

Frequency of *GBA* gene variants in complex disease patients in Southwestern Colombia

D. Arturo-Terranova^{1,4,5}, L.J.M. Giraldo^{1,2,3,4,5} and J.M. Satizábal^{1,2,4,5}

¹Universidad del Valle, Santiago de Cali, Colombia

²Universidad Santiago de Cali, Santiago de Cali, Colombia

³Universidad Libre, Santiago de Cali, Colombia

⁴Posgraduate Program in Biomedical Sciences – Universidad del Valle, Santiago de Cali, Colombia

⁵Congenital Diseases of Metabolism Research Group, Universidad del Valle, Santiago de Cali, Colombia

Corresponding author: D. Arturo-Terranova

E-mail: daniela.arturo@correounivalle.edu.co

Genet. Mol. Res. 20 (2): gmr18818

Received February 22, 2021

Accepted May 28, 2021

Published June 30, 2021

DOI <http://dx.doi.org/10.4238/gmr18818>

ABSTRACT. Gaucher Disease (GD) is an autosomal recessive genetic disorder, caused by a deficiency of the enzyme B-glucocerebrosidase (*GBA*). In Colombia, despite considerable research on GD, the frequency of the *GBA* gene variants in the population is unknown, making it difficult to determine the risk of occurrence based on carriers. To identify the variants of the *GBA* gene, a transversal, descriptive, non-experimental study was carried out with the results obtained from the sequencing of the complete exome of 320 patients with complex disease, without clinical suspicion of GD. Bioinformatics software was used to analyze the clinical significance of the different variants. The population frequency of each variant was calculated, and a network of interaction of the *GBA* gene was developed. As a result, 41 variants associated with the *GBA* gene were found; 21/41 of the variants reported have a benign significance, 5/41 of the variants reported were classified as pathogenic or probably pathogenic and 7/41 of the variants reported presented uncertain significance. The gene interaction network showed close associations between *GBA* and genes *PSAP*, *SCARB2*, *LAMP2*, all of them focused on functions of

vacuolar locations, lysosomal and vacuolar lumen instructions, vacuolar and lysosomal membranes. We conclude that the impact on the phenotype highly depends on the pathogenicity of the variants. In our sample, a high frequency of benign variants was found; however, pathogenic variants were detected, which should be the object of study in precision medicine associated with GD.

Key words: Bioinformatics tools; Exome sequencing; Gaucher disease; *GBA* gene; Lysosomal Storage Diseases; Variants

INTRODUCTION

Gaucher Disease (GD) (MIM # 230800, 230900, 231000; CIE-10 E75.2; ORPHA: 355) is the most common lysosomal deposition disease, with an autosomal recessive type inheritance, caused by a mutation or variants in the *GBA* gene located in the q21 region of chromosome 1. It contains 10 exons and 11 introns covering 7.5kb and encodes a protein of 497 amino acids with a molecular weight of 55.5 KDa. It is located within a complex locus containing seven genes and two pseudogenes, probably coming from a duplication event of this chromosomal region. In fact, a highly homologous pseudogene of 5.7 kb (*GBAP1*) is found approximately 16 kb from the functional *GBA* gene (Horowitz et al., 1993). The exonic region of *GBAP1* shares 96% of sequence homology with the coding region of the *GBA* gene, while the sequence homology reaches 98% in the region between the intron 8 and the 30 untranslated regions. The high homology and physical proximity between *GBAP1* and *GBA* allow both non-reciprocal homologous recombination and reciprocal events resulting in complex pseudogenic rearrangements (Zampieri et al., 2017).

This disease is characterized by a partial or total deficiency of the enzyme beta-glucocerebrosidase (β -acid glucosidase), which is necessary to degrade the type of lipid known as glucosylceramide, which, because of the enzymatic deficit, is accumulated in the body, predominantly in the liver, spleen, and bone marrow (Nguyen et al., 2019; Sun, 2018), leading to manifestations of multisystemic disease including splenomegaly, hepatomegaly, thrombocytopenia, anemia, and bone disease, therefore a timely and definitive diagnosis minimizes the impact of misdiagnosis and unnecessary and invasive diagnostic procedures and can aid the optimal management of symptomatic patients (Di Rocco et al., 2014; Simon et al., 2008).

Currently, GD is divided into three different subtypes: non-neuronopathic (GD1), acute neuronopathic (GD2), and chronic neuronopathic (GD3) (Table 1). GD1 is the most prevalent lysosomal storage disorder as well as the most common GD phenotype. Patients are usually identified by splenomegaly in childhood or early adulthood. Hepatomegaly is usually not as severe as splenomegaly. Thrombocytopenia and anemia are easily observed, and skeletal findings such as osteopenia, osteonecrosis, and bone pain are also seen. Additionally, parkinsonism is an established feature of GD and a higher frequency of *GBA* mutations has been reported in various ethnic series with sporadic Parkinson's disease (PD) (Lee et al., 2012); for this reason, an important step in the comprehension of the etiology of neuropsychiatric symptoms associated with GD-PD and the degree to which they overlap with each other will be the identification of risk factor genes associated with poor diagnosis related to these symptoms.

Table 1. Subtypes of Gaucher disease. Adapted from Van Rossum and Holsopple (2016).

Subtype	Type 1		Type 2			Type 3	
	Symptomatic	Asymptomatic	Infant	Neonatal	3a	3b	3c
Common genotype	N370S / other slight variants	N370S/N370S or 2 other mild variants	1 null variant and 1 severe variant	2 null or recombinant variants	D409H/D409H	L444p/L444P	None
Characteristics	Hepatosplenomegaly	None	Gaze paralysis	Congenital ichthyosis	Gaze paralysis	Gaze paralysis	Gaze paralysis
	Hypersplenism		Squint		Calcification in heart valves	Hepatosplenomegaly	myoclonic seizures
	Bleeding		Severe trismus			stunted growth	
	Bone pain						
	None	Parkinsonism ?	Severe	Lethal	gaze paralysis	Gaze paralysis	Gaze paralysis
CNS					Brachycephaly	Gradual cognitive decline	Progressive neurological deterioration .
Bones	Medium to severe	No	No	No	Minimal	Moderate to severe	Medium
Lungs	No (rarely)	No	Severe	Severe	Minimal	Moderate to severe	Medium to moderate
Life expectancy	Normal	Normal	Death before age 2	Neonatal death	Death near adulthood	Death in mid-adulthood	Death during childhood
Treatment	ERT*	No	Palliative	No	Valve replacement	ERT*	ERT*

* CNS Central Nervous System; Enzyme replacement therapy (ERT).

GD1 is most frequent in the United States, Europe, and Israel and occurs most frequently in the Ashkenazi Jewish population. It has been estimated that the standardized incidence of patients born with GD in the general population is between 0.4 and 5.8 per 100,000, with a prevalence of 0.7 to 1.8 per 100,000 (Nalysnyk et al., 2017). The frequency of carriers in the Ashkenazi Jewish population is approximately 1 in 12 (Acanda, 2012), and the frequency of genotypes associated with the disease is estimated at 1 in 850 (Hruska et al., 2008). The incidence of GD1 is much lower in non-Jewish populations, and it occurs at a frequency of approximately 1 in 40,000 to 1 in 86,000 live borns. GD2 (also called acute neuronopathic GD) has an estimated incidence of 1 in 150,000 (Sheth et al., 2019)

More than 460 variants of the *GBA* gene have been reported. However, four common variants - p.Asn370Ser (p.Asn409Ser, according to the new nomenclature), IVS2, 84GG, p.Leu444Pro (p.Leu483Pro, according to the new nomenclature) are responsible for approximately 96.5% of the GD in the Ashkenazi Jewish population in the Western Hemisphere and approximately 50-60% in non-Jewish populations (Nguyen et al., 2019).

Concerning Colombia, the first publication on GD was made by Cadavid in 1925 (Cadavid, 1925); Echavarría et al. (1966) reported two cases, one of them of GD1, and Rodríguez et al. (1970) presented three cases of the severe form. Thus, by 1972, a total of 12 reported cases were known in the literature for the country (Rodríguez et al., 1972). In 2006, 76 cases of GD had been reported, of which 55 were registered on the ClinicalTrials.gov platform; of these, 29 were adults (53%) and 26 children (47%). Fifty patients were of type I and 5 of type II and III (International Collaborative Gaucher Group, 2018). Lozano-Bernal (2006) described the clinical history of casuistic in the department of

Tolima about this pathology, he illustrated four cases diagnosed in the previous 10 years, the first of which was diagnosed on November 28th, 1996 and the last on February 9th, 2005.

Uribe et al. (2013) reported that, in Colombia, 9,112 people with suspected lysosomal storage disease (LSD) had been tested for enzyme activity, of which 117 were diagnosed with GD. According to the Colombian Association of Patients with Lysosomal Storage Diseases -ACOPEL (in its Spanish acronym), by 2017, the estimated prevalence was 1 for every 266,441 inhabitants (ADRES, 2017). However, even though GD is within the list of 2,198 Orphan Diseases (OD) in the country, according to Resolution 5265 of 2018, there are not still exact figures on the incidence and prevalence of this disease, since, due to lack of knowledge, clinical heterogeneity and/or difficult diagnosis, it is a sub-diagnosed disease in Colombia.

One of the difficulties in knowing the true incidence, prevalence and population burden of GD is its diagnosis, since, despite the availability of defined tests, such as the analysis of *GBA* activity in leukocytes and/or fibroblasts, and molecular analysis of the *GBA* gene, patients continue to be misdiagnosed and experience significant diagnostic delays (Mistry et al., 2007; Mistry et al., 2011). This is probably a consequence of the relative rarity of the disease, lack of knowledge about the pathology, and clinical heterogeneity (Thomas et al., 2013).

We looked for variants of the *GBA* gene in patients with various pathologies and without clinical suspicion of Gaucher disease.

MATERIAL AND METHODS

Type of Study

This was a transversal, descriptive, non-experimental study in which variants of the *GBA* gene were genomically characterized, calculating the population frequency in a sample of patients with various pathologies and not clinically diagnosed with Gaucher disease in Southwestern Colombia. This project had the approval of the Scientific Committee of Ethics and Bioethics of the University Santiago de Cali "CEB-USC" Faculty of Health in the session April 17th, 2020 according to act No. 07.

Population Survey

Results from complete exome sequencing of 320 patients with complex disease were used. Complex diseases are those that are difficult to diagnose, caused by the interaction of multiple genes and environmental factors. In disease research, pathogenetic variants in genes associated with monogenic diseases are likely to be found in patients who are unaffected, but are carriers, so this study focuses on finding *GBA*-associated variants.

Data on these patients is in the database of the Institute of Medical Genetics - GENOMICS (Cali) and the results of their exomes are maintained at the institute. All patients signed an informed consent for the use of their data.

Inclusion Criteria

Patient with complex disease, without clinical suspicion of Gaucher disease, with results of complete exome sequencing obtained from the database of the Institute of Medical Genetics - GENOMICS (Cali)

Exclusion criteria.

Patients previously confirmed for Gaucher disease or who had not signed the informed consent-assent.

Exome Sequencing

DNA extraction was made through Qiagen case; then, massive sequencing of NexteraTM libraries was performed with the Illumina platform with a coverage of 100X. Alignment with reference genome GRCh38/hg19 was performed and sequences were subsequently analyzed with the CRAVAT program which provides high performance services for researchers to annotate and prioritize genes and exome variants. Annotations are provided at the level of variants, genes, and bioinformatic scores to allow interpretation and prioritization of the variants identified in the sequencing studies, which include reading frame displacement, insertions/deletions, splice site, missense, and nonsense.

VCard File (VCF) Editor.

The exome data in VCF format were submitted to the program VCF Editor v.1.0 created for the elaboration of the doctoral work "Representation of the Genomic Variability of the Mucopolysaccharidosis Complex in Southwestern Colombia" which allows editing the VCF format of the exomes, and converting it into an Excel program, with the objective of applying the different search filters and obtaining a final file with the information of interest. Visualizing columns organized by chromosome, position, altered amino acid, clinvar record, among other features. (Moreno-Giraldo et al., 2020).

Bioinformatic analysis

The variants found in the *GBA* gene were analyzed through a search in population, genome, disease, and sequence databases (Table 2), with the objective of verifying if they have been previously described.

Table 2. Population, genome, disease, and sequence databases used to analyze variants of the *GBA* gene. Adapted from Richards et al. (2015).

Population databases		
1000 genomes	http://browser.1000genomes.org	Database of variants found in a low and high genomic sequencing coverage targeting 26 different populations. Provides diversity in data but some data is of low quality
Exome Aggregation Consortium (Exac)	http://exac.broadinstitute.org/	Database of variants found during exome sequencing of 61,486 unrelated individuals
Genome databases		
Ensembl Genome Browser	http://www.ensembl.org/index.html .	Vertebrate genome search engine supporting research in comparative genomics, evolution, sequence variation, and transcriptional regulation. It currently has about 50,000 genomes for study.
Disease databases		
OMIM	http://www.omim.org	Database of human genes and genetic conditions that also contains a representative sample of genetics associated with variant diseases.
Clinvar	http://www.ncbi.nlm.nih.gov/clinvar	Database of statements about clinical significance and phenotypic relationship of human variation.
Human Gene Mutation Database	http://www.hgmd.org	Database of variant annotations published in the literature. Requires paid subscription for much of the content
Sequence database		
Genome	NCBI	Source of full human genome reference sequences.

Its pathological implications were determined through bioinformatic tools, by means of *in-silico* technology, which allows predicting the consequences of genetic variants in the structure and function of proteins, besides studying characteristics such as changes in physical-chemical properties of substituted amino acids, evolutionary conservation degree, the environment of the sequence of an affected amino acid or the alteration in the structural properties of the protein. The following programs were used: Polymorphism Phenotyping v2 (Polyphen-2), Mutation Taster (<http://www.mutationtaster.org/>), Predictor UMD (<http://umd-predictor.eu/>), PROVEAN (<http://provean.jcvi.org/index.php>), SIFT (<http://sift.jcvi.org/>), Human Splicing Finder (<http://umd.be/Redirect.html>). For each of the variants the position, nucleotide change, amino acid change, clinical and clinical significance, and allelic frequency were tabulated.

Calculation of allele frequencies

The allele frequency was calculated by a simple count, dividing the number of copies of the allele found in the population by the total number of copies of the gene in the population: Casillas et al. (2017).

$$\frac{\text{\# of copies of the allele in the population}}{\text{\# total copies of the gene in the population}} \quad (\text{Eq. 1})$$

Gene Interaction Network

Interaction networks were made for the *GBA* gene, using STITCH 5 (<http://stitch.embl.de/>) and GeneMANIA software with associated proteins and small molecules, considering only the evidence found from experiments, databases, biological process, molecular function, path, or domain of the protein that was altered and co-expression with a confidence level of 0.900 to determine interaction between nearby genes.

RESULTS

Forty-one variants associated with the *GBA* gene were found in the 320 exomes of patients with various pathologies and not previously diagnosed with GD. Fourteen of these variants had not been previously reported in genetic, population and *in-silico* databases ([Supplementary 1](#)). Among the variants, 17/41 constituted changes of missense type and 4/35 were changes in intronic regions (Figure 1). When observing the traditional type changes (C>T), we found a prevalence of 17% in the variants.

Regarding the clinical significance given by the American College of Medical Genetics and Genomics (ACMG) and according to the prediction software used, 021/41 of the variants reported represented a benign significance, 5/41 of the variants reported obtained a classification of pathogenic or probably pathogenic and 8/41 of the variants reported presented uncertain significance; in this last case, by means of *in-silico* analysis it was possible to reclassify some of these variants, obtaining in this way three pathogenic and three benign variants. The remaining 35 have not been reported in a functional way, therefore their clinical significance is unknown now.

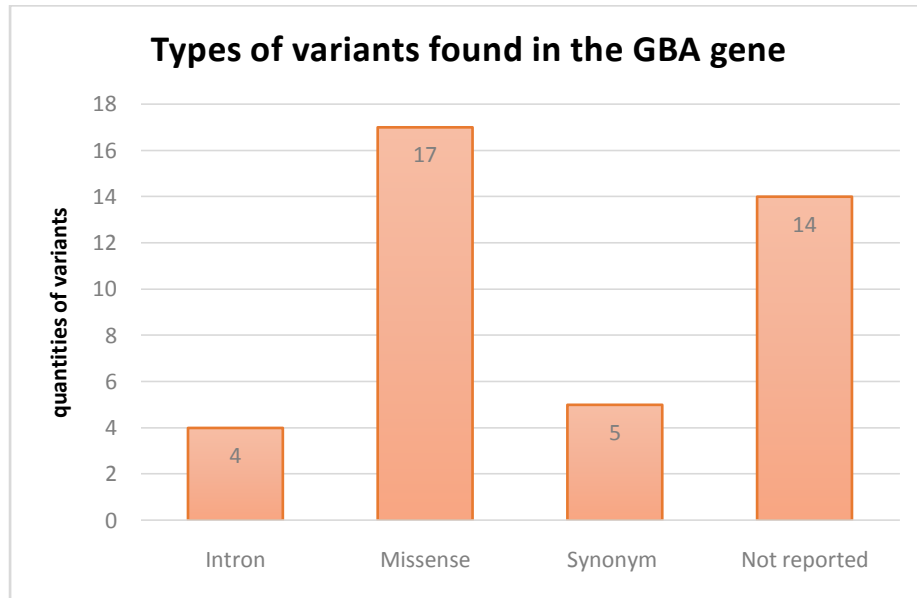


Figure 1. Types and quantities of variants found in the *GBA* gene in 320 exomes of patients from southwestern Colombia

The variant c.454+47G>A presented the highest allele frequency, being found in 37/41 of the samples; this variant, according to 5/6 bioinformatic predictors, presented a benign clinical significance; likewise, the variant p. Lys13Arg (c.38A>G) was found with a frequency of 22/41 in the samples and according to 6/6 predictors, it presents a benign clinical significance. The variants that according to the prediction software were classified as pathogenic or probably pathogenic (c.1444G>A p. Asp482Asn; c.259C>T p. Arg87Trp; c.247C>T p. Arg83Cys; c.669 G>T p. Trp223Cys; c.134C>T p. Pro45Leu) presented an allele frequency that ranged between 0.0062 and 0.0218, with 0.0062 being the lowest allele frequency reported in the variants found.

On the one hand, the gene interaction network in GeneMania, allowed us to observe close associations between *GBA*, *PSAP*, *SCARB2* and *LAMP2* genes and in turn 17 genes that interact slightly; all these genes have been related to functions associated with vacuolar locations, instructions concerning the lysosomal and vacuolar lumen, vacuolar and lysosomal membranes (Figure 2).

On the other hand, the network obtained with the STITCH 5 program allowed us to recognize biological processes associated with the activation and/or inactivation of the *GBA* gene and its nearby genes, such as catabolic and metabolic processes of glycosylceramide, and metabolic processes of sphingosines; molecular functions associated with the beta-galactosidase and hydrolase activity that affect cellular components such as the lysosomal lumen (Figure 3).

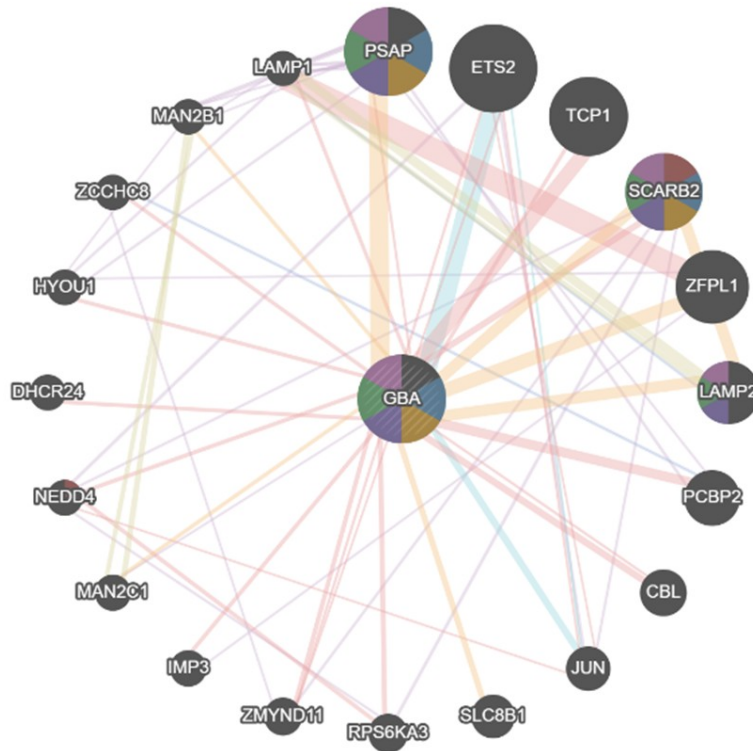


Figure 2. Interaction network between the *GBA* gene and associated genes. Generated with the GenMania program

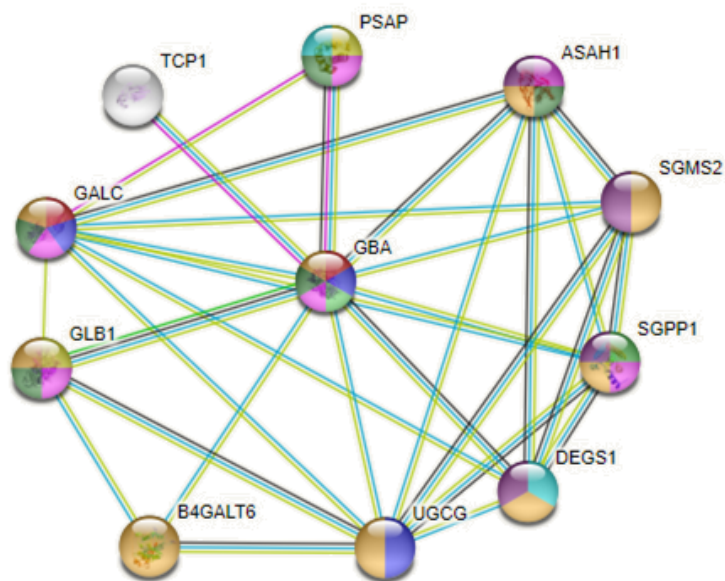


Figure 3. Diagram of the gene interaction network associated with EG using the STITCH 5 program

DISCUSSION

Considering that sometimes it is not possible to distinguish between healthy individuals and carriers, studies focused on the timely identification of the variants present in the *GBA* gene make it possible to discriminate between a benign and a pathogenic variant, predict the protein effect that occurs in the molecular conformation and determine the frequencies and probabilities of the appearance of the disease.

In the results of the complete exomes of 320 patients with different pathologies and not clinically diagnosed of GD, 41 variants of the *GBA* gene were found; 21 of them with clinical benign significance and 5 with clinical pathogenic significance; besides, 15 variants, that had not been previously informed within the bibliography, were reported. This highlights the relevance of this research, in favor of the pathological classification of the variants and the possible effects that can be found within the protein's functionality.

The variants described in the affected populations around the world seem to indicate that the region where a greater quantity of variants is located, is between exons 8, 9 and 10 (Beutler et al., 2005; Acanda, 2012). Of all the variants described in *GBA*, only about 120 have been associated with GD phenotypes; approximately 70 with type 1, 43 with type 2, and 8 with type 3 (Hruska et al., 2008), without establishing any gene region or domain where they occur preferentially and the ethnic and geographical distribution of each mutation may vary significantly (Bodamer et al., 2010; Franco-Ornelas, 2010). Approximately the 70% of all reported replacements, including the two most common pathogenic replacements, p. N370S and p. L444P, are found in exons 9-10 of the *GBA* gene (Asselta et al., 2014).

The different variants may be more represented ethnic groups, as well as phenotypes. The first variant reported in GD was p. Leu483Pro (Tsuji et al., 1987). The most common one identified in Romanian patients is p. Asn409Ser (Drugan et al., 2002). The most common among AJ patients is c.1226A <G (N370S; or N409S according to the new nomenclature) followed by c.84dupG (84GG), which is more uncommon (Riboldi et al., 2019). The change c.115 + 1G> A (IVS2 + 1), c.1504C> T (R463C) and c.1604G> A (R496H) are commonly found in AJ patients with GD1. In contrast, the N370S mutation is rarely found among Chinese and Japanese patients. (Hruska et al., 2008).

Among Asian ethnic groups, c.1448T> C (L444P or L483P according to the new nomenclature) and c.754T> A (F252I), generally associated with GD2 and 3, are more prevalent, and explain why among these populations, neuropathic forms of GD are more frequent (Pastores et al., 2000). c.1448T> C (L444P) is also the most frequent mutation among Caucasians of non-Ashkenazi Jewish descent (Grabowski et al., 1997).

The five variants of clinical pathogenic significance found in this work: the change c.1444 G>A p. Asp482Asn showed an allele frequency of 0.009; thus far it has not been reported in any patient with a clinical, enzymatic, or molecular diagnosis of GD, however it has been described by Neumann et al. (2009) in a sample of 790 British patients affected with Parkinson's disease, reporting it as an occasional mutation not described before in exons; Neumann et al. made a cross-species comparison of the affected amino acids, reporting that this variant is highly conserved in most mammalian species, but not in rats, zebrafish, *C. elegans* and *D. melanogaster*. These findings suggested that this variant in *GBA* not only causes an alteration in the amino acid sequence, but it is also very likely to have a pathogenic behavior. However, the precise functional effect and the effects on carrier

patients have not been revealed to the present. On the other hand, Lorraine (2018) conducted studies based on the report of genetic mutations that can be used to assess the risk of disease with Lewy bodies in 592 patients at the Alzheimer's Disease Research Center (ADRC) using the *GBA*, *SMPD1*, *HEXA*, and *MCOLN1* genes; they reported *GBA* mutations in 6.9% (41/592) of all subjects. When classifying subjects by the known phenotypic effect (in terms of Gaucher disease) of the mutation, as in the previously published studies, 2 silent variants (11191, Q226Q), 11 mild (all N370S), 3 severe (2 L444P and 1 R463C), and 25 of unknown phenotypic effect were observed where they included the variant c.1444G>A p. Asp482Asn found in 1/ 592 patients.

The pathogenic variant c.259C>T p. Arg87Trp was reported in this work with a frequency of 0.009; it has been registered in the population, genome and disease databases and it has not been reported so far in a homozygous or heterozygous state in any patient affected with GD. This variant was reported by Ortiz-Cabrera et al. (2016) in a study in which they collected blood samples from patients with GD for enzymatic and genetic analysis. The genetic analysis was performed on the DNA of 124 unrelated patients and 57 family members between 2007 and 2015. The two mutant alleles were found in 46 of 93 GD patients (49.5%). The most common variant was N370S (c.126A> G; current nomenclature p. Asp409Ser) (in 50.5% of patients), followed by L444P (c.1448T> C; current nomenclature p. Leu483Pro) (in 24.7%). The variant c.259C>T p. Arg87Trp was reported in compound heterozygosis together with c.1226A > G p.N409S (N370S) in an affected patient. Cullufi et al. (2019) reported the case of a family in which all members over two generations presented at least splenomegaly. *GBA* gene sequencing in the three first-generation siblings reported the presence of the same three heterozygous compound variants in each patient: c.259C> T (p. Arg87Trp; "R48W" according to historical nomenclature), c.1265_1319del (p. Leu422Profs *); and c.1342G> C (p.Asp448His; "D409H"). The consequences foreseen for the three pathogenic variants vary. While the deletion for Leu422Profs * 4 has been catalogued as a true allele of loss of function, the variants c.259C> T (p. Arg87Trp) and c.1226A> G (p. Asn409Ser) result in meaningless mutated proteins that can retain residual functionality. In addition, c.1226A> G (p. Asn409Ser) has proven to represent a particularly mild mutation.

The studies by Sheth et al. (2019) determined the spectrum of mutations and the molecular pathology of GD in India, investigating 100 unrelated patients (age range: 1 day to 31 years) who had splenomegaly, with or without hepatomegaly, cytopenia and bone abnormalities. Biochemical analysis reported a significant reduction in β -glucosidase activity in all patients. Sanger sequencing established 71 patients with homozygous mutation and 22 patients with heterozygous compound mutation in the *GBA1* gene. Overall, the study identified 33 mutations in 100 patients that also included four nonsense mutations (p. Ser136Leu, p. Leu279Val, p. Gly383Asp, p. Gly399Arg) not previously reported in patients with GD. The p. Leu483Pro mutation was identified as the most common Gaucher Disease mutation in the study (62% of patients) and additionally, the c.259C>T p. Arg87Trp variant was reported in compound heterozygosis in a patient diagnosed with GD2. Finally, Cullufi et al. (2020) characterized a cohort of 36 Albanian patients with GD, 31 with GD type 1 and 5 affected by GD types 2, 3, and an intermediate GD phenotype between type 2 and type 3. In their studies, 12 different GD alleles were detected, being p. Asn409 the most frequently observed, followed by p. Asp448His; His294Gln. Among the 12 alleles

found, c.259C>T p. Arg87Trp was reported in an affected patient in a state of compound heterozygosity.

Regarding the variant c.247C>T p. Arg83Cys (R44C) reported with a frequency of 0.0062 in this study ([Supplementary 1](#)), it has been registered in the population, genome, and disease databases. It has not been reported so far in patients affected with GD. Ruskey et al. (2019) reported this variant in a study conducted in 735 patients with AJ with PD and 662 AJ controls from Israel and New York. *GBA* variants were found in 18% of PD patients and 7.5% of controls. Curiously, they found a surprisingly high frequency of the c.247C>T p. Arg83Cys variant in the control population (1.06%). This variant presented a particular interest, since in control patients the frequency of this variant was approximately 4 times higher than in PD patients. Therefore, it will be important to examine the effect of the variant in additional populations of AJ and other ascendancies, proposing that, in this way, some of the controls (recruited at an early age without any clinical data), may develop PD in the future; Thus, although there is good evidence linking *GBA* mutations to a variety of neuropsychiatric phenotypes in PD and other diseases, and based on the loss and gain of function hypothesis to explain the mechanistic link between *GBA* mutations and the expression of synucleinopathies, the mechanisms underlying these associations are still unclear.

The variants c.669G>T p. Trp223Cys and c.134C>T p. Pro45Leu reported with a frequency of 0.0062 and 0.0218 respectively in this study, have been registered in the population, genome, and disease databases. According to Clinvar database, they present an uncertain clinical significance; however, when determining their functionality in bioinformatic databases, by means of in-silico technology, it was determined their possible pathogenic action in the protein. So far, no cases of patients affected with GD or PD have been reported, consequently it would be necessary to conduct studies of functionality associated to determine the protein action according to the corresponding zygosis.

On the other hand, when observing the variants that presented a benign clinical significance, it was found that the change c.454+47G>A was presented with the highest frequency in the samples (90%). This variant has been registered in the population, genome, and disease databases, and so far, it has not been reported in a homozygous or heterozygous state in any patient affected with GD. This variant was reported by Dutra et al. (2019) in a research carried out on mutations in *GBA* and *CHCHD2* genes in patients with Parkinson's disease in a sample of Brazilian population; where they identified eight intronic changes and the most incident was c.454 + 47G> A (rs3115534), found in 169/304 probands, in homozygosity in intron 4 of *GBA* gene, patients affected with PD presented this variant together with other variants of pathogenic clinical significance in the same gene. According to the studies conducted by Carvalho (2012), the *in-silico* analysis through programs such as NNSPLICE carried out in intronic alterations as c.454 + 47G> A, reports that there is not a condition in the recognition of donor sites and receptor sites. This variant is described in the SNPs database; however, it has not been associated by itself with any type of disease. Likewise, considering that this intronic variant does not seem to affect the recognition of the splicing site and according to the bioinformatic analyses performed, it has been concluded that this alteration is not pathogenic.

A second variant with benign clinical significance (c38A>G p. Lys13Arg) was found with a high frequency represented in 54% of the samples; this variant has been registered in the population, genome and diseases databases and it has not been reported so

far in homozygous state in patients affected with GD. This variant was first reported by Rozenberg et al. (2006) in a study conducted in 40 Brazilian patients not related to Gaucher disease type 1 through a combination of RFLP, dHPLC and DNA sequencing procedures; this study determined 12 new mutations including c.38A> G p.Lys13Arg (K-27R); this variant was found in compound heterozygosis together with N370S/W184R/K(-)27R (A1226G/T667C/A38G). According to Rozenberg et al. (2006) this alteration affects the signal peptide composed of 39 residues that is removed from the mature enzyme during processing. Mutations in this leading sequence have not been commonly found among patients with Gaucher disease. Since K (-27) R is the first nonsense mutation reported in the leader sequence, it would be interesting to investigate in the future whether it alters cytoplasmic transport; however, other mutations of this patient, N370S and W184R, may be the only causes of the phenotype in the patient reported in their studies with GD and this set of variants (Beutler, 2005).

The rest of the variants presented frequencies between 0.009 and 0.05; according to the prediction software used to interpret the functional effects of the sequence variants, they predict that these would be benign and well tolerated variants, reflecting the low probability of being associated with a clinical phenotype of GD; notwithstanding, some variants classified as probably benign or benign could function as disease modifiers: variants that do not cause the disease, but may worsen the outcome, therefore it is important to strictly monitor the possible *in-silico* re-classifications that can be addressed; likewise, it will be important to initiate studies on those variants that have not been previously reported and of which it will be necessary to initiate predictive studies to recognize their clinical significance.

Finally, from the gene interaction network based on *GBA*, it is possible to determine a close relationship between some of them, maintaining a high genetic interaction, which could possibly lead to specific phenotypes related to the observed phenotype and the clinical manifestations that occur in patients with GD. These networks allow evidencing the natural function of each of the genes, necessary for the development of biological processes such as catabolism and metabolism of lipids. That is how the close interaction of the *GBA* gene with nearby genes such as *PSAP*, *LAMP2* and *SCARB2* generates that a specific damage in some of these genes alters the activity of the others, resulting in side effects in the organism.

It has been widely discussed that the main activator of β -glucocerebrosidase, saposin C, is encoded by the prosaposin gene (*PSAP*) located at the locus 10q21-22; saposin C is an established activator for the hydrolysis of glycosylceramide by GCCase in lysosomes, however, the mechanism by which saposin C promotes lysosomal hydrolysis is not fully known (Tamargo et al., 2012).

Nowadays it is recognized that the conditions in *PSAP* cause alterations in the activity of *GBA* giving rise to typical signs such as hepatosplenomegaly and neurological involvement, being related to type 2 of the GD, which has led to the search for disease modifiers that can alter the phenotype of the GD. In humans, the deficiency of saposin C due to mutations in *PSAP* results in a phenotype like GD, despite having values of the normal glucocerebrosidase activity *in vitro*. It has also been shown that deficiency of C saposin modifies the phenotype in murine models with GD (Tamargo et al., 2012).

The transcription factors responsible for the differential expression of the *GBA* gene in specific tissues include proteins such as *OCTA*, *AP-1*, *PEA3* and *CAAT*. Deficiencies in

genes encoding membrane proteins associated to lysosomes (LAMP-1 and LAMP-2) and being involved in intracellular transport of glycosylcerebrosidase from the endoplasmic reticulum to lysosomes may also be present in patients with GD (Schnabel et al., 1991; Martinez-Arias et al., 2001, Arturo Terranova et al., 2020). Studies by Kurzawa-Akanbi et al. (2012) reported that carriers of pathogenic mutations in *GBA* demonstrated significantly reduced level of protein GCase and enzymatic activity and retention of glucocerebrosidase isoforms within the endoplasmic reticulum (ER), which was associated with increased expression of the lysosomal markers *LAMP1* and *LAMP2*. Understanding how genes close to *GBA* interact will allow a better understanding of phenotypic and genotypic characteristics in affected patients.

In conclusion, in the analyzed sample, a high frequency of benign variants was found, however find 5 variants with clinical pathogenic significance (c.1444G>A p. Asp482Asn; c.259C>T p. Arg87Trp; c.247C>T p. Arg83Cys; c.669G>T p.Trp223Cys; c.134 C>T p.Pro45Leu) circulating in the evaluated population, generates the need to make phenotype-genotype associations that allow an early and timely diagnosis to provide a specific and targeted treatment, thus improving the quality of life, based on the approach of precision medicine. Hence the importance of periodic reclassification of those variants of uncertain significance, not reported or benign for the correct determination of subsequent effects, allowing the issuance of new medical reports.

To increase the information reported on this pathology until the moment will open the way to further investigations, recognizing the functional impact that it can generate, contributing to the knowledge and creation of molecular bases on the GD in the Colombian Southwest and allowing to forecast the evolution of the disease according to the genotype of the affected individual.

ACKNOWLEDGMENTS

Project funded by the DGI-03-2020 call: Applied research projects, Universidad Santiago de Cali. Acknowledgements to the participating patients, Universidad del Valle, Faculty of Health, Research Group on Congenital Diseases of Metabolism.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- ADRES (2017). Análisis De Los Recobros Correspondientes A Los Principios Activos Imiglucerasa, Miglustato, Velaglucerasa Y Taliglucerasa Medicamentos Para La Enfermedad De Gaucher. Available from: http://www.med-informatica.net/TERAPEUTICA-STAR/EnfermedadGaucherVF_Recobros2017_InformeADRES_08ago18.pdf.
- Acanda de la Rocha A (2012). Aspectos bioquímicos, genéticos y comorbilidades de la enfermedad de Gaucher, diagnóstico molecular en Cuba. *Rev Cubana Genet Comunit.* 6(1): 8-19.
- Arturo-Terranova D, Moreno Giraldo LJ, Idrobo H and Satizabal JM (2020). Variantes del gen GBA en el Suroccidente Colombiano. *Rev. Asoc. Col. Cienc. Biol.* 1(32): 115-123. <https://doi.org/10.47499/revistaaccb.v1i32.214>.
- Asselta R, Rimoldi V, Siri C, Cilia R, et al. (2014). Glucocerebrosidase mutations in primary parkinsonism. *Parkinsonism Relat. Disord.* 20: 1215-1220 <https://doi.org/10.1016/j.parkreldis.2014.09.003>.
- Beutler E, Gelbart T and Scott CR (2005) Hematologically important mutations: Gaucher disease. *Blood Cells Mol Dis.* 35: 355- 64. <https://doi.org/10.1016/j.bcmd.2005.07.005>.

- Bodamer OA and Hung C (2010). Laboratory and genetic evaluation of Gaucher disease. *Wien Med Wochenschr.* 160: 600-4. <https://doi.org/10.1007/s10354-010-0814-1>.
- Cadavid F (1925). Un caso de enfermedad de Gaucher. *Rev. Med. Quirurg. Hosp. Bogotá.* 1(4): 167-169.
- Carvalho Guimaraes B (2012). Rastreamento de mutações no gene GBA como fator de risco ao desenvolvimento da doença de Parkinson na população brasileira. Masters of Science thesis. Universidade de Estado de Rio de Janeiro. Available at: http://www.bdt.uerj.br/tde_busca/arquivo.php?codArquivo=3690.
- Casillas S and Barbadilla A (2017). Molecular Population Genetics. *Genetics.* 205: 1003-1035 <https://doi.org/10.1534/genetics.116.196493>.
- Cullufi P, Tabaku M, Beetz C and Tomori S (2019). Comprehensive clinical, biochemical and genetic screening reveals four distinct GBA genotypes as underlying variable manifestation of Gaucher disease in a single family. *Mol. Genet. Metab. Rep.* 21: 100532 <https://doi.org/10.1016/j.ymgmr.2019.100532>.
- Cullufi P, Tabaku M, Velmishi V, Gjokopulli A, et al. (2020). Genetic characterization of the Albanian Gaucher disease patient population. *JIMD Reports.* 57: 52-57. <https://doi.org/10.1002/jmd2.12167>.
- Di Rocco M, Andria G, Deodato F, Giona F, et al. (2014). Early diagnosis of Gaucher disease in pediatric patients: proposal for a diagnostic algorithm. *Pediatr Blood Cancer.* 61(11): 1905-9. <https://doi.org/10.1002/pbc.25165>.
- Drugan C, Procopciuc L, Jebeleanu G, Grigorescu-Sido P, et al. (2002). Enfermedad de Gaucher en pacientes rumanos: incidencia de las mutaciones y manifestaciones fenotípicas más comunes. *Eur. J. Hum. Genet.* 10 (9): 511-5.
- Dutra D (2019). Investigação de mutações nos genes GBA e CHCHD2 em pacientes com doença de parkinson em uma amostra da população brasileira. Ph.D. thesis. Universidade do Grande Rio, Duque de Caxias. Available at: <http://localhost:8080/tede/handle/tede/367>.
- Echavarría A, Harry N, Bedoya V and López F (1966). Enfermedad de Gaucher. *Pediatra* (Bogotá). 8: 67.
- Franco-Ornelas S (2010). Consenso Mexicano de Enfermedad de Gaucher. *Rev. Med. Inst. Mex. Seguro Soc.* 48: 167-86.
- Grabowski GA and Horowitz M (1997). Gaucher's Disease: molecular, genetic and enzymological aspects. *Baillieres Clin Haematol.* 10(4): 635-56 [https://doi.org/10.1016/s0950-3536\(97\)80032-7](https://doi.org/10.1016/s0950-3536(97)80032-7).
- Hruska KS, LaMarca ME, Scott CR and Sidransky E (2008). Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum. Mutat.* 29(5): 567-83 <https://doi.org/10.1002/humu.20676>.
- Horowitz M, Tzuri G, Eyal N, Berebi A, et al. (1993). Prevalence of nine mutations among Jewish and non-Jewish Gaucher disease patients. *Am. J. Hum. Genet.* 53: 921-30.
- International Collaborative Gaucher Group (2018). Available at: www.gaucherregistry.com.
- Kurzawa-Akanbi M, Hanson PS, Blain PG and Lett D (2012). Glucocerebrosidase mutations alter the endoplasmic reticulum and lysosomes in Lewy body disease. *J. Neurochem.* 123(2): 298-309. <https://doi.org/10.1111/j.1471-4159.2012.07879.x>.
- Lee JY, Lee BH, Kim GH, Jung CW, et al. (2012). Clinical and genetic characteristics of Gaucher disease according to phenotypic subgroups. *Korean J. Pediatr.* 55(2): 48-53. <https://doi.org/10.3345/kjp.2012.55.2.48>.
- Lorraine C (2018). Predictive Value Of Combined Genetic Enzymatic And Lipidomic Data In Disease Risk For Lewy Body Disease. United States Patent. 20180051337. <https://www.freepatentsonline.com/y2018/0051337.html>.
- Lozano-Bernal JE (2006). Enfermedad de Gaucher: Casuística del Tolima. *Acta Med. Colomb.* 31: 416-421.
- Martínez-Arias R, Comas D, Mateu E and Bertranpetit J (2001). Glucocerebrosidase pseudogene variation and Gaucher disease: recognizing pseudogene tracts in GBA alleles. *Hum. Mutat.* 17(3): 191-198 <https://doi.org/10.1002/humu.4>.
- Mistry PK, Sadan S, Yang R, Yee J, et al. (2007). Consequences of diagnostic delays in type I Gaucher disease: the need for greater awareness among hematologists-oncologists and an opportunity for early diagnosis and intervention. *Am. J. Hematol.* 82(8): 697-701. <https://doi.org/10.1002/ajh.20908>.
- Mistry PK, Cappellini MD, Lukina E, Ozsan H, et al. (2011). A reappraisal of Gaucher disease-diagnosis and disease management algorithms. *Am. J. Hematol.* 86(1): 110-115 <https://doi.org/10.1002/ajh.21888>.
- Moreno-Giraldo LJ, Satizabal-Soto JM and Gomez A (2020). Genomic variability of the mucopolysaccharidosis complex in southwestern Colombia. *Genet. Mol. Res.* 19(2): GMR18502 <https://doi.org/10.4238/gmr18502>.
- Nalysnyk L, Rotella P, Simeone J, Hamed A, et al. (2017). Gaucher disease epidemiology and natural history: a comprehensive review of the literature. *Hematology.* 22: 65-73. <https://doi.org/10.1080/10245332.2016.1240391>.
- Neumann J, Bras J, Deas E, Sean S, et al. (2009). Glucocerebrosidase mutations in clinical and pathologically proven Parkinson's disease. *Brain.* 132: 1783-94. <http://doi.org/10.1093/brain/awp044>.
- Nguyen Y, Stirnemann J and Belmatoug N (2019). Gaucher disease: A review. *Rev. Med. Interne.* 18: 31185-8 <http://doi.org/10.1016/j.revmed.2018.11.012>.
- Ortiz-Cabrera NV, Gallego-Merlo J, Vélez-Monsalve C and de Nicolas R (2016). Nine-year experience in Gaucher disease diagnosis at the Spanish reference center Fundación Jiménez Díaz. *Mol. Genet. Metab. Rep.* 9: 79-85. <https://doi.org/10.1016/j.ymgmr.2016.06.008>.
- Pastores GM and Hughes DA (2000). Gaucher Disease. [Updated 2018 Jun 21]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. University of Washington, Seattle. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1269/>.

- Riboldi G and Di Fonzo A (2019). *GBA*, Gaucher Disease, and Parkinson's Disease: From Genetic to Clinic to New Therapeutic Approaches. *Cells*. 8(4): 364. <https://doi.org/10.3390/cells8040364>.
- Richards S, Aziz N, Bale S, et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 17: 405-423. <https://doi.org/10.1038/gim.2015.30>.
- Rodríguez G, Alvarado H, Yunis E, Sabogal A, et al. (1970). Enfermedad de Gaucher. Estudio de una familia. Trabajo presentado durante el XIII Congreso de la Sociedad Colombiana de Patología reunido en Popayán, Diciembre
- Rodríguez G and Cadena D (1972). Enfermedad de Gaucher en Colombia. *Rev. Fac. Med.* 38(1): 73-81.
- Rozenberg R, Fox D, Sobreira E and Pereira L (2006). Detection of 12 new mutations in Gaucher disease Brazilian patients. *Blood Cells Mol. Dis.* 37(3): 204-9 <https://doi.org/10.1016/j.bcmd.2006.09.004>.
- Ruskey J, Greenbaum L, Roncière L and Alam A (2019). Increased yield of full *GBA* sequencing in Ashkenazi Jews with Parkinson's disease. *Eur. J. Med. Genet.* 62(1): 65-69 <https://doi.org/10.1016/j.ejmg.2018.05.005>.
- Schnabel D, Schroder M and Sandhoff K (1991). Mutation in the sphingolipid activator protein 2 in a patient with a variant of Gaucher disease. *FEBS Lett.* 284(1): 57-59. [https://doi.org/10.1016/0014-5793\(91\)80760-z](https://doi.org/10.1016/0014-5793(91)80760-z).
- Simon G, Erdos M, Maródi L and Tóth J (2008). Gaucher-kór: a korai diagnózis és terápia jelentősége [Gaucher disease: importance of early diagnosis and therapy]. *Orvosi Hetilap.* 149(16): 743-50. <https://doi.org/10.1556/OH.2008.28337>.
- Sheth J, Bhavsar R, Mistri M, Pancholi D, et al. (2019). Gaucher disease: single gene molecular characterization of one-hundred Indian patients reveals novel variants and the most prevalent mutation. *BMC Med. Genet.* 20(1): 31. <https://doi.org/10.1186/s12881-019-0759-1>.
- Sun A (2018). Lysosomal storage disease overview. *Ann Transl Med.* 6(24): 6-12 <https://doi.org/10.21037/atm.2018.11.39>.
- Tamargo RJ, Velayati A, Goldin E and Sidransky E (2012). The role of saposin C in Gaucher disease. *Mol Genet Metab.* 106(3): 257-263. <https://doi.org/10.1016/j.ymgme.2012.04.024>.
- Thomas AS, Mehta ABv Hughes DA (2013). Diagnosing Gaucher disease: an on-going need for increased awareness amongst haematologists. *Blood Cells Mol. Dis.* 50(3): 212-217. <https://doi.org/10.1016/j.bcmd.2012.11.004>.
- Tsuji S, Choudary PV, Martin BM, Stubblefield BK, et al. (1987). A Mutation in the Human Glucocerebrosidase Gene in Neuronopathic Gaucher's Disease. *N. Engl. J. Med.* 316(10): 570-5.
- Uribe A, Ayala A, España M and Pacheco N (2013). Recopilación de 18 años de tamizaje de alto riesgo para desordenes lisosomales en Colombia (1995 – 2012) IX congreso latinoamericano de errores innatos del metabolismo y pesquisa neonatal. Resúmenes.
- Van Rossum A and Holsopple M (2016). Enzyme Replacement or Substrate Reduction? A Review of Gaucher Disease Treatment Options. *Hosp. Pharm.* 51(7): 553-563. <https://doi.org/10.1310/hpj5107-553>.
- Zampieri S, Cattarossi S, Bembí B and Dardis A (2017). *GBA* Analysis in Next-Generation Era Pitfalls, Challenges, and Possible Solutions. *J. Mol. Diagn.* 19: 5 <https://doi.org/10.1016/j.jmoldx.2017.05.005>.