

Genomic variability of the mucopolysaccharidosis complex in southwestern Colombia

L.J. Moreno-Giraldo^{1,2,3,4}, J.M. Satizábal-Soto^{1,2,3} and A. Sánchez-Gomez^{1,4}

¹ Universidad del Valle, Santiago de Cali, Colombia

² Universidad Santiago de Cali, Santiago de Cali, Colombia

³ Specialization in Pediatrics, Santiago de Cali, Colombia

⁴ Congenital Diseases of Metabolism Research Group, Universidad del Valle, Santiago de Cali, Colombia

Corresponding author: L.J. Moreno-Giraldo

E-mail: linajohannamoreno@yahoo.es

Genet. Mol. Res. 19 (2): gmr18502

Received October 16, 2019

Accepted May 31, 2020

Published June 30, 2020

DOI <http://dx.doi.org/10.4238/gmr18502>

ABSTRACT. Mucopolysaccharidosis (MPS) related information in Colombia is limited. A descriptive research involving results obtained in the full exome sequencing of 244 patients with different types of pathologies and not clinically diagnosed with MPS was carried out with the purpose of examining the genomic variability of genes associated with the MPS complex in South West Colombia, considering that it is a region where MPS cases are constantly reported. Bioinformatics software was applied with the purpose of analyzing the clinical significance of the different variants. The frequency of each of the variants was calculated, and an interaction network of the genes found within the MPS complex was developed. We found 509 different gene MPS complex variants, of which 262 were not previously reported. The most frequent genes were IDUA, GLB1 and GALNS, involving MPS I and MPS IV A/ B. In the entire MPS complex 9/244 variants were found associated with pathogenesis. Among the 509 variants, we found 27% missense, 2% non-coded 21% synonyms, 1.8% located within the splice region, 3% upstream, and 4.5% in the UTR region. These gene and allelic frequencies of the MPS complex will alert the medical community to the presence of the variants of the genes associated with MPS in the population in order to establish early diagnosis programs. This will

allow specific treatment for some of them, associated with transdisciplinary management that minimizes the morbidity and mortality attributed to this disease, including adequate genetic counseling.

Key words: Bioinformatics tools; Exome sequencing; Mucopolysaccharidosis Complex; Variability

INTRODUCTION

Mucopolysaccharidosis complex (MPS; OMIM252700) is a group of rare genetic disorders caused by specific somal enzyme deficiencies involved in the sequential degradation of the glycosaminoglycans (GAG) accumulated within different cells and tissues (Lin et al., 2019), leading to a progressive dysfunction of multiple organs. Seven different types of MPS disorders are recognized (I, II, III, IV, VI, VII and IX) with 11 specific lysosomal enzyme deficiencies (Table 1) (Zanetti et al., 2019). The clinical manifestations of MPS are progressive and chronic, with a wide range of clinical severity and prognosis of different varieties (Neufeld and Muenzer, 2001); The clinical presentation of patients with MPS, including implications of spreading, mainly affecting the liver, the spleen, the central nervous system, bones, cartilage, and eyes. When excessive amounts accumulate, it can affect development, skeleton abnormalities, cervical myelopathy as well as cord compression, defects within the immune and neural system and death at an early age (Manger, 2010; Giraldo et al., 2020).

All types of MPS have a mode of autosomal recessive inheritance, except for MPS II (The Hunter Syndrome), which is transmitted as a recessive form linked to the X and, therefore, affecting mainly males (Stapleton et al., 2017). MPS impact studies estimate frequencies of 2.04:100,000 in Asia and 1.2:100,000 in The United States (Lin et al., 2009).

The studies performed in Colombia by Gómez et al. (2012), proposed an incidence of 0.45:100,000 patients with MPS I, 0.17:100,000 with MPS III, 0.68:100,000 with MPS IV and 0.23:100,000 in patients with MPS VI; no impact of MPS VII and MPS IX were reported. However, these are the only figures reported for Colombia.

Information regarding the frequency of these diseases is scarce. Barrera (2009), reported that MPS IV is the most prevalent in the country, suggesting that Colombia might be one of the countries with the highest number of affected persons. Furthermore, research performed by Bernal and Briceño (2006) reported figures with characteristics suggesting MPS I and IV in ceramic studies of the Tumaco Culture - La Tolita and Pachajoa et al. (2014) which revealed evidence of possible presence of MPS VI in the Ceramics of indigenous peoples of the Americas since 2,000 years ago in the Colombian-Ecuadorian pacific coast belonging to the same culture. They may suggest a possible presence of said disease in our population since pre-Hispanic times as well as the presence of a founder's effect.

The MPS complex is classified with the Orphan Diseases in Colombia and form a wide and varied group of disorders which affect a reduced number of people. (Ministry of Public Health, 2018). The affected population faces difficulties in terms of lack of precise and timely diagnosis due to reduced prevalence.

Table 1. Approved classification of Mucopolysaccharidosis. Adapted from Albano et al. (2000)

MPS	Eponym	Enzyme	Mucopolysaccharide	Gene	Locus
I	Hurler's syndrome	α -L- iduronidase	Dermatan and heparan sulfate	IDUA	4p16.3
	Hurler Schei syndrome				
	Schei's syndrome				
II	Hunter syndrome	Iduronate-L- Sulfatase	Dermatan and heparan sulfate	IDS	Xq28
	Sanfilippo syndrome (A)	Heparan N-Sulfatase		SGSH	17q25.3
III	Sanfilippo syndrome (B)	α -N- Acetil glycosaminidase	Heparan sulfate	NAGLU	17q21
	Sanfilippo syndrome (C)	α - glucosamine acetyltransferase		HGSNAT	8p11.1
	Sanfilippo syndrome (D)	N-acetylglucosamine-6-sulfatase		GNS	12q14
	Morquio syndrome type A	Galactose-6-Sulfatase		GALNS	16q24.3
IV	Morquio syndrome type B	β - Galactosidase	Queratan sulfate	GLB1	3p21.33
		N-acetylglucosamine-4-sulfatase	Chondroitin sulfate, Dermatan Sulfate, N-acetylglucosamine	ARSB	5q14.1
VIII	Sly syndrome	β - Glucuronidase	Dermatan and heparan sulfate	GUSB	7q21.11
IX	Natowicz syndrome	Hyaluronidase	Hyaluronan	HYAL1	3p21.3-p21.2

Studies performed by the Colombian National Institute of Health reveal that in accordance with epidemiologic behavior of the Orphan Diseases in Colombia from 2016 up to the epidemiological week 05 of 2019, 115 cases of MPS IV were observed with a prevalence of 0.3:100,000 habitants (INS, 2019); The Colombian Ministry of Public Health and Social Protection indicated in 2017, that the total of users characterized MPS IV- A elevated to 82 and in terms of frequency analysis, Antioquia was the department with greatest concentration of patients (23 cases), followed by Bogota (15 cases) and Valle del Cauca (8 cases) (Ministry Of Public Health, 2017). Velasco et al. (2017) reported one case of MPSI, 2 of MPS III and 11 cases of MPS III in Boyacá, the latter revealing an incidence of 2.8:100,000 habitants, being this a higher number than expected in terms of the general population.

However, despite being aware of the aforementioned information, defining a disease prevalence rate of Colombia is complicated and has various information gaps. However, the studies performed by Gómez et al. (2012), suggested a Colombian incidence rate of 0.45:100,000 in patients with MPS I, 0.17:100,000 in MPS III, 0.68:100,000 in MPS IVA and 0.23:100,000 in patients with MPS VI, without specifying MPS subtypes referred to within the studies and researches having only been performed in the Altiplano Cundiboyacense. Therefore, it is not a nation-wide illustration, and, in consequence, the disease impact may be higher or lower for the entire country.

It has become known that there are indigenous communities in the Southwestern territory of Colombia which report a high frequency MPS type IV and VI due to inbreeding. It may be contemplated that the MPS prevalence in Colombia could be relatively high in comparison with other countries (Gómez et al., 2012). However, more research must be carried out in terms of incidence and prevalence of this disease in terms of geographical area as well as knowledge of the characteristics of the genetic variants that may appear in patients experiencing this syndrome. Thus, all this combined leads to the development of more direct and personalized care.

Molecular diagnosis is a dynamic field under permanent development which has revolutionized clinic diagnosis. The development of new, faster, and more precise genetic molecular technologies has transformed molecular diagnostic to become a key tool for clinical diagnosis (Farfan, 2015). This new research method for diseases has facilitated the set-up of large and structured genes networks that provide information about the genetic co-expression as well as important molecular functions of the different processes.

Considering that MPS information in Colombia is still limited, this type of research provides enhanced and more accurate information as well as offering solid foundations for the

support of this scientific context as well as for regulations, diagnosis and timely treatment of said disease. Therefore, this research aims to represent genomic variability of the MPS complex in the Southwestern territory of Colombia by means of analysis of the exome variants found in patients with different pathologies and who are not clinically diagnosed with MPS, by increasing the information reported and by opening the doors to future studies.

MATERIAL AND METHODS

Type of Study

A descriptive study was made of genetic variability of IDUA, IDS, SGSH, NAGLU, HGSNAT, GNS, GALNS, GLB1, ARSB, GUSB and HYAL1 genes, suggestive of MPS, by calculating the demographic frequency of each one in order to set up an outline of the Southwestern Territory of Colombia. This study is governed by the MPS studies approved in the institutional ethics committee No 016-016, Universidad del Valle.

Population Survey

244 patients of the Instituto de Genetica Medica (Institute of Medical genetics) GENOMICS (Cali), with no MPS associated clinical signs.

Inclusion Criteria

A different diagnostic from MPS, with suspected complex diseases, obtaining the database from the Instituto de Genetica Medica (Institute of Medical genetics) – GENOMICS (Cali). All patients signed an informed consent for data use.

Exome Sequencing

This technique is based on DNA extraction by means of a Qiagen kit; then, massive sequencing is undertaken by means of an Illumina platform in the Nextera libraries with a coverage of 100X. Genome reference alignment GRCh38/hg19.

Use of Bioinformatics Tools

An informatics program was set-up by means of the Python platform which allows for VCF format editing of the exomes and convert the format into Excel with the purpose of applying different search filters, thus obtaining a final file with the corresponding relevant information. With the VCF editor the organized columns may be visualized in accordance with the chromosome, position, altered amino acids, ClinVar records, among other exome characteristics.

Bioinformatics Analysis

A comprehensive search was performed in order to analyze the varieties observed, utilizing databases such as the Human Gene Mutation Database (HGMD) (<http://www.hgmd.cf.ac.uk/ac/index.php>), Leiden Open Variation Database (LOVD) (<http://www.lovd.nl/3.0/home/>), National Center for Biotechnology Information (NCBI)

(<http://www.ncbi.nlm.nih.gov/>), Ensembl Genome Browser (<http://www.ensembl.org/index.html>). The variants have already been listed as well as their clinical significance.

Nowadays, there are various available bioinformatics tools - *in silico* methods -, which may predict the consequences of the genetic variants within the structure and function of proteins. Such *in silico* methods study various features such as modifications in the physical-chemical properties of the supplanted amino acids, the degree of evolutionary conservation, the environment of the sequence of an affected amino acid or the alteration of structural protein properties. For the purpose of predicting, the following programs were applied: Polymorphism Phenotyping v2 (Polyphen-2), Mutation Taster (<http://www.mutationtaster.org/>), UMD PREDICTOR (<http://umd-predictor.eu/>), PROVEAN (<http://provean.jcvi.org/index.php>), SIFT (<http://sift.jcvi.org/>)

For each one of the variant outcomes revealed, the position, nucleotide modification, amino acid modification, clinical significance, genetic frequency, allelic frequency as well as NCBI frequency were tabulated.

Calculation of genetic and allelic frequencies

The genetic frequency was calculated by dividing the total number of alleles variants observed, by the total number of evaluated patients. Likewise, in terms of allelic frequency of the reported variants, existing data in the ClinVar and ExAC were taken and in term of unreported variants, the following equation was applied.

$$p^2+2pq+q^2 \quad (\text{Eq. 1})$$

in which: p^2 = alleles in homozygous subjects with a polymorphism n $2pq$ = Predicted homozygous frequency
 q^2 = alleles in homozygous subjects with a polymorphism m

Gene Interaction Network

In order to recognize the degree of interaction of each one of the genes found, the STITCH 5 (<http://stitch.embl.de/>) software was taken into use. (“Search Tool for Interacting Chemicals”) and Genemania (<https://genemania.org/>). The search/network was executed, taken into account only the experimental evidence found, processes of biological databases, molecular function, rote or domain of the protein that appeared altered and co-expression with a level of reliability of 0.900, applying this value due to its graphical representation of the phenomena and nonexclusive expressions.

RESULTS

We found 509 associated variants of genes affected by the MPS complex (Table 2); 262 of said variants had not previously been reported, based on the bioinformatics analysis performed regarding the exomes of 244 patients with different pathologies and not clinically

diagnosed with MPS. The genes that reported most variants were IDUA, GLB1 and GALNS, associated with MPS I and MPS IV A/ B accordingly. The genes that reported least variants were IDS and HYAL1, indicating 12 variants in both cases. 88 variants with benign clinical significance were found and 9/244 pathology associated variants, revealing the GALNS gene in 33.3% of all variants of the found pathology variants. Four variants were evidenced affecting the stop codon (0.78%) 24 types of Downstream (4.71%), 184 intronics (36%), 134 missenses (26%), 10 non-coded (1.96%), 106 synonyms (20.8%), 9 located within the splice region (1.76%), 15 upstream (2.94%) and 23 in the UTR region (4.51%).

Table 2. List of genetic variants found in 244 patients in southwestern Colombia with different pathologies and without clinical diagnosis of mucopolysaccharidosis.

Gene	# Variants reported	# New Variants	Total	# Variants with pathogenic effect
<i>IDUA</i>	54	64	118	1
<i>IDS</i>	4	8	12	1
<i>SGSH</i>	26	23	49	1
<i>NAGLU</i>	18	18	36	1
<i>HGSNAT</i>	8	19	27	0
<i>GNS</i>	14	5	19	0
<i>GALNS</i>	48	20	68	3
<i>GLB1</i>	52	39	91	1
<i>ARSB</i>	12	15	27	1
<i>GUSB</i>	21	29	50	0
<i>HYAL1</i>	6	6	12	0

Each of the observed variants were recorded in [Table Supplementary 1](#), including the position, nucleotide modification, amino acid modification, genetic frequency, allelic frequency, ClinVar reference as well as the corresponding predictions of the clinical significance in accordance with UMD PREDICTOR, POLYPHEN, PROVEAN and SIFT.

Some of the variants that reported high genetic frequency also revealed high allelic frequency, among which the p.H33Q in the gene IDUA, p.R737G in the gene NAGL and p.R521C and p.L12= in the GLB1 variants stood out. They all showed a frequency equal to or higher than 90% within the clinical prediction. All the variants were reported as benign. In the entire MPS complex 9/244 variants were found associated with pathogenesis, reporting 33.3% of pathogenic variants in the GALNS gene with respect to the total number of variants found.

The associated network of genes revealed an intensive interaction among the genes responsible for the disease (Figure 1). It is observed that they are all connected with biological processes as the catabolism and the metabolism of the glycosaminoglycan's. Some genes such as GNS, GLB1, ARSB, GALNS, IDUA, IDS and HAYAL1 are more connected with catabolic processes of the chondroitin sulfate. As of the molecular functions, all the genes except for HGSNAT, relate to hydrolase activity and some genes such as GALNS and ARSB are closely linked with the N- acetylgalactosamine 4-sulfatase activity. Without any exception, the genes associated with the MPS complex, perform important functions within the cellular components such as the lumen liposomal, vacuolar structures, lysosome as well as vesicles. Furthermore, each one of them takes part of the KEGG means of glycosaminoglycan's degradation as well as lysosomes means.

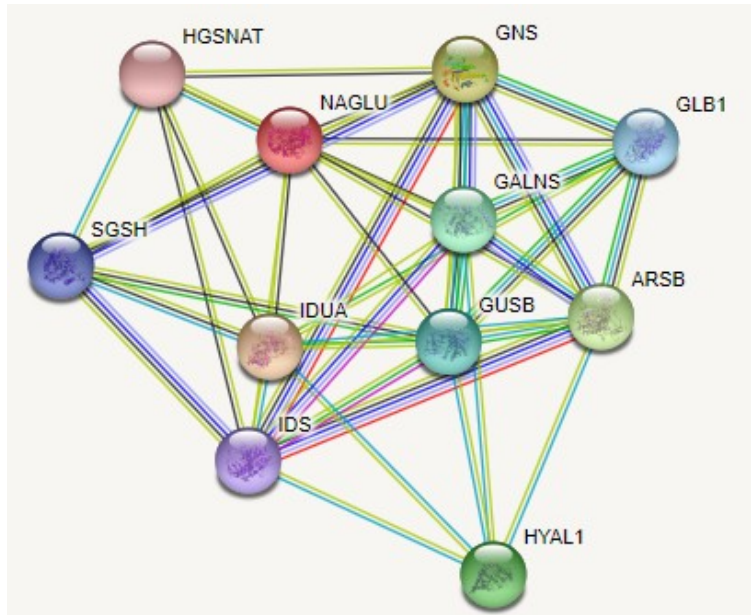


Figure 1. Scheme of the protein-protein interaction network associated with mucopolysaccharidosis, generated with the GenMania program.

Based on the results obtained by means of the network of interaction of genes and associated proteins, interaction modules between genes and various proteins-molecules were observed (Figure 2). Mostly, the interacting genes are responsible for similar functions such as catabolism of glycosaminoglycan, glycolipids metabolic processes, vacuole structures, sphingolipid metabolism, chondroitin sulfate, catabolism and heparan sulfate, cartilage development, skeletal system as well as glomerular filtration.

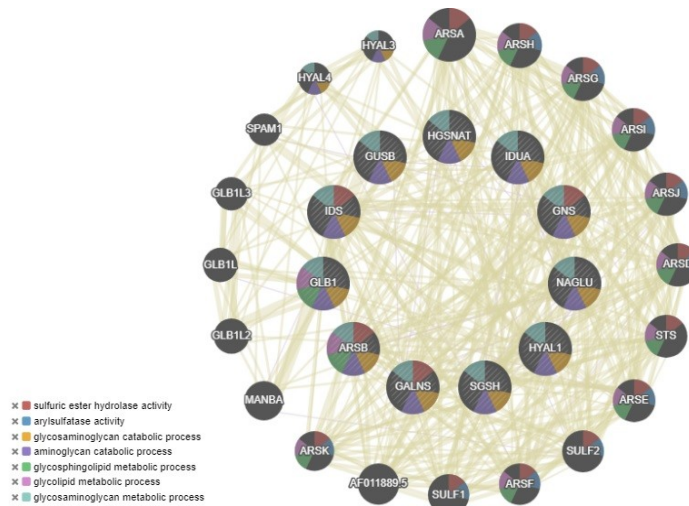


Figure 2. Interaction network between genes associated with the Elaboration, generated with the GenMania program.

DISCUSSION

There were found 509 variants in the 244 patients in the Southwestern Territory of Colombia revealing different pathologies and not clinically diagnosed with MPS. 88 with significant benign clinical reporting, nine pathogenic observations; furthermore, 262 variants were revealed which had not been reported beforehand within the bibliography. This highlights the importance of such research, classification as well as effects within the functionality of the protein and demographic expression.

Although this work only used data from patients with different types of pathologies and not clinically diagnosed with MPS, it was possible to confirm that in the population evaluated, the variants in the genes associated with MPS I, IV and VI present a higher frequency, helping understand the data reported by Gómez et al. (2012) in their studies of the population directly involved with the disease; suggesting in this way inbreeding processes and continuous genetic changes that will need to be evaluated later.

We reported 118 variants within our work in terms of the IDUA gene. Of those, the presence of 11 variants were observed, distributed within the exon 1 and 10 and the introns 5 and 7. Those were also found in the studies performed by Bertola et al. (2011) in patients confirmed with MPS I. The heterogeneity and the prevalence of mutations, highlight the importance of detection related studies in order to help elucidating genotype-phenotype in disorders as MPS I which are characterized by an extensive allelic heterogeneity.

In terms of variants found in the IDUA gene, the most frequently observed were: p.W402X, with a frequency of 50% in the North of Europe, the United Kingdom, North America and Spain (Beesley et al., 2001), p.P533R, which was probably originated in the North of Africa, expanding into Mediterranean countries, representing 13 and 10% of the IDUA mutant alleles in Italy and Spain, respectively (Alif et al., 1999; Venturi et al., 2002) and p.G51D, exclusively found in Italy with a relative frequency of 13% of the IDUA mutant alleles (Venturi et al., 2002).

The variant G409R observed in this study with a frequency of 2% and with a ClinVar conflict interpretation report, with pathogenic predictions, according to studies performed by Ribeiro et al. (2013), of 500 samples of anonymous people born in Portugal without MPS diagnosis, it was found in 4% of the patients. Likewise, Ou et al. (2017) reported a phenotype prediction of MPS type I analysis *in silico* in terms of G409R variant presence, confirming its pathogenicity.

Moreover, the benign variant Thr410= revealed in this study, had a frequency of 16% in the Ribeiro studies (2013), in which the same variant was observed, with an allelic frequency of 17%. This information is like the information reported in other databases, revealing that the consistent detection of alterations as well as reproducibility of the test.

Upon predicting possible deleterious SNPs of the IDUA gene, the number of SNPs chosen within the association of these diseases may reduce those who have more probabilities of altering the gene function. Upon setting up a prediction model of MPS I variants by means of the combination of bioinformatics tools, may benefit the diagnostic as well as the treatment allocation of patients with MPS I.

In terms of variants connected with MPS II, it is known that it is caused due to modification in the iduronato-2-sulfatasa (IDS) gene. Patients may be neuropathic with intellectual disability or non-neuropathic. However, few studies have revealed evidence

regarding the genotype-phenotype IDS relation and the molecular effects implicated (Vollebregt et al., 2017).

More than 500 mutations have been identified in the locus IDS (Xq28), including complex realignments, insertions, deletions, splicing defects and pointless truncations. Therefore, this molecular heterogeneity provides a wide range of phenotypes related diseases and combined with the rarity of individual mutations, make the genotype-phenotype correlations challenging to establish. A significant exception observed is, that patients storing significant structural modification in the IDS gene, caused by suppressions or gen-pseudogene recombination events, usually develop severe forms of the Hunter syndrome. The prediction of the clinical phenotype may be obstructed by other mutation effects in multiple genes, which may contribute to the spectrum of said disease (Zhang et al., 2011). In the variant analysis, the modification p.T409= revealed within this study was also evidenced in the work performed by Holmes (2017) in the comparative protein studies as well as iduronate 2-sulfatase genes (IDS) in vertebrates, granting support to the benign paper in individuals. The variant p.P156Q evidenced in this study with a 1.6% frequency reported by Decamargo et al. (2011) in patients with MPS II with a non-neuropathic phenotype. The variant p.F155L was reported with ClinVar interpretation conflict as well a Polyphen pathogenic probabilities. However, as of the other three predictors, it was defined as a benign variant. And lastly, within this study, eight variants associated with MPS II, not previously reported in any other studies, were evidenced.

In terms of the variants evidenced in the SGSH, NAGLU, HGSNAT, GNS genes linked with the MPS III, different results were revealed; initially in the SGSH gene, specific high frequency sulfamidase mutations have been found in different geographical locations. The missense p.R74C mutation is reported with a frequency of 56% within the Polish population. A p.R245H amino acid modification was evidenced in 31% of the allele's mutants in Australia, 35% in Germany and 58% in the Netherlands (Bunge et al 1997). Said residue is not stored among the human sulfatases and have little consequence in regard to this specific activity. The missense p.S66W mutation has been observed in 29% of the alleles in a cohort of Italian patients (Di Natale et al., 1998).

Our study evidenced 12 benign variants previously reported in patients with MPS III – Sanfilippo's syndrome type A. We reported the p.P350L variant in which we observed no ClinVar data. However, the predictors reported them as pathogenic. Furthermore, 23 variants were observed, not previously reported, within the ExAC, Human Genome Project 1000, Human Gene Mutation Data Base and the Single Nucleotide Polymorphism Database.

In 2014, The Mutation Database published 153 mutations in the NAGLU gene, associated with MPS IIIB, of which 90 mutations made no sense. The ExAC dataset cataloged 189 NAGLU missense mutations based on exome sequence data of approximately 60,000 persons of which 24 were aware of the association with said disease. Within our work, 11 modifications were evidenced which were previously reported in the list of genetic variants researched, for the purpose of genetic mucopolysaccharidosis III-B studies (ClinVar, 2019). It was evidenced that the p.S141T variant with clinical significance, probably benign, in 4 predictors, however, in the Polyphen predictor it was observed as probably pathogenic. Therefore, further research shall be required in order to evaluate the function and action of said registered modification. For the purpose of said task, we reported 18 variants in the NAGLU not previously registered within the ExAC,

Human Genome Project 1000, Human Gene Mutation Data Base and the Single Nucleotide Polymorphism Database.

Finally, in terms the MPS III group, 27 variants were evidenced with HGSNAT gene association, a low demographic frequency was reported with the exception of the p.Y583= modification found in 95% of the total population. 22% of the variants were previously reported as benign and the p.A15T variant presented conflicts of interpretation which were solved by means of the Provean and SIFT programs, defining the benignity of such variant. This confirms the importance of the use of significance predictors within such research.

Ultimately, in terms of the variables evidenced in the GNS gene connected with the MPSIIID, 15 previously registered variables were observed, mostly benign and others of which no concrete data were found due to being intronic variants in nature with difficult analysis. Of the 5 unreported variants observed, at least 2 belonged to introns and other 3 conformed missense type modification.

Within the GALNS gene analysis associated with MPS IV-A, there was evidence of the presence of 68 variants, 48 of which were previously reported, cataloged in other studies as benign modifications. In this study, the variant c.1431G>A was the most frequent indicating 60% of the alleles, followed in terms of frequency by the p.H36= benign mutations and the intronic variant 634-19G>A with 41% and 34% accordingly. Furthermore, 3 variants (p.V480F, p.R386C, p.G301C) outlined in all predictors as pathogenic. Especially, the variant V480F was reported for the first time by Cacioti et al. (2014), who reported said variant in patients with this disease. It has, however, not been found in healthy patients. Moreover, the variation was not reported within the Genome Project 1000.

During a study performed previously by our research team, the discovery of the previously reported mutation in the exon 11 c.1156C> T or p.R386C, in patients with MPS IV-A there was established a confirmation of 5% of these cases (Moreno et al. (2018)). In another similar study performed by Morrone et al. (2014), it was observed that the most frequent alleles were the modification in direction of c.1156C> T (p.R386C), reported 55 times (55/1,091; 5%). The second most frequent reported alleles was the modification of direction of c.337A> T (p.I113F), reported 52 times (52 /1,091; 5%); and the third most reported in terms of modification of direction was c.901G> T (p.G301C), reported 45 times (45 of 1,091; 4%). The same 3 alleles were reported as high frequency by Tomatsu et al. (2005) in which 48 single GALNS gene determined mutations were analyzed.

In comparison with other studies performed until now, this variant was evidenced in 0.08% of the samples, while in accordance with the Tomatsu et al. (2005) report, it was evidenced with 8,9% in patients with MPS IV and according to the Morrone et al. (2014) studies it was observed in 5% of the alleles. The variant c.337A> T (p.I113F) reported by Tomatsu. (2005) Morrone. (2014) of 5% and 5.7% of the alleles, accordingly, were not found in this research. Finally, the pathogenic variant c.901G> T (p.G301C) reported by Tomatsu et al. (2005) of 5.7% and by Morrone et al. (2014) of 4% of the reported alleles, it was evidenced in 0,08% of the samples of this study. The variants p.T312S, p.M391V, p.A291T, p.S287L, p.M318R, p.R253T, reported in Europe as frequent, were not evidenced in this study. This may suggest that there is little or no distribution of said variants in Latin-America (Montaño et al., 2008).

In terms of analysis of the GLB1 gene associated with MPS IV-B, it was evidenced that 52/91 of the variants previously reported in regards to the modification of p.R521C, a frequency of 95% was evidenced in the samples, followed by c.-37C>G and p.L12= reporting 81% and 79% of the samples accordingly. In 0.08% of the samples, the variant A457Q pathogenic connected with all the predictors. This variant was previously reported in a study of three Japanese patients with the Morquio B (Oshima et al., 1991) syndrome. Of the unreported database variants, 13 intronic variants were evidenced, thus highlighting this way, the importance of due interpretation and study.

Other alleles reported with gangliosidosis GM1 II / III as well as phenotype MPS IV-B. For example, the alleles R201H in homozygosity were reported in connection with phenotype MPS IV-B. However, the composed homozygosity with other GLB alleles, outlined in connection with the GM1 youth/adult gangliosidosis alleles, leading to a mixed phenotype demonstrating both characteristics of the Morquio dysostosis type as well as, simultaneous, pneumopathy characteristics as observed within the GM1 gangliosidosis. Patients with combined phenotypes have been outlined as atypical MPS IV-B or combined MPS IV-B / GM1 II / III. Although the W273 L variant caused a decrease in hydrolytic activity due to less affinity with the quatrain sulfate, most of the variants observed in patients with MBD / GM1 II / III are mutations making no sense causing instability and premature degradation in the enzymatic protein with an intact catalytic area (Bleier et al., 2018).

On the other hand, of the 27 unreported observed variants in the ARSB,15 gene and of which no clinical implications have been known, 12 have been previously reported connected with MPS VI, 11 associated with a benign phenotype, excluding p.V358M, p. P397= with 41% and 26% accordingly and a pathogenic variant (p. V376M). This latter variant was evidenced recently by Essawi et al. (2018) in 15 Egyptian families by means of sequencing of 8 exons codified with the ARSB gene. This research revealed pathogenic mutations in 13 patients. All exons observed were found in exons 1 and 2. Four new pathogenicity mutations were detected by means of functional analysis *in silico* by means of the use of different programs. The new mutations detected were c.189insA, c.257delA, p. S94L (c.281C> T), and p. L51P (c.152T> C). It was observed that p. R160X (c.478C> T) was the most frequent. The four mutations reported previously which were observed in the patients studied were p. L82R (c.245T> G), AR160X (c.478C> T), p. R160Q (c.479G> A), y p. S96R (c.288C> G). A functional analysis performed by Garrido et al. (2008) revealed that p. Val376Met showed equivalent enzymatic activity for the valine as well as for the methionine alleles. In our research, it was possible to evidence the V76M variant with a 29% of the samples. However, in contrast to the Essawi et al. (2018) study, the bioinformatics predictors used pathogenic indications, which shall be necessary for further studies in order to comprehend such variations.

The benign variant p.V358M, apart from being observed in the Essawi et al. (2018) studies, it was also reported in the works carried out by Yang et al. (2002), in which the clinic investigation and the analysis of mutation of 2 Taiwanese patients with severe phenotypes and MPS VI intermediacy were published. Three mutations with no sense were identified in said works as well as a polymorphism. It was evidenced that Case 1 had a new position from C to G within the nucleotide 1197 (F399L) as well as a heteroallele mutation Q239R. In Case 2, a heterozygous mutation was identifier to C192R as well as a benign

polymorphism V358M. Among those 3 mutations, the substitutions Q239R and F399L were reported only in the Taiwanese population.

The variant S84N was reported in the studies performed by Tomanin et al. (2018) and was found in our studies as well. Said benign variant had erroneously been informed as causing disease variant. It was evidenced some years ago that it was not a causing disease variant (Zanetti et al., 2014), but it had been informed as such in 9 alleles, where the reporting was incorrectly based on data collected from previously published literature.

As of the variants found for the GUSB gene associated with MPS VIII, we reported 50 variants in our work, 42% of them previously reported in the literature as benign. 29 non reported variants were found, of those 62% were intronic, Downstream or not encoding genes, only 2 missense types were found. The intronic modification T>TGA was reported in this study in 100% of the samples analyzed. Worldwide, said modification was found in 96% of the population (Lek et al., 2016)

In terms of the benign variants reported, the modification p.L649P obtained a demographic frequency of 72%, which is like the amplified data reported worldwide as Exac Browser, reporting a 0.5796 worldwide modification frequency. The modification 1065+27C>G reported in this study was found in 66% of the population. Furthermore, the variant p.Y580= was found in 20% of the samples in the Southwestern Territory of Colombia. The worldwide level shows a frequency of 0.1289 (Lek et al., 2016).

The variant L176F was found in several ethnic populations with a frequency of 20% in accordance with studies performed in 2011 by Tomatsu and his team, being this one of the most prevalent, while in the mutation p.PF415L and the double mutation p.PF415L / F408S, plus the mutation p. A619V which have only been detected in the Mexican and Japanese populations accordingly. The transition c.526C> TL176F, originally found in two Mennonite brothers, 21 alleles were identified in 11 patients from The United States of America, Brazil, The United Kingdom, Chile, France, Mexico, Poland, Spain, and Turkey. Our study did not reveal any worldwide proposed MPS VIII variants.

Lastly, 12 variants associated with the HYAL1 (MPS IX) gene was evidenced. Of those, 6 were already reported in the literature and 6 are new variants. Specifically, the modification in G256R is reported in the updated list of variants in the HYAL1 gene and was reported as a benign variant by ClinVar (Clinvar, 2019). The modification T>C Chr3:50334359 was reported in 10% of the samples. However, worldwide, it was reported in 0.00001867 and up to today, its clinical significance is unknown. The remainder of the variants reported had a low demographic frequency (between 0.0008 and 0.10) (ClinVar 2019).

The MPS type IX is the rarest form of MPS and has only been reported in four patients with HYAL1 malign associated variant pathologies to date. The first patient was reported by Kiykim et al. (1996). All patients with MPS type IX had joint and skeleton issues, therefore, MPS type IX may easily be erroneously diagnosed as juvenile idiopathic arthritis. There is no information in the literature regarding prevalence research of variants associated with phenotype and genotype MPS IX.

As of the interaction network of genes associated with the MPS complex, a close relation among each one of them was observed, maintaining a high genetic interaction which could possibly lead to specific phenotypes connected with the diagnose of MPS. The natural function of each one of the genes is evidenced and is necessary for the development of biological processes as catabolism and metabolism of the glycosaminoglycans. In

addition to the molecular functions as hydrolase activity and N- acetylgalactosamine 4-sulfatase activity, there are important functions of cellular components as the lumen liposomal, vacuolar components, lysosome and vesicles. The damage, the lack of or the mutation of said genes produce the known phenotype.

Likewise, the normal functioning of the genes allow for interaction and activation of related genes, so that within a network of related genes, the following may be observed, for example: HYAL and GLB1, found with the interaction of HYAL4, HAYAL 3, associated with HYAL1 and GLB1L3, GLB1L and GLB1L2, associated with GLB1. Activated genes such as MANBA, ARSK, with related genes SULF and ARSG were evidenced. Specifically, the influence in the GUSB, GLB1, IDUA, NAGLU and HYAL genes may lead to damage of MANBA, which, upon presenting some type of damage or inactivation, lead to an oligosaccharidosis known as B-Mannosidosis, which is a disease reflected in the lysosomes. Likewise, genes associated with sulfate (SULF) are in risk of being altered by other important genes as HGSNAT, GNS, SGSH, GALNS, ARSB and IDS. Given the above, a full description of the clinical characteristics is required as well as further evaluations in order to reach a conclusion for the genotype-phenotype correlations associated with genetic interaction.

The MPS complex is a group of rare or orphan disease, of low prevalence, characterized by enzyme deficiencies, which participate in the GAG metabolism at liposomal level. It is characterized by the accumulation of intracellular GAG, producing alterations of multiple organs and systems (Lampe et al. 2019). The identification non MPS diagnosed demographic variants may lead studies to evaluate possible founders' effects within our population.

The possible expansion of the genes is important for early detection with the purpose of minimizing the attributed pathology consequences, additionally to the importance of improving the adequate handling and counseling for such patients and their families.

Knowing the molecular genetics of each of the genes associated with the MPS complex, enable a genotype/phenotype correlation approach. Despite the elevated molecular heterogeneity of the mutations, in most studies performed in Colombia, some genotype- phenotype associations have been observed in which most of the particular mutations are associated with the attenuated phenotypes, while the mutation without sense, big deletions as well as alterations of the *splicing* area is found associated with severe phenotypes. In this study, the variant percentage observed in intronic areas, of *splicing* and UTR regions reached approximately 43% of the samples.

The bioinformatics analysis is basis for the pharmacogenomics and proteomics, leading to a complete approach in terms of the behavior of this disease in terms of the Colombian legislation as a high cost orphan disease and of special public health interest.

The current bioinformatics tools which allow for predictions of variants and probabilities of clinical significance, contribute to the understanding of the MPS complex, since it is a meeting point for the understanding of the predisposition within the population in order to develop a given variant. Most of the variants registered in our

study were found to be associated with genes for MPS I and MPS IV, being the latter one of the most frequent MPS in our country.

ACKNOWLEDGMENTS

The Instituto de Genética Médica-Genomics and the Genomic Technological Laboratory of the Faculty of Health of the Universidad del Valle supported this research.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Albano, L, Sugayama S, Bertola D, Andrade C, et al. (2000). Clinical and laboratorial study of 19 cases of mucopolysaccharidoses. *Revista do Hospital das Clínicas*. 55: 213.
- Alif N, Hess K, Straczek J, Sebbar S, et al. (1999). Mucopolysaccharidosis type I: characterization of a common mutation that causes Hurler Syndrome in Moroccan subjects. *Ann. Hum. Genet.* 63: 9-16.
- Barrera LA (2009). Estudios Bioquímicos de los errores innatos del metabolismo en Colombia, durante dos décadas. *Rev Acad. Colomb. Cienc. Exactas Fis. Natu.* 33: 377-97.
- Beesley CE, Meaney CA, Greenland G, Adams V, et al. (2001). Mutational analysis of 85 mucopolysaccharidosis type I mutations and in vitro expressions of missense mutations. *Hum. Genet.* 109: 503-11.
- Bernal JE and Briceño I (2006). Genetic and other diseases in the pottery of Tumaco-La Tolita culture in Colombia-Ecuador. *Clin. Genet.* 70: 188-91.
- Bertola F, Filocamo M, Casati G, Mort M, et al. (2011). IDUA Mutational profiling of a cohort of 102 European patients with mucopolysaccharidosis type I: identification and characterization of 35 novel alfa-L-iduronidase (IDUA) alleles. *Hum. Mutat.* 32: E2189-210.
- Bleier M, Yuskiv N, Priest T, Moisa M, et al. (2018). Morquio B patient/caregiver survey: First insight into the natural course of a rare GLB1 related condition. *Mol. Genet. Metab. Rep.* 16: 57-63.
- Bunge S, Ince H, Steglich C, Kleiker W, et al. (1997). Identification of 16 sulfamidase gene mutations including the common R74C in patients with mucopolysaccharidosis type IIIA (Sanfilippo A). *Hum. Mutat.* 10: 479-485.
- Caciotti A, Tonin R, Rigoldi M, Ferri L, et al. (2014). Optimizing the Molecular Diagnosis of GALNS: Novel Methods to Define and Characterize Morquio A Syndrome- Associated Mutations. *Hum. Mutat.* 36: 357-68.
- Chuang CK, Lin SP and Chung SF (2001). Diagnostic screening for mucopolysaccharidoses by the dimethylmethylene blue method and two dimensional electrophoresis. *Zhonghua Yi Xue Za Zhi (Taipei)*. 64: 15-22.
- Clinvar (2019). List of variants in gene NAGLU reported as pathogenic for Mucopolysaccharidosis, MPS-III-B. <https://clinvarminer.genetics.utah.edu/variants-by-gene/NAGLU/condition/Mucopolysaccharidosis%2C%20MPS-III-B/pathogenic>
- DeCamargo L, Maluf S, Leistner-Segal S, Zimmer da Silva C, et al. (2011). Are MPS II heterozygotes actually asymptomatic? A study based on clinical and biochemical data, X- inactivation analysis and imaging evaluations. *Am. J. Med. Genet. Part. A.* 155: 50-57.
- Di Natale P, Balzano N, Esposito S and Villani GR (1998). Identification of molecular defects in Italian Sanfilippo A patients including 13 novel mutations. *Hum. Mutat.* 11: 313-320.
- Essawi M, Elbagoury N, Sayed O, Aglan M, et al. (2018). Mutation analysis of the arylsulfatase B gene among Egyptian patients with Maroteaux-Lamy disorder. *Middle East J. Med. Genet.* 2: 96-103.
- Farfán B (2015). Biología Molecular Aplicada al Diagnóstico Clínico. *Rev. Méd. Clin. Condes.* 26: 788-793.
- Garrido E, Cormand B, Hopwood JJ, Chabas A, et al. (2008). Maroteaux-Lamy syndrome: functional characterization of pathogenic mutations and polymorphisms in the arylsulfatase B gene. *Mol. Genet. Metab.* 94: 305-312.
- Giraldo LJM, Arturo-Terranova D and Satizabal-Soto JM (2020). Otorhinolaryngological Findings in Patients from Southwestern Colombia with Clinical, Enzymatic and Molecular Diagnosis of Mucopolysaccharidosis II, IV-A and VI. *J. inborn errors metab. Screen.* 8: e20190006.
- Gómez AM, García-Robles R and Suárez-Obando F (2012). Estimación de las frecuencias de las mucopolisacaridosis y análisis de agrupamiento espacial en los departamentos de Cundinamarca y Boyacá. *Biomédica.* 32: 602-9.
- Holmes RS (2017). Comparative studies of vertebrate iduronate 2-sulfatase (IDS) gene and proteins: evolution of a mammalian X-linked gene. *3 Biotech.* 7: 22.

- Instituto Nacional de Salud (INS) (2016). Comportamiento epidemiológico de las enfermedades huérfanas. Colombia, hasta semana epidemiológica 05 de 2019. <https://www.ins.gov.co/buscadoreventos/BoletinEpidemiologico/2019%20Bolet%C3%ADn%20epidemiol%C3%B3gico%20semana%205.pdf>. Accessed 19 July 2019.
- Kiykim E, Barut K, Cansever MS, Zeybek CA, et al. (2016). Screening Mucopolysaccharidosis Type IX in Patients with Juvenile Idiopathic Arthritis. *JIMD Rep.* 25: 21-24.
- Lampe C, McNelly B, Gevorkian A, Hendriksz C, et al. (2019). Transition of patients with mucopolysaccharidosis from paediatric to adult care. *Mol. Genet. Metab. Rep.* 21: 100508.
- Lek M, Karczewski K, Minikel E, Samocha C, et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 536: 285-291.
- Lin HY, Chan WC, Chen LJ, Lee YC, et al. (2019). Ophthalmologic manifestations in Taiwanese patients with mucopolysaccharidoses. *Mol. Genet. Genomic Med.* 7: e00617.
- Lin HY, Lin SP, Chuang CK, Niu DM, et al. (2009). Incidence of the mucopolysaccharidoses in Taiwan, 1984–2004. *Am. J. Med. Genet. Part. A.* 149A: 960-964.
- Ministerio de salud y protección social (MINSALUD) (2017). Análisis Del Principio Activo Elosulfasa Alfa Vimizim 2016-2017 Medicamentopara El Síndrome De Morquio Tipo. https://consultorsalud.com/wpcontent/uploads/2018/04/reporte_sindrome_de_morquio_vimizim.pdf. Accessed 12 July 2019
- Ministerio de salud y protección social (MINSALUD) (2018). Resolución número 5265 de 2018. <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/resolucion-5265-de-2018.pdf>. Accessed 12 July 2019.
- Morrone A, Caciotti A, Atwood R, Davidson K, et al. (2014). Morquio a syndrome- associated mutations: A review of alterations in the GALNS gene and a new locus-specific database. *Hum. Mutat.* 35: 1271-9.
- Montaño AM, Tomatsu S, Brusius A, Smith M, et al. (2008). Growth charts for patients affected with Morquio A disease. *Am. J. Med. Genet. A.* 146A: 1286-1295.
- Moreno Giraldo LJ, Escudero A, Sanchez A and Satizabal JM (2018). Clinical and molecular characteristics of colombian patients with mucopolysaccharidosis IVA, and description of a new galns gene mutation. *Mol. Genet. Metab. Rep.* 16: 53-56.
- National Center for Biotechnology Information (2019). ClinVar; [VCV000346086.1], <https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000346086.1> (accessed Sept. 4, 2019).
- Neufeld E and Muenzer J (2001). The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. The metabolic and molecular basis of inherited disease. 7 th ed. New York: *McGraw-Hill.* 3421-3452.
- Oshima A, Yoshida K, Shimmoto M, Fukuhra Y, et al. (1991). Human B-Galactosidase gen mutation in Morquio B disease. *Am. J. Hum. Genet.* 49: 1091-3.
- Ou, L, Przybilla M and Whitley C (2017). Phenotype prediction for mucopolysaccharidosis type I by in silico analysis. *Orphanet J. Rare Dis.* 12: 125.
- Pachajoa H and Rodriguez CA (2014). Mucopolysaccharidosis type VI (Maroteaux-Lamy syndrome) in the pre-Columbian culture of Colombia. *Colomb. Med.* 45: 85-88.
- Ribeiro D, Cardoso A, Duarte A, Vieira L, et al. (2013). Efficient IDUA Gene Mutation Detection with Combined Use of dHPLC and Dried Blood Samples. *ISRN Mol. Biol.* 451298.
- Stapleton M, Kubaski F, Mason R, Yabe H, et al. (2017). Presentation and treatments for Mucopolysaccharidosis type II (MPS II; hunter syndrome) *Expert Opin. Orphan Drugs.* 295-307.
- Tomanin R, Karageorgos L, Zanetti A, Alsayaed M, et al. (2018). Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. *Human Mutat* 39:1788-1802
- Tomatsu S, Montaño A, Nishioka T, Gutierrez M et al. (2005) Mutation and polymorphism spectrum of the GALNS gene in mucopolysaccharidosis IVA (Morquio A). *Hum Mutat.* 26:500-512.
- Tomatsu S, Montaño A, Dung V, Grubb et al. (2009). Mutations and polymorphisms in GUSB gene in mucopolysaccharidosis VII (Sly Syndrome). *Hum. Mutat.* 30: 511-519.
- Velasco H, Martin A, Galvis J, Buelvas L, et al. (2016). Genética clínica comunitaria: Exploración de patología genética en Boyacá, Colombia. *Rev. salud pública.* 19: 32-38.
- Venturi N, Rovelli A, Parini R, Menni F, et al. (2002). Molecular analysis of 30 mucopolysaccharidosis type I patients: evaluation of the mutational spectrum in Italian population and identification of 13 novel mutations. *Hum. Mutat.* 20: 231.
- Vollebregt A, Hoogveen-Westerveld M, Kroos MA, Oussoren E, et al. (2017). Genotype-phenotype relationship in mucopolysaccharidosis II: predictive power of IDS variants for the neuronopathic phenotype. *Dev. Med. Child Neurol.* 59: 1063-1070.
- Yang F, Wu J, Shuan-Pei L and Fuu-Jen T (2002). Mucopolysaccharidosis type VI: Report of two Taiwanese patients and correlation of one novel mutation. *J. Formos Med. Assoc.* 100: 820-3.
- Zanetti A, Önenli-Mungan N, Elcioglu N, Özbek MN, et al. (2014). Molecular analysis of Turkish Maroteaux–Lamy patients and identification of one novel mutation in the arylsulfatase B(ARSB) gene. *JIMD Reports.* 14: 1-19.

- Zanetti A, D'Avanzo F, Rigon L, Rampazzo A, et al. (2019). Molecular diagnosis of patients affected by mucopolysaccharidosis: a multicenter study. *Eur. J. Pediatr.* 178: 739- 753.
- Zhang H, Li J, Zhang X, Wang Y, Qiu W, et al (2011). Analysis of the IDS Gene in 38 Patients with Hunter Syndrome: The c.879G>A (p.Gln293Gln) Synonymous Variation in a Female Create Exonic Splicing. *PLoS One.* 6(8): e2295.