SSM - Population Health 4 (2018) 301–306

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: www.elsevier.com/locate/ssmph

Article

Biogeographical ancestry is associated with socioenvironmental conditions and infections in a Latin American urban population

Thiago Magalhães da Silva[a,b,*], Rosemeire L. Fiaccone[c,d], Fernanda S.G. Kehdy[e], Eduardo Tarazona-Santos[f], Laura C. Rodrigues[g], Gustavo N.O. Costa[a,d], Camila A. Figueiredo[h], Neuza Maria Alcantara-Neves[h], Maurício L. Barreto[a,d]

[a] *Institute of Collective Health, Federal University of Bahia, Salvador, Bahia, Brazil*
[b] *Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié 45206-190, Bahia, Brazil*
[c] *Departamento de Estatística, Instituto de Matemática, Universidade Federal da Bahia, Salvador, Bahia, Brazil*
[d] *Center of Data and Knowledge Integration for Health, Instituto Gonçalo Muniz, Fundação Osvaldo Cruz, Salvador, Brazil*
[e] *Laboratório de Hanseníase, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil*
[f] *Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil*
[g] *Department of Infectious Disease Epidemiology, Faculty of Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom*
[h] *Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador, Bahia, Brazil*

ARTICLE INFO

Keywords:
Biogeographical ancestry
Racial inequalities
Socioeconomic status
Infections

ABSTRACT

Racial inequalities are observed for different diseases and are mainly caused by differences in socioeconomic status between ethnoracial groups. Genetic factors have also been implicated, and recently, several studies have investigated the association between biogeographical ancestry (BGA) and complex diseases. However, the role of BGA as a proxy for non-genetic health determinants has been little investigated. Similarly, studies comparing the association of BGA and self-reported skin colour with these determinants are scarce. Here, we report the association of BGA and self-reported skin colour with socioenvironmental conditions and infections. We studied 1246 children living in a Brazilian urban poor area. The BGA was estimated using 370,539 genome-wide autosomal markers. Standardised questionnaires were administered to the children's guardians to evaluate socioenvironmental conditions. Infection (or pathogen exposure) was defined by the presence of positive serologic test results for IgG to seven pathogens (*Toxocara spp*, *Toxoplasma gondii, Helicobacter pylori*, and hepatitis A, herpes simplex, herpes zoster and Epstein-Barr viruses) and the presence of intestinal helminth eggs in stool samples (*Ascaris lumbricoides* and *Trichiuris trichiura*). African ancestry was negatively associated with maternal education and household income and positively associated with infections and variables, indicating poorer housing and living conditions. The self-reported skin colour was associated with infections only. In stratified analyses, the proportion of African ancestry was associated with most of the outcomes investigated, particularly among admixed individuals. In conclusion, BGA was associated with socioenvironmental conditions and infections even in a low-income and highly admixed population, capturing differences that self-reported skin colour miss. Importantly, our findings suggest caution in interpreting significant associations between BGA and diseases as indicative of the genetic factors involved.

Introduction

Racial inequalities in health are observed for different outcomes, with race and ethnicity being traditionally considered a fundamental axis for the study of health inequalities in Europe and the United States, and more recently in Brazil (Chiavegatto Filho, Beltran-Sanchez & Kawachi, 2014; Hussey, Anderson & Berthelot, 2008). Although the socioeconomic status (SES) accounts for much of this difference, racial discrimination and putative genetic factors, in some cases, may also be involved (Collins, 2004). In Brazil, whose mixed population results from more than 500 years of physical, social and commercial contact between Europeans, Africans and Native Americans (Salzano, 2004), the classification based on skin colour is closely associated with the concept of race, and race and skin colour are used as interchangeable terms in epidemiological studies (Kabad, Bastos & Santos, 2012). Indeed, since the mid-20th century, the Brazilian Institute of Geography

---

[*] Corresponding author at: Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié 45206-190, Bahia, Brazil.
E-mail address: thiago@uesb.edu.br (T.M. da Silva).

https://doi.org/10.1016/j.ssmph.2018.03.006
Received 18 August 2017; Received in revised form 5 March 2018; Accepted 12 March 2018

and Statistics (IBGE) classifies individuals by self-reported “cor” (a Portuguese word used in Brazil) or race, a concept that involves a mix of perceptions that include ancestry, morphological traits (including skin colour) and cultural issues (Telles, 2002). However, the racial definition is dependent on a myriad of factors, such as self-perception, ascription by others, and institutional and cultural contexts (Brunsma, 2005; Jenkins, 1994). Moreover, the way individuals define themselves about their race/skin colour may not match the classification made by a third party (Telles & Lim, 1998).

Despite the issues involved with self-reported skin colour, in Brazil, as in other countries, it has traditionally been used in epidemiological studies as a proxy for biogeographical ancestry (BGA) (Pena, Di Pietro & Fuchshuber-Moraes, 2011), which is defined as proportions of the individual genome inherited from each of the continental groups that contribute to the formation of the admixed population (Royal, Novembre & Fullerton, 2010). In turn, the advent of high-throughput genotyping technologies and cataloguing of human genetic diversity in public databases have enabled the selection of ancestry informative markers (AIMs), which are used to estimate the geographical origins of one’s ancestors and the proportion of individual genome inherited from each these ancestors (Shriver, Parra & Dios, 2003). This approach provides a more objective estimate of the ancestral origins of individuals than self-reported skin colour. Furthermore, because BGA is related to self-reported skin colour, although the magnitude of this correlation varies among different populations (Lima-Costa, Rodrigues & Barreto, 2015; Parra, Kittles & Shriver, 2004), it can be used as a surrogate for genetic factors underlying the ethnoracial differences to disease risk. Indeed, in recent years, there has been increasing interest in association studies between individual ancestry inferred by genetic markers and complex diseases (Via, Ziv & Burchard, 2009). However, given the multifactorial nature of many of these outcomes, knowledge about the relationship between BGA and social determinants of health is essential to avoid being attributed to genetic variants differences to the risk of diseases that are due to non-genetic determinants that covariate with BGA (Risch, Burchard & Ziv, 2002). However, studies of the relationship between the individual BGA and SES, indicator variables of housing and living conditions and infections are scarce. Similarly, the comparison between BGA and self-reported skin colour as proxies for these factors has also been poorly investigated. The Social Changes, Asthma and Allergies in Latin America (SCAALA) cohort has a range of information on social and environmental factors that play an essential role in the development of asthma and allergies among children and adolescents (Barreto, Cunha & Alcantara-Neves, 2006). This information offers an opportunity to investigate and compare the association of BGA and self-reported skin colour with different variables. Thus, the present study aimed 1) to study the association of BGA with SES, environmental conditions and infection by pathogens; and 2) to compare the association between BGA and self-reported skin colour with these different factors.

Methods

Study population and data collection

This study was performed in the city of Salvador in Northeastern Brazil, which has 2.8 million inhabitants, of which more than 80% are self-declared admixed or black according to the last official census conducted in 2010 (SIDRA, 2016). The design of this study has been reported elsewhere (Barreto et al., 2006). Briefly, the study population consists of 1246 children previously studied to evaluate the effect of a sanitation programme and enrolled when they were 0–3 years old from 1996 to 2003 (Barreto, Genser & Strina, 2007). Standardised questionnaires were administered to the children’s guardians between 1997 and 2003 (baseline), and data were collected on housing, sanitation, and socioeconomic conditions (Strina, Cairncross & Barreto, 2003). Data collection was repeated in 2005 when the children were aged 4–11 years, at which time blood samples were collected, and serological and parasitological tests were performed. Concerning tap water and household sewage systems, we obtained data from two different time points: early life (i.e., < 3 years old) and later childhood (age range, 4–11 years).

Infection (or pathogen exposure) was defined by the presence of positive serologic test results for IgG to seven pathogens (T*oxocara spp*, *Toxoplasma gondii, Helicobacter pylori*, and hepatitis A, herpes simplex, herpes zoster and Epstein-Barr viruses) and the presence of intestinal helminth eggs in stool samples (*Ascaris lumbricoides* and *Trichiuris trichiura*). We used a threshold of 3 or fewer infections to distinguish light from heavy infection, as reported elsewhere (Janson, Asbjornsdottir & Birgisdottir, 2007).

To compare the observed associations of self-reported skin colour and BGA with the different outcomes evaluated in childhood, as well as to evaluate the concordance between BGA and self-reported skin colour, a subsample of 878 individuals with BGA information and aged between 12 and 19 years during the most recent SCAALA-Salvador survey (year 2013) was used. In this case, African BGA was categorised using tertiles. These 878 individuals answered the question “What is their race/colour?” with five response options (white, admixed, black, yellow, and indigenous or “branco”, “pardo”, “preto”, “amarelo” and “indígena” in Portuguese, respectively) and 69 self-reported white, 377 self-reported admixed, 432 self-reported blacks and no respondences self-reported yellow or indigenous. No statistically significant differences were observed between individuals with and without skin colour information regarding the proportions of ancestry as well as the different outcomes investigated (Supplementary Table 1). Ethical approval for this study was obtained from the Brazilian National Ethical Council, and written informed consent was obtained from the guardian of each child. Moreover, research has been conducted according to the principles outlined in the Declaration of Helsinki.

Genotyping and BGA estimation

DNA was extracted from peripheral blood using a commercial kit (Gentra® Puregene® Blood Kit (Qiagen)), and samples were successfully genotyped using Illumina platforms (San Diego, California) with the Omni 2.5 M array. To estimate the contribution from Africans, Europeans and Native Americans to each individual in the study population, we used the ADMIXTURE software (Alexander, Novembre & Lange, 2009). We performed unsupervised tri-hybrid ($k=3$) ADMIXTURE analyses based on 370,539 SNPs shared by samples from the HapMap Project, Human Genome Diversity Project (HGDP) (Li, Absher & Tang, 2008; Altshuler, Gibbs & Peltonen, 2010) and the study population. We did not perform the tests for association with Amerindian ancestry, as the proportion of Amerindian ancestry in our population was too low (median 5.92%, interquartile range 4.26% - 7.91%) for such tests to be meaningful.

Statistical analyses

Due to the measurement scale of BGA, all initial analyses were performed using a non-parametric approach, such as the Pearson chi-square test and Mann–Whitney test. Medians were used as a descriptive measure to characterise the distribution of BGA as a function of the different outcomes investigated. To estimate the magnitude of the association between the proportion of African/European ancestry and the different outcomes, bivariate analyses were performed using binary or multinomial logistic regression (for the politomic variables Mother Education and Income). Odds Ratio (OR) measures and their 95% confidence intervals were calculated for each 20% increase in individual BGA, using the formula $e^{(\beta*20)}$, where e is the base of the natural logarithms and β is the coefficient of binary or multinomial logistic regression. The concordance between the self-declared skin colour and African ancestry tertiles was estimated using the kappa