ANDES PEDIATRICA
REVISTA CHILENA DE PEDIATRÍA
www.scielo.cl

Andes pediatr. 2024;95(1): 41-52
DOI: 10.32641/andespediatr.v95i1.4752

ORIGINAL ARTICLE

Population characterization of mutations for sickle cell anemia and its treatment: One step towards personalized medicine for the disease

Caracterización poblacional de mutaciones relevantes para la Anemia Falciforme y su tratamiento: Un paso hacia la personalización de la enfermedad

Bernardita Cayupe[a,c], **Rafael Barra**[a,b]

[a]Centro de Investigación Biomédica y Aplicada (CIBAP), Escuela de Medicina, Facultad de Ciencias Médicas, Universidad de Santiago de Chile (USACH). Santiago, Chile.
[b]Bioquímico.
[c]Bióloga.

Received: April 6, 2023; Approved: November 8, 2023

What do we know about the subject matter of this study?

It has been established in the literature the relevance of the specific haplotypes related to the development of Sickle cell disease. These haplotypes are more frequent in different ethnic groups. Additionally, a second group of relevant mutations has been defined for the prognosis, diagnosis, and pharmacological treatment of Sickle Cell Disease, although it is yet to be defined whether it will be differentially expressed in different ethnic groups.

What does this study contribute to what is already known?

Using public databases, we analyzed and compared the population distribution of mutation relevant to the diagnosis, treatment, and prognosis of Sickle cell disease. Using an automated logistic correlation model, we found that many of the mutations are present more frequently in some ethnic groups. This information is relevant because it enlightens and supports the importance of genomic characterization in the treatment of a genetic pathology like sickle cell Disease.

Abstract

Sickle cell anemia (SCA) is the most common genetic disease worldwide. There are countries with massive public health programs for early detection of this condition. In the literature, several specific haplotypes or single-base polymorphic variants (SNPs) have been associated with the SCA prognosis. **Objective:** To demonstrate the significant correlation of SNPs relevant to the diagnosis and prognosis of SCA among different ethnic groups. **Methodology:** we analyzed population frequencies and correlations of several SNPs related to the prognosis of SCA (i.e., baseline fetal hemoglobin levels), response to hydroxyurea treatment, and response to other drugs used in the SCA treatment, collected from validated genomic databases among different ethnic groups. **Results:** The calculation of the Hardy-Weinberg equilibrium and the logistic regression was successful in classifying the ethnic groups as African (0 = 0.78, 1 = 0.89), and with a lower efficiency as Ameri-

Keywords:
Sickle Cell Anemia;
Single Base Polymorphisms;
Fetal Hemoglobin;
Ethnic Groups;
Bioinformatics.

Correspondence:
Rafael Barra
rafael.barra@usach.cl

Edited by:
Luisa Schonhaut Berman

How to cite this article: Andes pediatr. 2024;95(1):41-52. DOI: 10.32641/andespediatr.v95i1.4752

EDITORIAL IKU

can (AMR) (0 = 0.88, 1 = 0.00), East Asian (EAS) (0 = 0.80, 1 = 0.00), European (EUR) (0 = 0.79, 1 = 0.00), and South Asian (SAS) (0 = 0.80, 1 = 0.00). **Conclusions:** The results extend those from previous reports and show that the profile of most of the SNPs studied presented statistically significant distributions among general ethnic groups, pointing to the need to carry out massive early screening of relevant SNPs for SCA in patients diagnosed with this disease. It is concluded that the application of a broad mutation detection program will lead to a more personalized and efficient response in the treatment of SCA.

Introducción

Currently, sickle cell disease (SCD), most usually referred to as sickle cell anemia (SCA), is a common genetic disease distributed worldwide[1]. SCA is ten times more prevalent in Africa (1125 per 100000 live births) than globally (112 per 100000 live births) and less common in Europe (43.12 per 100000 live births)[2]. Usually, in adult patients, the most common form of hemoglobin (HB) is HBA, composed of two alpha chains and two beta chains[1]. Sickle cell anemia occurs due to the substitution of a Valine residue for a Glutamic Acid in the seventh position of the beta-globin chain of HB, and it is due to a mutation by nucleotide substitution of an Adenine to Thymine in the hemoglobin gene. This mutation generates a form of HB-denominated HBS and polymerizes in the deoxygenated state. The HBS polymerization generates a morphological change in the appearance of the erythrocytes, acquiring a sickle shape. It is noteworthy that besides the mutation etiologically responsible for the SCA, there is the co-occurrence of other allelic combinations, generating five described haplotypes[3]. The haplotypes have been associated to different prognosis of SCA[4,5].

Additionally, within the hemodynamic alterations of SCA, vaso-occlusive phenomena occur, including organ damage and pain[3]. Reoxygenation of HB recovers the standard shape of the erythrocyte. However, this process harms erythrocytes and leads to their degradation, causing anemia. The SCA pathology presents a pattern of autosomal recessive inheritance model, and this configuration implies that every cell must have both copies mutated[3]. The parents of a patient affected with SCA, each one, must carry one copy of the mutated allele, but they do not show or present symptoms[3].

The highest SCA incidences are observed in regions characterized by endemic malaria infections. For this reason, parts of sub-Saharan Africa and Southeast Asia have shown higher incidences of SCA[3,4]. However, migration processes have modified the ethnic matrix and have generated a new challenge for health systems and providers.

Usually, SCD is not considered lethal during childhood; however, in adulthood, it can cause heart attack and kidney damage, among other serious diseases[3]. SCD presents a very heterogeneous phenotypic range and can have complications ranging from very mild up to even the death of pediatric patients[5,6]. Particularly regarding the treatment of SCD, the main objective is to increase Fetal Hemoglobin (FH) levels in the patient. Currently, Hydroxyurea is used as a first-line antineoplasic drug to increase FH levels[7]; however, the response to treatment with Hydroxyurea presents high variability among patients. Interestingly, the literature has highlighted the importance of early diagnosis and identification of mutations associated with the prognosis of SCD to prevent or reduce its harmful effects[8]. Then, the importance of personalizing the SCA treatment has been highlighted, including the possibility of neonatal screening[7]. Different point mutations (SNPs) have been identified and related to diagnosis, prognosis, and treatment response. Mainly, FH levels are critical to determine the severity of SCA[9,10]. Identification of SNPs is an effort to develop personalized medicine, specifically in SCA, where the importance of genetic modifiers of variability has been recently highlighted[11,12].

The current work analyzed databases for a collection of mutations identified as relevant for the personalized treatment of SCA. Then, using bioinformatic tools for data analysis, we can recognize the relevant SNPs as predictors of various populations or ethnic groups. The dynamic databases generated will be available on the website https://github.com/barralabusach and will allow us to characterize and guide the development of personalized treatments for SCA in different populations.

Objective

Demonstrate the significant correlation of SNPs relevant for the diagnosis and prognosis of SCA with different ethnic groups.

Methodology

General workflow

We have reviewed the literature and selected significant SNPs for both diagnostic and prognosis of SCA. For every selected SNP, we have downloaded the individual genomic information for every studied population from the 1000 genome database[13]. Using this information, a new database was generated where the alleles and genotype information were highlighted. Once the database was created, Hardy-Weinberg equilibria (HWE) were calculated, and a logistic regression model characterizing the SNPs profile was performed.

Selection and search of SNPs

For the search of the SNPs, these were selected according to descriptions in the literature, closely related to FH levels and the response to treatment with Hydroxyurea. The selected SNPs were consistently validated in a clinical trial, identifying the highest correlated SNP with the clinical condition[14]. The inclusion criteria used for this clinical trial were similar to those previously published in the literature: the presence of HBs or HBs-B0 (thalassemia), the age in the range of 5-21 years. The exclusion criteria were: pregnancy, current or recent painful crisis, fever or acute illness within three weeks before evaluation, transfusion within the prior 100 days or active transfusion therapy, abnormally elevated serum creatinine or liver transaminases. The siblings were excluded to ensure genetic independence[14]. Also, population frequencies of significant SNPs for the pharmacological treatment and characterization of the haplotypes in SCD (Supplementary tables 1, 2, and 3 will be available to download) were analyzed. The discrete categories (diagnosis, prognosis, and FH levels) were assigned as shown in tables 1 and 2. These SNPs were identified in the free database "1000 genomes". With this information, a new dynamic database was generated. After selecting the SNPs of interest, the database of 1000 genomes (1000 Genomes Project Phase 3) was accessed, available in the Ensembl database (https://www.ensembl.org/). The search for each SNP was performed, and the information corresponding to their observed allele frequencies was collected in 5 categories, already described by the1000 genomes project, which corresponds to human populations: (i) African population, (ii) American population, (iii) East Asian population, (iv) European population, and (v) the population of South Asia. We obtained two databases; both sets were pre-processed with the Python 3.8.5 programming language. The first database was based on the frequency of allele and genotype to get a data set containing 80 rows and 28 columns, summarized in Table 4 of supplementary materials (available to download on request). The second database, which incorporates individual observations, was constructed using the "get_dummies pandas library" to convert a categorical variable into "dummy/indicator" variables to apply a logistic regression model. The dataset generated contains 87597 rows × 43 columns (See supplementary table 5, available to download).

Hardy Weinberg equilibrium calculation

Using Python 3.8.5 pandas library and the methodology described in McKinney et al.[15], the HWE was calculated for each SNP and broken down by each population using the first dataset generated (table 1).

Analysis of SNPs between populations

To analyze the distribution of SNPs in different populations, frequency histograms were made between the categorical populations described. Logistic regression was applied to the second dataset generated using the linear model module from the sklearn library. The model was submitted to an iterative automatic classification cycle for every specific subset population. Randomly, the 70% of the database was used for the training of the model, and the remaining 30% for automated diagnosis between the binary categories for every population.

Results

I. Database generation and HWE calculations.

After calculating the HWE equilibria in the SNPs of the generated database, it was noticeable that 23 of the 37 analyzed SNPs were not in HWE. Notably, 4 SNPs are consistently not in HWE in 4 of 5 studied populations (rs11886868, rs2182008, rs380620, rs7599488), and 6 SNP are not in HWE in one or two populations (East Asian population), rs2387634 (East Asian population), rs334 (American Population, African Population), rs61743453 (American Population, African Population), rs7557939 (East Asian Population American Population), rs9319428 (African population). The chi-value and p-significance are presented in Table 1 for every SNP and population analyzed.

Regarding the basal levels of FH, it is observed that the SNPs rs10128556, rs10189857, rs4671393, and rs4895441 are not in HWE in all populations. The SNP rs11886868 is only in equilibrium in the European population and rs7599488 in the South Asia population; the SNP rs7557939 is not in HWE in the American and East Asian populations.

II. Analysis of population variation of SNPs

Regarding the SNPs described for Basal FH Levels, after carrying out the characterization of the populations between the SNPs, several differences are obser-

Table 1. Chi-square values and p significance in every SNPs and population analyzed

SNP	Population	Description	chi_value	p_value
rs10128556	SAS	Basal Levels of FH	354,4552	1,07E-77
rs10128556	AMR	Basal Levels of FH	246,7468	2,63E-54
rs10128556	AFR	Basal Levels of FH	566,0709	1,2E-123
rs10128556	EAS	Basal Levels of FH	462,1858	4,3E-101
rs10128556	EUR	Basal Levels of FH	316,4231	1,95E-69
rs10189857	EAS	Basal Levels of FH	20,06321	4,4E-05
rs10189857	AMR	Basal Levels of FH	92,1718	9,66E-21
rs10189857	EUR	Basal Levels of FH	179,5613	1,02E-39
rs10189857	AFR	Basal Levels of FH	447,1257	8,09E-98
rs10189857	SAS	Basal Levels of FH	7,021048	0,029881
rs10494225	EUR	FH and FH response to Hydroxyurea	517,4806	4,3E-113
rs10494225	AMR	FH and FH response to Hydroxyurea	327,2085	8,86E-72
rs10494225	SAS	FH and FH response to Hydroxyurea	432,2749	1,36E-94
rs10494225	AFR	FH and FH response to Hydroxyurea	448,2387	4,64E-98
rs10494225	EAS	FH and FH response to Hydroxyurea	528,2404	2E-115
rs11886868	**AMR**	**Basal Levels of FH**	**80,34456**	**3,58E-18**
rs11886868	**AFR**	**Basal Levels of FH**	**28,48313**	**6,53E-07**
rs11886868	**EAS**	**Basal Levels of FH**	**511,0362**	**1,1E-111**
rs11886868	**SAS**	**Basal Levels of FH**	**210,5475**	**1,91E-46**
rs172652	AFR	FH and FH response to Hydroxyurea	49,54167	1,75E-11
rs172652	SAS	FH and FH response to Hydroxyurea	312,4771	1,4E-68
rs172652	EUR	FH and FH response to Hydroxyurea	207,1368	1,05E-45
rs172652	EAS	FH and FH response to Hydroxyurea	inf	0
rs172652	AMR	FH and FH response to Hydroxyurea	171,1094	6,98E-38
rs17599586	EAS	FH change, %	inf	0
rs17599586	SAS	FH change, %	477,6657	1,9E-104
rs17599586	EUR	FH change, %	481,9601	2,2E-105
rs17599586	AFR	FH change, %	615,3077	2,4E-134
rs17599586	AMR	FH change, %	356,733	3,44E-78
rs2182008	**EUR**	**FH and FH response to Hydroxyurea**	**inf**	**0**
rs2182008	**SAS**	**FH and FH response to Hydroxyurea**	**inf**	**0**
rs2182008	**AFR**	**FH and FH response to Hydroxyurea**	**201,9845**	**1,38E-44**
rs2182008	**AMR**	**FH and FH response to Hydroxyurea**	**283,9491**	**2,19E-62**
rs2295644	SAS	FH change, %	205,8865	1,96E-45
rs2295644	EUR	FH change, %	224,6866	1,62E-49
rs2295644	EAS	FH change, %	252,7167	1,33E-55
rs2295644	AMR	FH change, %	201,1632	2,08E-44
rs2295644	AFR	FH change, %	384,9919	2,51E-84
rs2310991	**EAS**	**FH levels in response to Hydroxyurea treatment**	**9,728914**	**0,007716**
rs2387634	**EAS**	**FH and FH response to Hydroxyurea**	**inf**	**0**
rs2693430	EUR	FH and FH response to Hydroxyurea	369,1575	6,89E-81
rs2693430	AFR	FH and FH response to Hydroxyurea	52,07308	4,93E-12
rs2693430	AMR	FH and FH response to Hydroxyurea	227,9695	3,14E-50
rs2693430	EAS	FH and FH response to Hydroxyurea	235,4426	7,49E-52
rs2693430	SAS	FH and FH response to Hydroxyurea	188,1194	1,41E-41
rs334	**AMR**	**sickle cell anemia**	**inf**	**0**

rs334	**AFR**	**sickle cell anemia**	**669,7165**	3,7E-146
rs380620	SAS	FH and FH response to Hydroxyurea	29,99445	3,07E-07
rs380620	EUR	FH and FH response to Hydroxyurea	53,55173	2,35E-12
rs380620	AFR	FH and FH response to Hydroxyurea	248,4805	1,1E-54
rs380620	AMR	FH and FH response to Hydroxyurea	106,0345	**9,44E-24**
rs4671393	EAS	Basal Levels of FH	27,43307	1,1E-06
rs4671393	AFR	Basal Levels of FH	24,83807	4,04E-06
rs4671393	SAS	Basal Levels of FH	179,2484	1,19E-39
rs4671393	AMR	Basal Levels of FH	32,2236	1,01E-07
rs4671393	EUR	Basal Levels of FH	136,5117	2,27E-30
rs4895441	EAS	Basal Levels of FH	351,4358	4,86E-77
rs4895441	EUR	Basal Levels of FH	367,4829	1,59E-80
rs4895441	SAS	Basal Levels of FH	463,6068	2,1E-101
rs4895441	AMR	Basal Levels of FH	305,5343	4,51E-67
rs4895441	AFR	Basal Levels of FH	661,9941	1,8E-144
rs5006884	SAS	Regulation of HbA2 level	354,6249	9,87E-78
rs5006884	AFR	Regulation of HbA2 level	407,6082	3,08E-89
rs5006884	EUR	Regulation of HbA2 level	372,5312	1,28E-81
rs5006884	AMR	Regulation of HbA2 level	268,4789	5,02E-59
rs5006884	EAS	Regulation of HbA2 level	461,0498	7,7E-101
rs61743453	**AFR**	**FH levels in response to Hydroxyurea treatment**	**728,4883**	**6,5E-159**
rs61743453	**AMR**	**FH levels in response to Hydroxyurea treatment**	**inf**	**0**
rs7309163	AMR	FH and FH response to Hydroxyurea	225,011	1,38E-49
rs7309163	AFR	FH and FH response to Hydroxyurea	44,29208	2,41E-10
rs7309163	EUR	FH and FH response to Hydroxyurea	268,9595	3,95E-59
rs7309163	SAS	FH and FH response to Hydroxyurea	234,6203	1,13E-51
rs7309163	EAS	FH and FH response to Hydroxyurea	272,7217	6,01E-60
rs7557939	**EAS**	**Basal Levels of FH**	**511,0362**	**1,1E-111**
rs7557939	**AMR**	**Basal Levels of FH**	**14,76837**	0,000621
rs7599488	AMR	Basal Levels of FH	87,10243	1,22E-19
rs7599488	AFR	Basal Levels of FH	451,5556	8,8E-99
rs7599488	EUR	Basal Levels of FH	184,4137	9,02E-41
rs7599488	EAS	Basal Levels of FH	18,26966	**0,000108**
rs816361	AFR	FH and FH response to Hydroxyurea	508,6557	3,5E-111
rs816361	AMR	FH and FH response to Hydroxyurea	251,5265	2,41E-55
rs816361	EAS	FH and FH response to Hydroxyurea	316,4722	1,9E-69
rs816361	SAS	FH and FH response to Hydroxyurea	378,6624	5,95E-83
rs816361	EUR	FH and FH response to Hydroxyurea	339,2567	2,14E-74
rs9319428	**AFR**	**FH and FH response to Hydroxyurea**	**6,132478**	**0,046596**
rs9693712	EAS	FH and FH response to Hydroxyurea	323,1362	6,79E-71
rs9693712	AFR	FH and FH response to Hydroxyurea	470,8223	5,8E-103
rs9693712	AMR	FH and FH response to Hydroxyurea	104,695	1,84E-23
rs9693712	SAS	FH and FH response to Hydroxyurea	48,71858	2,64E-11
rs9693712	EUR	FH and FH response to Hydroxyurea	13,89978	0,000959

The values of the Chi-square and p for HWE calculation of SNP in the studied populations. All presented SNP present statistical significance. The SNP rs11886868, rs2182008, rs380620, and rs7599488 present disequilibrium in 4 populations, and the rs2310991 (EAS), rs2387634 (EAS), rs334 (AMR, AFR), rs61743453 (AMR, AFR), rs7557939 (EAS, AMR), rs9319428 (AFR), present disequilibrium in one or two populations. Mentioned SNPs are presented in Bold. (FH: Fetal hemoglobin; AFR: African Population; SAS: Population of South Asia; AMR: American Population; EUR: European Population; EAS: East Asian Population)

ved between the different characterized populations (See figures 1, 2 and 3). As expected, the mutation rs 334, the main etiological component of SCA, is highly prevalent in the African population and has a much lower prevalence in the American population (figure 1).

Regarding the SNPs described for FH Levels in response to treatment with Hydroxyurea, it can be seen (figure 3) that the mutant allele of the SNP rs61743453 is present in the African population, unlike the rest of the populations. In purple, we can see that the SNP rs2310991 is found in the five categorical groups studied. However, it presents a higher frequency of appearance in the African population.

III. Logistic correlation and automatic classification

A logistic regression model was applied to classify the studied populations automatically. In this context, it is essential that the logistic regression successfully organized genomic information between African and

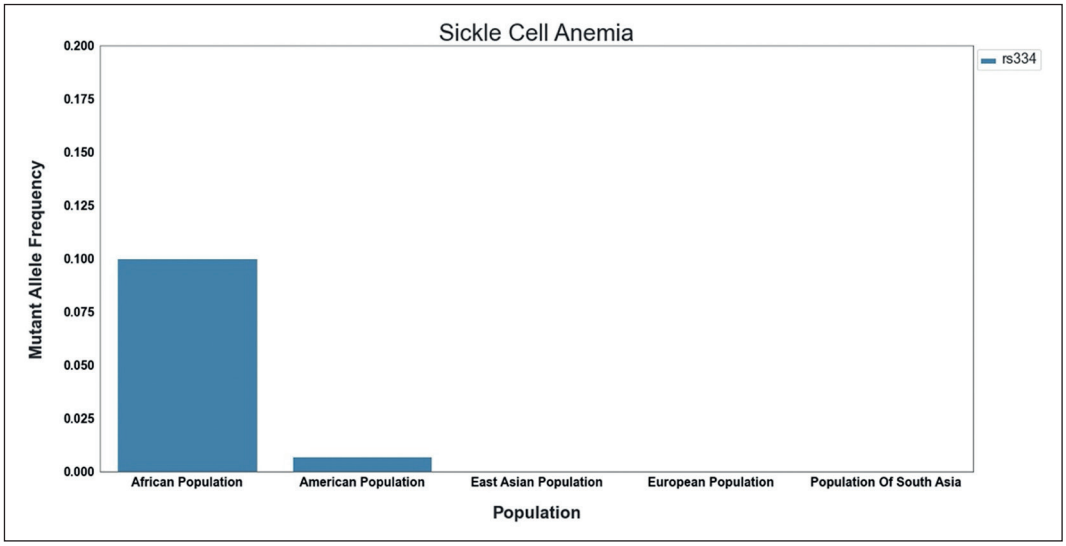

Figure 1. Frequency of the mutation rs 334 in the analyzed population.

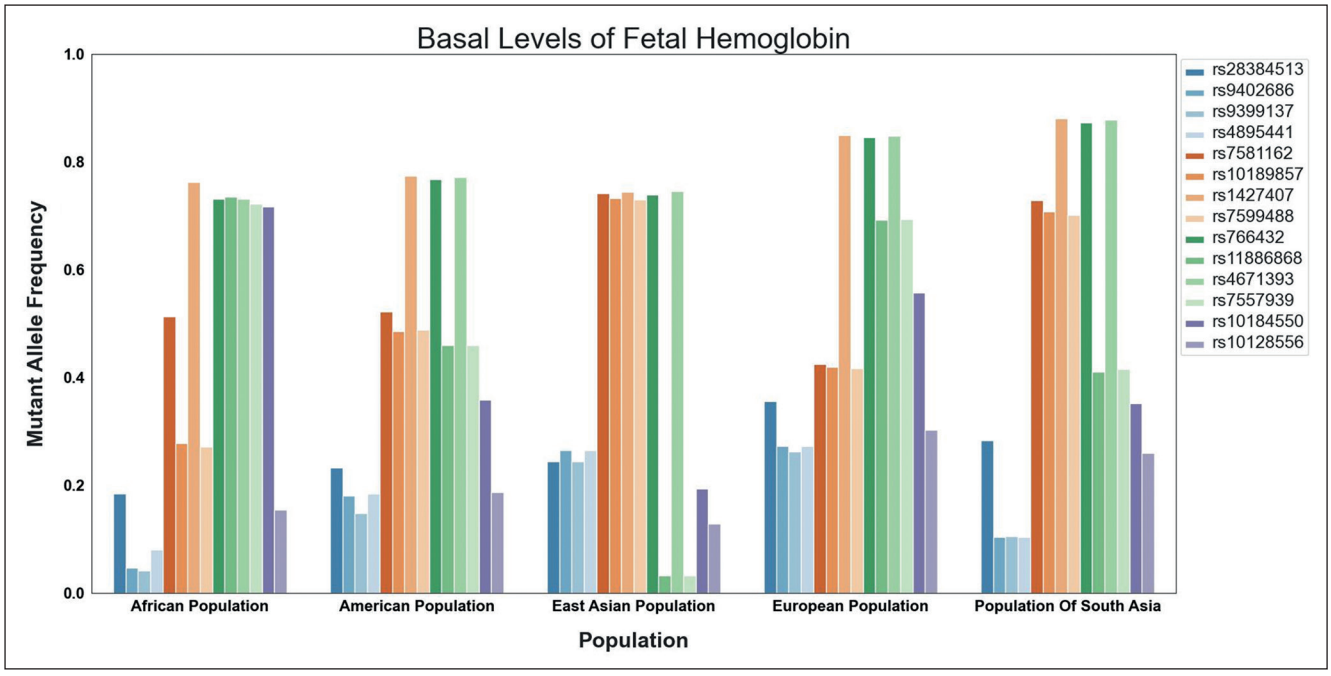

Figure 2. Frequency of the mutant alleles of the SNPs associated with basal levels of Fetal hemoglobin (FH).

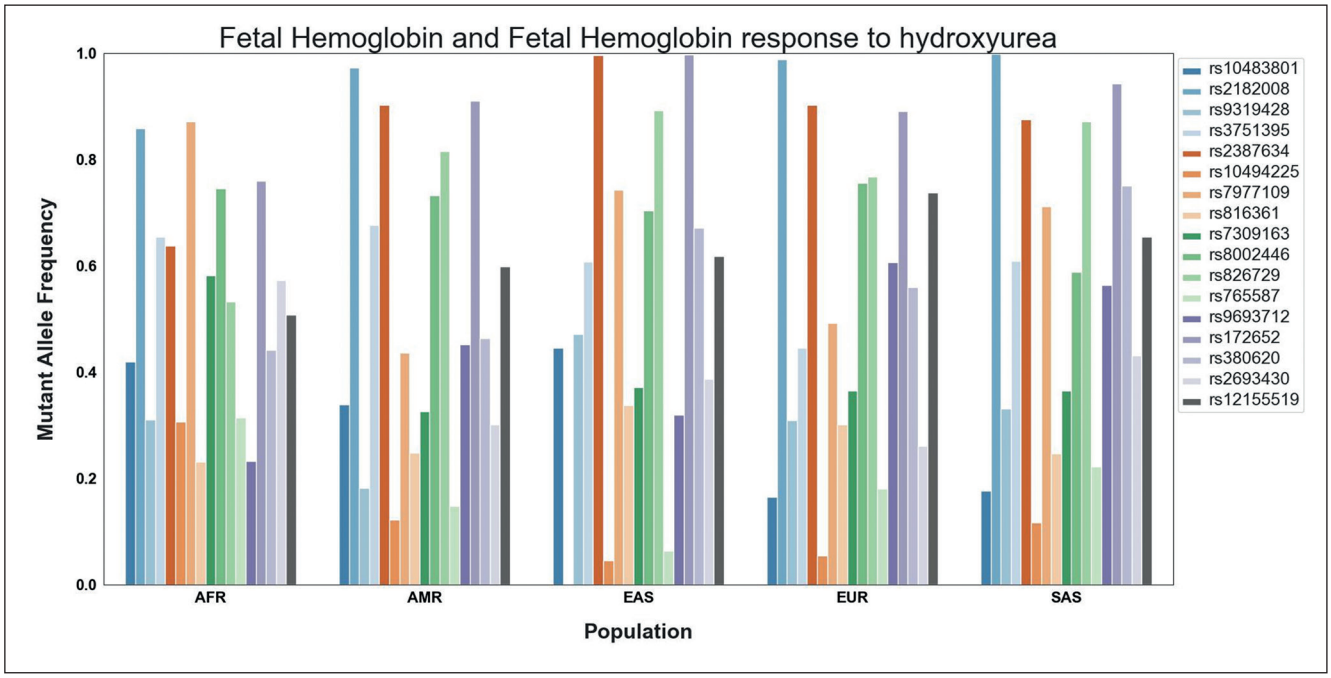

Figure 3. Frequency of the mutant allele of the SNPs associated with the response to treatment with Hydroxyurea.

non-African categories. In the other populations, the logistic regression could identify the correct absence of the definite population (0). Still, it could not correctly assign the category population in other different from the African group. The logistic regression was applied using SNPs related to the diagnosis, prognosis, and treatment associated with SCA.

In table 2 is possible to observe that the logistic regression model applied to data belonging to African population could be helpful to predict if a person with SNPs involved in response treatment with Hydroxyurea, basal levels of FH, regulation of the hemoglobin subunit alpha 2(HbA2) level and percentage of change of FH levels could be classified as an African (or with some genetic origin related to African population) with a precision of 89% to predict value 1. In the other cases, the model could not classify them as American Population, European Population, East Asian Population, or the Population of South Asia. However, the model was correct when the groups did not belong to one of these population groups. Seemingly, the presence of selected SNP; involved in a better prognosis of SCD according to previous descriptions, appears frequently related to a possible African ethnic origin, unlike other populations where malaria is also found. These SNPs could be a potential fitness improvement in ancestral populations with SCD or other hemoglobinopathies. It could be interesting to evaluate subpopulations already described in the 1000 genomes project (currently under work).

Discussion

Marked differences in the population frequency of the polymorphisms analyzed between different populations were observed. This differential frequency in clinically relevant SNPs is important for personalizing medical treatments. This information could be used to construct personalized public policies based on the population genomic information accessed in different databases[16]. Currently, there are abundant methods that will enable the reduction of costs associated with sequencing and characterizing the profile of polymorphisms in a patient. They have mainly been applied to the detection of mutations related to the diagnosis and prognosis of SCA[16], presenting higher indices of sensitivity and accuracy compared to classical techniques such as high performance liquid chromatography (HPLC). However, these technologies are not yet massive. Several advances in technology and techniques have made sequencing more affordable and accessible. These methods include next-generation sequencing (NGS) technologies and high-throughput genotyping platforms, which allow for more efficient and cost-effective analysis of genetic variations and polymorphisms in patients.

Additionally, there are databases with population information on SNP profiles. In this context, the literature has demonstrated in retrospective studies the benefit of polymorphism typing in patients diagnosed with SCA[17-19]. One study identified the possibility

Table 2. Classification of genomic information between populations: African (AFR), American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS) using a logistic regression model.

Population		Precision	Recall	f1-score	Support
AFR	0	0.78	0.99	0.87	21014
	1	0.89	0.25	0.39	7894
	Accuracy			0.79	28908
	Macro avg	0.83	0.62	0.63	28908
	Weighted avg	0.81	0.79	0.74	28908
AMR	0	0.88	1.00	0.94	25538
	1	0.00	0.00	0.00	3370
	Accuracy			0.88	28908
	Macro avg	0.44	0.50	0.47	28908
	Weighted avg	0.78	0.88	0.83	28908
EAS	0	0.80	1.00	0.89	22983
	1	0.00	0.00	0.00	5925
	Accuracy			0.80	28908
	Macro avg	0.40	0.50	0.44	28908
	Weighted avg	0.63	0.80	0.70	28908
EUR	0	0.79	1.00	0.88	22775
	1	0.00	0.00	0.00	6133
	Accuracy			0.79	28908
	Macro avg	0.39	0.50	0.44	28908
	Weighted avg	0.62	0.79	0.69	28908
SAS	0	0.80	1.00	0.89	23162
	1	0.00	0.00	0.00	5746
	Accuracy			0.80	28908
	Macro avg	0.40	0.50	0.44	28908
	Weighted avg	0.64	0.80	0.71	28908

AFR: African Population; SAS: Population of South Asia; AMR: American Population; EUR: European Population; EAS: East Asian Population

Table 3. Number of Male and Female distributed in every ethnic group analyzed

Populations	gender	snp	sample
AFR	F	12654	342
	M	11803	319
AMR	F	6549	177
	M	6290	170
EAS	F	9620	260
	M	9028	244
EUR	F	9731	263
	M	8880	240
SAS	F	8473	229
	M	9620	260
ALL	F	47027	1271
	M	45621	1233
		92648	2504

AFR: African Population; SAS: Population of South Asia; AMR: American Population; EUR: European Population; EAS: East Asian Population.

of better targeting SCA's pharmacological treatment by characterizing the SCA-associated mutations' profile. Regarding the treatment of SCA, personalization of drug treatment has been made difficult by the low number of drugs available. This means that all patients receive Hydroxyurea as the primary drug. In recent years, new drugs have appeared to treat SCA; for example, crizanlizumab and voxelotor have been approved by the FDA and EMA organizations. Voxelotor corresponds to an HB modulator that inhibits the polymerization of Hemoglobin S, favoring the oxygenated status of HB. After the finishing of the Clinical trial phase III, Voxelotor demonstrated the reduction of vaso-oclusive events and hemolotyc indices[20]. Voxelotor has been described as particularly useful for patients over 12 years old and with SCD resistant to hydroxyurea treatment[21,22]. Crizalizumab is a monoclonal antibody against P-selectin, a protein involved in the adhesion expressed in the endothelium. The partial blockade of P-selectin makes difficult the formation of aggregates between platelets and leukocytes, diminishing the for-

mation of vaso-oclusive events[23]. With the appearance of these new drugs, the importance of having tools that identify polymorphisms in population groups and improve the personalization of medical treatments for SCA and other pathologies is highlighted. However, efforts to personalize SCA treatment with Hydroxyurea have been described in the literature by calculating and adjusting individual specific doses[24,25].

Databases have been established that will allow information on SCA and other hemoglobinopathies to be stored. The mentioned databases are beneficial for establishing the personalization of treatments against SCA[26]. Some examples of personalization of SCA treatment have been carried out by characterizing polymorphisms associated with the response to treatment with Hydroxyurea in pediatric patients[27-30]. Several studies have tried to identify which population of patients will improve much in terms of FH levels. Interestingly, many patients already had higher Basal FH and presented many SNP in the *BCL 11A*, *HMIP* (HBS1L-MYB intergenic polymorphism), and *HBG2* genes[31]. The *BCL* 11A gen codifies for a zinc finger protein, transcription factor. The BCL11A gen is associated with the development of several lymphoid malignancies[32]. The HMIP names an array of SNPs in the intergenic region of the genes HBS1L (a G-protein/elongation factor) and the MYC oncogene[33]. In this context, one exciting approach corresponds to applying pharmacokinetic measurement to define a particular individual dose[34]. In this context, The SNPs related to the metabolization of Hydroxyurea and included in this study are significant to identify, and through this approach improved the pharmacogenetics profile of the individual doses' definition of Hydroxyurea[35,36].

Currently, processes such as migratory changes and further modifications in the ethnic matrix have generated the need to establish public policies for neonatal analysis for SCA in communities where they were not performed before[37]. For example, it has been described in the literature that some haplotypes of SCA could cosegregate and potentially be associated with different FH levels[38- 45].

Based on the evidence presented in our work, it would be desirable and oriented toward personalizing medicine to extend neonatal detection to a complete set of polymorphisms, mostly SNPs and potential haplotypes of interest, with diagnosis and treatment purposes. The early genetic profiling could be oriented to each population or ethnic group of interest, families with at least one member affected member, or children at risk. It is considered that there is a need to improve the detection technology of SNPs, allowing the simultaneous detection of multiple mutations relevant to SCA. Most public health efforts to establish neonatal studies only detect the etiological mutation of SCA[46].

One of the many complications of SCA is infections. Particularly in countries of high income, the mortality of children with SCA diagnosis is similar to patients without the diagnosis[47]. However, in African countries, only 50% of children with SCA survive[48]. Infections are mainly an environmental factor; interestingly, important differences are found in several SNP related to the metabolization of antibiotics and several drugs used to treat SCD (currently under work). Pharmacogenomics for other conditions concomitant with SCA has not been described. Some observational studies suggest that patients with SCA diagnosis suffer a particularly severe infection of COVID-19 and dengue fever[49-51]. Also, it is essential to personalize the SCA treatment and the complete characterization of the haplotype of the Hemoglobin subunit beta (HBB) gene[52]. Five different haplotypes of mutant HBB subunit beta (β S HBB) have been described, including Benin (BEN), Senegal (SEN), Arab/India (AI), Bantu (BAN), and Cameroon (CAM)[4]. ls of FH and a mild phenotype[5]. While patients with the BEN and CAM haplotypes had severe and mild phenotypes, respectively[52]. Interestingly, regarding SCA, many modifiers and cofactors have been described. One of the most important is the co-presence of α-thalassemia. Currently, it is estimated that 30-35% of the patients with an SCD diagnosis are heterozygous (αα/α-), with 3-5% homozygous for the deletion (α-/α-)[53,54] for α-thalassemia. The simultaneous presence of α-thalassemia is a fundamental issue because patients with SCA diagnosis and the α-thalassemia deletion had a better prognosis[7]. Regarding the logistic regression analysis, it is highlighted that the model could correctly assign really positive and negative characterize values when the African population was analyzed. In other words, it could be said that the SNPs profile analyzed shows a particular association with the African ethnic group.

In summary, the phenotype and severity of SCA have many multi-genic factors. Applying the earliest possible genomic screening is important for the adequate future application of precision medicine in diagnosing and treating SCA. The emergence of genomics as a central discipline for developing personalized medicine raises hope for better treatment of SCA and SCD in general. This viewpoint could be potentiated by incorporating "omic" information, including proteomics, lipidomics, metabolomics, transcriptomics, and microbiomics, in unraveling SCA's physiopathology. Currently, there are no studies in our country on the presence of relevant mutations for SCA; however, there are epidemiological reports in populations of Central America and South America. In the population of pediatric patients from Haiti, various studies have established the presence of SCA. Studies on immigrants from Haiti to the United States identified the presence

of hemoglobin S and hemoglobin C in pediatric patients in 8.0% and 4.7%, respectively[54]. Another study has established the prevalence of these mutations at 0.58% for the homozygous mutant and approximately 13% for heterozygous carriers. Interestingly, the evidence would point to establishing that the prevalence of SCA in the Haitian population would be twice that of ethnic groups in sub-Saharan Africa[54]. Currently, according to census information and the Jesuit migrant service, there are approximately 200,000 people of Haitian origin in Chile[55]. Very few studies have analyzed the prevalence of SCA in South American countries, particularly in Venezuela; it was estimated that the prevalence in new births with SCA would be approximately 0.10% and heterozygous carriers of the mutation would be approximately 3.5%[56]. Currently, in Chile, there are approximately 500,000 Venezuelan migrants. Therefore, it is important to highlight that the migratory processes will modify the population in our country, and in the future, there will be increasing pressure on the health systems related to the early detection of this disease.

Conclusion

Considering the statistical analysis performed and the result obtained, it is concluded that several significant SNPs related to the diagnosis, prognosis, and treatment of SCA are particularly correlated among different ethnic groups. It is our firm belief that with focalization in high-risk patients, also with a family history of hematologic alteration among other groups, the advance in the identification of essential SNPs, like the group mentioned in this study, and the characterization of the haplotype in patients with SCA and SCD will improve the responses to treatment and should be a priority, particularly in the public health system of our country, oriented to high priority ethnic groups or high-risk patients.

Ethical Responsibilities

Human Beings and animals protection: Disclosure the authors state that the procedures were followed according to the Declaration of Helsinki and the World Medical Association regarding human experimentation developed for the medical community.

Data confidentiality: The authors state that they have followed the protocols of their Center and Local regulations on the publication of patient data.

Rights to privacy and informed consent: The authors state that the information has been obtained anonymously from previous data, therefore, Research Ethics Committee, in its discretion, has exempted from obtaining an informed consent, which is recorded in the respective form.

Conflicts of Interest

Authors declare no conflict of interest regarding the present study.

Financial Disclosure

Project DICYT Clinico 022091BP_MED to RB.

Acknowledgments

Universidad de Santiago de Chile, USACH. Agradecimientos Proyecto InvClínica _DICYT, Código 022091BP_MED, Vicerrectoría de Investigación, Desarrollo e Innovación.

References

1. Khan H, Krull M, Hankins JS, et al. Sickle cell disease and social determinants of health: A scoping review. Pediatr Blood Cancer. 2023;70(2):e30089. doi: 10.1002/pbc.30089.
2. Wastnedge E, Waters D, Patel S, et al. The global burden of sickle cell disease in children under five years of age: a systematic review and meta-analysis. J Glob Health. 2018;8(2):021103. doi: 10.7189/jogh.08.021103.
3. Kavanagh PL, Fasipe TA, Wun T. Sickle Cell Disease: A Review. JAMA. 2022;328(1):57-68. doi: 10.1001/jama.2022.10233.
4. Nagel RL, Fleming AF. Genetic epidemiology of the beta s gene. Baillieres Clin Haematol. 1992;5(2):331-65. doi: 10.1016/s0950-3536(11)80023-5.
5. Adekile AD, Gu LH, Baysal E, et al. Molecular characterization of alpha-thalassemia determinants, beta-thalassemia alleles, and beta S haplotypes among Kuwaiti Arabs. Acta Haematol. 1994;92(4):176-81. doi: 10.1159/000204216.
6. Makani J, Cavazzana M, Gupta K, et al. Blood diseases in Africa: Redressing unjust disparities is an urgent unmet need. Am J Hematol. 2022;97(12):1505-6. doi: 10.1002/ajh.26754.
7. Kasai ET, Alworong'a Opara JP, Ntokamunda Kadima J, et al. Overview of current progress and challenges in diagnosis, and management of pediatric sickle cell disease in Democratic Republic of the Congo. Hematology. 2022;27(1):132-40. doi: 10.1080/16078454.2021.
8. Brousse V, Makani J, Rees DC.

Management of sickle cell disease in the community. BMJ. 2014;348:g1765. doi: 10.1136/bmj.g1765.
9. WATSON J. The significance of the paucity of sickle cells in newborn Negro infants. Am J Med Sci. 1948;215(4):419-23. doi: 10.1097/00000441-194804000-00008.
10. Rees DC, Brousse VAM, Brewin JN. Determinants of severity in sickle cell disease. Blood Rev. 2022;56:100983. doi: 10.1016/j.blre.2022.100983.
11. Kang HA, Barner JC, Lawson KA, et al. Impact of adherence to hydroxyurea on health outcomes among patients with sickle cell disease. Am J Hematol. 2023;98(1):90-101. doi: 10.1002/ajh.26765.
12. Elenga N, Celicourt D, Muanza B, et al. Dengue in hospitalized children with sickle cell disease: A retrospective cohort study in the French departments of America. J Infect Public Health. 2020;13(2):186-92. doi: 10.1016/j.jiph.2019.07.015.
13. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74. doi: 10.1038/nature15393.
14. Green NS, Ender KL, Pashankar F, et al. Candidate sequence variants and fetal hemoglobin in children with sickle cell disease treated with hydroxyurea. PLoS One. 2013;8(2):e55709. doi: 10.1371/journal.pone.0055709.
15. McKinney, W. Data structures for statistical computing in Python. In Proc. 9th Python in Science Conf. (eds van der Walt, S. & Millman, K. J.). 2010:56-61.
16. El Hoss S, El Nemer W, Rees DC. Precision Medicine and Sickle Cell Disease. Hemasphere. 2022;6(9):e762. doi: 10.1097/HS9.0000000000000762.
17. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. N Engl J Med. 1994;330(23):1639-44. doi: 10.1056/NEJM199406093302303.
18. Pincez T, Ashley-Koch AE, Lettre G, et al. Genetic Modifiers of Sickle Cell Disease. Hematol Oncol Clin North Am. 2022;36(6):1097-124. doi: 10.1016/j.hoc.2022.06.006.
19. WOOLF B. On estimating the relation between blood group and disease. Ann Hum Genet. 1955;19(4):251-3. doi: 10.1111/j.1469-1809.1955.tb01348.x.
20. Vichinsky E, Hoppe CC, Ataga KI, et al. A Phase 3 Randomized Trial of Voxelotor in Sickle Cell Disease. N Engl J Med. 2019;381(6):509-19. doi: 10.1056/NEJMoa1903212.
21. Food and Drug Administration (FDA). Oxbryta. https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/213137s000lbl. pdf. Accessed 02 Feb 2023 63
22. CHMP. Oxbryta, INN-Voxelotor. European Medicines Agency (EMA). Oxbryta. https://www. ema.europa.eu/en/documents/product-information/oxbryta-eparproduct-information_en.pdf. Accessed 02 Feb 2023
23. Ataga KI, Kutlar A, Kanter J, et al. Crizanlizumab for the Prevention of Pain Crises in Sickle Cell Disease. N Engl J Med. 2017;376(5):429-39. doi: 10.1056/NEJMoa1611770.
24. Adekile A, Akbulut-Jeradi N, Al Khaldi R, et al. Diagnosis of Sickle Cell Disease and HBB Haplotyping in the Era of Personalized Medicine: Role of Next Generation Sequencing. J Pers Med. 2021;11(6):454. doi: 10.3390/jpm11060454.
25. Gallaway KA, Sakon C, Ongeri J, et al. Opportunity for pharmacogenetics testing in patients with sickle cell anemia. Pharmacogenomics. 2022;23(17):925-31. doi: 10.2217/pgs-2022-0115.
26. Lanzkron S, Manwani D, Desai P, et al. GRNDaD: big data and sickle cell disease. Blood Adv. 2022;6(3):1088. doi: 10.1182/bloodadvances.2021005282.
27. Meier ER, Creary SE, Heeney MM, et al. Hydroxyurea Optimization through Precision Study (HOPS): study protocol for a randomized, multicenter trial in children with sickle cell anemia. Trials. 2020;21(1):983. doi: 10.1186/s13063-020-04912-z.
28. Dong M, McGann PT. Changing the Clinical Paradigm of Hydroxyurea Treatment for Sickle Cell Anemia Through Precision Medicine. Clin Pharmacol Ther. 2021;109(1):73-81. doi: 10.1002/cpt.2028.
29. Gueye Tall F, Martin C, Ndour EHM, et al. Influence of Oxidative Stress Biomarkers and Genetic Polymorphisms on the Clinical Severity of Hydroxyurea-Free Senegalese Children with Sickle Cell Anemia. Antioxidants (Basel). 2020;9(9):863. doi: 10.3390/antiox9090863.
30. Manu GP, Segbefia C, N'guessan BB, et al. Association Between Selected Single Nucleotide Polymorphisms in Globin and Related Genes and Response to Hydroxyurea Therapy in Ghanaian Children with Sickle Cell Disease. Pharmgenomics Pers Med. 2022;15:205-14. doi: 10.2147/PGPM.S351599.
31. Mohammad SNNA, Iberahim S, Wan Ab Rahman WS, et al. Single Nucleotide Polymorphisms in *XMN1-HBG2*, *HBS1L-MYB*, and *BCL11A* and Their Relation to High Fetal Hemoglobin Levels That Alleviate Anemia. Diagnostics (Basel). 2022;12(6):1374. doi: 10.3390/diagnostics12061374.
32. Satterwhite E, Sonoki T, Willis TG, et al. The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood. 2001;98(12):3413-20. doi: 10.1182/blood.v98.12.3413.
33. Menzel S, Jiang J, Silver N, et al. The HBS1L-MYB intergenic region on chromosome 6q23.3 influences erythrocyte, platelet, and monocyte counts in humans. Blood. 2007;110(10):3624-6. doi: 10.1182/blood-2007-05-093419.
34. Karimi M, Zarei T, Haghpanah S, et al. Relationship Between Some Single-nucleotide Polymorphism and Response to Hydroxyurea Therapy in Iranian Patients With β-Thalassemia Intermedia. J Pediatr Hematol Oncol. 2017;39(4):e171-6. doi: 10.1097/MPH.0000000000000779.
35. Allard P, Alhaj N, Lobitz S, et al. Genetic modifiers of fetal hemoglobin affect the course of sickle cell disease in patients treated with hydroxyurea. Haematologica. 2022;107(7):1577-88. doi: 10.3324/haematol.2021.278952.
36. Quinn CT, Niss O, Dong M, et al. Early initiation of hydroxyurea (hydroxycarbamide) using individualised, pharmacokinetics-guided dosing can produce sustained and nearly pancellular expression of fetal haemoglobin in children with sickle cell anaemia. Br J Haematol. 2021;194(3):617-25. doi: 10.1111/bjh.17663.
37. Zhou JR, Ridsdale R, MacNeil L, et al. The Alberta Newborn Screening Approach for Sickle Cell Disease: The Advantages of Molecular Testing. Int J Neonatal Screen. 2021;7(4):78. doi: 10.3390/ijns7040078.
38. Chebloune Y, Pagnier J, Trabuchet G, et al. Structural analysis of the 5' flanking region of the beta-globin gene in African sickle cell anemia patients: further evidence for three origins of the sickle cell mutation in Africa. Proc Natl Acad Sci U S A. 1988;85(12):4431-5. doi: 10.1073/pnas.85.12.4431.
39. Shriner D, Rotimi CN. Whole-Genome-Sequence-Based Haplotypes Reveal Single Origin of the Sickle Allele during the Holocene Wet Phase. Am J Hum Genet. 2018;102(4):547-56. doi: 10.1016/j.ajhg.2018.02.003.
40. Green NS, Fabry ME, Kaptue-Noche L, et al. Senegal haplotype is associated with higher HbF than Benin and Cameroon haplotypes in African children with sickle cell anemia. Am J Hematol. 1993;44(2):145-6. doi: 10.1002/ajh.2830440214.
41. Powars D, Hiti A. Sickle cell anemia. Beta s gene cluster haplotypes as genetic markers for severe disease expression. Am J Dis Child.

1993;147(11):1197-202. doi: 10.1001/archpedi.1993.02160350071011.
42. Perrine RP, Pembrey ME, John P, et al. Natural history of sickle cell anemia in Saudi Arabs. A study of 270 subjects. Ann Intern Med. 1978;88(1):1-6. doi: 10.7326/0003-4819-88-1-1.
43. Labie D, Pagnier J, Lapoumeroulie C, et al. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. Proc Natl Acad Sci USA. 1985;82(7):2111-4. doi: 10.1073/pnas.82.7.2111.
44. Nagel RL, Fabry ME, Pagnier J, et al. Hematologically and genetically distinct forms of sickle cell anemia in Africa. The Senegal type and the Benin type. N Engl J Med. 1985;312(14):880-4. doi: 10.1056/NEJM198504043121403.
45. Serjeant GR, Vichinsky E. Variability of homozygous sickle cell disease: The role of alpha and beta globin chain variation and other factors. Blood Cells Mol Dis. 2018;70:66-77. doi: 10.1016/j.bcmd.2017.06.004.
46. Ngole M, Race V, Mbayabo G, et al. DNA testing for sickle cell anemia in Africa: Implementation choices for the Democratic Republic of Congo. J Clin Lab Anal. 2022;36(5):e24398. doi: 10.1002/jcla.24398.
47. Telfer P, Coen P, Chakravorty S, et al. Clinical outcomes in children with sickle cell disease living in England: a neonatal cohort in East London. Haematologica. 2007;92(7):905-12. doi: 10.3324/haematol.10937.
48. Makani J, Cox SE, Soka D, et al. Mortality in sickle cell anemia in Africa: a prospective cohort study in Tanzania. PLoS One. 2011;6(2):e14699. doi: 10.1371/journal.pone.0014699.
49. Hoogenboom WS, Alamuri TT, McMahon DM, et al. Clinical outcomes of COVID-19 in patients with sickle cell disease and sickle cell trait: A critical appraisal of the literature. Blood Rev. 2022; 53:100911. doi: 10.1016/j.blre.2021.100911.
50. Telfer P, De la Fuente J, Sohal M, et al. Real-time national survey of COVID-19 in hemoglobinopathy and rare inherited anemia patients. Haematologica. 2020;105(11):2651-4. doi: 10.3324/haematol.2020.259440.
51. Panepinto JA, Brandow A, Mucalo L, et al. Coronavirus Disease among Persons with Sickle Cell Disease, United States, March 20-May 21, 2020. Emerg Infect Dis. 2020;26(10):2473-6. doi: 10.3201/eid2610.202792.
52. Powars DR, Chan L, Schroeder WA. Beta S-gene-cluster haplotypes in sickle cell anemia: clinical implications. Am J Pediatr Hematol Oncol. 1990;12(3):367-74. doi: 10.1097/00043426-199023000-00022.
53. Rotz S, Arty G, Dall'Amico R, et al. Prevalence of sickle cell disease, hemoglobin S, and hemoglobin C among Haitian newborns. Am J Hematol. 2013;88(9):827-8. doi: 10.1002/ajh.23510.
54. Saint Fleur R, Archer N, Hustace T, et al. Capacity building and networking to make newborn screening for sickle cell disease a reality in Haiti. Blood Adv. 2018;2(Suppl 1):54-5. doi: 10.1182/bloodadvances.2018GS111997.
55. Mercado Órdenes M, Figueiredo A. Construcciones identitarias de inmigrantes haitianos en Santiago de Chile desde una perspectiva interseccional. Migr Inter. 2023;13: rmiv1i12495.https://doi.org/10.33679/rmi.v1i1.2495
56. Huttle A, Maestre GE, Lantigua R, et al. Sickle cell in Latin America and the United States [corrected]. Pediatr Blood Cancer. 2015;62(7):1131-6. doi: 10.1002/pbc.25450. Erratum in: Pediatr Blood Cancer. 2015 Aug;62(8):1493.