

Relationship between Ancestry Inferred by Molecular Analysis, Self-Report and Hetero-Classification

Pablo Abdon da Costa Francez^{1,2*}, Adriano Ruiz Lima², Rivelton Riverson Pereira de Almeida² and Sidney Emanuel Batista dos Santos³

¹Department of Forensic Laboratories, Laboratory for Forensic Genetics, Technical and Scientific Police in Amapá, Macapá, Amapá, Brazil

²Faculdade SEAMA, Macapá, Amapá, Brazil

³Department of Pathology, Laboratory of Human and Medical Genetics, Federal University of Pará, Belém, Pará, Brazil

Abstract

Population stratification is the main cause of spurious results and the difficulty of replicating genetic association findings. In admixed populations, there is evidence that such proxies do not provide efficient control of stratification. We aimed to evaluate the statistical relationship between self-reported, hetero-classified and genetic ancestry, defined by categories of skin color and individual ancestry estimates and by the genotyping of 48 ancestry-informative indel markers in a sample of 130 volunteers from the city of Macapá in the Amazonian state of Amapá - Brazil. Five groups of self-reported skin color were defined: white, light brown, middle brown, dark brown and blacks, and three groups of predominant ancestry were defined: European, African and Native American. The results showed a significant correlation between skin color and predominant ancestry, self-reported and hetero-classified by volunteers, in comparison with the genetic ancestry, an important potential application of these AIMs in forensic investigations.

Keywords: Indel; Skin color; Genetic ancestry; Ancestry hetero-classified; Ancestry self-reported

Introduction

The genetic variations observed among different human populations are relevant in epidemiologic studies and can be used as a tool for the definition of information about the parental lineage in admixed populations. Although the biogeography of certain population groups in some parts of the world is culturally and genetically established, other groups have a relatively recent history of miscegenation, with ancestors coming from different continents. This is the case of the Brazilian population, characterized by the miscegenation of three philogeographic parental groups (Europeans, Africans and Native Americans), in regionally variable degrees of miscegenation [1,2].

Self-reported ancestry is a method described as having high correlation with the genetic structure of the population in well-defined and stratified philogeographic groups, such as the Europeans, Africans or Asians [3-5]. In highly miscegenated populations, however, both the declared ancestry or other anthropometric human features usually employed, such as skin color, prove controversial and are not very reliable for assuming ancestry [5-8]. In this context, the use of molecular genetic markers could be of great value to reduce the potential consequences of stratification in these populations [6,9,10].

Regardless of any sociological and anthropological context, in Brazil, according to the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística - IBGE), race and skin color are treated as equivalent in the demographic classification. In most Brazilian genetic association studies, control of the population structure is carried out based only on the self-reported skin color, the evaluation of anthropometric features, a genealogy analysis or the interviewer's subjective opinion [11-15].

While the grouping of individuals based on self-reported ancestry is a typical procedure for determining philogeographic groups in studies of the Brazilian population, the recent utilization of molecular markers for inferring genetic ancestry has revealed great genetic heterogeneity [16-22]. One of the problems of conducting association studies in

miscegenated populations by using only self-reported ancestry or skin color as equivalents of ethnicity is the possibility of a spurious association with false-positive or false-negative results [23-29].

An increasing number of publications have reported the use of ancestry-informative markers (AIMs), whose allelic frequencies vary significantly among populations of different geographic origins and can be employed to estimate individual admixture and to identify a population substructure. In most of the studies, single nucleotide polymorphisms (SNPs) were used [30-33], but insertion/deletion markers (INDELs) with small DNA fragments [29,34] and Short Tandem Repeats (STRs) [35] have also been employed.

Genetic variation studies using DNA polymorphisms distributed throughout the genome have allowed a better understanding of the history and diversity of human populations and have provided a genetic identification system of individuals. Insertion/deletion polymorphisms (INDELs) are length polymorphisms produced by insertions or deletions of one or more nucleotides in the genome. In the last years, the INDEL-type polymorphisms have been given attention in many studies.

In 2002, Weber et al. [36] identified and characterized 2000 bi-allelic polymorphisms of the INDEL type in humans, with a great variation in the length of the observed alleles, abundant in the genome and easy to analyze, bringing these markers into evidence for use in genetic investigations. Ever since, several studies have been conducted

***Corresponding author:** Pablo Abdon da Costa Francez, Human and Medical Genetics Laboratory; Institute for Biological Sciences; City University Prof. José da Silveira Netto; Rua Augusto Corrêa 01, Guamá, Box: 8615; CEP: 66.075-970, Belém, Pará, Brazil, Tel: +55 91 3201-7843; E-mail: pabdon@uol.com.br

Received July 22, 2015; **Accepted** August 21, 2015; **Published** August 24, 2015

Citation: Francez PADC, Lima AR, Almeida RRPD, Santos SEBD (2015) Relationship between Ancestry Inferred by Molecular Analysis, Self-Report and Hetero-Classification. J Forensic Res 6: 304. doi: [10.4172/2157-7145.1000304](https://doi.org/10.4172/2157-7145.1000304)

Copyright: © 2015 Francez PADC, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

using INDELs for a great variety of purposes, including investigations of relatedness, addressing the genetic structure of human populations and using them as genetic markers in natural populations [37-40].

The objective of this work was to evaluate the relationship between self-reported/hetero-classified ancestry, skin color and individual genetic ancestry estimated by using 48 insertion/deletion polymorphisms in an admixed population sample from the Brazilian Amazon region.

Material and Methods

Samples

Peripheral blood samples were collected from 130 healthy and unrelated individuals (63 male, 67 female, with a mean age of 37.5 years, varying from 18 to 69) seen for routine exams at the Clinical Analysis Laboratory (Laboratório de Análise Clínica UNILAB) in the city of Macapá (0° 02'20" N; 51° 03'59" W), State of Amapá, northern Brazil. After the blood collection, digital images (photographs) of all 130 volunteers were also taken and recorded. All participants in this study signed a free and informed consent form.

Molecular analysis

DNA was extracted from peripheral blood mononuclear cells, using the phenol-chloroform protocol [40]. DNA quantification was done with a NANODROP 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The PCR amplification conditions and primer sequences used were the same as described by Santos et al. [41] (Supplementary Table 1). The amplification products were submitted to capillary electrophoresis in an ABI 3130 genetic analyzer (Applied Biosystems^{RT}) and analyzed using Genescan and Gene Mapper software.

Enquiry

Self-report: In the interviews with the volunteer study were analyzed the skin color, predominant ancestry and percentages of ancestry (African, Native American and European) obtained by self-report. In other words, each of the 130 subjects indicated that skin color, predominant ancestry and percentage of European, African and Native American ancestry who believed present.

Hetero-classification: The recorded digital images of each study subject were presented to another group, composed of 40 volunteers (24 women and 16 men, with a mean age of 25.3 years, varying from 18 to 58), selected mostly among university students or graduates who signed a free and informed consent form by which they agreed to take part in the research. These volunteers evaluated issues regarding skin color, predominant ancestry and ancestry percentage of the 130 photographed persons (hetero-classification).

No specific period of time was determined for the evaluation of the digital images by the volunteers (hetero-classification), and most of the analyses took between 40 minutes and one and a half hour.

The hetero-classified ancestry and skin color of each one of the 130 subjects was defined based on the majority of statements obtained in the enquiry, i. e., the test subjects were included in a given skin color (white, light brown, medium brown, dark brown, and black) or predominant ancestry (African, European and Native American) category if at least 20 (50%) of the evaluators assigned them to the same category.

Regarding ancestry percentages by hetero-classification, each of forty volunteers indicated the percentage who believed that each subject had of African, Native American and European ancestry, based on their

physical characteristics inferred from the observed photographs. Then the percentage of hetero-classification ancestry was estimated for each individual using the arithmetic mean of the percentages mentioned by 40 volunteers. This methodology (hetero-classification ancestry) simulates an everyday situation in criminal cases is related to the description made by the eyewitnesses, the physical characteristics of criminals.

Statistical analysis

The genotypes obtained for the 130 samples investigated were analyzed with regard to their interethnic admixture using the Structure software http://pritch.bsd.uchicago.edu/software/structure2_1.html), and the FST matrix and WPGMA tree analysis were carried out with the GDA program. The tree was displayed by means of the TreeView software <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html> [42,43].

The results obtained by the enquiries were compared with the ancestry estimates based on the DNA of the 130 subjects. These comparisons were submitted to statistical treatments using the G test and linear regression, aiming to determine whether there was a correlation between the phenotypes (ancestry and skin color) indicated by self-report and by hetero-classification and the genetic ancestry inferred by the AIMs employed, using the BioStat 5.0 program (Ayres et al.).

Ethics Committee

This study was submitted for analysis to the Ethics Committee of SEAMA College and was given approval, under protocol: 133/2010.

Results and Discussion

DNA samples from the 130 study subjects were analyzed, comparing them with the genetic profile of the three parental groups (Africans, Europeans and Native Americans) using the 48 AIMs described earlier. The STRUCTURE program allowed estimating for the population of Macapá a mean ancestry of 21% African, 50% European and 29% Native American. These ancestry percentages are in accordance with those estimated in previous studies which employed genetic markers with bi-parental inheritance for the same population [42,43] and for other miscegenated Brazilian populations [44-48] (Supplementary Tables 2 and 3).

Comparing the genetic profile obtained for the AIM panel investigated in the population of Macapá with those obtained for other Brazilian populations (southern regions, Belém do Pará, and "quilombola" African descendants) and parental populations (Europeans, Africans and Native Americans), the population of Macapá was found to group to that of Belém, presenting a larger divergence from the other populations (Supplemental Figures 1-3). This result is in strict accordance with the geographic proximity of these two populations, their similar colonization history and the fact that a significant percentage of the population of Amapá originates from the neighboring State of Pará.

Enquiry (hetero-classification)

Comparing the study subjects' self-reported and hetero-classified (through the enquiries) skin color, overall 70% of the subjects self-reported as having white skin were also hetero-classified as having white skin. As for the other skin colors, no such evident relationship between self-report and hetero-classification was observed when evaluated isolatedly. However, when grouped into three categories (W-LB=white and light brown, MB=middle brown, and DB-B=dark

brown and black), it was observed that in 73.5% of cases the subjects self-reported as having white or light brown skin were also hetero-classified in the same manner. Likewise, 68.7% of the subjects self-reported as having dark brown or black skin were also hetero-classified similarly (Supplementary Table 4).

These results indicate that the majority of the participants of the study are convergent in their opinion with regard to skin color when evaluating persons with the most extreme skin colors (white/light brown and black/dark brown).

Comparing the predominant ancestry obtained by self-report and by hetero-classification, 63.6% of the subjects self-reported as having predominantly African ancestry were also hetero-classified as having predominantly African ancestry. With regard to predominantly European ancestry, 66.1% of the subjects were both self-reported and hetero-classified as having predominantly European ancestry. Among the subjects self-reported as having predominantly Native American ancestry, 44% were also hetero-classified as predominantly Native American (Supplementary Table 5).

A comparison between skin color and ancestry, both by self-report (Supplemental Table 6) and by hetero-classification (Supplementary Table 7), showed significant differences when the G test was applied to the ancestry percentages among the individuals with different skin colors. A European ancestry was predominant among the individuals who self-reported and/or were considered as having white or light brown skin, and an African ancestry was predominant among the individuals who self-reported and/or were considered as having dark brown or black skin.

Genetic Ancestry

A comparison between the ancestry percentages estimated by genetic analysis and the self-reported and hetero-classified ancestry data showed that individuals with European genetic ancestry values ranging from 60% to 70% were self-reported and hetero-classified as having predominantly European ancestry in 79% and 68.4% of cases, respectively. Among the individuals with a European genetic ancestry percentage superior to 70%, the predominant self-reported and hetero-classified ancestry percentages were 100% and 93.3%, respectively.

Similar results are also observed among the individuals with African genetic ancestry values ranging from 30% to 50% were self-reported and hetero-classified as having predominantly African ancestry in 43% and 53.6% of cases, respectively. Among the individuals with an African genetic ancestry percentage superior to 50%, who were self-reported and hetero-classified as having predominantly African ancestry in 100% of cases (Tables 1 and 2; Figure 1). However, as the sample size of the subgroup with percentage of African ancestry greater than 50% was very small, these results should be evaluated by employing additional studies with larger sample size.

When the ancestry percentages estimated by DNA analysis were compared with the self-reported and hetero-classified skin color, a correlation was found between the European and African ancestry percentages and skin color. The higher the European genetic ancestry percentages estimated, the greater the number of individuals self-reported or hetero-classified as having white or light brown skin, and similarly, the higher the African ancestry values, the greater the number of individuals self-reported or hetero-classified as having dark brown or black skin (Tables 3-5).

Figures 2-4 show graphs comparing the ancestry percentages (self-

reported, hetero-classified and genetic) and the skin color of the study subjects. There is a clear-cut decrease of European and increase of African ancestry as the skin color becomes darker. Linear regression analysis allowed disclosing statistically significant correlations between the genetic ancestry percentages and the ancestry percentages estimated by self-report and hetero-classification (Supplementary Figures 4-12).

An important point to be highlighted is the fact that, the observed correlation notwithstanding, the African and – to a lesser extent - the Native American contributions were overestimated and the European contribution was underestimated, comparing the self-reported/hetero-classified and the genetic ancestry. This result can be explained by the strong association observed between skin color and ancestry. A person with black or brown skin is automatically appointed as having a high degree of African ancestry, regardless of other characteristics. The same result was also reported by Santos et al. [39].

Some studies on ancestry [49] suggest that an attempt to correlate genetic ancestry based on AIMs with phenotypic characteristics such as skin color or another feature that is relevant in criminal investigations would not be applicable in Brazil, given the high degree of miscegenation of the population. However, many of these studies used AIMs which are not significantly different from the allele frequencies observed in the parental populations. These AIMs have often been selected for studies involving European populations and, when employed in Brazil, eventually overestimated the percentages of European ancestry and minimized the African and Native American contribution, making the Brazilian population seem much more homogeneous than it actually is.

Another recurrent problem in many studies is the utilization of modern African and Native American population samples as being representative of the parental populations that formed the Brazilian population, without taking into account the great genetic diversity of the current African populations and the several population bottlenecks undergone by the Native American populations along the colonization process, which substantially reduced the variability that existed in the past. In order to bypass these problems, it is fundamental to carry out a thorough investigation of the origin of the African populations which took part in the formation of the miscegenated population to be studied and to employ the largest possible sample of Native Americans, in an attempt to compensate for the loss of diversity observed in the modern populations.

The MULTI-INDELS panels described in this study were tested in different forensic samples and various concentrations of DNA and showed satisfactory results proving to be a promising technology in criminal investigations [50]. This molecular tool was first employed assisting a criminal investigation in the identification of two skeletons found in a small boat adrift located in the Atlantic Ocean near the north coast of the Brazilian Amazon region, in the state of Amapá. After speculation that the crew would be Africans, molecular analyzes using the AIMs MULTI-INDELS have been shown to treat admixed individuals, predominantly with African ancestry but with significant percentage of European and Native American ancestry, indicating that probably would be from the American continent, possibly in some region of northern South America or the Caribbean [51].

Data Access

Available upon request: pabdon@uol.com.br.

Conclusion

In conclusion, this AIM panel allowed inferring percentages of

Genetic Ancestry	Self-Reported Ancestry (%)			
	African	European	Native American	Total
African (30%-50%)	12 (43.0)	8 (28.5)	8 (28.5)	28 (100.0)
African (>50%)	5 (100.0)	0	0	5 (100.0)
European (50%-60%)	3 (11.1)	12 (44.4)	12 (44.4)	27 (100.0)
European (60%-70%)	2 (10.5)	15 (79.0)	2 (10.5)	19 (100.0)
European (>70%)	0	15 (100.0)	0	15 (100.0)
Native American (30%-40%)	5 (16.7)	8 (26.7)	17 (56.7)	30 (100.0)
Native American (40%-50%)	5 (29.4)	5 (29.4)	7 (41.7)	17 (100.0)
Native American (>50%)	2 (28.6)	1 (14.3)	4 (57.1)	7 (100.0)

Table 1: Relationship between self-reported and Genetic Ancestry.

Genetic Ancestry	Ancestry Hetero-Classified (%)				
	African	European	Native American	Other	Total
African (30%-50%)	15 (53.6)	6 (21.4)	4 (14.3)	3 (10.7)	28 (100.0)
African (>50%)	5 (100.0)	0	0	0	5 (100.0)
European (50%-60%)	7 (25.9)	8 (29.6)	10 (37.0)	2 (7.4)	27 (100.0)
European (60%-70%)	2 (10.5)	13 (68.4)	3 (15.8)	1 (5.3)	19 (100.0)
European (>70%)	0	14 (93.3)	0	1 (6.7)	15 (100.0)
Native American (30%-40%)	9 (30.0)	10 (33.3)	7 (23.3)	4 (13.3)	30 (100.0)
Native American (40%-50%)	4 (23.5)	6 (35.3)	6 (35.3)	1 (5.9)	17 (100.0)
Native American (>50%)	1 (14.3)	1 (14.3)	2 (28.6)	3 (42.9)	7 (100.0)

Table 2: Relationship between hetero-classified and Genetic Ancestry.

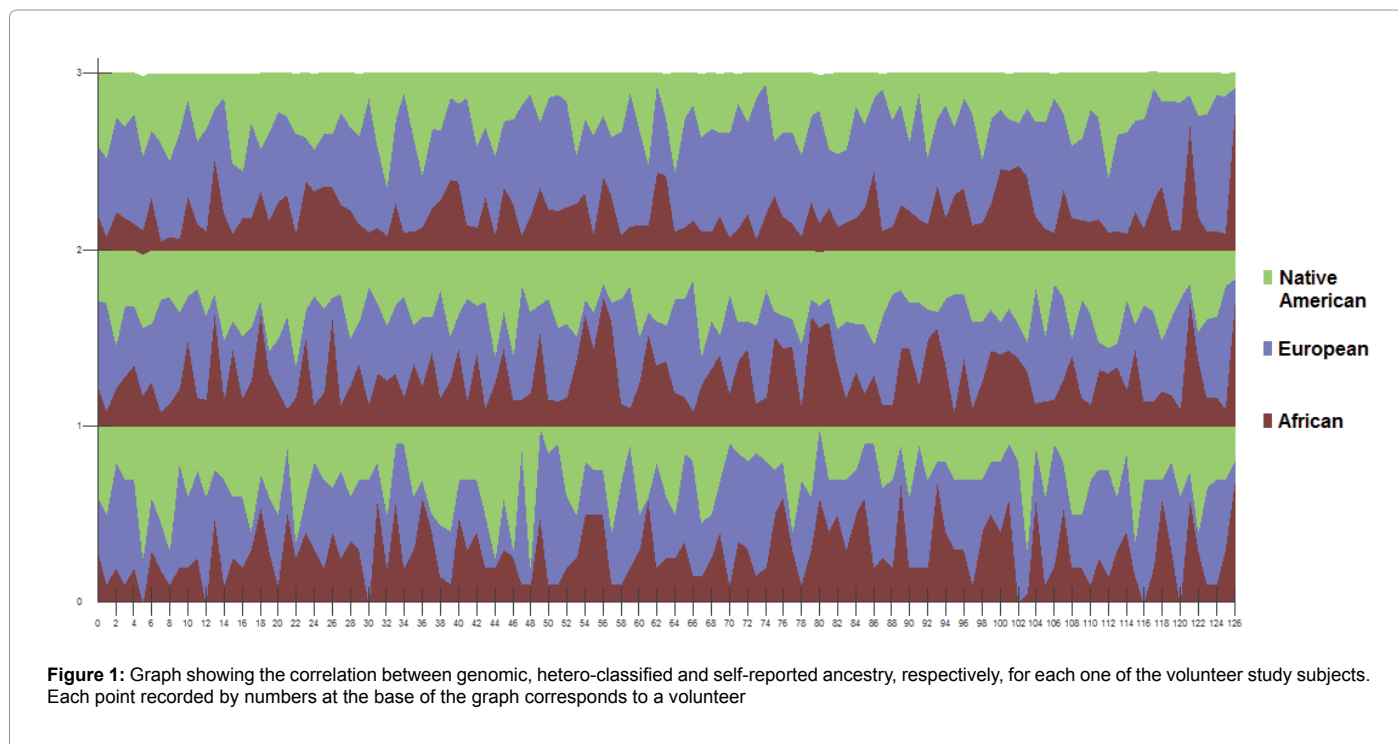


Figure 1: Graph showing the correlation between genomic, hetero-classified and self-reported ancestry, respectively, for each one of the volunteer study subjects. Each point recorded by numbers at the base of the graph corresponds to a volunteer

genetic ancestry which, after statistical analysis, proved significantly consistent with the ancestry estimates and the skin color classification of the volunteer study subjects obtained by hetero-classification and self-reports, indicating a potential application in crime investigations without suspects or in the identification of missing persons and unknown cadavers.

Furthermore, the use of MULTI-INDELS panels features easy

processing, fast reading results and does not require significant adjustments in Forensic Genetics laboratories already deployed, since it uses the same team chemistry and keeping the simple workflow is already established for STRs.

Acknowledgment

The authors greatly thank the volunteers who agreed to take part in this study and to Dr. Ricardo Ventura Santos for their corrections and suggestions regarding

Genetic Ancestry	Self-reported Skin Color (%)				
	W	LB	MB	DB/B	Total
African (30%-40%)	2 (10.0)	8 (40.0)	6 (30.0)	4 (20.0)	20 (100.0)
African (>40%)	0	3 (23.0)	4 (30.8)	6 (46.2)	13 (100.0)
European (50%-60%)	5 (18.5)	12 (44.4)	9 (33.3)	1 (3.7)	27 (100.0)
European (60%-70%)	9 (47.4)	9 (47.4)	1 (5.3)	0	19 (100.0)
European (>70%)	11 (73.3)	2 (13.3)	2 (13.3)	0	15 (100.0)
Native American (30%-40%)	4 (13.3)	14 (46.7)	7 (23.3)	5 (16.7)	30 (100.0)
Native American (40%-50%)	3 (17.7)	8 (47.1)	4 (23.5)	2 (11.7)	17 (100.0)
Native American (>50%)	1 (14.3)	5 (71.4)	0	1 (14.3)	7 (100.0)

Table 3: Relationship between self-reported skin color and Genetic Ancestry. Legend: W: White; LB: Light Brown; MB: Middle Brown; DB: Dark Brown; B: Black.

Genetic Ancestry	Hetero-Classified Skin Color (%)					
	W	LB	MB	DB/B	Other	Total
African (30%-40%)	3 (15.0)	3 (15.0)	3 (15.0)	3 (15.0)	8 (40.0)	20 (100.0)
African (>40%)	0	1 (7.6)	4 (30.8)	6 (46.2)	2 (15.4)	13 (100.0)
European (50%-60%)	6 (22.2)	3 (11.1)	6 (22.2)	3 (11.1)	9 (33.3)	27 (100.0)
European (60%-70%)	9 (47.4)	6 (31.6)	0	0	4 (21.1)	19 (100.0)
European (>70%)	10 (66.7)	0	0	0	5 (33.3)	15 (100.0)
Native American (30%-40%)	5 (16.7)	6 (20.0)	4 (13.3)	5 (16.7)	10 (33.3)	30 (100.0)
Native American (40%-50%)	4 (23.5)	3 (17.7)	2 (11.8)	3 (17.6)	5 (29.4)	17 (100.0)
Native American (>50%)	1 (14.3)	1 (14.3)	2 (28.6)	1 (14.3)	2 (28.6)	7 (100.0)

Table 4: Relationship between hetero-classified skin color and Genetic Ancestry. Legend: W: White; LB: Light Brown; MB: Middle Brown; DB: Dark Brown; B: Black.

Ancestry	Self-Reported Skin Color																			
	White				Light brown				Middle brown				Dark brown				Black			
	Mean	n	sd	se	Mean	n	sd	se	Mean	n	sd	se	Mean	n	sd	se	Mean	n	sd	se
Self-Reported																				
African	0.18	33	0.13	0.02	0.29	50	0.17	0.02	0.31	30	0.16	0.03	0.46	11	0.12	0.04	0.58	5	0.1	0.06
European	0.55	33	0.18	0.03	0.38	50	0.15	0.02	0.29	30	0.16	0.03	0.25	11	0.15	0.05	0.18	5	0.08	0.04
Native American	0.27	33	0.15	0.03	0.33	50	0.16	0.02	0.4	30	0.21	0.04	0.29	11	0.09	0.03	0.24	5	0.04	0.02
Hetero-Classified																				
African	0.15	33	0.05	0.01	0.29	50	0.12	0.02	0.36	30	0.14	0.03	0.5	11	0.17	0.05	0.68	5	0.04	0.02
European	0.54	33	0.12	0.02	0.35	50	0.16	0.02	0.21	30	0.1	0.02	0.15	11	0.08	0.02	0.09	5	0.02	0.01
Native American	0.31	33	0.1	0.02	0.36	50	0.1	0.01	0.43	30	0.1	0.02	0.35	11	0.11	0.03	0.23	5	0.06	0.03
Genetic																				
African	0.16	33	0.09	0.02	0.2	50	0.11	0.02	0.25	30	0.11	0.02	0.3	11	0.17	0.05	0.56	5	0.24	0.14
European	0.61	33	0.15	0.03	0.49	50	0.13	0.02	0.47	30	0.11	0.02	0.38	11	0.11	0.03	0.21	5	0.09	0.05
Native American	0.23	33	0.13	0.02	0.31	50	0.13	0.02	0.27	30	0.1	0.02	0.32	11	0.11	0.03	0.23	5	0.17	0.1

Table 5: Relationship between skin color and ancestry. Legend: n: Sample Size; sd: Standard Deviation; se: Standard Error.

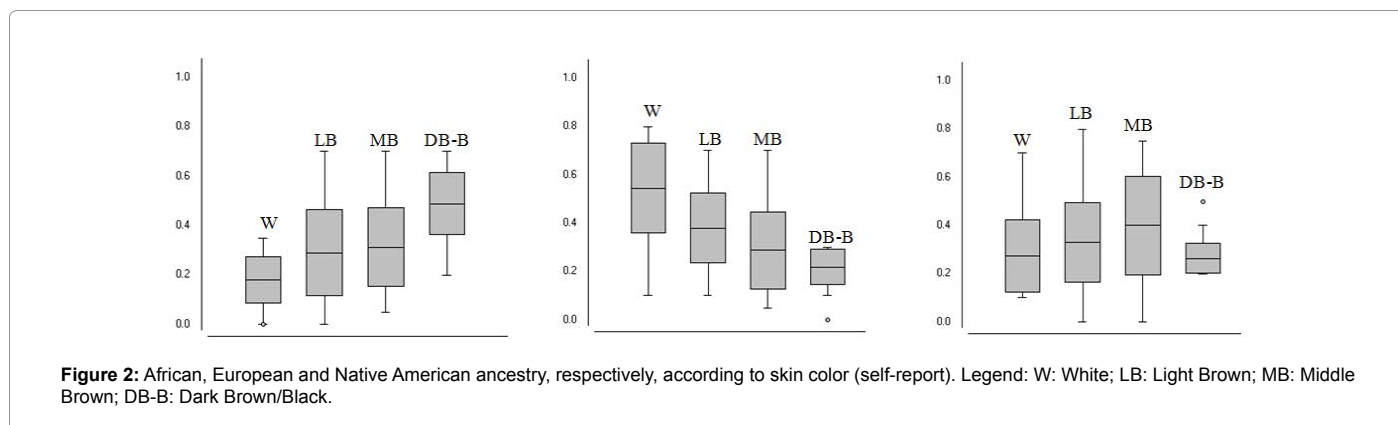


Figure 2: African, European and Native American ancestry, respectively, according to skin color (self-report). Legend: W: White; LB: Light Brown; MB: Middle Brown; DB-B: Dark Brown/Black.

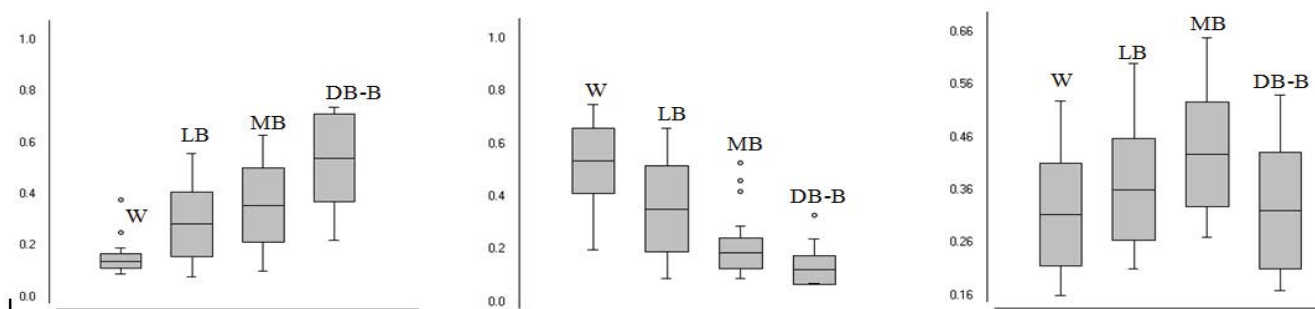


Figure 3: African, European and Native American ancestry, respectively, according to skin color (hetero-classification). Legend: W: White; LB: Light Brown; MB: Middle Brown; DB-B: Dark Brown/Black.

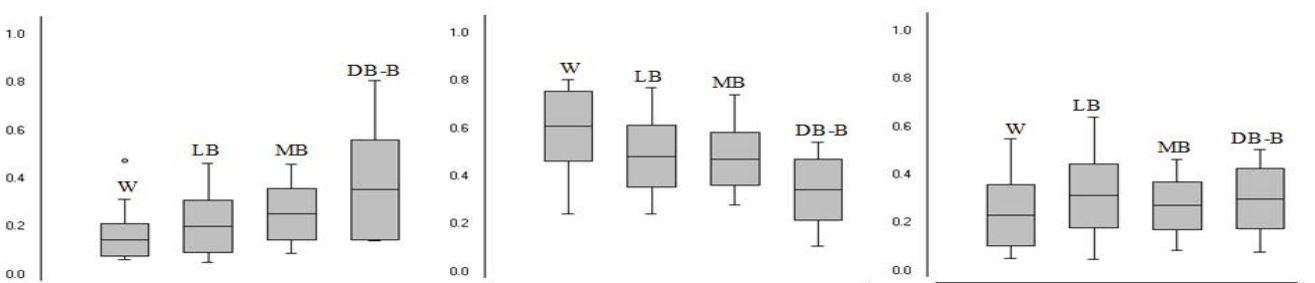


Figure 4: African, European and Native American ancestry, respectively, according to skin color (genetic). Legend: W: White; LB: Light Brown; MB: Middle Brown; DB-B: Dark Brown/Black.

this study.

References

- Salzano FM (2004) Interethnic variability and admixture in Latin America—social implications. *Rev Biol Trop* 52: 405-415.
- Sans M (2000) Admixture studies in Latin America: from the 20th to the 21st century. *Hum Biol* 72: 155-177.
- Liu XQ, Paterson AD, John EM, Knight JA (2006) The role of self-defined race/ethnicity in population structure control. *Ann Hum Genet* 70: 496-505.
- Tang H, Quertermous T, Rodriguez B, Kardia SL, Zhu X, et al. (2005) Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. *Am J Hum Genet* 76: 268-275.
- Yang N, Li H, Criswell LA, Gregersen PK, Alarcon-Riquelme ME, et al. (2005) Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum Genet* 118: 382-392.
- Barnholtz-Sloan JS, Chakraborty R, Sellers TA, Schwartz AG (2005) Examining population stratification via individual ancestry estimates versus self-reported race. *Cancer Epidemiol Biomarkers Prev* 14: 1545-1551.
- Burnett MS, Strain KJ, Lesnick TG, de Andrade M, Rocca WA, et al. (2006) Reliability of self-reported ancestry among siblings: implications for genetic association studies. *Am J Epidemiol* 163: 486-492.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, et al. (2003) Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci U S A* 100: 177-182.
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, et al. (2002) Genetic structure of human populations. *Science* 298: 2381-2385.
- Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, et al. (2003) Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 72: 1492-1504.
- Guzman VB, Morgun A, Shulzhenko N, Mine KL, Gonçalves-Primo A, et al. (2005) Characterization of CD28, CTLA4, and ICOS polymorphisms in three Brazilian ethnic groups. *Hum Immunol* 66: 773-776.
- Lazaretti-Castro M, Duarte-de-Oliveira MA, Russo EM, Vieira JG (1997) Vitamin D receptor alleles and bone mineral density in a normal premenopausal Brazilian female population. *Braz J Med Biol Res* 30: 929-932.
- Rezende VB, Barbosa F Jr, Montenegro MF, Sandrim VC, Gerlach RF, et al. (2007) An interethnic comparison of the distribution of vitamin D receptor genotypes and haplotypes. *Clin Chim Acta* 384: 155-159.
- Rossini A, Lima SS, Rapozo DC, Faria M, Albano RM, et al. (2006) CYP2A6 and CYP2E1 polymorphisms in a Brazilian population living in Rio de Janeiro. *Braz J Med Biol Res* 39: 195-201.
- Soares-Vieira JA, Billerbeck AE, Iwamura ES, Otto PA, Muñoz DR (1999) Gene and genotype frequencies for HLA-DQA1 in Caucasians and Mulattoes in Brazil. *J Forensic Sci* 44: 1051-1052.
- Abe-Sandes K, Silva WA Jr, Zago MA (2004) Heterogeneity of the Y chromosome in Afro-Brazilian populations. *Hum Biol* 76: 77-86.
- Callegari-Jacques SM, Grattapaglia D, Salzano FM, Salamoni SP, Crossetti SG, et al. (2003) Historical genetics: spatiotemporal analysis of the formation of the Brazilian population. *Am J Hum Biol* 15: 824-834.
- Carvalho-Silva DR, Santos FR, Rocha J, Pena SD (2001) The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet* 68: 281-286.
- Ferreira LB, Mendes-Junior CT, Wiezel CE, Luizon MR, Simões AL (2006) Genomic ancestry of a sample population from the state of São Paulo, Brazil. *Am J Hum Biol* 18: 702-705.
- Guerreiro-Junior V, Bisso-Machado R, Marrero A, Hünemeier T, Salzano FM, et al. (2009) Genetic signatures of parental contribution in black and white populations in Brazil. *Genet Mol Biol* 32: 1-11.
- Lins TC, Vieira RG, Abreu BS, Grattapaglia D, Pereira RW (2010) Genetic composition of Brazilian population samples based on a set of twenty-eight ancestry informative SNPs. *Am J Hum Biol* 22: 187-192.

22. Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, Peres LM, Kommers TC, et al. (2005) Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol* 17: 496-506.
23. Burchard EG, Ziv E, Coyle N, Gomez SL, Tang H, et al. (2003) The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med* 348: 1170-1175.
24. Cardon LR, Palmer LJ (2003) Population stratification and spurious allelic association. *Lancet* 361: 598-604.
25. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, et al. (2006) Population stratification confounds genetic association studies among Latinos. *Hum Genet* 118: 652-664.
26. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, et al. (2006) Population stratification confounds genetic association studies among Latinos. *Hum Genet* 118: 652-664.
27. Ziv E, Burchard EG (2003) Human population structure and genetic association studies. *Pharmacogenomics* 4: 431-441.
28. Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, et al. (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet* 112: 387-399.
29. Tsai HJ, Kho JY, Shaikh N, Choudhry S, Naqvi M, et al. (2006) Admixture-matched case-control study: a practical approach for genetic association studies in admixed populations. *Hum Genet* 118: 626-639.
30. Ribeiro-Rodrigues EM, dos Santos NP, dos Santos AK, Pereira R, Amorim A, et al. (2009) Assessing interethnic admixture using an X-linked insertion-deletion multiplex. *Am J Hum Biol* 21: 707-709.
31. Benn-Torres J, Bonilla C, Robbins CM, Waterman L, Moses TY, et al. (2008) Admixture and population stratification in African Caribbean populations. *Ann Hum Genet* 72: 90-98.
32. Choudhry S, Burchard EG, Borrell LN, Tang H, Gomez I, et al. (2006) Ancestry-environment interactions and asthma risk among Puerto Ricans. *Am J Respir Crit Care Med* 174: 1088-1093.
33. Kosoy R, Nassir R, Tian C, White PA, Butler LM, et al. (2009) Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum Mutat* 30: 69-78.
34. Bedoya G, Montoya P, García J, Soto I, Bourgeois S, et al. (2006) Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. *Proc Natl Acad Sci U S A* 103: 7234-7239.
35. Pimenta JR, Zuccherato LW, Debes AA, Maselli L, Soares RP, et al. (2006) Color and genomic ancestry in Brazilians: a study with forensic microsatellites. *Hum Hered* 62: 190-195.
36. Weber JL, David D, Heil J, Fan Y, Zhao C, et al. (2002) Human diallelic insertion/deletion polymorphisms. *Am J Hum Genet* 71: 854-862.
37. Yang N, Li H, Criswell LA, Gregersen PK, Alarcon-Riquelme ME, et al. (2005) Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine *Hum Genet* 118: 382-392.
38. Rosenberg NA, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, et al. (2005) Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genet* 1: e70.
39. Santos RV, Fry PH, Monteiro S, Maio MC, Rodrigues JC, et al. (2009) Color, race, and genomic ancestry in Brazil: dialogues between anthropology and genetics. *Curr Anthropol* 50: 787-819.
40. Sambrook J, Fritsch EF, Maniatis T (1989) Isolations of DNA from mammalian cells. In: Ford N, Nolan C, Ferguson M (eds) *Molecular Cloning*, Cold Spring Harbor Laboratory Press, New York pp. 916-919.
41. Santos NPC, Ribeiro-Rodrigues EM, Ribeiro-dos-Santos AKC, Pereira R, Gusmão L, Amorim A et al (2010) Assessing Individual Interethnic Admixture and Population Substructure Using a 48bp Insertion-Deletion (INSEL) Ancestry-Informative Marker (AIM) Panel. *Human Mutation* 31: 184-190.
42. da Costa Francez PA, Rodrigues EM, Frazão GF, Dos Reis Borges ND, Dos Santos SE (2011) Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. *Genet Mol Biol* 34: 35-39.
43. Francez PA, Ribeiro-Rodrigues EM, dos Santos SE (2012) Allelic frequencies and statistical data obtained from 48 AIM INDEL loci in an admixed population from the Brazilian Amazon. *Forensic Sci Int Genet* 6: 132-135.
44. Godinho NMO (2008) O impacto das migrações na constituição genética de populações Latino-Americanas. Tese de Doutorado do curso de Ciências Biológicas da Universidade de Brasília, DF 160.
45. Góes AC, da Silva DA, Fonseca Gil EH, da Silva MT, Pereira RW, et al. (2004) Allele frequencies data and statistic parameters for 16 STR loci-D19S433, D2S1338, CSF1PO, D16S539, D7S820, D21S1, D18S5, D13S317, D5S818, FGA, Penta E, TH0, vWA, D8S1179, TPOX, D3S1358-in the Rio de Janeiro population, Brazil. *Forensic Sci Int* 140: 131-132.
46. Grattapaglia D, Schmidt AB, Costa e Silva C, Stringher C, Fernandes AP, et al. (2001) Brazilian population database for the 13 STR loci of the AmpFISTR Profiler Plus and Cofiler multiplex kits. *Forensic Sci Int* 118: 91-94.
47. Rodrigues EM, Palha Tde J, dos Santos SE (2007) Allele frequencies data and statistic parameters for 13 STR loci in a population of the Brazilian Amazon Region. *Forensic Sci Int* 168: 244-247.
48. So-Bento M, Carvalho M, Andrade L, Lopes V, Serra A, et al. (2008) STR data for the 15 AmpFISTR1 Identifier™ loci in the Brazilian population of So Paulo State. *Forensic Sci Int Genetics Supp*: 367-369.
49. Pena SD, Bastos-Rodrigues L, Pimenta JR, Bydlowski SP (2009) DNA tests probe the genomic ancestry of Brazilians. *Braz J Med Biol Res* 42: 870-876.
50. da Costa Francez PA, Rodrigues EM, de Velasco AM, dos Santos SE (2012) Insertion-deletion polymorphisms--utilization on forensic analysis. *Int J Legal Med* 126: 491-496.
51. Avila EF, Jacques GS, Chemale G, Francez PAC (2013) Genética Forense. In: Velho JA, Geiser GC, Espindula A (eds) *Ciências Forenses "Uma introdução principais áreas da Criminalística Moderna. (2ª edição)*, Editora Millenium, Campinas-SP 227-256.