INTRODUCTION

Plasma Factor IX protein deficiency causes a coagulation disorder known as Hemophilia B (HB) (OMIM 306900), characterized by spontaneous or secondary bleedings in response to traumatic or surgical events. It is caused by alterations in the DNA sequence of the Factor IX gene (F9). This gene is 32 Kilobases in size, with 8 exons and 7 introns, and it is located on the X chromosome (Xq27.1). HB presents an X linked recessive inheritance pattern (Goodeve, 2015). HB is classified based on the percentage of residual activity of FIX in plasma in: severe (<1%), moderate (1%–5%) and mild (5%–40%), and it has a high phenotypic heterogeneity (Srivastava et al., 2013).

There is a high allelic heterogeneity in HB, and it is possible to find almost any type of variant. The point variants are observed more frequently (73% of the cases), followed by deletions (16%) and, in minor proportion, insertions, duplications, small indels and large rearrangements (Belvini et al., 2005).

Sixty-five percent of patients with HB have missense variants, approximately 83% of the pathogenic variants...
are located in exonic regions, and 12% in intronic regions (Rallapalli, Kemball-Cook, Tuddenham, Gomez, & Perkins, 2013).

The 2017 Epidemiological Report, shows a total of 362 men with HB in Colombia, out of which 38.1% had a severe phenotype, 36.4% moderate phenotype, and 23.4% mild phenotype (Fondo Colombiano de enfermedades de alto costo, 2018).

Different studies in countries worldwide, such as Spain, United States, Italy, Argentina, and Brazil, have identified the pathogenic variants of the F9 in their population, these countries have reference centers for diagnosis as well, (Belvini et al., 2005; Li, Miller, Payne, & Craig Hooper, 2013; Meireles et al., 2017; Radic, 2010). This study is the first molecular analysis of Hemophilia B in Colombian population.

2 | MATERIALS, SUBJECTS, AND METHODS

2.1 | Subjects

Twenty unrelated men samples (ages between 5 and 68 years) with a confirmed HB diagnosis and receiving treatment in a comprehensive management program for Hemophilia B, were analyzed. The study was previously approved by the Ethics Committees of the Faculty of Medicine of the National University of Colombia and the participating health care entities.

Table 1 shows patients' data such as: age, severity of HB, percentage of factor IX activity (coagulometric analysis), history of inhibitor development, and whether they have a family history of HB or not.

2.2 | Collection of blood sample and DNA isolation

Four milliliters of peripheral blood were collected from each patient in EDTA tubes. Genomic DNA extraction was carried out with 200 μl of the collected blood, using the “DNA Mini kit Quiagen” (Quiagen Corporation) following the manufacturer's instructions; the quantification of the isolated DNA was performed by spectrometry in a “Nanodrop 2000”. Subsequently, the DNA was stored at −4°C.

2.3 | F9 sequencing

The most important F9 regions were amplified: promoter, eight exons, exon-intron junction regions, and the 3’ UTR, in 12 PCR reactions, with primers previously reported (Radic, 2010). In each PCR reaction, we used the sample of a normal individual as a positive control.

PCR was carried out with 3 μl of DNA, 5 μl Buffer 5X GoTaQ (Promega), 2 μl dNTPS (200 μM), 1.5/2.0 μl MgCl2 (Stock), 8.375 μl of H2O, 0.125 μl of GoTaQ Polymerase, and 5 μl of each primer pair, with a final reaction volume of 25 μl. Thermocycling conditions for the 12 PCR reactions were: Hot start 94°C for three minutes, followed by 30 cycles of: denaturation at 94°C for 30 s, annealing at 55°C for 30 s, elongation at 72°C for 60 s, and a final extension at 72°C for five minutes (Radic, 2010). PCR amplification was confirmed through electrophoresis in Nusieve agarose gel 3:1 at 2% at 100 V for one hour.

Purification of the amplified regions was performed using "Purelink Quick PCR Purification Kit" (Invitrogen), following the manufacturer's instructions. Quantification was performed by spectrometry with “Nanodrop 2000”.

The sequence of each fragment was amplified with the “BigDye 3.1 Terminator Cycle Sequencing Kit” (Applied Biosystems), using 3.5 ng/μl of the PCR products purified with 1.2 μl of the primer forward or reverse (10 pmol/μl); capillary electrophoresis was performed in a “ABI3500” (Applied Biosystems).

2.4 | Sequence analysis

For sequence quality analysis and variants identification, the “Chromas 2.6.5” (BioEdit) and “SeqScape V5.4” (Applied Biosystems) software packages were used applying the reference sequence. The Genebank accession numbers are: NG_007994.1 (Homo sapiens coagulation Factor IX (F9), RefSeqGen LRG 556 on chromosome X), NM_000133.3 and NP_000124.1.

The variants were reported following the guidelines of the Human Genome Variation Society (HGVS) (den Dunnen et al., 2016), the international F9 databases were consulted (http://www.factorix.org/; https://www.cdc.gov/ncbd/dd/hemophilia/champs; http://www.hemobase.com/) to verify if they were previously reported.

2.5 | Bioinformatics analysis

The pathogenicity and stability of the protein in the missense variants found were analyzed through the use of different software and online bioinformatics tools, such as: “POLYPHEN2,” “I-Mutant2.0,” “PANTHER,” “Mutation Taster,” and “iStable.” “Project HOPE” (http://www.cmbi.ru.nl/hope/input) allowed us to make a 3D modeling of the protein, and simultaneously showed us both mutated and wild residue was (Capriotti, Calabrese, & Casadio, 2006;
Chen, Lin, & Chu, 2013; Ramensky, 2002; Schwarz, Cooper, Schuelke, & Seelow, 2014; Tang & Thomas., 2016).

The analysis for variants located in intron 5, and 3’ UTR, were performed with the “Human Splicing Finder HSF 3.0” tools which combines 12 different algorithms to predict the impact of variants in acceptor sites, splicing donors, branch points, possible breaking points, and auxiliary sequences of potentiators or silencers of both: exonic and intronic splicing (Tang, Prosser, & Love, 2016), another tool used to analyze these variants was NNSplicing (Reese, Eeckman, Kulp, & Haussler, 1997).

3 | RESULTS

The pathogenic variant was identified in all samples tested. We identified 14 unique pathogenic variants, of which, 7 were missense, 3 nonsense, 2 large deletions, 1 substitution affecting splicing, and 1 substitution at the 3’ UTR (Table 2). Of these variants, only the large deletion of exon 3 and 4 have not been previously reported. Eighteen patients presented a single variant and two had a double variant (HB_07 and 14).

### 3.1 Missense variants

In seven patients, seven missense variants were identified, two of these were found in exon 6, three in exon 8, one in exon 5, and one in exon 3 (See Figure 1). Only two of these variants, p.Arg294Gln and p.Arg191Cys, are located in CpG dinucleotides considered variant hot spots. The p.Arg191Cys variant was found in two unrelated patients.

Two patients (HB_07 and 14) had a double variant. Patient HB_07 has both variants in exon 8, g.35917G>A and g.36288G>C, located in the catalytic domain of the protein (p.Glu323Lys; p.Lys446Asn). The two variants found in patient HB_14 are located in exon 6, g.25527C>A and g.25386G>A (p.Gln241Lys and p.Thr194Ala). One of them, p.Thr194Ala, is a benign variant (HB_14* in Figure 1).

All these variants, according to the in-silico analysis and literature, are classified as deleterious or probably deleterious, and affecting the stability of the protein.

### 3.2 Nonsense variants

Nonsense variants were found in 15% of the patients, two variants located in exon 8: g.36085 C>T identified in patient HB_03, and g.35842 C>T identified in patient HB_13;
one variant in exon 2: g.11409 C>T was observed in patient HB_04. All three are in CpG dinucleotides, for which there are numerous reports of patients with these same variants in the literature.

### 3.3 Substitution in the intron 5 and 3′ UTR

In seven apparently unrelated patients (35%), the same variant was observed in intron 5, 13 nucleotides from the donor splice site (g.22769A>G; c.520+13A>G). The bioinformatics analysis with “NNSplicing,” predicted a new splicing donor site with a score of 0.41 compared to the wild-type sequence of 0.2; with “Human Splicing Finder, HSF”, we identified the formation of a new potential splicing donor site with a +50.25% variation, in favor of the mutant sequence versus the wild sequence; additionally, two new splice enhancers binding sites are also formed. This variant presumably generates a mRNA with the inclusion of 12 bp in exon 5 that will introduce 4 additional codons, altering protein translation by introducing 4 new amino acids (Gly, His, Asn, and Leu), also in the exon 6 the first amino acid Val would change to Met, resulting in the possible decrease of FIX activity.

A variant in exon 8 was detected in patient HB_02, g.37493 A>G, which corresponds to the 3′ UTR region of the F9. Bioinformatics analysis with “Human Splicing Finder, HSF” showed that this variant generates a new cryptic donor splicing site, and additionally, generates an exonic splicing enhancer (ESE).

### 3.4 Large deletions

Two patients (HB_08 and 09), 10% of the analyzed sample, had large deletions (>50 pb), evidenced by the absence of PCR amplicon bands in agarose gel electrophoresis from the amplified exons.
Patient HB_08 showed the absence of fragments E and F amplification, corresponding to exon 5 and exon 6, respectively.

A deletion of exons 3 and 4 (fragments C and D, respectively) was observed in patient HB_09, who presents moderate HB. This variant has not been reported in the literature to date. There are deletions reports that include these two exons with other additional exons. The exact location of the break in the F9 is not known since the Primers used for the amplification are found in adjacent intronic regions.

These results were confirmed in duplicate with a new DNA isolation for each patient, where the amplicon absence of the fragments was evidenced again. Each PCR was carried out simultaneously with the DNA sample from a normal individual (+ control), that showed successful amplification of those fragments, in addition, multiplex PCR was performed on both cases with primers involved in the deletion sequence (HB_08: exons 5 and 6; HB_09: exons 3 and 4) and the primers non involving (F9: exon 1 and beta globin gene NM 000518.4), the result confirmed the absence of amplification of the deleted exons (data not shown), therefore we ruled out the lack of amplifications due to defective primers.

### 3.5 | Genotype-phenotype correlation

Five patients (25%) had a severe phenotype. From these, two have a nonsense variant, two a missense variant and one a deletion of exons 5 and 6. Nine patients (45%) have moderate HB. Three of them presented a variant in intron 5 that affects splicing, one has a double missense variant, two have a missense variant, one has a nonsense variant, one has a variant in the 3' UTR, and one presented the newly detected variant, a deletion of exons 3 and 4. Six patients (30%) were reported to have mild HB, four of them have the same variant that affects splicing in intron 5, one has a double missense, and one has a missense variant.

None of the 20 patients have developed inhibitors to date. Eighty-five percent of the patients have a family history of HB. From the three patients without a family history of HB, one has severe, one moderate, and one mild phenotype. Seven patients (35%) received treatment in the prophylactic scheme and 13 (65%) received on demand treatment.

### 4 | DISCUSSION

This is the first molecular HB study in Colombia. The pathogenic variant for HB was identified in 100% of the analyzed cohort, similar to other reports in the literature (Abla et al., 2018; Khan et al, 2018). Ninety percent of the studied patients presented single nucleotide variants, which is slightly higher than what is reported in other studies (Abla et al., 2018; Khan et al, 2018). Ten percent of patients had large structural variants, similar to previous studies (Li et al., 2014), and to the European database (EAHAD-CFDB).

The pathogenic variant frequencies found in this series are different from those reported in the literature. For example, missense variants represent 35% of the variants in this study, compared to 65%–70% from other studies (Li et al., 2014; Mårtensson, Letelier, Hallldén, & Ljung, 2016; Rallapalli et al., 2013). The frequency of nonsense variants (15%) are similar to other studies and databases (Li et al., 2014; Radic, 2010). The largest difference in mutational frequencies was found in variants that affect splicing, with a frequency of 35% of the patients analyzed. These results contrast strongly with other reports. Splice variants had been found in frequencies of 3 and 8% in the Argentinean
and Brazilian cohort, respectively (Li et al., 2014; Mårtensson et al., 2016; Meireles et al., 2017; Radic et al., 2013; Rallapalli et al., 2013). This great difference is due to the presence of the same pathogenic variant (g.22769A>G) in seven patients, who are unrelated based on genetic interviews and genealogies. These unexpected findings suggest the possible existence of a founder effect. Further single nucleotide variant studies of X chromosome would allow us to determine if the variant is associated with the same haplotype to confirm this hypothesis.

The variant found in patient HB_02 in the 3’ UTR region was previously described by Vielhaber and collaborators in four individuals of different ethnicities. They pointed out the high possibility that this new splicing site generates a bypass in the polyadenylation signal which would produce an unstable mRNA that could easily be degraded in the cytoplasm (Vielhaber, Jacobson, Ketterling, Liu, & Sommer, 1993).

The only pathogenic variant not reported previously in the literature corresponds to a large deletion of exons 3 and 4. This deletion is responsible for the absence of the Gla and EGF1 protein domains involved in calcium-dependent platelet surface adhesion and cellular interactions (Thompson, 1986).

Some variants identified in the present study have been reported previously in Argentinian (Radic, 2010), and Brazilian populations (Meireles et al., 2017). Variants p.Arg75* and p.Arg298* have been identified in Brazil and Argentina, while p.Arg191Cys variant, p.Arg294Gln, p.Arg379*, and p.Arg298* have been identified in Brazil and Argentina, populations (Meireles et al., 2017). Variants p.Arg75* and p.Arg191Cys variant affect the possible existence of a founder effect. Further single nucleotide variant studies of X chromosome would allow us to determine if the variant is associated with the same haplotype to confirm this hypothesis.

In patient HB_16, with a mild phenotype, the genotype identified (a missense variant in exon 3) differs from that reported in the literature: three patients, two with moderate phenotype and one with unknown phenotype.

5 | CONCLUSIONS

We have performed molecular analysis for the identification of pathogenic variants in HB patients for the first time in Colombia. The pathogenic variants were identified in 100% of the studied cohort. Remarkably, 35% (7/20) of apparently non related patients carried the same intronic variant affecting splicing (g.22769A>G; c.520+13A>G), which suggests a possible founder effect. In addition, a new variant involving a large exon deletion (exons 3 and 4), was found.

ACKNOWLEDGMENTS

We wish to thank each of the patients and their families for agreeing to participate in this study; to the Comprehensive Hemophilia Treatment Center of Colsubsidio Children’s Clinic, to the Clinical Research Center of Colsubsidio; to the Institute of Human Genetics of the National University of Colombia where the processing and analysis of the samples was carried out; to the Sequencing Service of the Institute of Genetics of the National University of Colombia SSIGMOL; to the nurses Claudia Suarez and Yadira Valderrama for their contribution collecting the samples and their assistance in the evaluation of patient’s pedigrees. This project was financed entirely by the National University of Colombia HERMES Code 34708 directed by Juan J. Yunis L.

ORCID

Yolima A. Parrado Jara https://orcid.org/0000-0001-5666-0524

Luz K. Yunis Hazbun https://orcid.org/0000-0003-3365-5196

Adriana Linares https://orcid.org/0000-0001-7429-6850

Juan J. Yunis Londoño https://orcid.org/0000-0001-7688-9094

REFERENCES

Abla, Z., Mouloud, Y., Hejer, E. L. M., Emna, G., Meriem, A., Yamina, O., & Naouel, S. (2018). Mutations causing hemophilia B in Algeria: Identification of two novel mutations of the factor 9 gene. Biodiversitas, 19(1), 52–58. https://doi.org/10.13057/biodiv/d190109

Awidi, A., Alhattab, D., Bsoul, N., Magablah, A., Mefleh, R., Dweiri, M., & Fauori, A. S. (2011). FIX mutation spectrum in haemophilia B patients from Jordan: Identification of three novel mutations. Haemophilia, 17(1), 162–163. https://doi.org/10.1111/j.1365-2516.2010.02365.x

Belvini, D., Salviato, R., Radossi, P., Pierobon, F., Mori, P., Castaldo, G., & Tagariello, G. (2005). Molecular genotyping of the Italian cohort of patients with hemophilia B. Haemotologica, 90(5), 635–642.

Capriotti, E., Calabrese, R., & Casadio, R. (2006). Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics, 22(22), 2729–2734. https://doi.org/10.1093/bioinformatics/btl423

Chen, C. W., Lin, J., & Chu, Y. W. (2013). iStable: Off-the-shelf predictor integration for predicting protein stability changes. BMC Bioinformatics, 14(S5).https://doi.org/10.1186/1471-2105-14-S2-S5

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., … Taschner, P. E. (2016). HGVS recommendations for the description of sequence variants: 2016 Update. Human Mutation, 37, 564–569. https://doi.org/10.1002/humu.22981

Fondo colombiano de enfermedades de alto costo, C. de alto costo. (2018). Situación de la hemofilia en Colombia 2017. Bogotá Colombia. Retrieved from www.cuentadealcosto.org

Goodove, A. C. (2015). Hemophilia B: Molecular pathogenesis and mutation analysis. Journal of Thrombosis and Haemostasis, 13(7), 1184–1195. https://doi.org/10.1111/jth.12958

Khan, M. T. M., Naz, A., Ahmed, J., Shamsi, T., Ahmed, S., … Taj, A. S. (2018). Mutation spectrum and genotype – Phenotype analyses in a Pakistani cohort with hemophilia B. Clinical and Applied Thrombosis/Hemostasis, 24, 741–748. https://doi.org/10.1177/1076029617721011
Li, T., Miller, C. H., Driggers, J., Payne, A. B., Ellingsen, D., & Hooper, W. C. (2014). Mutation analysis of a cohort of US patients with hemophilia B. *American Journal of Hematology, 89*(4), 375–379. https://doi.org/10.1002/ajh.23645

Li, T., Miller, C. H., Payne, A. B., & Craig Hooper, W. (2013). The CDC hemophilia B mutation project mutation list: A new online resource. *Molecular Genetics & Genomic Medicine, 1*(4), 238–245. https://doi.org/10.1002/mgg3.30

Mårtensson, A., Letelier, A., Halldén, C., & Ljung, R. (2016). Mutation analysis of Swedish haemophilia B families - High frequency of unique mutations. *Haemophilia, 22*(3), 440–445. https://doi.org/10.1111/hae.12854

Meireles, M. R., Pantoja, A. G., Ornaghi, A. P. M., Vieira, G. F., Salzano, F. M., & Bandinelli, E. (2017). Molecular characterization of haemophilia B patients in southern Brazil. *Haemophilia, 5*, 457–461. https://doi.org/10.1111/hae.13277

Radic, C. P. (2010). *Genética molecular de hemofilia: caracterización de mutaciones en hemofilia B, expresión de hemofilia en mujeres y desarrollo de nuevos métodos de análisis de inversiones*. Tesis Doctoral, Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales. Retrieved from http://hdl.handle.net/20.500.12110/tesis_n4794_Radic

Radic, C. P., Rossetti, L. C., Abelleyro, M. M., Candela, M., Bianco, R. P., Pinto, M. D. T., ... Brasi, C. D. D. (2013). Assessment of the F9 genotype-specific FIX inhibitor risks and characterisation of 10 novel severe F9 defects in the first molecular series of Argentinian patients with haemophilia B. *Thrombosis and Haemostasis, 109*(1), 24–33. https://doi.org/10.1160/TH12-05-0302

Rallapalli, P. M., Kemball-Cook, G., Tuddenham, E. G., Gomez, K., & Perkins, S. J. (2013). An interactive mutation database for human coagulation factor IX provides novel insights into the phenotypes and genetics of hemophilia B. *Journal of Thrombosis and Haemostasis, 11*(7), 1329–1340. https://doi.org/10.1111/jth.12276

Ramensky, V. (2002). Human non-synonymous SNPs: Server and survey. *Nucleic Acids Research, 30*(17), 3894–3900. https://doi.org/10.1093/nar/gkf493

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in genie. *Journal of Computational Biology, 4*(3), 311–323. https://doi.org/10.1089/cmb.1997.4.311

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nature Methods, 11*(4), 361–362. https://doi.org/10.1038/nmeth.2890

Srivastava, A., Brewer, A. K., Mauser-Bunschoten, E. P., Key, N. S., Kitchen, S., Llinas, A., & Street, A. (2013). Guidelines for the management of hemophilia. *Haemophilia, 19*(1), e1–e47. https://doi.org/10.2217/fmeb2013.13.148

Tang, H., & Thomas, P. D. (2016). PANTHER-PSEP: Predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics, 32*(14), 2230–2232. https://doi.org/10.1093/bioinformatics/btw222

Tang, R., Prosser, D. O., & Love, D. R. (2016). Evaluation of bioinformatic programmes for the analysis of variants within splice site consensus regions. *Advances in Bioinformatics, 2016*, 5614058.

Thompson, A. (1986). Structure, function, and molecular defects of factor IX. *Blood, 67*(3), 565–572. https://doi.org/10.1182/blood.V67.3.565.565

Vielhaber, E., Jacobson, D. P., Ketterling, R. P., Liu, J. Z., & Sommer, S. S. (1993). A mutation in the 3′ untranslated region of the factor IX gene in four families with hemophilia B. *Human Molecular Genetics, 2*(8), 1309–1310. https://doi.org/10.2217/fmeb2013.13.148

---

**How to cite this article:** Parrado Jara YA, Yunis Hazbun LK, Linares A, Yunis Londoño JJ. Molecular characterization of hemophilia B patients in Colombia. *Mol Genet Genomic Med*. 2020;8:e1210. [https://doi.org/10.1002/mgg3.1210](https://doi.org/10.1002/mgg3.1210)