



Genomic ancestry evaluated by ancestry-informative markers in patients with sickle cell disease

A.F. Nascimento¹, J.S. Oliveira², J.C. Silva Junior² and A.A.L. Barbosa²

¹Programa de Pós-Graduação em Genética Biodiversidade e Conservação, Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié, BA, Brasil

²Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié, BA, Brasil

Corresponding author: A.F. Nascimento
E-mail: alana20007@hotmail.com

Genet. Mol. Res. 15 (1): gmr.15017604

Received September 9, 2015

Accepted November 17, 2015

Published March 31, 2016

DOI <http://dx.doi.org/10.4238/gmr.15017604>

ABSTRACT. The β^s mutation is responsible for the most aggressive form of sickle cell disease, has a predominantly African origin, and arrived in Brazil through the slave trade. However, the Brazilian population is highly miscegenated, underscoring the importance of ancestry-informative markers (AIMs) for the identification of the genetic structure of a population. In this study, we have estimated the genetic contributions of various ethnicities in individuals with sickle cell disease in the microregion of Jequié, Bahia, in Brazil, by using AIMs, and compared the findings to those from a phenotypic characterization. Eight AIMs were analyzed: *AT3* (rs3138521), *DRD2* (rs1079598), *APO* (rs3138522), *PV92*, *Sb19.3* (rs3138524), *CKM* (rs4884), *LPL* (rs285), and *CCR5 Δ 32* (rs333). Samples were subjected to polymerase chain reaction (PCR) and PCR-restriction fragment length polymorphism. The amplified products were electrophoresed on agarose gels, and the data were statistically analyzed using Genepop, FSTAT 2.9, and Admix3. Phenotypic classification showed a high frequency of mulattos

(85%) in the Brazilian population; however, ancestry-informative markers indicated that 44, 42, and 11% of the population had European, African, and native American ancestries, respectively. The phenotypic classification is justified as a complementary method for the characterization of the genetic ancestry in patients with sickle cell disease, as it confirms the molecular findings regarding ancestry.

Key words: Sickle cell disease; Ancestry; Genetic markers; Miscegenation; PCR-RFLP

INTRODUCTION

The population of Brazil is one of the most miscegenated in the world, composed of genetic contributions predominantly from three groups dating back to the colonization period: European colonizers, African slaves, and native American (Daniel et al., 2008). Ancestry-informative genetic markers (AIM) allow for an accurate discrimination between miscegenated populations (Shriver et al., 2003), thereby allowing us to generate precise estimates of the ethnic composition of the trihybrid Brazilian population (Parra et al., 1998).

The patterns of disease prevalence and mortality differ widely among populations. Subpopulations in a single country often show heterogeneous patterns of susceptibility and resistance. Diseases are distributed heterogeneously across different geographic regions, reflecting genetic adaptations to environmental factors. This is well exemplified by some monogenic diseases (Pena, 2005) such as sickle cell disease (caused by a mutation in the β -globin gene - β^S) in malaria-endemic regions, which confers a higher chance of survival to heterozygous, asymptomatic individuals during the acute phase of malaria (Bortolini and Salzano, 1999). The origin of the mutation in the β -globin gene is almost completely limited to Africa; of the five different β^S mutations, named after their geographic origins (Senegal, Cameroon, Benin, Banto, and Indo-Arabic haplotypes) (Bortolini and Salzano, 1999), the first four originated in Africa, whereas the fifth occurs in Minor Asia and India. Despite this, the occurrence of sickle cell disease in populations of non-African descent and in non-African regions suggests that this disease is not a “black disease” or an “African disease”, but an eminently geographic disease; the product of a well-succeeded evolutionary strategy of humans to cope with malaria, caused by *Plasmodium falciparum* (Haemosporida: Plasmodiidae) (Hiernaux, 1975). It is difficult to make such an association in Brazil, given that the correlation between phenotypic characterization and genetic ancestry is low (Pena, 2005). Therefore, in this study, we attempted to estimate the ethnic composition of sickle cell patients in the microregion of Jequié (BA) by assessing the genetic contributions of the disease through AIMS. The results were then compared to those of a phenotypic characterization to analyze the concordance between the results of racial classification from phenotypic analyses and those employing ancestry markers.

MATERIAL AND METHODS

Sample characterization

All cases of sickle cell disease recorded by the Health Secretary of Jequié between the years 2009 and 2012 were surveyed; additionally, a search was conducted for new cases.

Subsequently, 20 individuals (aged 4-20 years) with sickle cell disease were included in this study. The project was approved by the Committee for Ethics in Research of Universidade Estadual do Sudoeste da Bahia - UESB (Protocol number 077/2011, CAAE: 0057.0.454.000-11). The participants were informed of all research procedures, and an informed consent form was obtained from each patient.

Genomic DNA extraction

Biological material was collected from all patients at the Laboratory of Human Genetics at the UESB. Blood samples (10 mL) were obtained from all individuals by venous puncture and stored in Vacutainer tubes containing EDTA. DNA was extracted with the QIAamp DNA Investigator kit (Qiagen, Venlo, Netherlands).

Phenotypic characterization

The patients were also subjected to an intensive phenotypic characterization during sample collection, based on known criteria such as skin color, hair type, and lip and nose shape, as described by Azevêdo et al. (1982).

AIM analysis

Eight AIMs - *AT3* (rs3138521), *DRD2* (rs1079598), *APO* (rs3138522), *PV92*, *Sb19.3* (rs3138524), *CKM* (rs4884), *LPL* (rs285), and *CCR5Δ32* (rs333) - were analyzed in this study. The samples were subjected to a polymerase chain reaction (PCR) and PCR-restriction fragment length polymorphism (RFLP), and the amplified products were electrophoresed on agarose gels (1 and 3%).

Statistical analysis

The allelic frequencies at each locus were estimated by direct counting with FSTAT v.2.8 (Goudet, 1995). Deviations from the Hardy-Weinberg equilibrium, as well as pairwise allelic and genotypic comparisons between loci, were determined by using the GENEPOP program (Raymond and Rousset, 1995). Estimates of ethnic mixtures were obtained using the genetic identity method (Chakraborty, 1985) with the programs ADMIX 2.0 and ADMIX 3. The model fit was evaluated by determining the coefficient of multiple correlation (R^2) between allelic frequencies in the hybrid populations and those in the ancestral populations (Chakraborty, 1986). The frequencies of parental populations previously described in the literature were used to calculate the corresponding frequencies in the ethnic mixture.

RESULTS

Analysis of the 8 AIMs allowed us to identify the correlation between molecular parameters and ethnicity based on phenotypic classification and/or self-declared ethnicity. Based on phenotypic classification, the cultural ancestry of the 20 patients with sickle cell disease (SS and SC) in Southwestern Bahia was as follows: 10% Caucasians, 85% Mulattos, and 5% Blacks; ethnic

contribution (indicated by AIMs) showed that the contribution of the European parental population was the highest (approximately 44%), followed by those from African (42%) and native American (11%) populations.

The frequency of the *1 allele of the markers studied in patients with sickle cell disease was as follows: 47.3% for CKMM, 81.2% for LPL, 57.8% for AT3I/D, 35.2% for PV92, 5.5% for DRD2, 90.0% for APO, 67.5% for Sb19.3, and 100.0% for CCR5.

Table 1 shows a comparison between the allele *1 frequencies of studied AIMs in parental populations, from our studied sickle cell disease population and a normal population from Bahia.

Table 1. Frequency of allele *1 of the AIMs in parental populations, the population with sickle cell disease (analyzed in this study), and in a normal population from Bahia.

Marker	CKMM	LPL	AT3I/D	PV92	DRD2	APO	Sb19.3	CCR5
African	0.164*	0.971*	0.858*	0.225*	0.063*	0.420*	0.415*	0.998 ^a
European	0.313*	0.492*	0.282*	0.152*	0.144*	0.925*	0.903*	0.931 ^b
Amerindian	0.904*	0.442*	0.061*	0.792*	0.665*	0.977*	0.645*	1.000 ^c
S.C.D.	0.473	0.812	0.578	0.352	0.055	0.900	0.675	1.000
Normal	0.298	0.692	0.517	0.273		0.746	0.667	

*Shriver et al., 2003; ^aMartinson et al., 1997; ^bLibert et al., 1998; ^cLeboute et al., 1999. S.C.D = frequency of allele *1 in patients with sickle cell disease in this study. Normal = frequency of allele *1 in normal individuals in Bahia. Source: Teló (2010).

DISCUSSION

Owing to the highly miscegenated Brazilian population, phenotypic classification is not a precise proxy for the determination of genomic ancestry (Parra et al., 2003). However, it can be used to complement molecular parameters, as shown in the present study.

In order to compare the molecular parameters with ethnicity (based on the phenotypic classification), we attempted to assess the degree to which the physical appearance of patients with sickle cell disease from Southwestern Bahia was determined by African genomic ancestry, as sickle cell disease has a predominantly African origin.

By the analysis of Table 1, with regards to the parental populations, our results indicated that the markers LPL, AT3I/D, PV92, and DRD2 showed the closest frequency to African ancestry, CKMM and APO to European ancestry, and Sb19.3 and CCR5 to native American ancestry. Interestingly, although the frequency values showed a greater proximity to one of the parental populations, some markers showed a relative equidistance/similarity to more than one ancestry (e.g., AT3I/D to the African and European ancestries; PDB to the European and native American ancestries; and CCR5 to all three ancestries). Therefore, these results suggest miscegenation in the analyzed population. Additionally, the marker frequencies in individuals with sickle cell disease were also similar to those seen in normal subjects from Bahia.

The results of molecular analyses, in addition to the results from phenotypic classification, exhibited a high frequency of mulattos (85%). This led us to infer that patients with sickle cell disease in Bahia are as miscegenated as those without the disease; that is, despite being affected with a disease of African origin, a majority of the population is not of African descent. AIM analysis also revealed such miscegenation; both African (42%) and European (44%) ancestries were high in the population.

Therefore, we propose that the results of the phenotypic classification complement those of the molecular studies; that is, African and European ancestries are distributed evenly in the study population. None of the patients was of a single (100%) ancestry; rather, we observed heterogeneous

ethnic contributions, in accordance with the historical data on miscegenation in Brazil.

Brazilian history, comprising of native American, the European colonization, black slave trade, and the opening of Brazilian ports to friendly nations and consequent European immigration, explains the highly miscegenated ethnic composition of Brazilians. This allows us to suggest that Brazilians are generally not of a single ancestry (African, European, or native American). This has also been observed in other studies carried out in Bahia; Machado (2008) analyzed nine AIMs (*FY-Null*, *LPL*, *AT3-ID*, *Sb19.3*, *APO*, *PV92*, *GC 1S*, *GC 1F*, and *CYP3A4*) to estimate the ethnic mixture in more than 1200 individuals from different neighborhoods of Salvador, and reported a contribution of 49.2, 36.3, and 14.5% from African, European, and native American ancestries, respectively.

Isolated communities, such as quilombo remnants, are exceptions to the estimates of genetic contribution (Pedrosa, 2006); ancestry can be almost exclusive as these communities are endogamic. For instance, the Valango quilombo community in Southern Brazil is almost entirely of African ancestry (97.3%) (Souza and Culpi, 2005). Despite this, they display more heterogeneity, when compared to other isolated communities.

Marker-based classification showed that European ethnicity made a higher contribution than the African and native American ethnicities. However, the difference between the European and African contributions was not statistically significant. The data obtained in this study diverge from those of another study on genomic ancestry in sickle cell patients from Minas Gerais using AIMs (Silva, 2011). In the latter, African ancestry was more prevalent (47.3%) than the European (39.7%) and native American (13.0%) ancestries. These results were expected owing to the almost exclusive origin of the mutation responsible for the disease.

Nonetheless, the results of a general analysis of sickle cell patients focusing on individual-level miscegenation agreed with those of this study, indicating that the European genomic ancestry (15.58%) was more prevalent than the African ancestry (11.05%); despite this, most individuals showed intermediate levels of miscegenation (73.37%).

However, the observed miscegenation is in agreement with the theory that sickle cell disease is not just a "Black disease" in Brazil, specifically in Bahia. This thought has been ingrained into the subconscious of a large portion of the Brazilian population. During this study, we observed that many people were unable to understand how they or their children could have this disease, as they had a Caucasian phenotype. This, as reported by Silva (2011), is an incorrect view that is oftentimes reinforced by the general media, as well as publicity material regarding the disease.

Therefore, in light of such results and the undeniable historical context, it can be concluded that sickle cell disease in Brazil, particularly in Bahia, is not restricted to black individuals or to those with a higher genetic contribution of African descent, in spite of originally being an African disease. This contribution relates only to the probability of carrying the disease, as a person with low proportion of African ancestry could also be afflicted.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank all the patients and family members for their participation. Research supported by Universidade Estadual do Sudoeste da Bahia (UESB), and Fundo Nacional de Desenvolvimento da Educação (FNDE).

REFERENCES

- Azevêdo ES, Fortuna CMM, Silva KMC, Sousa MGF, et al. (1982). Spread and diversity of human populations in Bahia, Brazil. *Hum. Biol.* 54: 329-341.
- Bortolini MC and Salzano FM (1999). Beta S haplotype diversity in Afro-Americans, Africans and Euro-Asiatics - An attempt at a synthesis. *Cienc. Cult.* 51: 175-180.
- Chakraborty R (1985). Gene identity in racial hybrids and estimation of admixture rates. In: Genetic Micro-differentiation in man and other animals (Neel JV and Ahuja Y, eds.). American Journal of Physical Anthropology, New Delhi, 171-180.
- Chakraborty R (1986). Gene admixture in human populations: Models and predictions. *Am. J. Phys. Anthropol.* 29: 1-43. <http://dx.doi.org/10.1002/ajpa.1330290502>
- Daniel R, Sanchez JJ, Nassif NT, Hernandez A, et al. (2008). SNPs associated with physical traits-A valuable tool for the inference of biogeographical ancestry. *Forensic Sci. Int.* 1: 538-540.
- Goudet J (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* 86: 484-486.
- Hiernaux J (1975). The People of Africa. Charles Scribner's Sons, New York.
- Leboute APM, de Carvalho MWP and Simões AL (1999). Absence of the *deltaccr5* mutation in indigenous populations of the Brazilian Amazon. *Hum. Genet.* 105: 442-443. <http://dx.doi.org/10.1007/s004390051128>
- Libert F, Cochaux P, Beckman G, Samson M, et al. (1998). The *deltaccr5* mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum. Mol. Genet.* 7: 399-406. <http://dx.doi.org/10.1093/hmg/7.3.399>
- Machado TMB (2008). Estimativa de mistura genética em uma amostra da população de Salvador. Master's thesis, Centro de Pesquisas Gonçalo Moniz Fundação Oswaldo Cruz, Fiocruz, Salvador.
- Martinson JJ, Chapman NH, Rees DC, Liu YT, et al. (1997). Global distribution of the CCR5 gene 32-basepair deletion. *Nat. Genet.* 16: 100-103. <http://dx.doi.org/10.1038/ng0597-100>
- Parra EJ, Marcini A, Akey J, Martinson J, et al. (1998). Estimating African American admixture proportions by use of population-specific alleles. *Am. J. Hum. Genet.* 63: 1839-1851. <http://dx.doi.org/10.1086/302148>
- Parra FC, Amado RC, Lambertucci JR, Rocha J, et al. (2003). Color and genomic ancestry in Brazilians. *Proc. Natl. Acad. Sci. USA* 100: 177-182. <http://dx.doi.org/10.1073/pnas.0126614100>
- Pedrosa MAF (2006). Composição genética de quatro populações remanescentes de quilombos do Brasil com base em microssatélites e marcadores de ancestralidade. Master's thesis, Universidade de Brasília, Salvador.
- Pena SDJ (2005). Razões para banir o conceito de raça de medicina brasileira. *Hist. Cienc. Saude Manguinhos* 12: 321-346.
- Raymond M and Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. *J. Hered.* 86: 248-249.
- Shriver MD, Parra EJ, Dios S, Bonilla C, et al. (2003). Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum. Genet.* 112: 387-399.
- Silva MCF (2011). Miscigenação e doença falciforme em Minas Gerais. *J. Hemominas* 26: 6.
- Souza IR and Culp L (2005). Valongo, genetic studies on an isolated afro-Brazilian community. *Genet. Mol. Biol.* 28: 402-406. <http://dx.doi.org/10.1590/S1415-47572005000300012>
- Teló EP (2010). Estimativa de mistura étnica avaliada por Marcadores Informativos de Ancestralidade (AIMs) e Microssatélites (STRs). Master's thesis, Centro de Pesquisas Gonçalo Moniz Fundação Oswaldo Cruz, Fiocruz, Salvador.