004(1997)10:6<479:AID-HUMU10>3.0.CO;2-X

Celik, B., Tomatsu, S. C., Tomatsu, S., and Khan, S. A. (2021). Epidemiology of mucopolysaccharidoses update. *Diagnostics* 11 (2), 273. doi:10.3390/diagnostics11020273

Chien, Y. H., Lee, N. C., Chen, P. W., Yeh, H. Y., Gelb, M. H., Chiu, P. C., et al. (2020). Newborn screening for morquio disease and other lysosomal storage diseases: results from the 8-plex assay for 70,000 newborns. *Orphanet J. rare Dis.* 15 (1), 38–47. doi:10.1186/s13023-020-1322-z

Chih-Kuang, C., Shuan-Pei, L., Shyue-Jye, L., and &Tuen-Jen, W. (2002). MPS screening methods, the berry spot and acid turbidity tests, cause a high incidence of false-negative results in sanfilippo and morquio syndromes. *J. Clin. laboratory analysis* 16 (5), 253–258. doi:10.1002/jcla.10051

Davidson, B. A., Hassan, S., Garcia, E. J., Tayebi, N., and Sidransky, E. (2018). Exploring genetic modifiers of gaucher disease: the next horizon. *Hum. Mutat.* 39 (12), 1739–1751. doi:10.1002/humu.23611

De Pasquale, V., and Pavone, L. M. (2019). Heparan sulfate proteoglycans: the sweet side of development turns sour in mucopolysaccharidoses. *Biochimica Biophysica Acta (BBA)-Molecular Basis Dis.* 1865 (11), 165539. doi:10.1016/j.bbadis.2019.165539

Fedele, A. O. (2015). Sanfilippo syndrome: causes, consequences, and treatments. *Appl. Clin. Genet.* 8, 269–281. doi:10.2147/TACG.S57672

Feldhammer, M., Durand, S., Mrázová, L., Boucher, R. M., Laframboise, R., Steinfeld, R., et al. (2009). Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene. *Hum. Mutat.* 30 (6), 918–925. doi:10.1002/humu.20986

Feldhammer, M., Durand, S., Mrazova, L., van Diggelen, O. P., Hrebicek, M., Kmoch, S., et al. (2009). a. 54. Sanfilippo syndrome type C: novel mutations in the HGSNAT gene. *Mol. Genet. Metabolism* 96 (2), S24S24.

Gul, R., Firasat, S., Schubert, M., Ullah, A., Peña, E., Thuesen, A. C., et al. (2023). Identifying the genetic causes of phenotypically diagnosed Pakistani mucopolysaccharidoses patients by whole genome sequencing. *Front. Genet.* 14, 1128850. doi:10.3389/fgene.2023.1128850

Hendriksz, C. J., Harmatz, P., Beck, M., Jones, S., Wood, T., Lachman, R., et al. (2013). Review of clinical presentation and diagnosis of mucopolysaccharidosis IVA. *Mol. Genet. metabolism* 110 (1-2), 54–64. doi:10.1016/j.ymgme.2013.04.002

Hussain, R., and Bittles, A. H. (1998). The prevalence and demographic characteristics of consanguineous marriages in Pakistan. *J. Biosoc. Sci.* 30 (2), 261–275. doi:10.1017/s0021932098002612

Jain, E., Flanagan, M., Sheth, S., Patel, S., Gan, Q., Patel, B., et al. (2020). Biodegradable polyethylene glycol hydrogels for sustained release and enhancedstability of rhGALNS enzyme. *Drug Deliv. Transl. Res.* 10 (5), 1341–1352. doi:10.1007/s13346-020-00714-7

Khan, S. A., Peracha, H., Ballhausen, D., Wiesbauer, A., Rohrbach, M., Gautschi, M., et al. (2017). Epidemiology of mucopolysaccharidoses. *Mol. Genet. metabolism* 121 (3), 227–240. doi:10.1016/j.ymgme.2017.05.016

Lin, H. Y., Lee, C. L., Chang, C. Y., Chiu, P. C., Chien, Y. H., Niu, D. M., et al. (2020). Survival and diagnostic age of 175 Taiwanese patients with mucopolysaccharidoses (1985–2019). *Orphanet J. Rare Dis.* 15 (1), 314–411. doi:10.1186/s13023-020-01598-z

Lotfi, P., Tse, D. Y., Di Ronza, A., Seymour, M. L., Martano, G., Cooper, J. D., et al. (2018). Trehalose reduces retinal degeneration, neuroinflammation and storage burden caused by a lysosomal hydrolase deficiency. *Autophagy* 14 (8), 1419–1434. doi:10.1080/15548627.2018.1474313

Matalonga, L., Arias, A., Coll, M. J., Garcia-Villoria, J., Gort, L., and Ribes, A. (2014). Treatment effect of coenzyme Q 10 and an antioxidant cocktail in fibroblasts of patients

with Sanfilippo disease. *J. Inherit. metabolic Dis.* 37 (3), 439–446. doi:10.1007/s10545-013-9668-1

Martell, L., Lau, K., Mei, M., Burnett, V., Decker, C., and Foehr, E. D. (2011). Biomarker analysis of morquio syndrome: identification of disease state and drug responsive markers. *Orphanet J. Rare Dis.* 6 (1), 84–10. doi:10.1186/1750-1172-6-84

Mohamed, F. E., Al-Gazali, L., Al-Jasmi, F., and Ali, B. R. (2017). Pharmaceutical chaperones and proteostasis regulators in the therapy of lysosomal storage disorders: current perspective and future promises. *Front. Pharmacol.* 8, 448. doi:10.3389/fphar.2017.00448

Morquio, L. (1929). Sur une forme de dystrophie osseuse familiale. *Bull. Soc. Pediatr. Paris* 27, 145–152.

Morrone, A., Tylee, K. L., Al-Sayed, M., Brusius-Facchin, A. C., Caciotti, A., Church, H. J., et al. (2014). Molecular testing of 163 patients with Morquio A (Mucopolysaccharidosis IVA) identifies 39 novel GALNS mutations. *Mol. Genet. metabolism* 112 (2), 160–170. doi:10.1016/j.ymgme.2014.03.004

Nelson, J., Crowhurst, J., Carey, B., and Greed, L. (2003). Incidence of the mucopolysaccharidoses in western Australia. *Am. J. Med. Genet. Part A* 123 (3), 310–313. doi:10.1002/ajmg.a.20314

Okada, S., and O'Brien, J. S. (1968). Generalized gangliosidosis: beta-galactosidase deficiency. *Science* 160 (3831), 1002–1004. doi:10.1126/science.160.3831.1002

Ozkinay, F., Emecen, D. A., Kose, M., Isik, E., Bozaci, A. E., Canda, E., et al. (2021). Clinical and genetic features of 13 patients with mucopolysaccarhidosis type IIIB: description of two novel NAGLU gene mutations. *Mol. Genet. Metabolism Rep.* 27, 100732. doi:10.1016/j.ymgmr.2021.100732

Penon-Portmann, M., Blair, D. R., and Harmatz, P. (2023). Current and new therapies for mucopolysaccharidoses. *Pediatr. Neonatol.* 64, S10–S17. doi:10.1016/j.pedneo.2022.10.001

Peracha, H., Sawamoto, K., Averill, L., Kecskemethy, H., Theroux, M., Thacker, M., et al. (2018). Molecular genetics and metabolism, special edition: diagnosis, diagnosis and prognosis of mucopolysaccharidosis IVA. *Mol. Genet. metabolism* 125 (1-2), 18–37. doi:10.1016/j.ymgme.2018.05.004

Piraud, M., Boyer, S., Mathieu, M., and Maire, I. (1993). Diagnosis of mucopolysaccharidoses in a clinically selected population by urinary glycosaminoglycan analysis: a study of 2,000 urine samples. *Clin. Chim. acta* 221 (1-2), 171–181. doi:10.1016/0009-8981(93)90031-x

Pispa, J. (1965). Animal biotinidase. *Ann. Med. Biol.* 43, 1–39.

Qian, Y., Lee, I., Lee, W. S., Qian, M., Kudo, M., Canfield, W. M., et al. (2010). Functions of the alpha, beta, and gamma subunits of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. *J. Biol. Chem.* 285 (5), 3360–3370. doi:10.1074/jbc.M109.068650

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17 (5), 405–424. doi:10.1038/gim.2015.30

Saha, S., Panigrahi, D. P., Patil, S., and Bhutia, S. K. (2018). Autophagy in health and disease: A comprehensive review. *Biomed. Pharmacother.* 104, 485–495. doi:10.1016/j.biopha.2018.05.007

Sanfilippo, S. J., Podosin, R., Langer, L., and Good, R. A. (1963). Mental retardation associated with acid mucopolysacchariduria (heparitin sulfate type). *J. Pediatr.* 63, 837–838. doi:10.1016/s0022-3476(63)80279-6

Scerra, G., De Pasquale, V., Scarcella, M., Caporaso, M. G., Pavone, L. M., and D'Agostino, M. (2022). Lysosomal positioning diseases: beyond substrate storage. *Open Biol.* 12 (10), 220155. doi:10.1098/rsob.220155

Seker Yilmaz, B., Davison, J., Jones, S. A., and Baruteau, J. (2021). Novel therapies for mucopolysaccharidosis type III. *J. Inherit. Metabolic Dis.* 44 (1), 129–147. doi:10.1002/jimd.12316

Shahid, M., Firasat, S., Satti, H. S., Satti, T. M., Ghafoor, T., Sharif, I., et al. (2020). Screening of the FANCA gene mutational hotspots in the Pakistani fanconi anemia patients revealed 19 sequence variations. *Congenit. Anomalies* 60 (1), 32–39. doi:10.1111/cga.12331

Shahzadi, M., Firasat, S., Kaul, H., Afshan, K., Afzal, R., and Naz, S. (2020). Genetic mapping of autosomal recessive microspherophakia to chromosome 14q24. 3 in a consanguineous Pakistani family and screening of exon 36 of LTBP2 gene. *JPMA. J. Pak. Med. Assoc.* 70 (3), 515–518. doi:10.5455/JPMA.302440

Shapiro, E. G., Jones, S. A., and Escolar, M. L. (2017). Developmental and behavioral aspects of mucopolysaccharidoses with brain manifestations—Neurological signs and symptoms. *Mol. Genet. metabolism* 122, 1–7. doi:10.1016/j.ymgme.2017.08.009

Sheth, J., Nair, A., and Jee, B. (2022). Lysosomal storage disorders: from biology to the clinic with reference to India. *Lancet Regional Health-Southeast Asia* 2022.

Tiede, S., Muschol, N., Reutter, G., Cantz, M., Ullrich, K., and &Braulke, T. (2005). Missense mutations in N-acetylglucosamine-1-phosphotransferase alpha/beta subunit gene in a patient with mucolipidosis III and a mild clinical phenotype. *Am. J. Med. Genet. Part A* 137 (3), 235–240. doi:10.1002/ajmg.a.30868

Tomatsu, S., Montaño, A. M., Nishioka, T., Gutierrez, M. A., Peña, O. M., Trandafirescu, G. G., et al. (2005). Mutation and polymorphism spectrum of the GALNS gene in mucopolysaccharidosis IVA (Morquio A). *Hum. Mutat.* 26 (6), 500–512. doi:10.1002/humu.20257

Tomatsu, S., Nishioka, T., Montaño, A. M., Gutierrez, M. A., Pena, O. S., Orii, K. O., et al. (2004). Mucopolysaccharidosis IVA: identification of mutations and methylation study in GALNS gene. *J. Med. Genet.* 41 (7), e98e98. doi:10.1136/jmg.2003.018010

Ullah, I., Nasir, A., Mehmood, S., Ahmed, S., Ullah, M. I., Ullah, A., et al. (2017). Identification and *in silico* analysis of GALNS mutations causing Morquio A syndrome in eight consanguineous families. *Turkish J. Biol.* 41 (3), 458–468. doi:10.3906/biy-1607-81

Valencia, M., Kim, S. R., Jang, Y., and Lee, S. H. (2021). Neuronal autophagy: characteristic features and roles in neuronal pathophysiology. *Biomol. Ther.* 29 (6), 605–614. doi:10.4062/biomolther.2021.012

Valstar, M. J., Neijs, S., Bruggenwirth, H. T., Olmer, R., Ruijter, G. J., Wevers, R. A., et al. (2010). Mucopolysaccharidosis type IIIA: clinical spectrum and genotype-phenotype correlations. *Ann. neurology* 68 (6), 876–887. doi:10.1002/ana.22092

Valstar, M. J., Ruijter, G. J. G., Van Diggelen, O. P., Poorthuis, B. J., and Wijburg, F. A. (2008). Sanfilippo syndrome: a mini-review. *J. Inherit. metabolic Dis.* 31 (2), 240–252. doi:10.1007/s10545-008-0838-5

Wolf, B., Grier, R. E., SecorMcVoy, J. R., and Heard, G. S. (1985). Biotinidase deficiency: a novel vitamin recycling defect. *J. Inherit. metabolic Dis.* 8 (1), 53–58. doi:10.1007/BF01800660

World Medical Association, (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama* 310 (20), 2191–2194. doi:10.1001/jama.2013.281053

Zhao, H. G., Li, H. H., Bach, G., Schmidtchen, A., and Neufeld, E. F. (1996). The molecular basis of Sanfilippo syndrome type B. *Proc. Natl. Acad. Sci. U. S. A.* 93, 6101–6105. doi:10.1073/pnas.93.12.6101

Zhao, H. G., Aronovich, E. L., and Whitley, C. B. (1998). Genotype-phenotype correspondence in Sanfilippo syndrome type B. *Am. J. Hum. Genet.* 62 (1), 53–63. doi:10.1086/301682

Zubaida, B., Hashmi, M. A., Cheema, H. A., and Naeem, M. (2018). Identification of a novel GLB1 mutation in a consanguineous Pakistani family affected by rare infantile GM1 gangliosidosis. *J. Genet.* 97 (5), 1445–1449. doi:10.1007/s12041-018-1002-0

ASSOCIATE PROFESSOR
Department of Zoology
Quaid-i-Azam University
Islamabad

Turnitin Originality Report

Molecular Analysis of Lysosomal Storage Disorders in Pakistani Population by Rutaba Gul .

From PhD (PhD DRSML)

- Processed on 20-Jan-2023 08:21 PKT
- ID: 1995783155
- Word Count: 22500

Similarity Index
16%
Similarity by Source

Internet Sources:
11%
Publications:
13%
Student Papers:
4%

sources:

1 < 1% match (student papers from 23-Apr-2017)
Submitted to Higher Education Commission Pakistan on 2017-04-23

2 < 1% match (student papers from 18-Mar-2019)
Submitted to Higher Education Commission Pakistan on 2019-03-18

3 < 1% match (student papers from 29-May-2017)
Submitted to Higher Education Commission Pakistan on 2017-05-29

4 < 1% match (student papers from 22-Jun-2018)
Submitted to Higher Education Commission Pakistan on 2018-06-22

5 < 1% match (student papers from 04-Jan-2016)
Submitted to Higher Education Commission Pakistan on 2016-01-04

6 < 1% match (student papers from 20-Apr-2022)
Submitted to Higher Education Commission Pakistan on 2022-04-20

7 < 1% match (student papers from 07-Sep-2016)
Submitted to Higher Education Commission Pakistan on 2016-09-07

8 < 1% match (student papers from 03-Aug-2019)
Submitted to Higher Education Commission Pakistan on 2019-08-03

9 < 1% match (student papers from 13-Sep-2018)
Submitted to Higher Education Commission Pakistan on 2018-09-13

10 < 1% match (student papers from 01-Feb-2016)
Submitted to Higher Education Commission Pakistan on 2016-02-01

11 < 1% match (student papers from 15-Aug-2022)
Submitted to Higher Education Commission Pakistan on 2022-08-15

12 < 1% match (Internet from 15-Jan-2023)
https://clinvarminer.genetics.utah.edu/variants-in-conflict-by-significance/likely%20pathogenic/pathogenic

13 < 1% match (Internet from 15-Jan-2023)
https://clinvarminer.genetics.utah.edu/variants-by-submitter/500026/significance/any

14 < 1% match (Lysosomal Storage Disorders, 2007.)
Lysosomal Storage Disorders, 2007.

15 < 1% match (Internet from 10-Jan-2023)
https://pure.manchester.ac.uk/ws/portalfiles/portal/157710117/FULL_TEXT.PDF

16 < 1% match (Rubab Raza, Asmat Ullah, Nighat Haider, Jai Krishin et al. "Exome Sequencing Revealed the First Intragenic Deletion in ABCA5 Underlying Autosomal Recessive Hypertrichosis", Clinical and Experimental Dermatology, 2022)
Rubab Raza, Asmat Ullah, Nighat Haider, Jai Krishin et al. "Exome Sequencing Revealed the First Intragenic Deletion in ABCA5 Underlying Autosomal Recessive Hypertrichosis", Clinical and Experimental Dermatology, 2022

17 < 1% match ()
Saima Ghafoor, Karina da Costa Silveira, Raheel Qamar, Maleeha Azam, Pelar Kannu, "Exome Sequencing Identifies a Biallelic GALNS Variant (p.Asp233Asn) Causing Mucopolysaccharidosis Type IVA in a Pakistani Consanguineous Family", Genes

18 < 1% match ()
Alessandra Zanetti, Francesca D'Avanzo, Moegenaldeen AlSayed, Ana Carolina Brusius-Facchin et al. "Molecular basis of mucopolysaccharidosis IVA (Morquio A syndrome): A review and classification of gene variants and reporting of 68 novel variants", Human Mutation

19 < 1% match (Internet from 18-Nov-2013)
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351003/

20 < 1% match (Internet from 02-Apr-2014)
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3195295/

21 < 1% match (Internet from 24-May-2021)

Focal Person (Turnitin)
Quaid-i-Azam University
Islamabad