

# Relação entre a etnia auto-declarada, haplogrupo mitocondrial e ancestralidade genômica em indivíduos do sudeste brasileiro\*

## Relationship between self-declared ethnicity, mitochondrial haplogroup and genomic ancestry in individuals from southeast of Brazil

Cardena, MMSG<sup>1</sup>; Ribeiro-dos-Santos, A<sup>2</sup>; Santos, S<sup>2</sup>; Mansur, AJ<sup>3</sup>; Pereira, AC<sup>3</sup>; Fridman, C<sup>1</sup>

---

Cardena, MMSG; Santos, AR; Santos, S; Mansur, AJ; Pereira, AC; Fridman, C. Relação entre a etnia auto-declarada, haplogrupo mitocondrial e ancestralidade genômica em indivíduos do sudeste brasileiro. Saúde, Ética & Justiça. 2013;18(Ed. Especial):62-6.

**RESUMO:** Em populações onde há um alto grau de miscigenação, como no Brasil, o uso exclusivo de informações da etnia auto-declarada **não é um bom método** de classificação étnica. Neste trabalho, nós **avaliamos a relação entre as** etnias auto-declaradas com ancestralidade genômica e haplogrupos mitocondriais em 492 indivíduos do Sudeste Brasileiro. Haplogrupos mitocondriais foram obtidos pela análise das regiões hipervariáveis do DNA mitocondrial (mtDNA) e a ancestralidade genômica foi obtida utilizando 48 marcadores autossômicos informativos de ancestralidade (AIM). Dos 492 indivíduos, 74,6% se auto-declararam brancos, 13,8% pardos e 10,4% pretos. Em relação aos haplogrupos mitocondriais, 46,3% apresentaram mtDNA Africano e a maior contribuição de ancestralidade genômica foi Europeia (57,4%). Quando realizamos a distribuição do mtDNA e ancestralidade genômica de acordo com as etnias auto-declaradas, dos 367 indivíduos auto-declarados brancos, encontramos 37,6% com mtDNA Africano, sendo observado maior contribuição de ancestralidade Europeia (63,3%). Dos 68 indivíduos auto-declarados pardos, 25% apresentaram mtDNA Ameríndio e pouca diferença na contribuição de ancestralidade Europeia e Africana. Dos 51 indivíduos auto-declarados pretos, 80,4% apresentaram mtDNA Africano e maior contribuição de ancestralidade Africana (55,6%). A população brasileira apresenta uma uniformidade de ancestralidade genômica Ameríndia, e apenas o uso de marcadores genéticos (autossômico e mitocondrial) foi capaz de capturar essa informação. Sugerimos que estudos epidemiológicos façam o uso associado destes métodos, pois poderiam fornecer informações complementares.

**PALAVRAS-CHAVE:** Etnia auto-declarada; DNA mitocondrial; Ancestralidade genômica.

---

---

\* Este trabalho faz parte da Dissertação de Mestrado de Mari Maki S. G. Cardena. O artigo completo foi aceito para publicação na Plos One (2013).

<sup>1</sup> Departamento de Medicina Legal, Ética Médica e Medicina Social e do Trabalho, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brasil.

<sup>2</sup> Laboratório de Genética Humana e Medicina, Universidade Federal do Pará, Belém, Brasil.

<sup>3</sup> Laboratório de Genética e Cardiologia Molecular, Instituto do Coração, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brasil.

**Endereço para correspondência:** Dra. Cintia Fridman. Rua Teodoro Sampaio, 115. São Paulo, SP. CEP: 05405-000. e-mail: cfridman@usp.br

## INTRODUCTION

In the last few years, literature has proposed various applications to ethnicity information, as in forensics science, epidemiological studies, in clinical and pharmacological trials. However, in highly admixed populations, such as in Brazil, this personal information cannot provide the same robust estimative as in less diverse populations<sup>1,2</sup>.

Brazilian population is one of the most heterogeneous in the world. Consequently, the genotypic and phenotypic characteristics of various groups have been added to our native population<sup>3,4,5</sup>. This high rate of admixture made physical appearance characteristics such as skin and eye color, hair, shape of lips and nose not good indicators of the geographical origin of a Brazilian individual's ancestors<sup>6</sup>.

Ancestry Informative Markers (AIMs) are autosomal markers, which have been used to estimate the genomic ancestry of a population or individual, since they show differences in allele frequencies between two or more distinct populations<sup>7,8</sup>. These markers have a great advantage with respect to physical features because they are constant throughout life<sup>6,9</sup>.

Mitochondrial DNA (mtDNA) has also proved to be a good marker for the inference of maternal ethnicity. Several studies have indicated the feasibility of inferring the probable geographic origin of an individual from the sequence of the hypervariable regions (HV) of mitochondrial genome. These studies clearly demonstrate that the mitochondrial sequence alone does not determine one's ethnicity, by presenting exclusively maternal inheritance<sup>1,10</sup>.

## OBJECTIVE

The aim of this study was to evaluate the relationship between self-declared ethnicity, genomic ancestry and mitochondrial haplogroups in 492 individuals from Southeast Brazil.

## MATERIALS AND METHODS

### Population Samples

We studied 492 individuals living in São Paulo (Southeast of Brazil). The volunteers answered a questionnaire that included a multiple-choice question on self-declared ethnicity, based on the method used by the Brazilian Institute of Geography and Statistics (IBGE) national census survey, which

classify individuals as "Branços" (i.e. white), "Pardos" (i.e. brown), "Pretos" (i.e. black), "Amarelos" (i.e. yellow) and "Indígenas" (i.e. indigenous). All individuals signed an informed consent form, and the research protocol was approved by the Ethics Committee of the Clinical Hospital from Medical School, University of São Paulo.

### Mitochondrial DNA Analysis

DNA was extracted from peripheral blood leukocytes following standard salting out techniques<sup>11</sup>.

The DNA were amplified by a single PCR reaction, using primers designed using the Primer3 program. The primers used were L15879 (5'-AATGGGCCTGTCCTTG TAGT-3') and H727 (5'-AGGGTGA ACTCACTGGAACG-3'). The amplified segment refers to the nucleotide sequence 15879-727, containing 1417 base pairs (bp), comprising the three hypervariable regions of interest (HV1, HV2 and HV3). Samples were sequenced forward and reverse. Capillary electrophoresis was performed in a sequencer ABI3130 and results were analyzed using BioEdit program.

Individual sequences were compared with the Cambridge Reference sequence (rCRS)<sup>12,13</sup> using the ClustalW software. Differences found in each sequence regarding the rCRS were typed following the nomenclature recommendations<sup>14,15</sup>. Classification of mtDNA haplogroup was done using mtDNAmanager and Phylotree<sup>16</sup> programs.

### Ancestry Informative Markers Analysis

The evaluation of genomic ancestry was conducted using forty-eight biallelic AIMs type insertion-deletion (INDELS) from autosomal chromosomes, 16 markers of African ancestry, 16 markers of European ancestry and 16 Native American. They have been successfully used to assess parental genetic contribution in different populations<sup>17,18</sup>.

## RESULTS

Our sample was composed of 297 (60.4%) men and 195 (39.6%) women. The mean age of the individuals was 58 years (Standard Deviation (SD)  $\pm 14.4$ ). From a total of 492 individuals evaluated in this study, 74.6% (n=367) self-declared as white, 13.8% (n=68) as Brown, 10.4% (n=51) as Black, 1.0% (n=5) as Yellow and 0.2% (n=1) as Indigenous.

The results of the distribution of self-declared

ethnicities, mitochondrial haplogroups and genomic ancestry are present in the Table 1.

**TABLE 1.** Distribution of mtDNA and genomic ancestry according to the self-declared ethnicities

Variables	Total (n=492)	Self-Reported Ethnicities					
		White (n=367)	Brown (n=68)	Black (n=51)	Yellow (n=5)	Indigenous (n=1)	
<b>mt-DNA, n (%)</b>	<b>African</b>	228 (46.3)	138 (37.6)	47 (69.1)	41 (80.4)	1 (20.0)	1 (100)
	<b>Amerindian</b>	141 (28.7)	116 (31.6)	17 (25.0)	6 (11.8)	2 (40.0)	-
	<b>European</b>	123 (25.0)	113 (30.8)	4 (5.9)	4 (7.8)	2 (40.0)	-
<b>Genomic Ancestry, mean±SD</b>	<b>African</b>	0.283±0.217	0.220±0.169	0.410±0.212	0.556±0.236	0.204±0.154	0.562
	<b>European</b>	0.574±0.221	0.633±0.190	0.454±0.205	0.321±0.211	0.557±0.193	0.404
	<b>Amerindian</b>	0.143±0.101	0.147±0.101	0.136±0.106	0.113±0.077	0.239±0.159	0.034

## DISCUSSION

Of the 492 individuals who reported their ethnicity, 74.6% (n=367) were defined as white, suggesting a social process of “whitening”. According to the IBGE, the Brazil is composed of 48.2% of individuals self-declared as white and in the Southeast is 56.7%<sup>19</sup>. Various genetics studies with different Brazilian sub-populations had demonstrated the discrepancy between the self-declaration information and the individual’s genetic background<sup>4,20</sup>, including this one. These facts demonstrate the existence of problems in the methodology used for the acquisition of information on ethnicity in the Brazilian population, resulting in a major difficulty in using these data in epidemiological studies.

Regarding mitochondrial analysis, our results showed that 46.3% (n=228) of all of the individuals and 37.6% (n=138) of the individuals who self-reported as white presented African mtDNA (Table 1). When comparing the mtDNA in each ethnicity, there are some disparities: in individuals who self-declared as white 37.6% (n=138) had African mtDNA, while 25% (n=17) in individuals who self-declared as brown had Amerindian mtDNA (Table 1). This emphasizes the fact that in a population where there is a high degree of admixture and where superficial physical traits can vary with age and environmental factors, only the self-declaration of ethnicity is not a good method for ethnic classification<sup>6,20</sup>.

Genomic ancestry found in this study showed that Brazilians have major contribution of European ancestry (57.5% ±22.2%) (Table 1). Although these data corroborate with previous studies in relation of major contribution of genomic ancestry<sup>2,20</sup>, it is possible to observe that there were differences between our data and these studies because we found a slightly larger percentage of African genomic ancestry and a slightly lower percentage of European genomic ancestry in the black and brown individuals. These differences could be mainly due to the socio-economic bias that exists in the Brazilian population.

Interestingly, using only genetic markers (autosomal or mitochondrial) we could capture the information about the Amerindian component of the Brazilian population, even with only one individual self-declared as indigenous. We observed 14.2% (±10%) of genomic contribution of Amerindian ancestry and 28.7% (n=141) of subjects had Amerindian mtDNA when we look at the sample as a whole (Table 1).

Some studies have demonstrated that there is no strong correlation between self-declared ethnicity and genetic ancestry<sup>6,8,20</sup>, however in the present study we suggested that increased information content can be brought through the using of combined information from self-declared ethnicity with the different types of inheritance, uni-parental and bi-parental. Because using mtDNA, which provides a clear pattern of historical events that are not obscured by factors of recombination<sup>21</sup>, and

information of AIMS, which can estimate individual and population interethnic admixture<sup>7</sup>, one can have an accurate reconstitution of the genetic ancestry of a certain population or individual. Nevertheless, self-declared ethnicity can bring different information to the researcher since it is the result of visual traits like skin color combined with socio-economic and cultural aspects and is determined by factors that may not be captured by genetic markers of admixture or geographical ancestry<sup>22</sup>.

## CONCLUSION

In conclusion, this study revealed that the

genomes of most Brazilians are mixed and genetic markers are also capable of providing new valuable insights about the current structure of Brazilian population. Regardless of their skin color, the majority of individuals from Southeast of Brazil presented African matrilineages, they have a high degree of European genomic ancestry, and very uniform degree of Amerindian genomic ancestry. Due to the overall low correlation between ethnic, genetic and matrilinear information we propose that, especially in admixed populations, the use of a three dimensional construct may provide more predictive power to the study of the association between ethnicity and complex human phenotypes.

**Acknowledgements:** We thank Dayse Oliveira Alencar for technical assistance during the ancestry markers analysis.

---

Cardena, MMSG; Santos, AR; Santos, S; Mansur, AJ; Pereira, AC; Fridman, C. Relationship between self-declared ethnicity, mitochondrial haplogroup and genomic ancestry in individuals from southeast of Brazil. *Saúde, Ética & Justiça*. 2013;18(Ed. Especial):62-6.

**ABSTRACT:** In populations where there is a high degree of admixture, as in Brazil, the sole use of ethnicity self-declaration information is not a good method of ethnic classification. We evaluate the relationship between self-declared ethnicities with genomic ancestry and mitochondrial haplogroups in 492 individuals from Southeastern Brazil. Mitochondrial haplogroups were obtained by analyzing the hypervariable regions of mitochondrial DNA (mtDNA) and genomic ancestry was obtained using 48 autosomal ancestry informative markers (AIM). Of the 492 individuals, 74.6% self-declared as white, 13.8% as Brown and 10.4% as Black. In relation of mtDNA haplogroups, 46.3% presented African mtDNA and the major genomic ancestry was European (57.4%). When we performed the distribution of mtDNA and genomic ancestry according to the self-declared ethnicities, from 367 individuals self-declared white, 37.6% showed African mtDNA, and had a higher contribution of European ancestry (63.3%). The 68 individuals self-declared brown, 25% showed Amerindian mtDNA and few differences in the averages contribution of European and African ancestries. Those 51 subjects self-declared black, 80.4% had African mtDNA and the main contribution of African ancestry (55.6%). The Brazilian population had a very uniform degree of Amerindian genomic ancestry, and only by using genetic markers (autosomal and mitochondrial) we were able to capture this information. Epidemiological studies should use the association of these methods to provide complementary information.

**KEYWORDS:** Self-declared ethnicity; Mitochondrial DNA; Genomic ancestry.

---

## REFERENCES

1. Lee C, Mändouli II, Nelson CE. Inferring ethnicity from mitochondrial DNA sequence. *BMC Proc*. 2011;5:S11.
2. Pena SD, Di Pietro G, Fuchshuber-Moraes M, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One*. 2011;6:e17063.
3. Alves-Silva J, da Silva Santos M, Guimarães PE, Ferreira AC, et al. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 2000;67:444-61.
4. Giolo SR, Soler JM, Greenway SC, Almeida MA, et al. Brazilian urban population genetic structure reveals a high degree of admixture. *Eur J Hum Genet*. 2012;20:111-6.
5. Ribeiro D. O povo brasileiro: formação e sentido do Brasil. São Paulo: Companhia das Letras; 1995. p.25-257.
6. Pena SDJ. Razões para banir o conceito de etnia da medicina brasileira. *História Ciên Saúde* –

- Manguinhos. 2005;12:321-46.
7. Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE. Caucasian genes in American blacks: new data. *Am J Hum Genet.* 1992;50:145-55.
  8. Parra FC, Amado RC, Lambertucci JR, Rocha J, et al. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA.* 2003;100:177-82.
  9. Enoch MA, Shen PH, Xu K, Hodgkinson C, et al. Using ancestry-informative markers to define populations and detect population stratification. *J Psychopharmacol.* 2006;20:19-26.
  10. Egeland T, Bøvelstad HM, Storvik GO, Salas A. Inferring the most likely geographical origin of mtDNA sequence profiles. *Ann Hum Genet.* 2004;68:461-71.
  11. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16:1215.
  12. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, et al. Sequence and organization of the human mitochondrial genome. *Nature.* 1981;290:457-65.
  13. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.* 1999;23:147.
  14. Lee HY, Yoo JE, Park MJ, Chung U, et al. Mitochondrial DNA control region sequences in Koreans: identification of useful variable sites and phylogenetic analysis for mtDNA data quality control. *Int J Legal Med.* 2006;120:5-14.
  15. Salas A, Carracedo A, Macaulay V, Richards M, et al. A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics. *Biochem Biophys Res Commun.* 2005;335:891-9.
  16. van Owen M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* 2009;30:386-94.
  17. Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, Pereira R, et al. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat.* 2010;31:184-90.
  18. Francez PA, Ribeiro-Rodrigues EM, dos Santos SE. Allelic frequencies and statistical data obtained from 48 AIM INDEL loci in an admixed population from the Brazilian Amazon. *Forensic Sci Int Genet.* 2012;6:132-5.
  19. IBGE. Síntese de indicadores sociais. Uma análise das condições de vida da população brasileira. [citado em 10 nov. 2012]. Disponível em: [http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsoais2010/SIS\\_2010.pdf](http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsoais2010/SIS_2010.pdf).
  20. Lins TC, Vieira RG, Abreu BS, Grattapaglia D, et al. Genetic composition of Brazilian population samples based on a set of twenty-eight ancestry informative SNPs. *Am J Hum Biol.* 2010;22:187-92.
  21. Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. *Gene.* 1999;30:211-38.
  22. Sucheston LE, Bensen JT, Xu Z, Singh PK, et al. Genetic ancestry, self-reported race and ethnicity in African Americans and European Americans in the PCaP cohort. *PLoS One.* 2012;7:e30950.

“Financial Support: FAPESP and LIM 40”