haematologica
Journal of the Ferrata Storti Foundation

Genomic ancestry, F8 variants, and immune tolerance in hemophilia A patients with inhibitors: exome sequencing insights

by Hanaisa Sant'Anna, Rafael Tou, Lucas Faria-Costa, Julia Duarte, Bruno Miwa, Renan Pedra de Souza, Ricardo Mesquita Camelo, Daniel Gonçalves Chaves, Claudia Santos Lorenzato, Tânia Hissa Anegawa, Andrea Gonçalves de Oliveira, Clarissa Barros Ferreira, Luany Elvira Mesquita Carvalho, Vivian Karla Brognoli Franco, Monica Hermida Cerqueira, Maria do Rosário Ferraz Roberti, Fabia Michelle Rodrigues de Araujo Callado, Leina Yukari Etto, Maria Aline Ferreira de Cerqueira, Ieda Solange de Souza Pinto, Andrea Aparecida Garcia, Doralice Marvulle Tan, Daniele Campos Fontes Neves, Maíse Moreira Dias, Luciana Werneck Zuccherato, Eduardo Tarazona-Santos and Suely Meireles Rezende

Received: December 20, 2024.
Accepted: April 30, 2025.

Citation: Hanaisa Sant'Anna, Rafael Tou, Lucas Faria-Costa, Julia Duarte, Bruno Miwa, Renan Pedra de Souza, Ricardo Mesquita Camelo, Daniel Gonçalves Chaves, Claudia Santos Lorenzato, Tânia Hissa Anegawa, Andrea Gonçalves de Oliveira, Clarissa Barros Ferreira, Luany Elvira Mesquita Carvalho, Vivian Karla Brognoli Franco, Monica Hermida Cerqueira, Maria do Rosário Ferraz Roberti, Fabia Michelle Rodrigues de Araujo Callado, Leina Yukari Etto, Maria Aline Ferreira de Cerqueira, Ieda Solange de Souza Pinto, Andrea Aparecida Garcia, Doralice Marvulle Tan, Daniele Campos Fontes Neves, Maíse Moreira Dias, Luciana Werneck Zuccherato, Eduardo Tarazona-Santos and Suely Meireles Rezende. Genomic ancestry, F8 variants, and immune tolerance in hemophilia A patients with inhibitors: exome sequencing insights.
Haematologica. 2025 May 15. doi: 10.3324/haematol.2024.287232 [Epub ahead of print]

Genomic ancestry, F8 variants, and immune tolerance in hemophilia A patients with inhibitors: exome sequencing insights

SHORT RUNNING TITLE: Genomic Ancestry in Hemophilia A

Hanaisa Sant'Anna,[1]* Rafael Tou,[1]* Lucas Faria-Costa,[1] Julia Duarte,[1] Bruno Miwa,[1] Renan Pedra de Souza,[1] Ricardo Mesquita Camelo,[2] Daniel Gonçalves Chaves,[3] Claudia Santos Lorenzato,[4] Tânia Hissa Anegawa,[5] Andrea Gonçalves de Oliveira,[3] Clarissa Barros Ferreira,[6] Luany Elvira Mesquita Carvalho[7], Vivian Karla Brognoli Franco,[8] Monica Hermida Cerqueira,[9] Maria do Rosário Ferraz Roberti,[10,11] Fabia Michelle Rodrigues de Araujo Callado,[12] Leina Yukari Etto,[13,14] Maria Aline Ferreira de Cerqueira,[15] Ieda Solange de Souza Pinto,[16,17] Andrea Aparecida Garcia,[18,19] Doralice Marvulle Tan,[20] Daniele Campos Fontes Neves,[21,22] Maíse Moreira Dias,[2] Luciana Werneck Zuccherato,[1,2] Eduardo Tarazona-Santos,[1,&] Suely Meireles Rezende,[2, &]

1. Department of Genetics, Ecology and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais, 31270-901, Belo Horizonte, Minas Gerais, Brazil

2. Department of Internal Medicine, Faculty of Medicine, Federal University of Minas Gerais, 30130-100, Belo Horizonte, Minas Gerais, Brazil

3. HEMOMINAS, Hemocentro de Belo Horizonte, 31630-903, Belo Horizonte, Minas Gerais, Brazil

4. HEMEPAR - Paraná Hemocenter, 80045-145, Curitiba, Paraná, Brazil

5. HEMEPAR - Regional Hemocenter of Londrina, 86038-670, Londrina, Paraná, Brazil

6. HEMORGS, Hemocenter of the State of Rio Grande do Sul, 90650-001, Porto Alegre, Rio Grande do Sul, Brazil

7. Ceará Hematology and Hemotherapy Center, 60440-261, Fortaleza, Ceará, Brazil

8. HEMOSC - Santa Catarina Hematology and Hemotherapy Center, 88015-240, Florianópolis, Santa Catarina, Brazil

9. HEMORIO, 20211-030, Rio de Janeiro, Rio de Janeiro, Brazil

10. Faculty of Medicine, Federal University of Goiás, 74605-050, Goiânia, Goiás, Brazil

11. HEMOGO, Hemocenter of Goiás, 74535-010, Goiânia, Goiás, Brazil

12. HEMOPE, 52011-000, Recife, Pernambuco, Brazil

13. Department of Internal Medicine, Medical Sciences Center, Federal University of Paraíba, 58051-900, João Pessoa, Paraíba, Brazil

14. Hemocenter of Paraíba, 58013-420, João Pessoa, Paraíba, Brazil

15. Hematology and Hemotherapy Center of Piauí, 64001-430, Teresina, Piauí, Brazil

16. João de Barros Barreto University Hospital, Federal University of Pará, 66073-000, Belém, Pará, Brazil

17. Department of Clinical Hematology, Hemotherapy and Hematology Center of Pará, 66033-000, Belém, Pará, Brazil

18. Regional Faculty of Medicine Foundation of São José do Rio Preto, 15090-000, São José do Rio Preto, São Paulo, Brazil

19. Hemocenter of São José do Rio Preto, 15091-240, São José do Rio Preto, São Paulo, Brazil

20. Hemocenter of Marília, 17519-050, Marília, São Paulo, Brazil

21. Department of Medicine, Federal University of Rondônia, 76801-974, Porto Velho, Rondônia, Brazil

22. Hemocenter of Rondônia, 78905-160, Porto Velho, Rondônia, Brazil

* These authors equally contributed to the present article.

& These authors equally contributed to the present article.

Correspondence: Professor Suely Meireles Rezende, Faculty of Medicine, Universidade Federal de Minas Gerais. Avenida Professor Alfredo Balena, 190, room 255, ZIP 30130-100, Belo Horizonte, Minas Gerais, Brazil. Email: srezende@ufmg.br; suely.rezende@uol.com.br

Data sharing statement: The data that support the findings of this study are available from the corresponding author upon request (srezende@ufmg.br; suely.rezende@uol.com.br).

Disclosures: Hanaisa Sant'Anna, Rafael Tou, Lucas Faria-Costa, Julia Duarte, Bruno Miwa, Renan Pedra de Souza, Ricardo Mesquita Camelo, Daniel Gonçalves Chaves, Claudia Santos Lorenzato, Tânia Hissa Anegawa, Andrea Gonçalves de Oliveira, Clarissa Barros Ferreira, Luany Elvira Mesquita Carvalho, Vivian Karla Brognoli Franco, Monica Hermida Cerqueira, Maria do Rosário Ferraz Roberti, Fabia Michelle Rodrigues de Araujo Callado, Leina Yukari Etto, Maria Aline Ferreira de Cerqueira, Ieda Solange de Souza Pinto, Andrea Aparecida Garcia, Doralice Marvulle Tan, Daniele Campos Fontes Neves, Maíse Moreira Dias, Luciana Werneck Zuccherato, Eduardo Tarazona-Santos, and Suely Meireles Rezende declare no conflicts of interest.

Contributions: S.M.R. conceived the study. S.M.R. and R.M.C. supervised the study and provided funding. H.S. and R.T., E.T.-S and S.M.R. designed the experiments/analyses and wrote the manuscript. H.S. R.T., L.F., B.M., L.W.Z. conducted analyses. R.P.S. provided statistical input. All other authors coordinated the study in each of the centers, contributed to the acquisition of data, and contributed to the critical reading of the manuscript. All authors approved the final version of the manuscript.

Acknowledgments: We thank the participants of BrazIT, their guardians, and the clinical teams from the Hemophilia Treatment Centers. We are also grateful to Eduardo Paiva for the insightful discussions on Brazil's demographic history and to Marcos Nunes for his support with the graphical work. We gratefully acknowledge the funding agencies that supported this study.

Funding: This work was supported by the Brazilian National Health Fund (17217.9850001-15-006), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, RED-00089-23, APQ-04228-24, PPM-00366-18), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 407046/2023-2, 406913/2022-6, 420008/2018-7, 440238/2022-6), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, 88887.939455/2024-00).

Letter

Hemophilia A is an X-linked recessive bleeding disorder caused by factor VIII gene (*F8*) mutations, affecting 1 in 5,000-10,000 male births globally. The primary treatment is FVIII protein replacement, but ~30% of patients develop FVIII-neutralizing inhibitors that reduce treatment efficacy[1]. Immune tolerance induction (ITI), effective in about two-thirds[2], is the only eradication treatment. While *F8* genotype and inhibitor titers are established ITI outcome predictors[3], the role of genomic ancestry remains underexplored, particularly in admixed populations such as Latin-Americans and North American-Hispanics.

We investigated the influence of autosomal and X-chromosome genetic ancestry on ITI outcomes in 193 Brazilian patients with hemophilia A and high-responding inhibitors, enrolled in the multicenter Brazilian Immune Tolerance Study (BrazIT). Participants were treated with low-dose ITI (50 IU/kg FVIII, 3x/week) using either recombinant (ADVATE, Takeda, Japan) or plasma-derived (mostly Octavi, Octapharma, United States) FVIII concentrates. ITI outcomes were classified as complete success, partial success, or failure, following the Hay and DiMichele criteria[4]. Clinical and genetic methodologies are detailed in Camelo et al.[2,5] and Zuccherato et al.[6]. Statistical approaches for population structure and genetic ancestry inference are provided in Table S1. The study was approved by institutional ethical committees (CAAE 52812415.8.0000.5149).

Demographic, clinical, and genetic characteristics of the BrazIT cohort are summarized in Table 1 and Figure 1. Median age at diagnosis was 0.9 years; 95%

had severe hemophilia. ITI outcomes: 64 (33%) complete success, 62 (32%) partial success, 67 (35%) failure. Exome sequencing showed high-quality data across ~34 Mb of targeted regions, with a median coverage of 123.66X and a transition/transversion ratio (Ts/Tv ratio, i.e., a measure of sequencing data quality) of 3.11, indicative of reliable variant calls. We identified 222,314 variants, including 91,599 singletons and 7,088 novel variants, reflecting genetic diversity (Table 2). *F8* haplotypes[7] frequencies, defined by R503H (rs35383156), D1260E (rs1800291) and M2257V (rs1800297), were: H1 (N=133, 68.5%), H2 (N=37, 19.2%), H3 (N=10, 5.2%), H4 (N=1), H5 (N=0), and missing haplotypes (N=13, 6.7%).

We confirmed minimal inbreeding (mean individual coefficient: 0.004, 95% CI= [-0.028,0.044]) and relatedness (mean kinship coefficient: 0.00333, 95% CI=[0.000132,0.0092]). Still, we confirmed five cases of known relatedness between pairs of individuals sharing the same types of *F8* mutation and an unreported second-degree relatedness.

Population genetics analyses confirmed the BrazIT cohort is admixed, reflecting Brazil's post-Columbian ancestry, predominantly European (64%), African (17%), and Native American (13%; Table 1, Figure 1). Ancestry varied between individuals, Brazilian states, and mainly between Brazilian geographic regions (nested-ANOVA for autosomes and X-chromosome, F values always > 4.065, p always <0.003, Table S2). European ancestry was highest in the South (73% autosomal), Native American in the Northeast (20% autosomal), and African in the Southeast (30% autosomal).

In Latin-American populations, sex-ancestry bias is the rule: during the last five centuries, European admixture has been preferentially mediated by males, while African and Native-American admixture by females, producing higher European ancestries in autosomes than in X-chromosomes. As expected, BrazIT shows Native-American sex-ancestry bias (Autosome - X-chromosome bias: 0.06, $p<0.001$), but exceptionally, no African sex-ancestry bias (-0.005, not-significant). BrazIT shows the lowest value of African sex-ancestry bias among 21 studied Latin American populations[8,9,10]. This result suggests that X-chromosomes of predominant African origin are underrepresented among hemophilia A patients with high-responding inhibitors. However, the composition of the BrazIT cohort does not allow us to extend this result to hemophilia A patients in general.

Different scenarios, possibly concurrent, may explain this result. First, there may be genetic factors mapped in African X-chromosomes negatively associated with being a hemophilia A patient with high-responding inhibitor (or a hemophilia A patient in general), partially preventing the inclusion of X-chromosomes of African origin in the BrazIT cohort and eliminating the pervasive population-based sex-African ancestry bias. Although *F8*, located on the X-chromosome, is an obvious candidate, exome data do not have enough Single Nucleotide Variants (SNVs, i.e., unique changes in the DNA sequence) density to perform an admixture mapping to identify the causing locus. Second, the lack of African sex-ancestry bias in the BrazIT cohort may result from the X-linked inheritance of hemophilia A (with 30-60% of *de novo* mutations[11]), interacting with the demographic history of Brazilians

during the last 500 years, characterized by intensive immigration of Europeans and Africans. Indeed, all the DNA fragments of European and African origins of the Brazilian genomes were introduced by these immigrants, with the Native-Americans being the only population present in the Americas before 1492. Historical demography suggests that the male/female ratio among immigrants to the Americas was higher for Africans than for Europeans, and a ratio of 2-3 males for each female from Africa is broadly accepted (https://www.slavevoyages.org/). Because affected males were likely not included in the slave trade (or maybe, patients with severe disease have died before reaching adulthood), few X-chromosomes of African origins carrying hemophilia A mutations were brought to the Americas, mostly limited to the small proportion of carrier female immigrants. Therefore, this may have contributed to the absence of the African sex-ancestry bias in a hemophilia A cohort such as BrazIT.

F8 large deletions are prevalent in about 3-5% of individuals with severe hemophilia A[12]. We observed that *F8* large deletions are less frequent among individuals with higher autosomal (not for X-chromosome) Native-American ancestry, even after adjusting for inhibitor titers, geographical regions, and kinship coefficients (β=-0.081, 95% CI -0.144,-0.018; p=0.011). We also observed an association between H2 and H3 haplotypes and African ancestry, even after adjusting for peaks of inhibitor titers (X-chromosome: β=2.95-2.96, p always <0.001; Autosomes: β=3.02-3.66, p always <0.04), consistent with observations in US racial groups[3].

Several studies have investigated hemophilia-related traits using self-reported ethnoracial categories [13,14] (Table S3); however, genomic ancestry has not yet been explored in this context. While ethnoracial classification (a categorical social construct) and genomic ancestry (a continuous biological variable) are conceptually distinct, they are statistically correlated in the BrazIT cohort. Specifically, we observed an adjusted R^2 of 0.32 between European genomic ancestry and the White ethnoracial group, and an adjusted R^2 of 0.39 between African genomic ancestry and the Black ethnoracial group (both $p<0.001$). Given the lack of comparable genomic ancestry-based studies in Brazilian studies on hemophilia-related traits, we contextualize our findings with US studies that use ethnoracial classification and benefit from larger sample sizes, enhancing statistical power. It is also important to note that African Americans tend to have a higher proportion of African ancestry compared to admixed Brazilian individuals[8].

Unlike prior studies using self-reported race/ethnicity, we inferred genomic ancestry from exome data. We found no association between autosomal or X-chromosome ancestry (including African ancestry) and ITI outcomes, aligning with prior studies using self-identified race/ethnicity[13,14] (Table S3). Similarly, no association was observed between inhibitor peak titers and genomic ancestry or *F8* haplotypes (p always > 0.262). Genomic ancestry explained part of the variability in ethnoracial classification (adjusted-R^2: 0.32 for White, 0.39 for Black; $p<0.001$), but

replacing genomic ancestry with reported race (White, Brown, Black) confirmed the lack of association with *F8* large deletions, inhibitor titers, or ITI outcomes.

The relationship between reported race/ethnicity and genomic ancestry is complex, sensitive, and controversial. It varies in different geographic and social contexts, particularly in countries with a history of social inequalities and admixture, such as Brazil. A caveat in testing the association between health-related outcomes and reported racial categories or genomic ancestry is that in the US and Latin-America, Native-American and African ancestries (and the associated ethnoracial classes) are associated with poorer socio-economic conditions, a potential confounder. Thus, socioeconomic conditions should be included as covariates when testing the association between phenotypes and racial categories/ancestry. A limitation of BrazIT is that we do not have a variable for socioeconomic conditions.

Our study has several strengths. First, this is the largest study of Latin American individuals with hemophilia A and high-responding inhibitors treated with ITI to date, and the first study to perform exome sequencing of this population. The study is a well-characterized cohort that assessed relevant clinical, immunological, and genetic factors with little missing data. All BrazIT participants were followed up until the end of ITI, which was performed according to a national standardized ITI protocol.

In conclusion, exome sequencing of BrazIT contributes to the need for genomic data on patients with rare diseases and more diverse ancestries. Our data, harmonized and integrated with other cohorts from different ancestries to be studied, will allow us to gain statistical power and identify genetic variants and genes associated with hemophilia-related outcomes such as inhibitor titers and ITI success. We did not observe an association between genomic ancestries and inhibitor titers or ITI response. Still, Native American ancestry was negatively associated with *F8* large deletion, and H2-H3 haplotypes were associated with African ancestry. An intriguing result is that X-chromosomes of predominant African origins are underrepresented among Brazilian hemophilia A patients with high-responding inhibitors, exemplifying the unexplored issue of how sex-ancestry bias in post-Columbian migrations from Europe and Africa to the Americas may have differently shaped the patterns of genetic diversity of X-chromosome and autosomal Mendelian diseases in different populations of the Americas.

REFERENCES

1. Srivastava A, Santagostino E, Dougall A, et al. WFH Guidelines for the Management of Hemophilia, 3rd edition. Haemophilia. 2020;26 Suppl 6:1-158.

2. Camelo RM, Chaves DG, Zuccherato LW, Rezende SM, BrazIT Study Team. Predictors of the outcome of immune tolerance induction in patients with haemophilia A and inhibitors: The Brazilian Immune Tolerance (BrazIT) Study protocol. PLoS One. 2021;16(8):e0256265.

3. Pratt KP, Gunasekera D, Vir P, et al. Anti-FVIII antibodies in Black and White hemophilia A subjects: do F8 haplotypes play a role? Blood Adv. 2023;7(17):4983-4998.

4. Hay CR, DiMichele DM; International Immune Tolerance Study. The principal results of the International Immune Tolerance Study: a randomized dose comparison. Blood. 2012;119(6):1335-1344.

5. Camelo RM, Dias MM, Caram-Deelder C, et al. Time between inhibitor detection and start of immune tolerance induction: Association with outcome in the BrazIT study. J Thromb Haemost. 2022;20(11):2526-2537.

6. Zuccherato LW, Souza RP, Camelo RM, et al. Large deletions and small insertions and deletions in the factor VIII gene predict unfavorable immune tolerance induction outcome in people with severe hemophilia A and high-responding inhibitors. Thromb Res. 2024;242:109115.

7. Lambert C, Lannoy N, Meité N, Sanogo I, Eeckhoudt S, Hermans C. Inhibitor epidemiology and genetic□related risk factors in people with haemophilia from Côte d'Ivoire. Haemophilia. 2020;26(1):79-85.

8. Kehdy FSG, Gouveia MH, Machado M, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. Proc Natl Acad Sci U S A. 2015;112(28):8696-8701.

9. Fortes-Lima C, Gessain A, Ruiz-Linares A, et al. Genome-wide ancestry and demographic history of African-descendant maroon communities from French Guiana and Suriname. Am J Hum Genet. 2017;101(5):725-736.

10. Ongaro L, Molinaro L, Flores R, et al. Evaluating the impact of sex-biased genetic admixture in the Americas through the analysis of haplotype data. Genes (Basel). 2021;12(10):1580.

11. Santana MAP, Chaves DG, Camelo RM, et al. Prevalence of sporadic haemophilia A. Haemophilia. 2023;29(2):668-670.

12. Jourdy Y, Chatron N, Fretigny M, Dericquebourg A, Sanlaville D, Vinciguerra C. Comprehensive analysis of F8 large deletions: Characterization of full breakpoint junctions and description of a possible DNA breakage hotspot in intron 6. J Thromb Haemost. 2022;20(10):2293-2305.

13. Kempton CL, Payne AB, Fedewa SA. Race, ethnicity, and immune tolerance induction in hemophilia A in the United States. Res Pract Thromb Haemost. 2023;7(8):102251.

14. Fedewa SA, Kempton CL. Race and ethnicity and the success of immune tolerance induction among people with severe haemophilia A in the United States. Haemophilia. 2024;30(3):628-637.

Table 1. Demographic, clinical, and genetic characteristics of the BrazIT cohort.

A. General Characteristics of the BrazIT Cohort (n = 193)

Male, n (%)		192 (99%)
Age at diagnosis of hemophilia A in years, median (IQR)		0.90 (0.55,1.44)
	Missing Data, n (%)	10 (5%)
Ethnoracial classification, n (%)	Black	15 (8%)
	Mixed[1]	73 (38%)
	White	105 (54%)
	Yellow (Asian)	0 (0%)
	Indigenous (Native-American)	0 (0%)
Continental genomic ancestry by autosomes, mean (IQR)	European	0.65 (0.53,0.77)
	African	0.21 (0.09,0.29)
	Native-American	0.14 (0.08,0.19)
Continental genomic ancestry of X-chromosomes, mean (IQR)	European	0.59 (0.37,0.82)
	African	0.21 (0.03,0.30)
	Native-American	0.20 (0.02,0.33)
Continental sex-ancestry bias (mean in autosomes - mean in X-chromosomes)	European	0.065
	African	- 0.005
	Native-American	-0.060
Distribution of BrazIT patients by Brazilian geographic region	West-Central	13 (7%)
	Northeast	47 (24%)
	North	8 (4%)
	Southeast	53 (27%)
	South	72 (37%)
Hemophilia A severity, n (%)[2]	Severe	173 (90%)
	Moderately-severe[3]	8 (4%)
	Moderate[4]	4 (2%)
	Missing Data, n (%)	8 (4%)
***F8* variant type, n (%)**	Large Deletion	18 (9%)
	Frameshift	26 (13%)
	Inversions	
	Inversion intron 22	85 (44%)
	Inversion intron 1	4 (2%)
	Nonsense	34 (18%)
	Missense	7 (4%)
	Splice donor	6 (3%)
	Missing Data, n (%)	13 (7%)
ITI outcomes, n (%)	Failure	67 (35%)
	Partial success	62 (32%)
	Complete success	64 (33%)

B. Inhibitor Parameters and ITI

Age at inhibitor diagnosis in years, median (IQR)	Median (IQR)	2.16 (1.28,6.68)
	Missing Data, n (%)	3 (1.55%)
Age at ITI start (years)	Median (IQR)	6.48 (2.23,18.81)
Interval between diagnosis of hemophilia A and inhibitor development (years)	Median (IQR)	0.87 (0.40,3.71)
	Missing Data, n (%)	13 (7%)
Historical inhibitor peak (BU/mL)	Median (IQR)	42.00 (14.80,119.40)
Inhibitor titer immediately before ITI starts (BU/mL)	Median (IQR)	7.20 (3.00,14.40)
Inhibitor peak during ITI (BU/mL)	Median (IQR)	30.00 (6.00,140.00)
ITI duration (years)	Median (IQR)	2.55 (1.67,3.13)

Note: BrazIT, Brazilian Immune Tolerance Study; F8, factor VIII; IQR, interquartile range; ITI, immune tolerance induction; IU, International Units. [1] The "Mixed" category corresponds to "pardo" in official Portuguese terminology. [2] The lowest plasma FVIII activity before inhibitor development was considered to classify hemophilia severity. [3] Defined when FVIII levels are between 0.01-0.02 IU/mL; [4] Defined when FVIII levels are between >0.02 and 0.05 IU/mL.

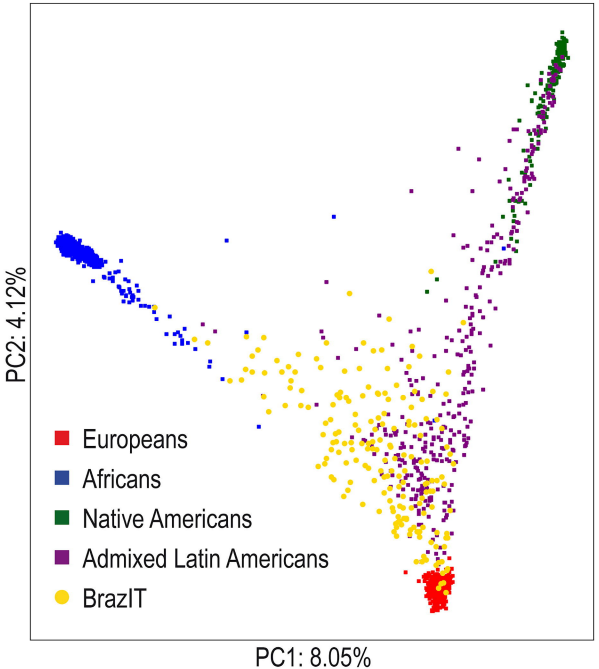
Table 2. Exome diversity of the target region (34,156,490 bp) of the BrazIT cohort.

SNVs	INDELs	Variants not present in dbSNP	Singletons	Doubletons	Transition/ Transversion ratio (Ts/Tv)	Total (SNV and INDEL)
216,531	5,783	7,088	93,657	25,761	3.110	222,314

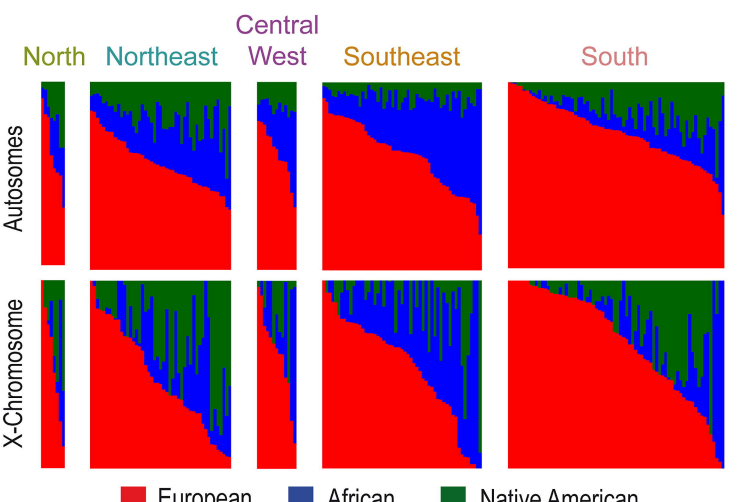
Note: SNV, single nucleotide variants; INDELS, insertions or deletions of nucleotides of <50bp in DNA; Singletons, variants present at an absolute allele frequency of 1; Doubletons, variants present at an absolute allele frequency of 2; The transition/transversion ratio (Ts/Tv) is used to assess mutation patterns and genomic integrity in sequencing data. It reflects the number of base substitutions that are transitions (A↔G, C↔T) divided by those that are transversions (e.g., A↔C, A↔T, G↔C, G↔T). A Ts/Tv ratio of 3.110 in this cohort indicates a predominance of transitions, consistent with a stable genomic profile and high-quality variant calls.

Figure 1. Geographic Distribution, Genetic Structure, and Admixture Patterns Across Brazilian Regions of the BrazIT Cohort. (A) The first and second Principal Components (PC1 and PC2, respectively) of autosomal genotypes distinguish individuals from European, African, Native American, and Asian parental populations. Individuals from the BrazIT cohort are scattered across the Principal Component Analysis (PCA) plot, indicating admixture. The numbers in parentheses next to PC1 and PC2 indicate the percentage of variance explained by each Principal Component. **(B)** Geographic distribution of BrazIT participants across the five Brazilian regions. Different colors represent the regions. States that include BrazIT individuals are shaded darker in each region. Below are vertical barplots depicting the individual proportions of continental ancestry in autosomes and X-chromosomes, estimated by the ADMIXTURE method. Each of the ten blocks of vertical bars (e.g., autosomal ancestry for the North region in the upper left) consists of thin, adjacent vertical bars. Each vertical bar represents the percentage of Native American (green, at the top), African (blue, in the middle), and European (red, below) ancestries for each individual. The width of each block of vertical bars is proportional to the number of individuals. **(C)** Sex-ancestry bias (mean autosomal ancestry - mean X-chromosome ancestry) is divided by region (Central-West, North, Northeast, Southeast, South) and by BrazIT as a total. We used 42,542 and 654 unlinked (linkage disequilibrium estimator $r^2 < 0.4$) Single Nucleotide Variants (SNVs) in the autosomes and X-chromosomes for PCA and ADMIXTURE analyses.

A



B



C

Sex-ancestry bias

	N	European	African	Native American
Central-West	13	0.011	0.001	-0.013
North	8	0.047	0.027	-0.073
Northeast	47	0.085	0.030	-0.115
Southeast	53	0.072	-0.021	-0.050
South	72	0.059	-0.022	-0.037
BrazIT	193	0.065	-0.005	-0.060

Table S1. Overview of Statistical Analyses for Population Structure and Genetic Ancestry.

Analyses	Statistical Analyses and Software	Observation
Autosomal population structure and continental ancestry. We included 2,199 reference individuals for autosomes	We used 42,542 unlinked ($r^2 < 0.4$) autosomal SNVs derived from exome sequencing, and inferred population structure and genomic ancestry using Principal Component Analysis (Patterson et al. 2006), and ADMIXTURE (Alexander et al. 2009, with K=3 European, African, and Native-American parental populations)	-
Ancestry of the X-chromosome in the 192 males of BrazIT	We used ADMIXTURE (K=3) and 654 unlinked ($r^2 < 0.4$) non-pseudoautosomal SNVs on the X-chromosome (Table S2, Supplemental Section 12 for a discussion). We used as reference 1,555 males for the X-chromosome, representing European, African, and Native-American populations (Table S3)	-
Nested analysis of variance (nested-ANOVA) to estimate the apportionment of European, African, and Native-American ancestries among geographic regions and states (within those regions), including kinship coefficients as a covariate	'aov' function in R	We could not quantify the distribution of ancestry among individuals within states due to insufficient degrees of freedom

Inbreeding coefficient estimation for each individual	We estimated inbreeding coefficients for each individual using VCFtools (Danecek et al. 2011)	We estimated as F_{ind} *= (O-E)/(L-E)*, where, for each biallelic locus, *O* is the observed number of loci in homozygosity in an individual, $E = \sum 1 - 2p_i q_i$ is the expected number of loci in homozygosity based on allele frequencies p_i and q_i, and *L* is the number of valid loci for that individual. Considering the admixture of the BrazIT cohort, we introduced a novelty in the estimation of the individual inbreeding coefficients F_{ind}. For a homogeneous population, $2p_i q_i$ is expected to be the same for all individuals. However, populations that are a product of admixture between populations with different levels of diversity $2p_i q_i$ are expected to be higher for those individuals with higher ancestry of the more diverse population (i.e., Africans, Campbell et al. 2014). In Brazilians, using the same $2p_i q_i$ values for all individuals underestimates F_{ind} for individuals with more African ancestry and overestimates F_{ind} for individuals with less African ancestry. To avoid this artifact, we stratified the estimation of F_{ind} by considering six non-overlapping bins of African ancestry of 32-33 individuals each
Association between genetic ancestry (proportions of European, African, and Native-American ancestries) and F8 variant types: large deletions, frameshifts, inversions, nonsense, missense, and splice donor mutations	Generalized linear regression, 'glm' function in R	Covariates: the kinship matrix, Brazilian geographic regions as categorical variables, and historical peak inhibitor titer
Association between genetic ancestry (proportions of European, African, and Native-American ancestries) and inhibitor titer, considering separately the highest inhibitor titer before ITI (historical peak), immediately before ITI starts, and the highest inhibitor titer during ITI	Generalized linear regression, 'glm' function in R	Covariates: kinship matrix, Brazilian geographic regions, and F8 variant types as categorical variables, including large deletions, frameshifts, inversions, nonsense, missense, and splice donor variants
Association between response to ITI (failure, partial success, and complete success) and genomic ancestry (European, African, and Native-American)	Ordinal logistic regression ('polr' function in R and MASS v. 7.3.51.6, Venables and Ripley 2002)	Covariates: the kinship matrix, Brazilian geographical regions, F8 variants (large deletions, frameshifts, inversions, nonsense, missense, and splice donor variants), and a historical inhibitor peak
Association between F8 haplotypes and genomic ancestry (proportions of European, African, and Native-American ancestries), using ancestry estimates from X-chromosome and autosomal variants	Binomial logistic regression model, 'glm' function in R	Covariates: the kinship matrix, Brazilian geographical regions, F8 variants (large deletions, frameshifts, inversions, nonsense, missense, and splice donor variants), inhibitor titer, considering separately the highest inhibitor titer before ITI (historical peak), immediately before ITI starts, and the highest inhibitor titer during ITI

Table S2. The proportion of individual continental ancestry between Brazilian geographic regions and states was estimated by nested-ANOVA (n=193).

Ancestry	Genetic Marker	Source	Degrees of freedom	Sum of squares	Mean squares	F value	P-value
European	**Autosomal**	Between regions	4	0.889	0.22231	9.692	4.03e-07
		Between states within regions	8	0.463	0.05781	2.520	0.013
		Residuals	179	4.129	0.02294	-	-
	X-Chromosome	Between regions	4	1.161	0.29025	4.065	0.003
		Between states within regions	8	1.311	0.16385	2.295	0.023
		Residuals	179	12.781	0.07140	-	-
African	**Autosomal**	Between regions	4	1.213	0.30318	18.856	5.77e-13
		Between states within regions	8	0.159	0.0199	1.238	0.279
		Residuals	179	2.894	0.01608	-	-
	X-Chromosome	Between regions	4	1.163	0.29065	5.713	2.38e-04
		Between states within regions	8	0.240	0.03006	0.591	0.785
		Residuals	179	9.106	0.05087	-	-
Native-American	**Autosomal**	Between regions	4	0.363	0.09072	15.539	6.14e-11
		Between states within regions	8	0.271	0.03385	5.798	1.39e-06
		Residuals	179	1.051	0.00584	-	-
	X-Chromosome	Between regions	4	0.899	0.22479	5.616	2.79e-04
		Between states within regions	8	0.836	0.10456	2.612	0.010
		Residuals	179	7.165	0.04003	-	-

Table S3. Cited studies on Hemophilia A with racial/ethnic classification and association with hemophilia-related traits.

Study	Sample Size	Population Description	Main Results	Data Type	Genomic Ancestry Analysis	Association Tested	Association Result
Kempton et al. 2023 (Cross-Sectional Study)	614	Severe hemophilia A patients in the U.S., focusing on ITI practices and outcomes with an analysis of racial/ethnic disparities.	No significant racial disparities in ITI outcomes were found.	Clinical	No	Race/Ethnicity vs. ITI outcome	No significant association found
Fedewa and Kempton 2024 (Observational Study)	559	Hemophilia A patients (White, Black, Hispanic, Asian) analyzing ITI success rates.	Found comparable ITI success rates across racial/ethnic groups, contradicting the hypothesis that race influences ITI response.	Clinical	No	Race/Ethnicity vs. ITI outcomes	No significant association found
Sant'Anna et al. 2024 BrazIT Cohort (Cross-Sectional Study)	193	Brazil cohort of hemophilia A patients with inhibitor history undergoing ITI, analyzing genomic ancestry for its association with *F8* mutations and inhibitors.	Demonstrated a link between genomic ancestry and *F8* mutation patterns.	Clinical & Genetic (whole exome sequencing)	Yes	Race/Ethnicity vs. ITI outcomes	No significant association found
						Race/Ethnicity vs. inhibitor titers	No significant association found
						Race/Ethnicity vs. *F8* large deletions	No significant association found
						Genomic ancestry vs. inhibitor titers	No significant association found
						Genomic ancestry vs. ITI outcomes	No significant association found
						Genomic ancestry vs. *F8* mutation type	Significant association found (Native-American ancestry and F8 Large Deletion: β=-0.081, 95% CI -0.144,-0.018; p=0.011)
						Genomic ancestry vs. *F8* haplotype	Significant association found (*F8* H2 and H3 haplotypes and African ancestry, X-chromosome: β=2.95-2.96, p always <0.001; Autosomes: β=3.02-3.66, p always <0.04)