Molecular characterization of a large cohort of Mucopolysaccharidosis patients: Iran Mucopolysaccharidosis RE-diagnosis Study (IMPRESsion)

Saeed Ghaffari[1], Maryam Rafati[1], Mahdi Shadnoush[2], Shokooh Pourbabaee[3], Mohammad Aghighi[3], Siamak Mirab Samiee[3], Jamshid Kermanchi[3], Mohammad Reza Alaei[2], Shadab Salehpoor[2], Davoud Amirkashani[4], Aria Sotoodeh[5], Peymaneh Sarkhail[1], Reza Shervin Badv[6], Majid Aminzadeh[7], Siamak Shiva[8], Peyman Eshraghi[9], Hossein Moravej[10], Mahin Hashemipour[11], Noushin Rostampour[12], Amir Ali Hamidieh[6], Bibi Shahin Shamsian[2], Sedigheh Shams[6], Daniel Zamanfar[13], Ayoub Ebrahimi[8], Ali Otadi[14], Seyedeh Zahra Tara[15], Zeinab Barati[16], Laya Fakhri[1], Azadeh Hosseini[1], Hosna Amiri[1], Somayeh Ramandi[1], Niusha Mostofinezhad[1], Zahra Pahlevani Kani[17], Elham Mohammadyari[17], Mahsa Khosravi[17], Masoome Saadati[18], Fatemeh Hoseininasab[16], Hamid Reza Khorram Khorshid[1], and Younes Modaberisaber[1]

[1]Hope Generation Foundation
[2]Shahid Beheshti University of Medical Sciences
[3]Iran Ministry of Health and Medical Education
[4]Iran University of Medical Sciences
[5]Department of Pediatrics, Tehran University of Medical Sciences, Tehran, Iran
[6]Tehran University of Medical Sciences
[7]Ahvaz Jundishapur University
[8]Tabriz University of Medical Sciences
[9]Mashhad University of Medical Sciences
[10]Shiraz University of Medical Sciences
[11]Isfahan University of Medical Sciences
[12]Isfahan University of Medical Sciences Isfahan Endocrine and Metabolism Research Center
[13]Mazandaran University of Medical Sciences
[14]The University of British Columbia Faculty of Science
[15]The Children's Hospital of Philadelphia
[16]Avicenna Research Institute, Academic Center for Education, Culture and Research (ACECR)
[17]Islamic Azad University Tehran Medical Sciences
[18]Islamic Azad University Tehran North Branch

July 22, 2021

Abstract

Mucopolysaccharidoses (MPSs) are rare, heterogeneous inborn errors of metabolism (IEM) diagnosed through a combination of

1

clinical, biochemical, and genetic investigations. The aim of this study was molecular characterization of the largest cohort of Iranian MPS patients (302 patients from 289 unrelated families), along with tracking their ethnicity and geographical origins. 185/289 patients were studied using an IEM-targeted NGS panel followed by complementary Sanger sequencing, which led to the diagnosis of 154 MPS patients and 5 non-MPS IEMs (diagnostic yield: 85.9%). Furthermore, 106/289 patients who were referred with positive findings went through re-analysis and confirmatory tests which confirmed MPS diagnosis in 104. Among the total of 258 MPS patients, 225 were homozygous, and 9 had copy number variations, and 90 had novel variants. MPS IV was the most common type (34.8%) followed by MPS I (22.7%) and MPS VI (22.5%). Geographical origin analysis unveiled a pattern of distribution for frequent variants in ARSB (c.430G>A, c.962C>T, c.281C>A), GALNS (c.319G>A, c.860C>T, c.1042A>G), and IDUA (c.1A>C, c.1598C>G). Our extensive patient cohort reveals the genetic and geographic landscape of MPS in Iran, which provides insight into the genetic epidemiology of MPS and can facilitate a more cost-effective, time-efficient diagnostic approach based on the region-specific variants.

Molecular characterization of a large cohort of Mucopolysaccharidosis patients: Iran Mucopolysaccharidosis RE-diagnosis Study (IMPRESsion)

Authors:

Saeed Reza Ghaffari[*^1,2,3] , Maryam Rafati[*^1,2], Mahdi Shadnoush[4], Shokooh Pourbabaee[5], Mohammad Aghighi[6], Siamak Mirab Samiee[7,8], Jamshid Kermanchi[6], Mohammad Reza Alaei [9], Shadab Salehpour[9], Davoud Amirkashani [10], Aria Setoodeh [11], Peymaneh Sarkhail [12], Reza Shervin Badv [13], Majid Aminzadeh[14], Siamak Shiva [15], Peyman Eshraghi[16], Hossein Moravej [17], Mahin Hashemipour[18], Noushin Rostampour [19]Amir Ali Hamidieh [20], Bibi Shahin Shamsian[21], Sedigheh Shams [22], Daniel Zamanfar [23], Ayoub Ebrahimi [15], Ali Otadi [24], Seyedeh Zahra Tara[25], Zeinab Barati [1,2], Laya Fakhri[1,2], Azadeh Hoseini[2], Hosna Amiri[2], Somayeh Ramandi[2], Niusha Mostofinezhad[2], Zahra Pahlevani Kani[26], Elham Mohammadyari [26], Mahsa Khosravi[26], Masoome Saadati [27], Fatemeh Hoseininasab[2], Hamid Reza Khorram Khorshid[1] , Younes Modaberisaber[1]

*These authors contributed equally to this work.

Affiliations:

1. Fetal Health Research Center, Hope Generation Foundation, Tehran, Iran
2. Reproductive Biotechnology Research Center, Avicenna Research Institute, ACECR, Tehran, Iran
3. Gene Clinic, Tehran, Iran
4. Department of Clinical Nutrition, Faculty of Nutrition & Food Technology, Shahid Beheshti University of Medical Sciences, Tehran, Iran
5. Deputy of Curative Affairs, Ministry of Health and Medical Education, , Tehran, Iran
6. Transplantation and Diseases Administrationw Deputy of Curative Affairs, Ministry of Health and Medical Education, Tehran, Iran
7. Food and Drug Laboratory Research Center, Tehran, Iran
8. General Reference Health Laboratory, Ministry of Health and Medical Education, Tehran, IR
9. Shahid Beheshti University of Medical Sciences, Tehran, IR
10. Division of Endocrinology and Metabolism, Ali-Asghar Children's Hospital, Iran University of Medical Sciences, Tehran, IR
11. Children's Medical Center, Growth and Development Research Center, Tehran University of Medical Sciences, Tehran, Tehran, IR
12. Personalized Medicine and Genometabolomics Research Center, Hope Generation Foundation, Tehran, Iran
13. Department of Pediatrics, Children's Medical Center, Pediatrics Center of Excellence, Tehran University of Medical Sciences, Tehran, Tehran, IR
14. Diabetes Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Khuzestan, IR
15. Pediatric Health Research Center, Tabriz University of Medical Sciences, Tabriz, East Azerbaijan, IR
16. Mashhad University of Medical Sciences, Mashhad, Khorasan Razavi, IR
17. Neonatal Research Center, Shiraz University of Medical Sciences, Shiraz, Fars, IR

2

18. Metabolic Liver Disease Research Center, Isfahan University of Medical Sciences, Isfahan, Isfahan, IR
19. Isfahan Endocrine and Metabolism Research Center, Isfahan University of Medical Sciences Isfahan, Isfahan, Isfahan, IR
20. Pediatric Cell and Gene Therapy Research Center , Tehran University of Medical Sciences, Tehran, Tehran, IR
21. Pediatric Congenital Hematologic Disorders Research Center, Research Institute for Children's Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran
22. Department of Pathology, Children's Medical Center, Pediatrics Center of Excellence, Tehran University of Medical Sciences, Tehran, Tehran, IR
23. Diabetes Research Center, Mazandaran University of Medical Sciences, Sari, IR
24. The University of British Columbia, Faculty of Science, Vancouver, BC, Canada
25. Pediatric Endocrinology Department, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
26. Islamic Azad University Tehran Medical Sciences, Tehran, Tehran, IR
27. Islamic Azad University Tehran North Branch, Tehran, Tehran, IR

^**Correspondence:**

Saeed Reza Ghaffari

Postal address: Gene Building, No. 344, Keshavarz Blvd, Tehran, Iran

Postal code: 1418894161

Tel: +98 21 66935655

Fax: +98 21 66933336

Email: *saeed@ghaffari.org*

Maryam Rafati

Postal address: Avicenna Research Institute, Shahid Beheshti University, Shahid Shahriari Square, Daneshjou Boulevard, Shahid Chamran Highway, Tehran, Iran

Postal code: 1983969411

Tel: +98 21 22432020

Fax: +98 21 22432021

Email: *rafati76@gmail.com*

Data Availability Statement

The data that support the findings of this study are openly available in *Iran Variome MPS Database* at *https://variome.ir/lovd*

Abstract

Mucopolysaccharidoses (MPSs) are rare, heterogeneous inborn errors of metabolism (IEM) diagnosed through a combination of clinical, biochemical, and genetic investigations. The aim of this study was molecular characterization of the largest cohort of Iranian MPS patients (302 patients from 289 unrelated families), along with tracking their ethnicity and geographical origins. 185/289 patients were studied using an IEM-targeted NGS panel followed by complementary Sanger sequencing, which led to the diagnosis of 154 MPS patients and 5 non-MPS IEMs (diagnostic yield: 85.9%). Furthermore, 106/289 patients who were referred with positive findings went through re-analysis and confirmatory tests which confirmed MPS

diagnosis in 104. Among the total of 258 MPS patients, 225 were homozygous, and 9 had copy number variations, and 90 had novel variants. MPS IV was the most common type (34.8%) followed by MPS I (22.7%) and MPS VI (22.5%). Geographical origin analysis unveiled a pattern of distribution for frequent variants in *ARSB* (c.430G>A, c.962C>T, c.281C>A), *GALNS* (c.319G>A, c.860C>T, c.1042A>G), and *IDUA*(c.1A>C, c.1598C>G). Our extensive patient cohort reveals the genetic and geographic landscape of MPS in Iran, which provides insight into genetic epidemiology of MPS and can facilitate a more cost-effective, time-efficient diagnostic approach based on the region-specific variants.

Keywords

Mucopolysaccharidosis, Targeted Next-Generation Sequencing, Lysosomal Storage Disorders, Copy Number Variation, Single Nucleotide Variant, Geographical Distribution

Introduction

Mucopolysaccharidoses (MPSs) are a group of lysosomal storage disorders (LSDs) caused by deficiency in the activity of each of 11 lysosomal enzymes required for degradation of glycosaminoglycans (GAGs). The 11 genes coding for the lysosomal enzymes are *IDUA, IDS, SGSH, NAGLU, HGSNAT, GNS, GALNS, GLB1, ARSB, GUSB* , and *HYAL1* . Deficiency in these enzymes results in abnormal accumulation of GAGs in lysosomes of most cells leading to progressive cellular damage and multiple organ failure, which is the phenotypic hallmark of most MPS types. There is a high degree of phenotypic overlap between the seven types of MPSs which are inherited in either autosomal recessive (MPS types I, III, IV, VI, VII, and IX) or X-linked (MPS type II) (Brusius-Facchin et al. 2019).

MPSs are diagnosed through a combination of clinical phenotype and biochemical testing, such as quantitation of urinary GAGs and measurement of enzyme activity in leukocytes or fibroblasts. Measurement of GAGs using tandem mass spectrometry is a plausible way to screen for MPS patients, but it cannot provide a definitive differential diagnosis for the disease type. Biochemical characterization of the GAGs using two-dimensional electrophoresis can help with achieving a more specific differential diagnosis, but it requires enzymatic confirmation. However, performing several enzymatic assays in screen-positive patients could be time-consuming and not cost-effective. Molecular characterization of MPS genes using Sanger or next-generation sequencing is another diagnostic method that is considered as accurate as of the enzymatic assay, especially when the assignment of patients to the costly enzyme replacement therapies is taken into account. Such molecular approaches could also be beneficial for predicting the prognosis and providing genetic counseling to the family and their at-risk relatives about carrier screening and prenatal or preimplantation genetic diagnosis (Brusius-Facchin et al. 2019; Filocamo et al. 2018; Kadali et al. 2019).

Sanger sequencing of the MPS genes has been widely used for molecular diagnosis in clinically and/or biochemically diagnosed patients over the last two decades. This individualized approach, which requires exon-by-exon sequencing of the corresponding genes, might be time-consuming and less cost-effective. Furthermore, when it comes to the differential diagnoses of each MPS type, which encompass a wide range of other types of MPSs and the other LSDs, Sanger sequencing cannot be considered an effective tool. In a handful of recent research articles, next-generation sequencing (NGS) has therefore been suggested as a quick, efficient, and reliable alternative to Sanger sequencing. Next-Generation

Sequencing allows achieving a comprehensive molecular investigation of all MPS types as well as their differential diagnosis in a single test, which could be of paramount importance, particularly in patients with non-specific overlapping phenotypes. Whole exome sequencing (WES) and targeted capture NGS are two NGS-based methods that have been applied for molecular diagnosis of genetic disorders, and each has its own capabilities and limitations. While the low sensitivity of the Whole-Exome Sequencing (WES) resulting from its limited coverage could be considered a drawback, employing targeted capture NGS approach makes the technology much less error-prone due to the higher coverage and mean depth achievable in targeted regions. Additionally, the application of target panels with a larger number of genes encompassing MPS genes and

4

their differential diagnosis can facilitate reaching a conclusive diagnosis (Cobos et al. 2015; Costa-Motta et al. 2014; Filocamo et al. 2018; Kadali et al. 2019; Wood et al. 2013).

Here, we report on the genetic investigation of a cohort of 302 patients from 289 unrelated families with the clinical and/or biochemical diagnosis of various types of MPS which is the largest cohort of Iranian patients reported so far. This report is part of a national project on "Iran MucoPolysaccharidosis REdiagnosis Study (IMPRESsion)", which has been carried out to determine the genetic profile of MPSs in Iran. The recruited patients have been investigated using either Sanger sequencing or NGS panel, targeting 594 genes associated with inborn errors of metabolism. Moreover, in search of possible founder mutations and/or region-specific pathogenic variants, the geographical origin of the studied patients has also been collected. Such data can facilitate mutation detection of patients originating from the designated areas, and propose a stepwise diagnostic approach starting from testing region-specific pathogenic variants, followed by more comprehensive testing by NGS-based methods.

Methods

Patient Recruitment

Patients recruited in this study were part of a national project on "Iran MucoPolysaccharidosis REdiagnosis Study (IMPRESsion)", which has been carried out through a collaboration between "Fetal Health Hope Generation Research Center" (FHHGRC) and "National Scientific Committee on Inborn Errors of Metabolism" (NSCIEM) from September 2018 to February 2020. Patients with the clinical and/or biochemical diagnosis of MPS who were registered by NSCIEM participated in this study. Patients were recalled for re-examination and blood sampling through 6 designated provincial centers located in the cities of Tehran, Tabriz, Mashhad, Ahvaz, Isfahan, and Shiraz. Families were assigned to the most accessible provincial center based on their place of residence.

Patients who were assigned to Tehran provincial centers were directly referred to FHHGRC. For these patients, pre-test and post-test genetic counseling was provided and medical photo and video documentation were carried out. Also, peripheral blood samples were drawn from the probands, their parents, affected/unaffected siblings, and other affected relatives, if available. For patients referred from centers other than Tehran, collected peripheral blood samples from the probands and their parents/siblings/other affected relatives coupled with their clinical data and signed consent forms were sent to the FHHGRC by mail. In these patients, history taking, and pedigree drawing was performed through phone contacts with families by genetic counselors.

The ethnicity and geographical origin of the probands and their parents (including the country, province, city, and village, if applicable) were recorded in detail. Probands or parents born in large cities were inquired about the geographical origin of their previous generations (e.g. grandfather, grandmother) to avoid misrepresentation of the data with regard to the fact that the proband's birthplace could have been affected by immigration of the previous generations to the large cities. Written informed consents were obtained from the patients or their parents/guardians. The study was approved by the research center review board and all procedures were performed in accordance with the ethical standards of the local committee and the revised version of the Helsinki Declaration in 2013. Details of the pedigrees, clinical data and the results of biochemical tests or previous genetic tests were completely reviewed, summarized and tabulated by a genetic counselor. There were two groups of families based on their genetic testing status:

Families who had positive findings in their previous genetic tests (Group A): in these patients, the previously reported variants were re-analyzed according to the American College of Medical Genetics and Genomic (ACMG) guideline. Sanger verification and segregation analysis were then carried out if indicated.

Families with no previous genetic tests or no finding in their tests (Group B): these patients went through our

comprehensive genetic approach to MPSs, starting from targeted NGS, followed by complementary Sanger sequencing of poorly covered regions, Sanger verification and/or segregation analysis. Details of the applied approach are discussed below.

DNA Preparation and Next-Generation Sequencing

Venous blood sample was used for DNA extraction using the QIAamp DNA Mini Kit (Qiagen, Germany). Targeted capture NGS was performed using Ion Ampliseq Inborn Errors of Metabolism Research Panel V2 (Life Technologies). The capture panel consisted of 9681 amplicons (1.05 Mbp) covering coding regions and flanking exon/intron boundaries of 594 genes associated with inborn errors of metabolism with an average gene coverage of 99.17%. Amplicon length and average insert size were 254 bp and 207 bp respectively.

The library preparation for targeted capture was carried out according to the manufacturer's protocol. Briefly, 100 ng of DNA quantified by Qubit DNA HS assay (Life Technologies) was used. The amplicons were then ligated to the barcode-labelled X adapter and the Ion Sphere Particle (ISP) compatible P adapter. The library was purified with AMPure XP Beads (Beckman Coulter, Pasadena, CA) and the final library concentration was measured using the Ion Library TaqMan® Quantitation Kit (Life Technologies). Following pooling equimolar libraries, automated template preparation and chip loading were carried out using Ion Chef System (Life Technologies). The Ion Proton Sequencing 200 kit (Life Technologies) was used for sequencing on a PI chip V2 following the manufacturer's protocol.

Variant Identification

Single Nucleotide Variant (SNV) analysis

Data were aligned to the hg19/GRCh37 reference genome and were quality trimmed via Ion Torrent Suite (Life Technologies). Variants were then called against the panel BED file (Ion_AmpliSeq_Inborn_Errors_of_-Metabolism_v2.20160503 BED file), filtered and analyzed using the Ion Reporter Pipeline (Life Technologies). Analysis was firstly focused on 11 genes involved in MPSs and if no variant was identified, further analyses were performed on the other 583 genes associated with inborn errors of metabolism. We firstly removed intergenic, 3'/5' UTR, non-splice related intronic variants and variants with high allele frequency (mean allele frequency >1%) in known genome reference databases including The Genome Aggregation Database (gnomAD), 1000 Genomes, 5000 Exomes, UCSC genome browser, Exome Aggregation Consortium (ExAC), dbSNP and an in-house database. The remaining variants were then classified according to the recommendations of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP)[5] using VarSome (*https://varsome.com/*). Variants were considered putatively causal if: they were classified as pathogenic/likely pathogenic (P/LP) according to the triggered ACMG rules; or they were variants of unknown significance (VUS) in which at least half of the in-silico predictors predicted a deleterious effect. Prioritized variants in each family were individually reviewed in terms of the proband's clinical findings and pedigree analysis to be matched to the clinically diagnosed MPS type.

Copy Number Variation (CNV) analysis

Patients with no positive findings in SNV analysis or those who had monoallelic variants were further analyzed for CNV events using the Ion Reporter CNV caller tool. Depth of coverage data were retrieved from BAM files and normalized against a reference set of 100 normal controls. CNVs overlapping MPS genes were individually reviewed and prioritized if: they had deletion status; passed quality metrics set by the software; and were not shared by several patients. Identified CNVs were validated using multiplex PCRs targeting exons overlapping CNVs as well as non-deleted exons flanking the CNV.

Reviewing the quality of reads by IGV

Raw sequencing data were reviewed for the detected variants using Integrative Genomic Viewer (IGV) software (*https://software.broadinstitute.org/software/igv/*) according to its instructions. Briefly, BAM files

were imported to IGV and the quality of sequencing reads, coverage and number of the reads (depth) for the given variants were reviewed.

Variant Confirmation and Segregation Analysis

The reported positive findings in group A were reviewed with regard to the test setting (research or clinical), applied technique (NGS-based methods or Sanger sequencing), and the variant characteristics, to indicate whether further confirmatory tests were required. As a result, patients that were studied in a research setting, or had INDEL variants identified by WES, were investigated by Sanger verification. Although group B patients were studied by targeted-capture panel with high coverage depth, a subset of variants with depressed allele fraction (<30% of alternate allele), particularly deletion/insertion type variants, underwent additional validation with Sanger sequencing which was performed on an ABI3130 (Life Technologies) according to the standard procedures. Amplification primers were either selected from published papers (if available) or designed using Primer 3 software. Segregation analysis was carried out when additional family members were available and segregation data was helpful for classification and interpretation of the variants, especially variants with unknown clinical significance.

Complementary Sanger Sequencing

In case of no positive finding in target NGS, raw sequencing data were reviewed in each patient using IGV software, and low coverage regions were then sequenced by Sanger sequencing.

Results

In this study, 302 MPS patients were recruited, of whom 13 were found to be related to at least one other patient based on an extensive review of their documents and pedigrees and were therefore excluded from further analyses. As shown in Figure 1, among the remaining 289 unrelated families, 106 (36.3%) had positive results in previous genetic tests (Group A), while 183 (63.7%) had no prior tests or no positive results (Group B). Using ACMG guidelines, we reviewed previously reported variants in Group A patients to verify causal variants and determine their pathogenicity. This approach confirmed the pathogenicity of the reported variants in 105 patients, while disproved a variant in MPS-69 (c.314G>A in *IDUA*) due to its population frequency of 15.72% and 13.74% in gnomAD Exomes and Genomes respectively, which resulted in moving this patient to Group B. The positive findings in the remaining 105 Group A patients were reviewed with regard to the test setting, the applied technique, and the variant characteristics. As a result, patients that were studied in a research setting, or had INDEL variants identified by WES, were investigated by Sanger verification. Among them, all variants except two (in MPS-215 and MPS-237) were confirmed. In MPS-215, who had the clinical diagnosis of MPS II, Sanger verification did not confirm the previously reported variants, but revealed a known in-frame hemizygous deletion (c.121_123del3, p.Asp35Gly1) in *IDS* . Sanger sequencing did not confirm c.1104_1105delCT in the second patient (MPS-237), in whom, biallelic variants c.1104_1105delCT and c.656G>A had previously been reported by another lab using Whole Exome Sequencing. Therefore, MPS-237 was allocated to Group B and the NGS test was repeated using the IEM panel, which again showed the same variants in the patient. Reappearance of the same findings led us to believe that Sanger analysis had failed to call c.1104_1105delCT, likely as a result of allelic dropout in a GC rich region. Among the reported variants, there were two CNVs

In Group B, a total of 185 patients went through target capture NGS using the Ion Ampliseq IEM panel, consisting of 594 genes, and high-quality sequencing results were obtained. On average, 463,693,908 bases were called, 2,557,476 reads were mapped, and a mean depth of coverage of 304X over the targeted regions were achieved. Moreover, 87.2% of all the bases were covered by 20 or more reads, and an average uniformity of 83.3% was obtained per sample, indicating that the coverage of target regions was evenly distributed.

7

NGS analysis of Group B patients discovered causal MPS variants in 124/185 (67.02%), and non-MPS IEM variants in 5/185 (2.7%) patients. Details of the non-MPS IEMs are shown in Table 1. For the remaining 55 negative patients, IGV was employed to manually review the NGS sequence data of the clinically suggested MPS genes. This approach showed sufficient depth of coverage for the exons of the corresponding genes in 10 patients, low to zero depth of coverage in 41 patients, and non-annotated variants in 4 patients. The breakdown of the corresponding genes for 41 patients with low-depth coverage is as follows: exon 1 of *ARSB* in 13 patients, exon 1 of *GALNS* in 16 patients, exon 1 of *IDUA* in 11 patients, and exon 1 of *NAGLU* in 1 patient. Complementary Sanger sequencing for low coverage regions of the aforementioned genes resulted in discovering causal variants in 26 additional patients (26/41, 63.4%), which in turn improved the diagnostic yield by 14% (26/185) (Figure 1).

Regarding the 4 non-annotated variants, manual IGV review revealed the following variants, which were then confirmed by Sanger analysis: c.1279del in *GALNS* , g.997263G>A (c.1650+5G>A) in *IDUA* , c.544G>C in *IDUA* , and g.43033381A>G (c.1012+4A>G) in *HGSNAT*. As a result, the diagnostic yield was further improved by 2.2% (4/185), leading to a total improvement of 16.2% (30/185).

The IGV data corresponding to 124 positive MPS patients and 5 non-MPS patients in Group B were manually inspected to determine whether further confirmatory tests were required. Eighty-six variants with depressed allele fraction (<30% of alternate allele), mostly INDELs, were investigated by Sanger verification and/or segregation analysis, which all were confirmed. Overall, among 185 studied patients in Group B, 154 patients (83.2%) were found to harbor putative causal variants in MPS genes, five patients (2.7%) were diagnosed with a non-MPS IEM, while in 26 patients (14.1%) no pathogenic variant could be identified in the studied genes. Hence, the clinical diagnosis of MPS was not genetically confirmed in 16.8% of group B patients. 142 of these 154 patients were clinically assigned an MPS type, among them, the molecular findings were concordant in 97.2% (138/142). The four discordant cases comprised two MPS VI patients with the clinical diagnosis of MPS I (MPS-154) and IV (MPS-32), as well as two MPS IV patients with the clinical diagnosis of MPS I (MPS-213) and VI (MPS-142).

Further analyses on a total of 258 MPS patients (104 confirmed Group A Patients and 154 positive Group B MPS patients) were then carried out regarding pedigree analysis, geographical origin, and the distribution of variants in MPS genes (Table 2). In 258 genetically confirmed patients, 34.9% had MPS IV, 22.9% MPS I, 22.5% MPS VI, 12.9% MPS III, and 7.1% MPS II. In total, 191 unique putatively casual variants in 8 MPS genes were discovered in 258 MPS patients, among them 97.4% (186/191) were pathogenic/likely pathogenic (P/LP), and 2.6% (5/191) were variants of unknown significance (VUS) according to the ACMG/AMP guidelines. Furthermore, according to the public version of the HGMD® database (Stenson et al. 2014), and an intensive literature review, 90/258 patients (34.9%) harbored novel variants with no pathogenicity report, among which, 21/90 (23.3%) had variants sharing their location with a known pathogenic variant, but the nucleotide and/or amino acid changes were different. The remaining 168 of 258 patients (65.1%) possessed known causal variants with pathogenicity report.

Three patients had two homozygous variants in their corresponding MPS gene: MPS-95 had c.469T>C and c.523_543del in *IDUA* ; MPS-161 harbored c.46_57del and c.58G>T in *IDU* ; and MPS-254 had c.1493C>T and c.1496C>T in GALNS. Moreover, MPS-213 was homozygous for c.860C>T in *GALNS* which was a predominant known variant among our patients and also harbored a homozygous VUS variant in *IDUA* (c.806C>G). Notably, the clinical manifestations of this patient were in favor of MPS IV, while a single enzyme assay on DBS was suggestive of MPS I. This case was grouped in MPS IV considering the known pathogenic variant in *GALNS* . There was another patient with dual molecular diagnosis who had a homozygous P/LP variant in *GALNS* (c.722C>A) as well as a homozygous known nonsense variant in *HOGA1* (c.763C>T) which is associated with primary hyperoxaluria, type III.

In terms of pedigree analysis, 32.6% were familial, and 66.3% were sporadic, and 81.8% of the parents had a consanguineous marriage. A breakdown of the consanguineous relations is depicted in Table 2. Moreover, a majority of the patients were homozygous for the MPS variants (87.2%) (Table 2). The mutational spectrum included 9 (3.5%) CNVs which were then verified by IGV review and confirmed by multiple PCRs targeting

8

the exons within the deleted regions, as well as those encompassing the CNVs. The *GALNS* gene had the highest number of CNVs (6/9) accounting for 6.7% of all variants identified in this gene (6/89). Details for the detected CNVs are shown in Table 3.

Furthermore, The IMPRESsion study recorded the clinical findings of all the patients with a heatmap of the frequency of symptoms illustrated in Figure 2. The most common symptoms in patients with MPS I, II, and VI recorded in at least 80% of the patients were coarse facies, joint contractures, umbilical hernia, short stature, sleep apnea, and macrocephaly. Among MPS IV patients, the most observed symptoms were umbilical hernia, short stature, corneal clouding, kyphosis, visual impairment, joint hypermobility, and pectus carinatum. MPS III patients displayed symptoms of coarse facies and umbilical hernia, but they also had developmental delay, intellectual disability, developmental regression, and behavioral disorders, which were not recorded in the rest of the MPS patients. (Figure 2)

In terms of geographical origin, 250 patients were Iranian, and 3 were Afghan (five with unknown geographical origin). The most common ethnicities among the Iranians were Fars, Azeri, Lor, and Kurd, with Fars being the most common (117/253, 46.2%), followed by Azeri (43/253, 17%) (Figure 3). We used parental geographical origin to depict the exact geographical origin of the family since the proband's birthplace could have been affected by immigration of the previous generations to the large cities.

Details of the variants identified in this study are available at Iran Variome MPS Database (Ghaffari and Rafati 2020). The most common variants discovered in MPS VI patients were c.430G>A (15/58, 25.9%), c.962T>C (7/58, 12%), c.281C>A (4/58, 6.8%), c.275C>A (3/58, 5.1%), and c.753C>G (3/58, 5.1%). All 15 patients harboring c.430G>A originated from 3 neighboring provinces located in the southeast of Iran; 7 patients from Yazd, 4 from Sistan-and-Baluchistan, and 4 from Kerman. Furthermore, variant c.962T>C was detected in 6 patients from northwestern provinces including Gilan (3/7), Ardabil (1/7), West Azerbaijan (1/7), and East Azerbaijan (1/7). Likewise, variant c.281C>A was only detected in Khuzestani patients. Overall, the two most prevalent variants in MPS VI account for 37.9% (22/58) of the patients, while the five most prevalent variants account for 55.1% (32/58) of the patients. On the other hand, variant c.430G>A was present in 88.2% (15/17) of all MPS VI patients from Sistan-and-Baluchistan, Kerman, Yazd, Fars, and Hormozgan province, making it a significant variant in southeastern region of Iran. The same feature was observed in c.962T>C and c.281C>A variants, accounting for 85.7% (6/7) of northwestern and 66.7% (4/6) of Khuzestan MPS VI patients respectively.

The most frequent variants among MPS I patients were c.1A>C (11/59, 18.7%) and c.1598C>G (4/59, 6.8%). These two variants alone account for 25.5% of MPS I patients. Variant c.1A>C originated from two distinct regions of Iran, 5 patients from Gilan province, and 6 patients from provinces in the southwestern regions of Iran. Three of the patients carrying the c.1598C>G variant were from the Khorasan region of Iran, while one patient was from Lorestan.

The 4 most common variants in the *GALNS* gene, c.860C>T, c.319G>A, c.1042A>G, and c.29G>A were responsible for 18.9% (17/90) of the MPS IV patients. The c.319G>A variant was found in 4 patients (4.5%, 4/90) all from the East Azerbaijan province with Azeri ethnicity. This variant was present in 50% (4/8) of all the MPS IV patients from East Azerbaijan. The 4 families harboring the c.1042A>G variant (4.5%, 4/90) originated from Khorasan-Razavi (3/4) and Mazandaran (1/4). The c.29G>A variant originated from two central neighbor provinces in Iran, Chaharmahal and Bakhtiari, and Isfahan.

Discussion

This study is the largest cohort of Iranian MPS population with precise tracking of ethnicity and geographical origin which unveils patterns of the distribution of frequent mutations in the MPS patients as well as redirecting MPS to their appropriate treatment options. Here, we discuss the logistics behind the algorithm used to optimize the re-diagnosis process and the inevitable limitations. In this study, the genetic profiling

of 289 clinically diagnosed MPS patients confirmed MPS diagnosis in 258 (89.3%), and identified 5 (1.7%) with non-MPS inborn errors of metabolism.

In total, we discovered novel causal variants in 90 patients (34.9%), which is a significant proportion of our cohort, and could be due to the fact our study is the first extensive MPS study of the Iranian population. Although there have been a few studies published in recent years that have contributed to our understanding of the molecular causes of MPS in Iran, the patient pool in those studies was considerably smaller than that of ours. Overall, 18 studies have reported molecular characterization of Iranian MPS patients, comprising 46 MPS I patients, 34 MPS VI patients, 19 MPS III patients, 4 MPS II patients, 4 MPS IV patients, and 1 MPS VII patient (Abbasi et al. 2017; Alaei et al. 2020; Aminzadeh, Malekpour, and Ghandil 2019; Jafaryazdi et al. 2018, 2019; Kamranjam and Alaei 2019; Khorrami et al. 2019; Malekpour, Vakili, and Hamzehloie 2018; Mansour, Rasoul, and B 2007; Martins et al. 2019; Mosallanejad et al. 2020; Nouri et al. 2012; Salehi 2004; Seyedhassani et al. 2015; Shafaat et al. 2020; Taghikhani et al. 2019; Marlies J. Valstar et al. 2010; Yassaee et al. 2017).

Multiethnic nature of the Iranian population has regionalized a number of MPS variants. Additionally, c.962C>T in *ARSB*(c.962C>T), which was the dominant MPS VI variant in the northwestern provinces of Iran, has also been reported in the neighboring country, Turkey. Detailed documentation of geographic origins, pedigrees, and clinical findings have allowed us to uncover these patterns, which will be discussed in detail in order of prevalence.

Mucopolysaccharidosis IV (MPS IV): Most Common MPS Type

MPS IV was discovered to be the most common type in the current Iranian cohort accounting for 34.9% of the patients, even though MPS IV has never been reported as the most common type in previous studies. Geographical origin of MPS IV patients along with the distribution pattern of 3 frequent variants in *GALNS* gene are depicted in Figure 4. c.860C>T was the most common variant in the*GALNS* gene discovered in 6.6% (6/90) of MPS IV patients. This variant has been classified as one of the top 10 most common variants in the *GALNS* gene with reports from the Polish, American Caucasian, Austrian, Macedonian, Turkish, Greek, and Indian population (Bidchol et al. 2014; Morrone et al. 2014). The distribution of this variant among the Iranian population was rather scattered across the country, but 4 of the 6 patients harboring this variant were from the southern and southwestern provinces.

The second most common variant among MPS IV patients was c.319G>A with a 4.4% (4/90) frequency, which has been previously reported in Italy, China, and Dubai (Al-Jasmi et al. 2013; Tomatsu et al. 2004; Wang et al. 2010). All the patients diagnosed with this variant were from the East Azerbaijan province in the northwest of Iran, with parents from the same city, Tabriz. Moreover, the paternal side of MPS-235 and MPS-142 were both from a small village near Tabriz, suggesting that a founder effect may have resulted in distribution of c.319G>A in East Azerbaijan. Since c.319G>A is responsible for 50% of MPS IV patients in East Azerbaijan (4/8), a stepwise genetic testing considering the prevalence of this variant can be developed, suggesting genetic analysis of c.319G>A for MPS IV patients originated from this region before performing a more comprehensive analysis.

The c.1042A>G variant, reported for the first time in this study, was found in 3 patients and was the third most common variant in MPS IV patients. Notably, two of these patients were from the same city in the northeastern province of Khorasan-Razavi, and the other patient was from the northern province of Mazandaran.

Mucopolysaccharidosis I (MPS I)

MPS I was the second most common type of MPS in the IMPRESsion study with 59 patients (22.7%). The two most common variants in the*IDUA* gene, c.1A>C and c.1598C>G, accounted for 25.4% (15/59) of the patients. Five out of the eleven patients harboring c.1A>C were from the city of Rasht in Gilan province, making this city a focal point for this variant. The rest of the patients were from the southwestern regions of Iran spanning across 5 provinces (Figure 5). This variant has previously been reported in studies conducted

by Atceken et al. (2016) and Bertola et al. (2011), in patients from Turkey and Spain respectively. However, there has been no report suggestive of a founder effect for this variant.

c.1598C>G was discovered in 4 patients, three of them were from the Khorasan region in the northeast, and one was from Lorestan in the southwest of Iran (Figure 5). This variant has been discovered in Tunisian and Morroccan patients in multiple studies (Alif et al. 1999; Chkioua et al. 2011; Laradi et al. 2005), and according to Bertola et al. (2011) it has spread from North African countries to the mediterrinean regions accounting for 13% and 10% of the *IDUA* mutant alleles in Italy and Spain respectively.

Mucopolysaccharidosis VI (MPS VI)

MPS VI is known as one of the rare types with prevalence ranging from 0.7% to 15% in most of the studied populations. The only exceptions are Brazil with 30% prevalence, and Saudi Arabia where MPS VI is the most common type, accounting for 46.4% of the MPS patients. The demographics of our study were similar to that of Brazil and Saudi Arabia concerning the frequency of MPS VI patients. MPS VI was the third most common type among the Iranian patients, accounting for 22.5% (58/258) of the total number of patients. The high prevalence of MPS VI in this cohort, may be attributed to the high level of consanguinity, which results in the expression of otherwise rare recessive disorders (Costa-Motta et al. 2014). It is also noteworthy that the five frequent variants in *ARSB* accounted for 55.2% (32/58) of our MPS VI patients, which can support the possible role of founder variants in reinforcing MPS VI in the iranian population.

Geolocalization of the three most common variants in the *ARSB* gene, c.430G>A, c.962T>C, and c.281C>A, revealed that they are concentrated in certain regions or provinces of Iran, likely to be founder variants (Figure 6). c.430G>A was the most common variant across all MPS types in this study, discovered in 15 patients, whose parents were from the southeastern provinces of Iran. Also, 88.2% (15/17) of MPS VI patients in this region harbored this variant (Supp. Table 1). As a result, a tailored diagnostic approach to MPS VI patients originating from southeast of Iran is hereby suggested, beginning with genetic testing of c.430G>A before employing more comprehensive genetic analysis to decrease the total cost and improve the diagnostic yield.

Six patients diagnosed with c.962C>T, the second most common MPS VI variant, were originally from the northwestern provinces, which are neighboring Turkey and Azerbaijan. This variant has also been reported in the Turkish population with a high prevalence and is believed to be a founder variant in that area (Kantaputra et al. 2014), suggesting the possibility that the origin of this variant in the Iranian and Turkish patients track back to the same ancestor. Moreover, 85.7% (6/7) of MPS VI patients from this region were harboring c.962C>T (Supp. Table 1). The significant prevalence of c.962C>T in the northwestern region justifies checking for this variant in MPS VI patients whose origin track back to these provinces before employing more labor-intensive and demanding techniques.

Four patients with c.281C>A participating in our cohort, were all previously reported in a study carried out on 10 MPS VI Iranian families (Aminzadeh et al. 2019). They suggested that c.281C>A can be "the most common mutation in the Arab Iranian patients with MPS VI". These patients were born in the Khuzestan province in the southwest of Iran in villages within the same county, and beside the Karun River. These findings are in favor of a founder mutation, especially because this variant has not been reported in other regions of Iran, or other countries. Additionally, genetic analysis of c.281C>A can prove helpful in making a final diagnosis for patients suspected of MPS VI in Khuzestan province using less resources and less-demanding techniques due to the 66.7% (4/6) prevalence of this variant among MPS VI patients in Khuzestan.

Mucopolysaccharidosis II and III (MPS II and MPS III)

As shown in Figure 7, there were no frequent mutations exhibiting a geographical pattern of distribution present in a significant number of patients.

Pedigree Analysis

In total, autosomal recessive MPS types (I, III, IV, VI) encompassed 93% of genetically confirmed cases in

IMPRESsion study, which could be explained by the observed consanguinity rate of 81.2% in the studied families. Identifying the variant in these families allows them to seek prenatal diagnosis or other preventive measures after comprehensive genetic counseling.

A universal study carried out by Morrone et al. (2014) on the alterations in the *GALNS* gene listing 277 unique variants in this gene, discovered that 48% of the MPS IV patients were homozygous, while 39% harbored two different variants in the *GALNS* gene, hence deemed as compound heterozygous. On the contrary, 94.4% (85/90) of the MPS IV patients in the IMPRESsion study were homozygous. This staggering difference is likely to be the result of high prevalence of consanguinity in the Iranian population. To further illustrate, only 8 MPS IV patients had non-consanguineous parents, resulting in 91.1% (82/90) consanguinity among MPS IV patients. Similarly, 54.8% of MPS VI patients in Tomanin et al.'s (2018) comprehensive study were homozygous for the *ARSB* gene variants, while our cohort consisted of 96.6% (56/58) homozygous MPS VI patients with a 84.4% (49/58) rate of consanguinity, in which the only compound heterozygous MPS VI patient had non-consanguineous parents. Overall, the high prevalence of consanguinity in the Iranian population led to higher frequencies of homozygous patients compared to the western populations.

Technical Perspective

Next-Generation Sequencing

Implementation of NGS for discovering MPS variants has offered a number of advantages in this study. Firstly, as a low-throughput technique, Sanger sequencing works on a gene-dependent exon-by-exon basis. Therefore, lack of a clinically designated MPS type would lead to rounds of labor-intensive gene-by-gene analyses. Due to the wide range of symptoms shared by MPS types, a misdiagnosis of the MPS type is also probable. Secondly, CNV analysis, which conventionally required separate genetic investigations, was integrated within our NGS workflow that resulted in an improvement of the diagnostic yield at a fraction of time and cost.

As recorded in the IMPRESsion study, NGS could overcome these limitations by sequencing all MPS genes in a single test per patient. This capability was also expanded to screen for other MPS differential diagnoses, which led to the identification of 4 cases with other lysosomal storage disorders including sialidosis I/II, galactosialidosis, mucolipidosis II/III alpha/beta and mucolipidosis III gamma and one case of L-2-Hydroxyglutaric Aciduria. Also, such comprehensive analysis diagnosed 2 cases with dual molecular diagnosis; the first with homozygous variants in *GALNS* and *IDUA* and the second with homozygous variants in *GALNS* and *HOGA1* genes and the dual diagnoses of MPS IV and primary hyperoxaluria type III. These variant combinations could support the recently proposed concept of complex genetic models where disease-causing variants in multiple genes cause or modify the development of one or several disease phenotypes, and challenge the classic concept of one gene leading to a particular phenotype that now appears to be an oversimplification (Papadimitriou et al. 2019).

Overall, the applied genetic workflow in group B cases could not confirm MPS in 16.8%, including 26 negative patients and 5 non-MPS IEMs. Lack of a clinically designated MPS type was evident in 34.6% (9/26) of negative cases, which was more than four times higher than that of positive cases (7.8%, 12/154). Such uncertainty could be as a result of Inconclusive biochemical assays and/or non-specific clinical symptoms. The most common symptoms in these negative patients were intellectual disability, developmental delay, and short stature which are not unique to MPS and overlap with other genetic or non-genetic disorders. In addition, the lack of finding variants in these patients might be due to an NGS shortcoming in sequencing GC rich regions. Although complementary sanger sequencing has been performed to compensate for these limitations, the possibility of uncovered pathogenic variants in these regions, as well as upstream, downstream, deep intronic regions and even in unknown MPS associated genes should still be considered.

The workflow also revealed an incompatibility between the genetic and clinical type of MPS in four patients of group B (2.8%). To further elaborate on the four discordant cases, MA-032 had inconclusive enzyme assay results, and MPS-154 was not tested by an enzyme assay, and so their diagnosis was based only on the clinical features. Since MPS VI shares a variety of symptoms with MPS I and MPS IV, making a differential diagnosis

based solely on the clinical signs and symptoms could be error prone. Additionally, in MPS-213 and MPS-142, clinical diagnoses were based on the enzymatic assays conducted on dried blood samples (DBS), once for each patient. These limitations may have led to a discordant diagnosis as the standard diagnostic guidelines strongly recommend confirmation of DBS findings by a second independent assay on DBS, since the number of cells in DBS is lower than in other methods, and the sample is more susceptible to environmental factors (Wood et al. 2013). Moreover, diagnosis of MPS I with DBS is even more unreliable since α-iduronidase becomes less stable in DBS compared to the other enzymes (Cobos et al. 2015; Khan et al. 2017; Tomatsu et al. 2011). As two of these discordant cases were under enzyme replacement therapy based on their clinical diagnosis, this finding was of paramount importance in providing appropriate treatment for the patients.

Copy Number Variations

Until recently, CNVs could only be detected through non-high throughput techniques such as Sanger sequencing and MLPA, both of which can analyze a limited region of the genome. But in recent years, discovering CNVs with NGS has been made possible through quantitative readout of high-coverage NGS data (Eisenberger et al. 2013). The value of this technique is shown in cohort studies like ours where performing the old techniques for each patient takes more time and effort. In our study, NGS analysis was able to detect 9 CNVs in the MPS patients (Table 3). The gene with the highest number of CNVs was *GALNS* , with a 6.7% frequency of CNVs among MPS IVA patients. This is in line with the previous reports on the contribution of CNVs to 2-3% of the reported*GALNS* variants (Regier, Oetgen, and Tanpaiboon 1993). Furthermore, a whole *IDS* gene deletion was detected among the 18 MPS I patients, accounting for 5.5% which is compatible with the previous studies reporting a 9% frequency of CNVs in *IDS* (Scarpa 1993). We also identified two patients with CNVs in *HGSNAT* and*IDUA* , but there is no report in the literature about the proportion of patients with CNVs in these two genes.

Sanger Sequencing

Confirmation of NGS findings by Sanger sequencing was first suggested in guidelines as a result of low coverage and high false positive error rate of the early NGS instruments (Consortium et al. 2010; Fox et al. 2014; Harismendy et al. 2009; Quail et al. 2008). With the advancement of NGS technology, particularly targeted panel sequencing with higher depth of coverage, the sensitivity and specificity of NGS variant calling has improved. Therefore, validation of variants detected by targeted-capture panels that usually have sufficient coverage, could be waived, while it can be helpful in reducing false positives for variants with depressed allele fraction (<30% of alternate alleles) or located in poorly covered regions. Also, performing manual inspection of the NGS data using IGV obviates the need to carry out further validation studies. However, we employed a hybrid approach utilising Sanger sequencing and IGV review for a subset of variants on a case-by-case basis.

Complementary Sanger Sequencing

IGV review of the gene/genes matched with the clinical MPS type was carried out for each of the cases with no primary IEM-panel findings. This approach led to the identification of zero or poor coverage in exon 1 of *ARSB* , *IDUA* , *GALNS* , and *NAGLU* genes. Complementary Sanger sequencing of these poorly covered regions resulted in discovering 26 additional patients, improving the diagnostic yield by 14%. This finding is in accordance with the study reported by Brusius-Facchin, et.al (2019) in which they observed failure of NGS in sequencing exon 1 of *IDUA* , *ARSB* , *NAGLU* ,*HGSNAT* , and *GALNS* genes likely as a result of high GC content and repeat regions. As a result, complementary Sanger sequencing for exon 1 of the aforementioned genes is hereby suggested to ensure that a pathogenic variant is not overlooked because of this shortcoming.

Furthermore, based on the same approach, the IGV review led to the identification of 4 variants not previously annotated by the Ion Reporter®. MPS-23 was found to have a single-nucleotide deletion in exon 12 of *GALNS* (c.1279del) overlooked by Ion Reporter®, which was confirmed by Sanger sequencing. Additionally, two homozygous splicing variants in *IDUA* (g.997263G>A, c.1650+5G>A), and *HGSNAT*(g.43033381A>G, c.1012+4A>G) were uncovered in MPS-144 and MPS-155 respectively. It is likely that the Ion Reporter®

13

failed to annotate these variants because they were more than 3 base pairs downstream of the exons. Review of the IGV data also revealed that MPS-166 harbored a pathogenic variant adjacent to a polymorphism. The pathogenic variant was not annotated because these two adjacent variants were grouped together as one multi-nucleotide variant which was then filtered out due to its high population frequency. Overall, complementary Sanger Sequencing, carried out following the IGV review, allowed us to rule out false positive and false negative results with a significant reduction in time and resources.

Clinical Interpretations

As part of the IMPRESsion study, we included a list of clinical features associated with the different types of MPS, which clinicians were asked to fill out. The heatmap of the clinical features associated with each MPS type is depicted in Figure 2. In general, the trends and patterns observed in the heatmap are in line with previous findings. As with other literature reports, our analysis revealed that MPS I, MPS II, and MPS VI share a wide range of symptoms, and the heatmap indicates that symptoms are more concentrated in the same areas for these three types. However, there were no reports of corneal opacity, visual impairment, and cardiac valve defect among MPS II patients participating in our study, while these clinical features have been reported in previous studies (Martin et al. 2008; Montano et al. 2016; Muenzer et al. 2009, 2011). High occurrence of developmental and neurological problems among our MPS III patients was also suggested by the literature, as well as the rarity of skeletal abnormalities (M. J. Valstar et al. 2010). Moreover, the symptoms of MPS IV patients in our study fell in line with the known symptoms of this disease, with a high frequency of skeletal abnormalities and no cognitive impairment reported in our patients (Northover, Cowie, and Wraith 1996; Tomatsu et al. 2011).

Treatment of MPS is highly contingent on the type of the disease, which can be conclusively determined by molecular studies. While NGS provides a way to analyze all MPS genes with high specificity, it is an expensive technique, which requires well-equipped labs with high standards and specialized technicians. On the other hand, Sanger sequencing is a less complicated and more commonly available technique. By cross-referencing geographical patterns of frequent variants, it is possible to develop a stepwise diagnostic approach that can be used on clinically diagnosed MPS patients before they undergo costly biochemical assays or NGS analysis. This method can be helpful since we were able to identify variants that account for a significant proportion of the MPS patients originating from specific geographical regions as well as a few possible founder mutations. To illustrate, the c.430G>A in *ARSB* is concentrated in the southeastern regions of Iran, accounting for 88.2% of MPS VI patients, while c.962C>T is responsible for 85.7% of MPS VI cases in northwestern regions. Such patterns justify checking for c.430G>A and c.962C>T in MPS VI patients originating respectively from southeast or northwest before going through further diagnostic approaches, which could lead to a conclusive result at a fraction of the cost and effort. The same approach is suggested for MPS IV patients originating from East Azerbaijan, where c.319G>A in *GALNS* is responsible for 50% of MPS IV patients. Therefore, a stepwise genetic testing starting from c.319G>A analysis can be considered before performing a more comprehensive analysis.

Conclusion

In summary, this study presents the genetic profiling of the largest cohort of Iranian MPS patients consisting of 289 unrelated families. Our comprehensive diagnostic approach was able to molecularly characterize 258 patients harboring putatively causal variants in 8 MPS genes, in addition to identification of five non-MPS IEM patients. Notably, 87.2% of the diagnosed patients were homozygous, besides a high proportion of consanguineous marriages among this cohort (83%). MPS IV was the most common type in this study followed by MPS I and MPS VI, contrary to previous reports where MPS VI contributed to a considerably low proportion of the patients. Furthermore, we identified a region-specific pattern for several predominant pathogenic variants (eg. *ARSB* : c.430G>A, c.962C>T and *GALNS* : c.319G>A) which provides insight into genetic epidemiology of MPS and can facilitate a more cost-effective, time-efficient diagnostic approach.

Conflict of Interest Statement:

The authors have no conflict of interest to declare.

References

Abbasi, S., M. Noruzinia, O. Bashti, M. Ahmadvand, A. R. Salehi Chaleshtori, and L. Mahootipou. 2017. "Another Novel Missense Mutation in ARSB Gene in Iran." *Acta Medica Iranica* 55(9):585–90.

Alaei, Mohammad Reza, Meghdad Kheirkhahan, Saeed Talebi, Elham Davoudi-Dehaghani, and Mohammad Keramatipour. 2020. "Once in a Blue Moon, a Very Rare Coexistence of Glutaric Acidemia Type I and Mucopolysaccharidosis Type IIIB in a Patient." *Iranian Biomedical Journal* 24(3):201–5. doi: 10.29252/ibj.24.3.201.

Alif, N., K. Hess, J. Straczek, S. Sebbar, A. N'Bou, P. Nabet, and B. Dousset. 1999. "Mucopolysaccharidosis Type I: Characterization of a Common Mutation That Causes Hurler Syndrome in Moroccan Subjects." *Annals of Human Genetics* 63(Pt 1):9–16. doi: 10.1046/j.1469-1809.1999.6310009.x [doi].

Al-Jasmi, F. A., N. Tawfig, A. Berniah, B. R. Ali, M. Taleb, J. L. Hertecant, F. Bastaki, and A. K. Souid. 2013. "Prevalence and Novel Mutations of Lysosomal Storage Disorders in United Arab Emirates : LSD in UAE." *JIMD Reports* 10:1–9. doi: 10.1007/8904_2012_182 [doi].

Aminzadeh, M., N. Malekpour, and P. Ghandil. 2019. "Identification of Arylsulfatase B Gene Mutations and Clinical Presentations of Iranian Patients with Mucopolysaccharidosis VI." *Gene* 706:1–5. doi: S0378-1119(19)30403-2 [pii].

Atceken, N., R. K. Ozgul, D. Yucel Yilmaz, A. Tokatli, T. Coskun, H. S. Sivri, A. Dursun, and M. Karaca. 2016. "Evaluation and Identification of IDUA Gene Mutations in Turkishpatients with Mucopolysaccharidosis Type I." *Turkish Journal of Medical Sciences* 46(2):404–8. doi: 10.3906/sag-1411-160 [doi].

Bertola, F., M. Filocamo, G. Casati, M. Mort, C. Rosano, A. Tylki-Szymanska, B. Tuysuz, O. Gabrielli, S. Grossi, M. Scarpa, G. Parenti, D. Antuzzi, J. Dalmau, M. Di Rocco, C. Dionisi Vici, I. Okur, J. Rosell, A. Rovelli, F. Furlan, M. Rigoldi, A. Biondi, D. N. Cooper, and R. Parini. 2011. "IDUA Mutational Profiling of a Cohort of 102 European Patients with Mucopolysaccharidosis Type I: Identification and Characterization of 35 Novel Alpha-L-Iduronidase (IDUA) Alleles." *Human Mutation* 32(6):2189. doi: 10.1002/humu.21479 [doi].

Bidchol, A. M., A. Dalal, H. Shah, S. S, S. Nampoothiri, M. Kabra, N. Gupta, S. Danda, K. Gowrishankar, S. R. Phadke, S. Kapoor, M. Kamate, I. C. Verma, R. D. Puri, V. H. Sankar, A. R. Devi, S. J. Patil, P. Ranganath, S. J. Jain, M. Agarwal, A. Singh, P. Mishra, P. M. Tamhankar, P. M. Gopinath, H. A. Nagarajaram, K. Satyamoorthy, and K. M. Girisha. 2014. "GALNS Mutations in Indian Patients with Mucopolysaccharidosis IVA." *American Journal of Medical Genetics.Part A* 164A(11):2793–2801. doi: 10.1002/ajmg.a.36735 [doi].

Brusius-Facchin, A. C., M. Siebert, D. Leao, D. R. Malaga, G. Pasqualim, F. Trapp, U. Matte, R. Giugliani, and S. Leistner-Segal. 2019. "Phenotype-Oriented NGS Panels for Mucopolysaccharidoses: Validation and Potential Use in the Diagnostic Flowchart." *Genetics and Molecular Biology* 42(1 suppl 1):207–14. doi: S1415-47572019000200207 [pii].

Chkioua, L., S. Khedhiri, A. Kassab, A. Bibi, S. Ferchichi, R. Froissart, C. Vianey-Saban, S. Laradi, and A. Miled. 2011. "Molecular Analysis of Mucopolysaccharidosis Type I in Tunisia: Identification of Novel Mutation and Eight Novel Polymorphisms." *Diagnostic Pathology* 6:39–39. doi: 10.1186/1746-1596-6-39 [doi].

Cobos, P. N., C. Steglich, R. Santer, Z. Lukacs, and A. Gal. 2015. "Dried Blood Spots Allow Targeted Screening to Diagnose Mucopolysaccharidosis and Mucolipidosis." *JIMD Reports* 15:123–32. doi: 10.1007/8904_2014_308 [doi].

Consortium, 1000 Genomes Project, G. R. Abecasis, D. Altshuler, A. Auton, L. D. Brooks, R. M. Durbin, R. A. Gibbs, M. E. Hurles, and G. A. McVean. 2010. "A Map of Human Genome Variation from Population-Scale Sequencing." *Nature* 467(7319):1061–73. doi: 10.1038/nature09534 [doi].

Costa-Motta, F. M., F. Bender, A. Acosta, K. Abe-Sandes, T. Machado, T. Bomfim, T. Boa Sorte, D. da Silva, A. Bittles, R. Giugliani, and S. Leistner-Segal. 2014. "A Community-Based Study of Mucopolysaccharidosis Type VI in Brazil: The Influence of Founder Effect, Endogamy and Consanguinity." *Human Heredity* 77(1–4):189–96. doi: 10.1159/000358404 [doi].

Eisenberger, T., C. Neuhaus, A. O. Khan, C. Decker, M. N. Preising, C. Friedburg, A. Bieg, M. Gliem, P. Charbel Issa, F. G. Holz, S. M. Baig, Y. Hellenbroich, A. Galvez, K. Platzer, B. Wollnik, N. Laddach, S. R. Ghaffari, M. Rafati, E. Botzenhart, S. Tinschert, D. Borger, A. Bohring, J. Schreml, S. Kortge-Jung, C. Schell-Apacik, K. Bakur, J. Y. Al-Aama, T. Neuhann, P. Herkenrath, G. Nurnberg, P. Nurnberg, J. S. Davis, A. Gal, C. Bergmann, B. Lorenz, and H. J. Bolz. 2013. "Increasing the Yield in Targeted Next-Generation Sequencing by Implicating CNV Analysis, Non-Coding Exons and the Overall Variant Load: The Example of Retinal Dystrophies." *PloS One* 8(11):e78496. doi: 10.1371/journal.pone.0078496 [doi].

Filocamo, M., R. Tomanin, F. Bertola, and A. Morrone. 2018. *Biochemical and Molecular Analysis in Mucopolysaccharidoses: What a Paediatrician Must Know* . Vol. 44.

Fox, E. J., K. S. Reid-Bayliss, M. J. Emond, and L. A. Loeb. 2014. "Accuracy of Next Generation Sequencing Platforms." *Next Generation, Sequencing & Applications* 1:10.4172/jngsa.1000106. doi: 1000106 [pii].

Ghaffari, Saeed Reza, and Maryam Rafati. 2020. "Iran Variome MPS Database." Retrieved June 26, 2021 (https://variome.ir/lovd).

Harismendy, O., P. C. Ng, R. L. Strausberg, X. Wang, T. B. Stockwell, K. Y. Beeson, N. J. Schork, S. S. Murray, E. J. Topol, S. Levy, and K. A. Frazer. 2009. "Evaluation of next Generation Sequencing Platforms for Population Targeted Sequencing Studies." *Genome Biology* 10(3):R32-r32. Epub 2009 Mar 27. doi: 10.1186/gb-2009-10-3-r32 [doi].

Jafaryazdi, Rokhsareh, Sedigheh Shams, Anna Isaian, Aria Setoodeh, and Shahram Teimourian. 2019. "Identification of Eleven Different Mutations Including Six Novel, in the Arylsulfatase B Gene in Iranian Patients with Mucopolysaccharidosis Type VI." *Molecular Biology Reports* 46(3):3417–26. doi: 10.1007/s11033-019-04804-9.

Jafaryazdi, Rokhsareh, Sediqheh Shams, Anna Isaian, Emel Ebadi, Reza Safaralizadeh, and Shahram Teimourian. 2018. "A Novel Compound Heterozygote Mutation in the ARSB Gene in a Patient with Maroteaux-Lamy Syndrome and Its Insilico Evaluation." *Meta Gene* 18:127–31. doi: https://doi.org/10.1016/j.mgene.2018.08.013.

Kadali, S., S. M. Naushad, A. Radha Rama Devi, and V. L. Bodiga. 2019. *Biochemical, Machine Learning and Molecular Approaches for the Differential Diagnosis of Mucopolysaccharidoses* . Vol. 458.

Kamranjam, M., and M. Alaei. 2019. "Mutation Analysis of the IDUA Gene in Iranian Patients with Mucopolysaccharidosis Type 1: Identification of Four Novel Mutations." *Genetic Testing and Molecular Biomarkers* 23(8):515–22. doi: 10.1089/gtmb.2019.0022 [doi].

Kantaputra, P. N., H. Kayserili, Y. Guven, W. Kantaputra, M. C. Balci, P. Tanpaiboon, N. Tananuvat, A. Uttarilli, and A. Dalal. 2014. "Clinical Manifestations of 17 Patients Affected with Mucopolysaccharidosis Type VI and Eight Novel ARSB Mutations." *American Journal of Medical Genetics. Part A* 164A(6):1443–53. doi: 10.1002/ajmg.a.36489 [doi].

Khan, S. A., H. Peracha, D. Ballhausen, A. Wiesbauer, M. Rohrbach, M. Gautschi, R. W. Mason, R. Giugliani, Y. Suzuki, K. E. Orii, T. Orii, and S. Tomatsu. 2017. "Epidemiology of Mucopolysaccharidoses." *Molecular Genetics and Metabolism* 121(3):227–40. doi: S1096-7192(17)30206-8 [pii].

Khorrami, Mehdi, Manijeh Mahdavi, Fatemeh Fakhr, and Majid Kheirollahi. 2019. "In Silico Analysis of A Novel Pathogenic Variant in an Iranian Mucopolysaccharidosis IIIB Patient Identified by Targeted Next-Generation Sequencing." *Iranian Journal of Child Neurology* 13(4):173–83. doi: 10.22037/ijcn.v13i4.19894.

Laradi, S., T. Tukel, M. Erazo, J. Shabbeer, L. Chkioua, S. Khedhiri, S. Ferchichi, M. Chaabouni, A. Miled, and R. J. Desnick. 2005. "Mucopolysaccharidosis I: Alpha-L-Iduronidase Mutations in Three Tunisian Families." *Journal of Inherited Metabolic Disease* 28(6):1019–26. doi: 10.1007/s10545-005-0197-4 [doi].

Malekpour, Nasrin, Rahim Vakili, and Tayebeh Hamzehloie. 2018. "Mutational Analysis of ARSB Gene in Mucopolysaccharidosis Type VI: Identification of Three Novel Mutations in Iranian Patients." *Iranian Journal of Basic Medical Sciences* 21(9):950–56. doi: 10.22038/IJBMS.2018.27742.6760.

Mansour, Salehi, Salehi Rasoul, and Nasr Esfahani B. 2007. "MOLECULAR DETECTION AND IDENTIFICATION OF ALFA- L-IDURONIDASE GENE MUTATIONSJN 5 IRANIAN FAMILIES SUSPECTED FOR MULLER SYNDROME (MUCOPOLYSACCHARIDOSIS I)." 15(2):57–63.

Martin, R., M. Beck, C. Eng, R. Giugliani, P. Harmatz, V. Munoz, and J. Muenzer. 2008. "Recognition and Diagnosis of Mucopolysaccharidosis II (Hunter Syndrome)." *Pediatrics* 121(2):377. doi: 10.1542/peds.2007-1350 [doi].

Martins, Carla, Paula Frassinetti V. de Medeiros, Sandra Leistner-Segal, Larbi Dridi, Nursel Elcioglu, Jill Wood, Mahdiyeh Behnam, Bilge Noyan, Lucia Lacerda, Michael T. Geraghty, Damian Labuda, Roberto Giugliani, and Alexey V. Pshezhetsky. 2019. "Molecular Characterization of a Large Group of Mucopolysaccharidosis Type IIIC Patients Reveals the Evolutionary History of the Disease." *Human Mutation* 40(8):1084–1100. doi: https://doi.org/10.1002/humu.23752.

Montano, A. M., N. Lock-Hock, R. D. Steiner, B. H. Graham, M. Szlago, R. Greenstein, M. Pineda, A. Gonzalez-Meneses, M. Coker, D. Bartholomew, M. S. Sands, R. Wang, R. Giugliani, A. Macaya, G. Pastores, A. K. Ketko, F. Ezgu, A. Tanaka, L. Arash, M. Beck, R. E. Falk, K. Bhattacharya, J. Franco, K. K. White, G. A. Mitchell, L. Cimbalistiene, M. Holtz, and W. S. Sly. 2016. "Clinical Course of Sly Syndrome (Mucopolysaccharidosis Type VII)." *Journal of Medical Genetics* 53(6):403–18. doi: 10.1136/jmedgenet-2015-103322 [doi].

Morrone, A., A. Caciotti, R. Atwood, K. Davidson, C. Du, P. Francis-Lyon, P. Harmatz, M. Mealiffe, S. Mooney, T. R. Oron, A. Ryles, K. A. Zawadzki, and N. Miller. 2014. "Morquio A Syndrome-Associated Mutations: A Review of Alterations in the GALNS Gene and a New Locus-Specific Database." *Human Mutation* 35(11):1271–79. doi: 10.1002/humu.22635 [doi].

Mosallanejad, A., M. Alaei, S. R. Ghaffari, M. Rafati, and H. Saneifard. 2020. "Non-Progressive Nonimmune Hydrops Fetalis Caused by a Novel Mutation in GUSB Gene." *Iranian Journal of Child Neurology* 14(2):101–6. doi: jcn-14-101 [pii].

Muenzer, J., M. Beck, R. Giugliani, Y. Suzuki, A. Tylki-Szymanska, V. Valayannopoulos, A. Vellodi, and J. E. Wraith. 2011. "Idursulfase Treatment of Hunter Syndrome in Children Younger than 6 Years: Results from the Hunter Outcome Survey." *Genetics in Medicine : Official Journal of the American College of Medical Genetics* 13(2):102–9. doi: 10.1097/GIM.0b013e318206786f [doi].

Muenzer, J., J. E. Wraith, L. A. Clarke, International Consensus Panel on Management, and I. Treatment of Mucopolysaccharidosis. 2009. "Mucopolysaccharidosis I: Management and Treatment Guidelines." *Pediatrics* 123(1):19–29. doi: 10.1542/peds.2008-0416 [doi].

Northover, H., R. A. Cowie, and J. E. Wraith. 1996. "Mucopolysaccharidosis Type IVA (Morquio Syndrome): A Clinical Review." *Journal of Inherited Metabolic Disease* 19(3):357–65. doi: 10.1007/BF01799267 [doi].

Nouri, Nayerossadat, Nargesossadat Nouri, Omid Aryani, Behnam Kamalidehghan, and Massoud Houshmand. 2012. "Identification of a Novel Arylsulfatase B Gene Mutation in Three Unrelated Iranian Mucopolysaccharidosis Type-VI Patients with Different Phenotype Severity." *Iranian Biomedical Journal* 16(3):169–71. doi: 10.6091/ibj.1049.2012.

Papadimitriou, Sofia, Andrea Gazzo, Nassim Versbraegen, Charlotte Nachtegael, Jan Aerts, Yves Moreau, Sonia Van Dooren, Ann Nowe, Guillaume Smits, and Tom Lenaerts. 2019. "Predicting Disease-Causing Variant Combinations." *Proceedings of the National Academy of Sciences of the United States of America* 116(24):11878–87. doi: 10.1073/pnas.1815601116.

Quail, M. A., I. Kozarewa, F. Smith, A. Scally, P. J. Stephens, R. Durbin, H. Swerdlow, and D. J. Turner. 2008. "A Large Genome Center's Improvements to the Illumina Sequencing System." *Nature Methods* 5(12):1005–10. doi: 10.1038/nmeth.1270 [doi].

Regier, Debra S., Matthew Oetgen, and Pranoot Tanpaiboon. 1993. "Mucopolysaccharidosis Type IVA." in *GeneReviews(r)* , edited by M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. Bean, G. Mirzaa, and A. Amemiya. Seattle (WA): University of Washington, Seattle.

Salehi, R. 2004. "REPORT OF A NOVEL MUTATION IN A-L-IDURONIDASE (IDUA) IN AN IRANIAN FAMILY WITH HURLER SYNDROME (MUCOPOLYSACCHARIDOSIS I)." 17(1):80–88.

Scarpa, M. 1993. "Mucopolysaccharidosis Type II." *GeneReviews((R))* , edited by M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, and A. Amemiya. Seattle (WA): University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved.

Seyedhassani, Seyed Mohammad, Feyzollah Hashemi-Gorji, Mahdieh Yavari, and Reza Mirfakhraie. 2015. "Novel Missense Mutation in the GALNS Gene in an Affected Patient with Severe Form of Mucopolysaccharidosis Type IVA." *Clinica Chimica Acta* 450:121–24. doi: https://doi.org/10.1016/j.cca.2015.08.006.

Shafaat, Seyed Mehdi, Mehrdad Hashemi, Ahmad Majd, Maryam Abiri, and Sirous Zeinali. 2020. "Genetic Assessment of Mucopolysaccharidosis Type IV and the First Pathological Mutation of c.313A>G in the Iranian Population." *The Horizon of Medical Sciences* 26(2):182–91. doi: 10.32598/hms.26.2.3275.1.

Stenson, Peter D., Matthew Mort, Edward V. Ball, Katy Shaw, Andrew Phillips, and David N. Cooper. 2014. "The Human Gene Mutation Database: Building a Comprehensive Mutation Repository for Clinical and Molecular Genetics, Diagnostic Testing and Personalized Genomic Medicine." *Human Genetics* 133(1):1–9. doi: 10.1007/s00439-013-1358-4.

Taghikhani, Mohammad, Shohreh Khatami, Mohammad Abdi, Mohammad Said Hakhamaneshi, Mohammad Reza Alaei, Daniel Zamanfar, and Rahim Vakili. 2019. "Mutation Analysis and Clinical Characterization of Iranian Patients with Mucopolysaccharidosis Type I." *Journal of Clinical Laboratory Analysis* 33(8):e22963. doi: https://doi.org/10.1002/jcla.22963.

Tomanin, R., L. Karageorgos, A. Zanetti, M. Al-Sayed, M. Bailey, N. Miller, H. Sakuraba, and J. J. Hopwood. 2018. "Mucopolysaccharidosis Type VI (MPS VI) and Molecular Analysis: Review and Classification of Published Variants in the ARSB Gene." *Human Mutation* 39(12):1788–1802. doi: 10.1002/humu.23613 [doi].

Tomatsu, S., M. Filocamo, K. O. Orii, W. S. Sly, M. A. Gutierrez, T. Nishioka, O. P. Serrato, P. Di Natale, A. M. Montano, S. Yamaguchi, N. Kondo, T. Orii, and A. Noguchi. 2004. "Mucopolysaccharidosis IVA (Morquio A): Identification of Novel Common Mutations in the N-Acetylgalactosamine-6-Sulfate Sulfatase (GALNS) Gene in Italian Patients." *Human Mutation* 24(2):187–88. doi: 10.1002/humu.9265 [doi].

Tomatsu, S., A. M. Montano, H. Oikawa, M. Smith, L. Barrera, Y. Chinen, M. M. Thacker, W. G. Mackenzie, Y. Suzuki, and T. Orii. 2011. "Mucopolysaccharidosis Type IVA (Morquio A Disease): Clinical Review and

Current Treatment." *Current Pharmaceutical Biotechnology* 12(6):931–45. doi: 1389-2010/11 $58.00+.00 [doi].

Valstar, M. J., S. Neijs, H. T. Bruggenwirth, R. Olmer, G. J. Ruijter, R. A. Wevers, O. P. van Diggelen, B. J. Poorthuis, D. J. Halley, and F. A. Wijburg. 2010. "Mucopolysaccharidosis Type IIIA: Clinical Spectrum and Genotype-Phenotype Correlations." *Annals of Neurology* 68(6):876–87. doi: 10.1002/ana.22092 [doi].

Valstar, Marlies J., Aida Bertoli-Avella, Marja W. Wessels, George J. G. Ruijter, Bianca de Graaf, Renske Olmer, Peter Elfferich, Sanne Neijs, Roxana Kariminejad, Fatih Suheyl Ezgü, Aysegul Tokatli, Barbara Czartoryska, Ad N. Bosschaart, Feikje van den Bos-Terpstra, Hugues Puissant, Friederike Bürger, Heymut Omran, D. Eckert, Mirella Filocamo, Emil Simeonov, Patrick J. Willems, Ron A. Wevers, Martinus F. Niermeijer, Dicky J. J. Halley, Ben J. H. M. Poorthuis, and Otto P. van Diggelen. 2010. "Mucopolysaccharidosis Type IIID: 12 New Patients and 15 Novel Mutationsb." *Human Mutation* 31(5):E1348–60. doi: https://doi.org/10.1002/humu.21234.

Wang, Z., W. Zhang, Y. Wang, Y. Meng, L. Su, H. Shi, and S. Huang. 2010. "Mucopolysaccharidosis IVA Mutations in Chinese Patients: 16 Novel Mutations." *Journal of Human Genetics* 55(8):534–40. doi: 10.1038/jhg.2010.65 [doi].

Wood, T. C., K. Harvey, M. Beck, M. G. Burin, Y. H. Chien, H. J. Church, V. D'Almeida, O. P. van Diggelen, M. Fietz, R. Giugliani, P. Harmatz, S. M. Hawley, W. L. Hwu, D. Ketteridge, Z. Lukacs, N. Miller, M. Pasquali, A. Schenone, J. N. Thompson, K. Tylee, C. Yu, and C. J. Hendriksz. 2013. "Diagnosing Mucopolysaccharidosis IVA." *Journal of Inherited Metabolic Disease* 36(2):293–307. doi: 10.1007/s10545-013-9587-1 [doi].

Yassaee, Vahid Reza, Feyzollah Hashemi-Gorji, Mohammad Miryounesi, Alireza Rezayi, Zeinab Ravesh, Fakhrolmolouk Yassaee, and Shadab Salehpour. 2017. "Clinical, Biochemical and Molecular Features of Iranian Families with Mucopolysaccharidosis: A Case Series." *Clinica Chimica Acta* 474:88–95. doi: https://doi.org/10.1016/j.cca.2017.08.017.

Tables

Table 1. Patients diagnosed with non-MPS inborn errors of metabolism.

ID	Clinical Diagnosis	Pedigree information	Pedigree information	Pedigree information	Genetic Findings	Genetic Findings	Genetic Findings	Genetic Findings	Genetic Findings	Genetic Findings	Genetic Findings	
		Parental Consanguinity	Sporadic	Familial/Inheritance	Gene	Nucleic Acid Alteration	Amino Acid Alteration	RefSeq 1 (NM)	Coding Impact	Genotype	Known/Novel	
IEM-1	MPS	FC	Sporadic	AR	NEU1	c.727G>A	p.Gly243Arg	NM_000434.3	Missense	Homozygous	Known	
IEM-2	MPS IV	FC once removed	Familial	AR	GNPTG	c.426del	p.Cys142fs	NM_032520.4	Frameshift deletion	Homozygous	Novel	
IEM-3	MPS II	FC	Sporadic	AR	GNPTAB	c.1212C>G	p.Tyr404Ter	NM_024312.4	Nonsense	Homozygous	Known	

19

IEM-4	MPS II	FC	Sporadic	AR	L2HGDH	c.1015delA	p.Arg339fs	NM_-024884.2	Frameshift deletion	Homozygous	Known	P
IEM-5	MPS IV	NC	Sporadic	AR	CTSA	c.1090G>A	p.Val364Met	NM_-000308.2	Missense	Heterozygous	Novel	L
						c.1396G>A	p.Ala466Thr	NM_-000308.2	Missense	Heterozygous	Novel	V
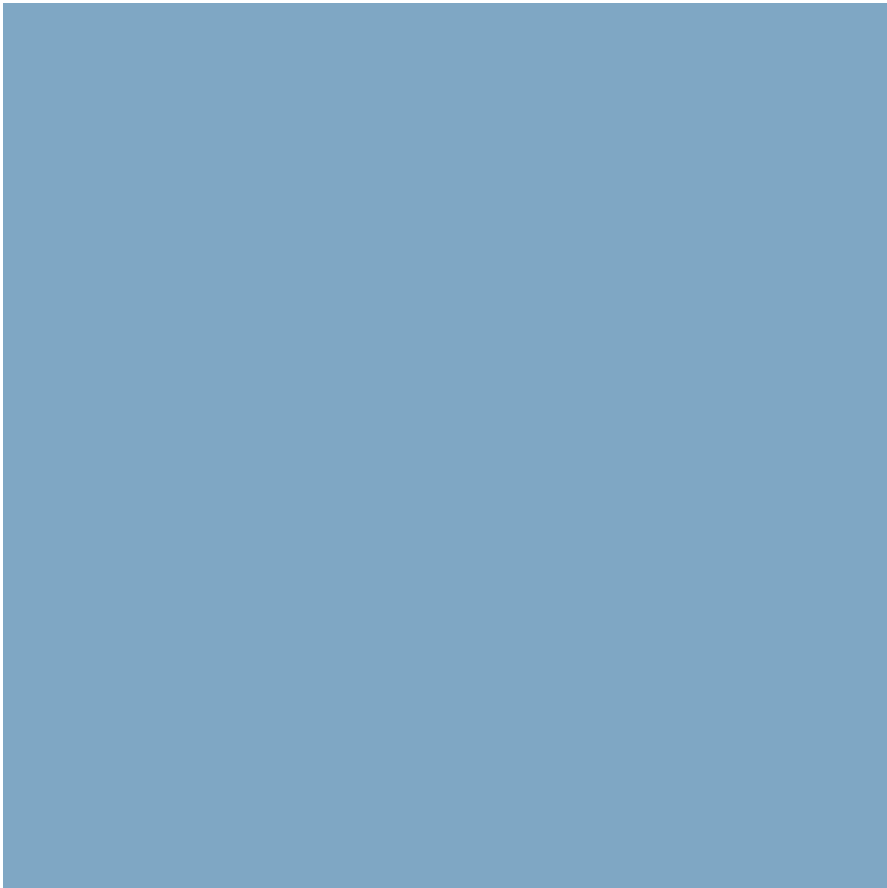
IEM: Inborn Errors of Metabolism, FC: First Cousin, NC: No Consanguinity, AR: Autosomal Recessive, P: Pathogenic, LP: Likely Pathogenic, VUS: Variant of Uncertain Significance

Table 2. Demographics and molecular characteristics of patients with MPS causal variants within each subtype of the disease.

Variable	Total (N=258)	MPS I (N=59)	MPS II (N=18)	MPS III (N=33)	MPS IV (N=90)	MPS VI (N=58)
Familial/Sporadic						
Familial	84 (32.6%)	19 (32.2%)	6 (33.3%)	8 (24.2%)	31 (34.4%)	20 (34.5%)
Sporadic	171 (66.3%)	39 (66.1%)	12 (66.7%)	25 (75.8%)	59 (65.6%)	36 (62%)
Unknown	3 (1.1%)	1 (1.7%)	-		-	2 (3.5%)
Consanguinity						

DFC	4 (1.6%)	-	-	-	2 (2.2%)	2 (3.4%)
FC/FCOR	179 (69.4%)	43 (72.9%)	4 (2.2%)	24 (72.8%)	64 (71.1%)	44 (75.9%)
SC or further	28 (10.9%)	2 (3.4%)	3 (16.7%)	4 (12.1%)	16 (17.8%)	3 (5.1%)
Non-consanguineous	42 (16.3%)	11 (18.6%)	11 (61.1%)	5 (15.1%)	8 (8.9%)	7 (12%)
Unknown	5 (1.9%)	3 (5.1%)	-	-	-	2 (3.4%)
Genotype						
Homozygous	225 (87.2%)	54 (91.5%)	-	30 (90.1%)	85 (94.4%)	56 (96.6%)
Compound Heterozygous	12 (4.7%)	4 (6.8%)	-	3 (9.9%)	4 (4.5%)	1 (1.7%)
Hemizygous	18 (7%)	-	18 (100%)	-	-	-
Heterozygous	3 (1.2%)	1 (1.7%)	-	-	1 (1.1%)	1 (1.7%)
Coding Impact						
Missense	168 (65.1%)	33 (55.9%)	8 (4.4%)	19 (57.6%)	66 (73.2%)	42 (72.4%)
Nonsense	31 (12%)	5 (8.5%)	2 (1.1%)	6 (18.2%)	6 (6.7%)	12 (20.7%)
Frameshift Insertion	5 (1.9%)	3 (5.2%)	1 (5.6%)		1 (1.1%)	-
Frameshift Deletion	15 (5.8%)	8 (1.4%)	1 (5.6%)		5 (5.6%)	1 (1.7%)
Non-frameshift Insertion	2(0.8%)	1 (1.7%)	-	1 (3%)		-
Non-frameshift Deletion	9 (3.5%)	1 (1.7%)	3 (1.6%)	4 (12.1%)	1 (1.1%)	-
Splice-site	18 (7%)	7 (11.9)	2 (1.1%)	2 (6.1%)	5 (5.6%)	2 (3.5%)
Stop-loss	1 (0.4%)	-				1 (1.7%)
CNV	9 (3.5%)	1 (1.7%)	1 (5.6%)	1 (3%)	6 (6.7%)	-
Novel/Known						
Known	168 (65.1%)	36 (61%))	9 (50%)	18 (54.6%)	55 (61.1%)	50 (86.2%)
Novel	90 (34.9%)	23 (39%)	9 (50%)	15 (45.4%)	35 (38.9%)	8 (13.8%)

Table 3. Patients with Copy Number Variations in MPS genes.

ID	Clinical Diagnosis	Pedigree information Parental Consanguinity	Pedigree information Sporadic/Familial	Genetic Findings Gene	Genetic Nucleic
MPS-12	MPS IV	FC	Sporadic	GALNS	Exons 5-8
MPS-19	MPS IV	FC	Sporadic	GALNS	Exon 5 de
MPS-61	MPS I	FC	Sporadic	IDUA	Exons 1-8
MPS-97	MPS IV	FC	Sporadic	GALNS	Exon 5 de
MPS-106	MPS IV	FC	Familial	GALNS	Exons 2-4
MPS-132	MPS III	FC	Sporadic	HGSNAT	Exons 9-1
MPS-207	MPS	FC once removed	Sporadic	GALNS	Exon 13 c
MPS-209	MPS IV	SC	Sporadic	GALNS	Exons 1-3
MPS-270	MPS II	FC once removed	Sporadic	IDS	Whole ger

FC: First Cousins, SC: Second Cousins, AR: Autosomal Recessive, XLR: X-Linked Recessive P: Pathogenic

Figure Legends

Figure 1. Schematics of study workflow. n: the number of patients in the group. MPS: Mucopolysaccharidosis. IEM: Inborn Errors of Metabolism. IGV: Integrative Genomics Viewer.

Figure 2. Clustering heatmap representing the clinical manifestation of genetically confirmed MPS patients.
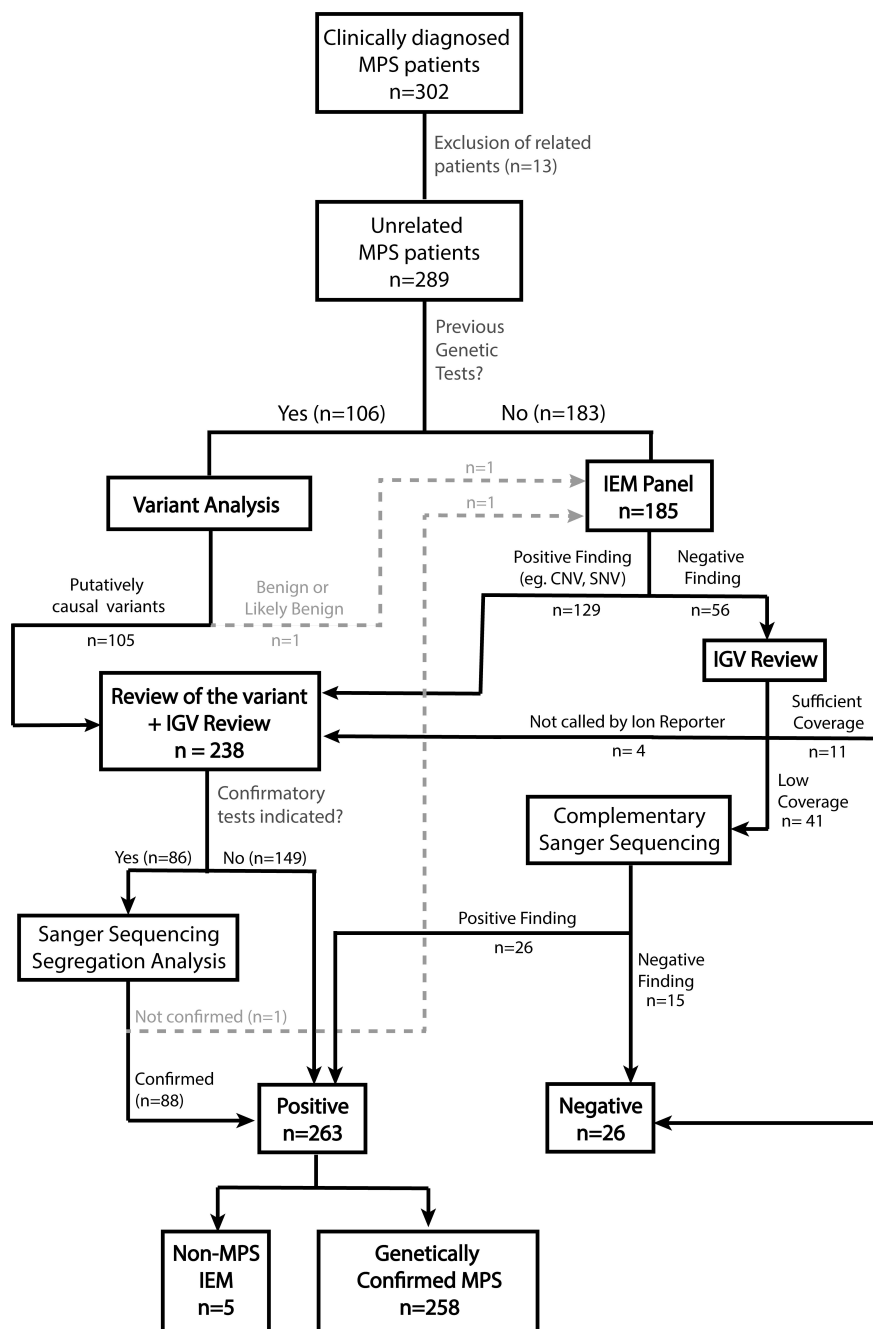
Figure 3. MPS patients' ethnicity and age demographics. A) Number of MPS patients among different ethnicities of the Iranian population and relative distribution of MPS types among them. B) Distribution of MPS patients based on their age at diagnosis divided by each MPS type.
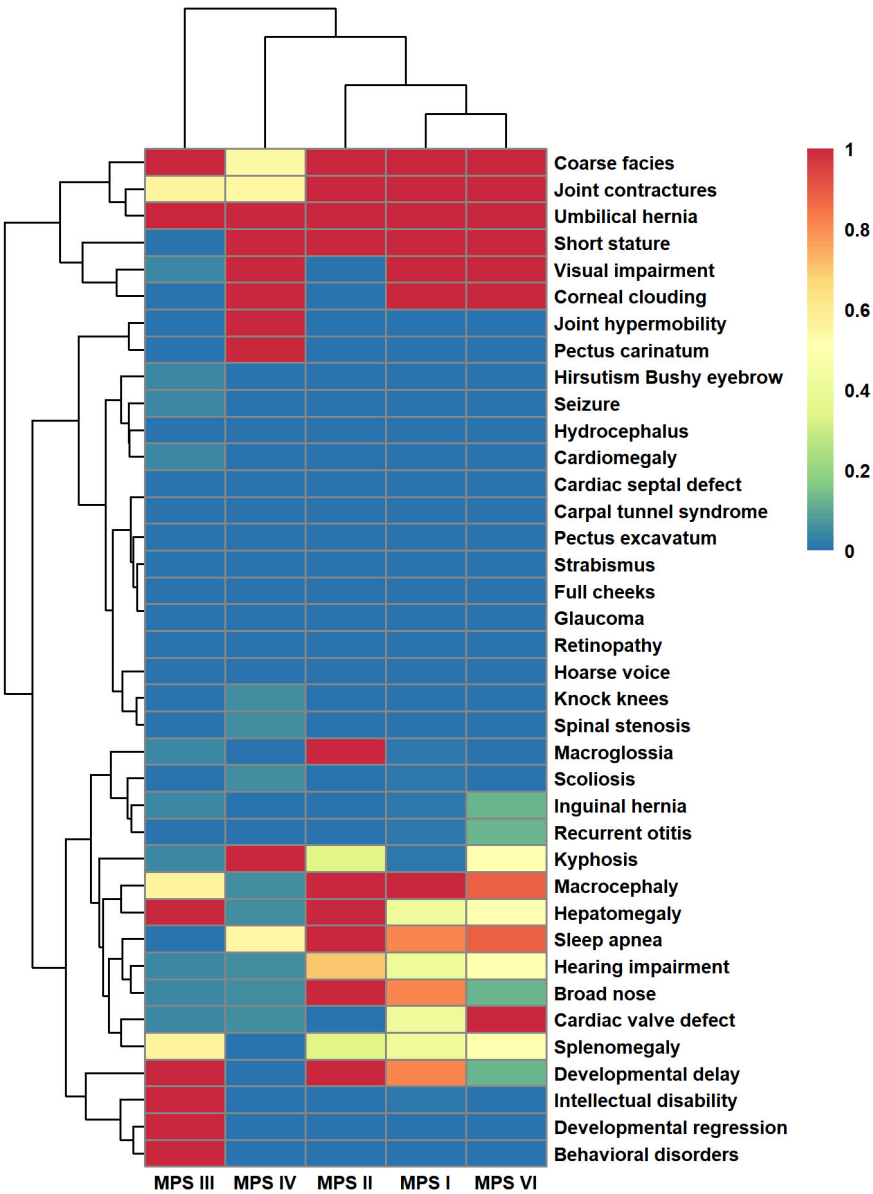
Figure 4. Geographical distribution of pathogenic variants associated with MPS IV. Total number of patients = 90. Geographical distribution was determined based on the parental origin of the affected individuals. Variants with a concentrated pattern of distribution in a region are shown underneath.
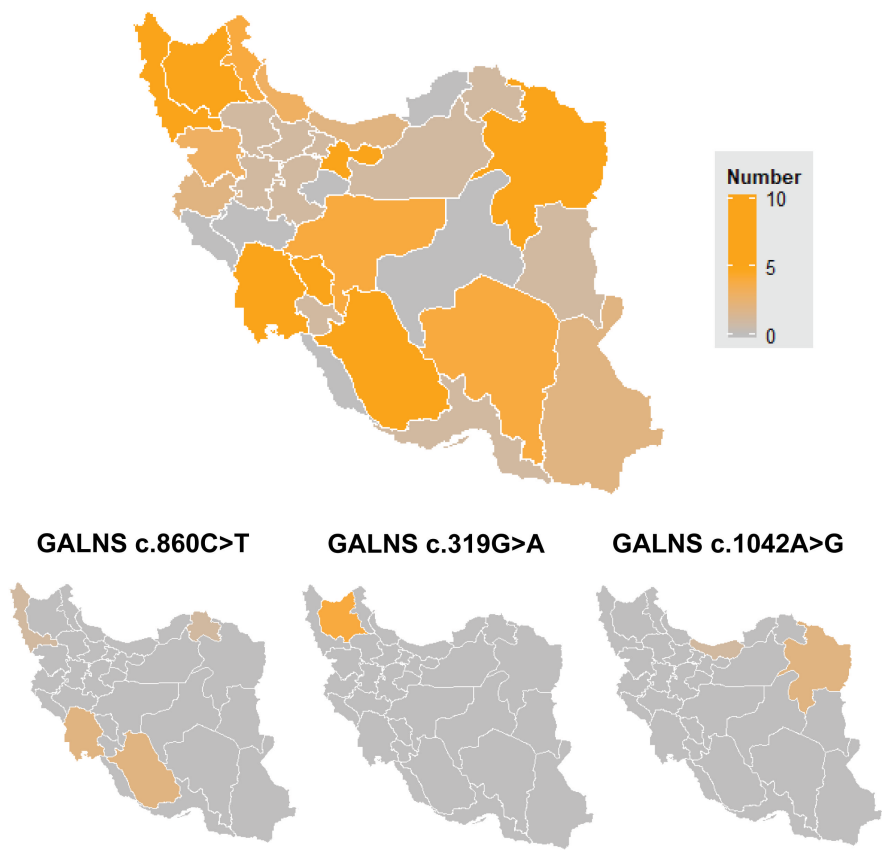
Figure 5. Geographical distribution of pathogenic variants in the Iranian cohort associated with MPS I. Total number of patients = 59. Geographical origin was determined based on the parental origin of the affected individuals. Variants with a concentrated pattern of distribution in a region are shown underneath.

Figure 6. Geographical distribution of pathogenic variants in the Iranian cohort associated with MPS VI. Total number of patients = 59. Geographical distribution was determined based on the parental origin of the affected individuals. Variants with a concentrated pattern of distribution in a region are shown underneath.

Figure 7. Geographical distribution of pathogenic variants in the Iranian cohort associated with MPS II and III. Total number of MPS II patients = 18. Total number of MPS III patients = 33. Geographical origin was determined based on the parental origin of the affected individuals.
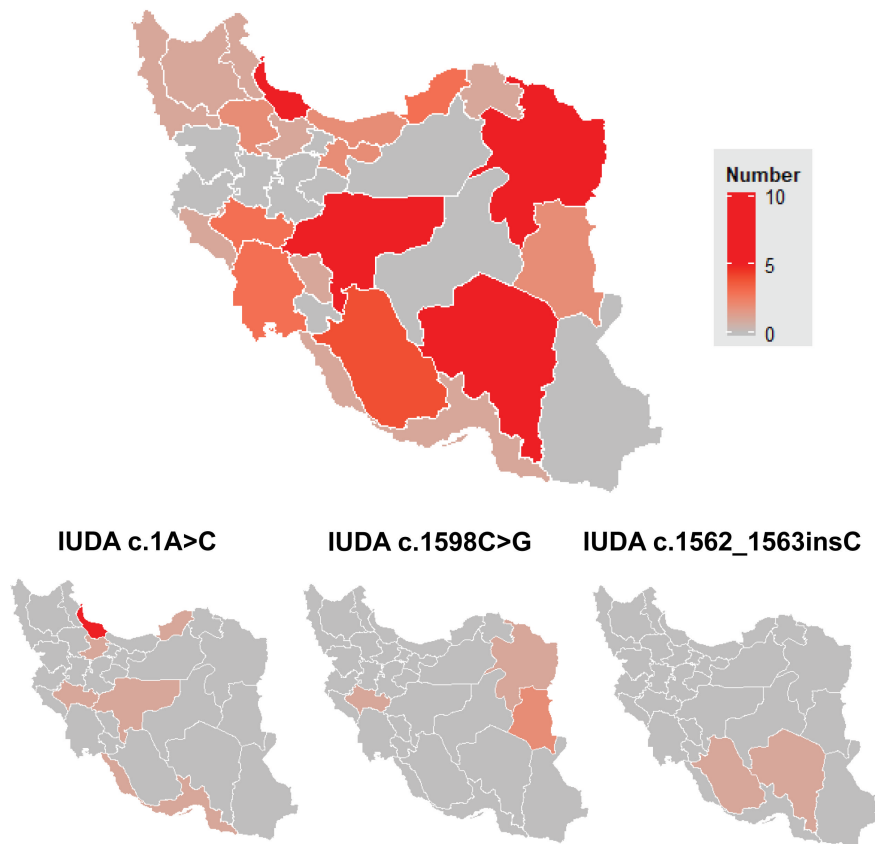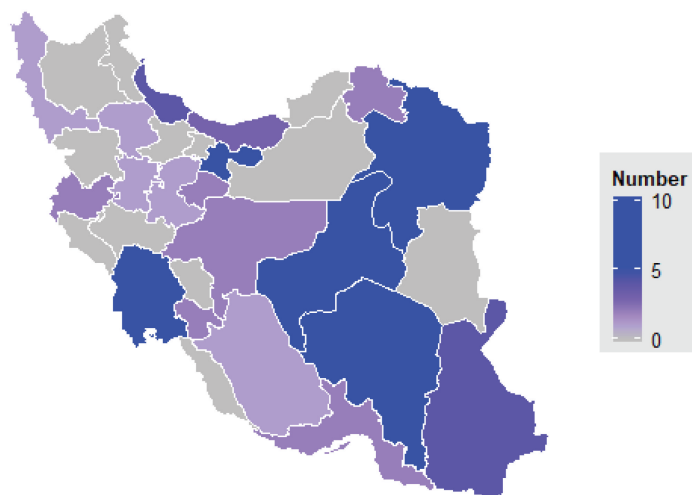
A MPS Ethnicity Distribution

B Age at Diagnosis

MPS IV Distribution

GALNS c.860C>T   GALNS c.319G>A   GALNS c.1042A>G

MPS I Distribution

IUDA c.1A>C   **IUDA c.1598C>G**   **IUDA c.1562_1563insC**
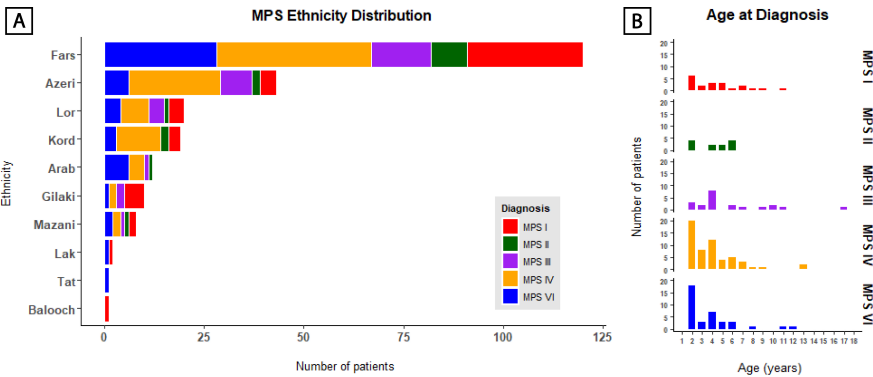
MPS VI Distribution

ARSB c.430G>A   **ARSB c.962T>C**   **ARSB c.281C>A**

A   MPS II Distribution   **B**   MPS III Distribution