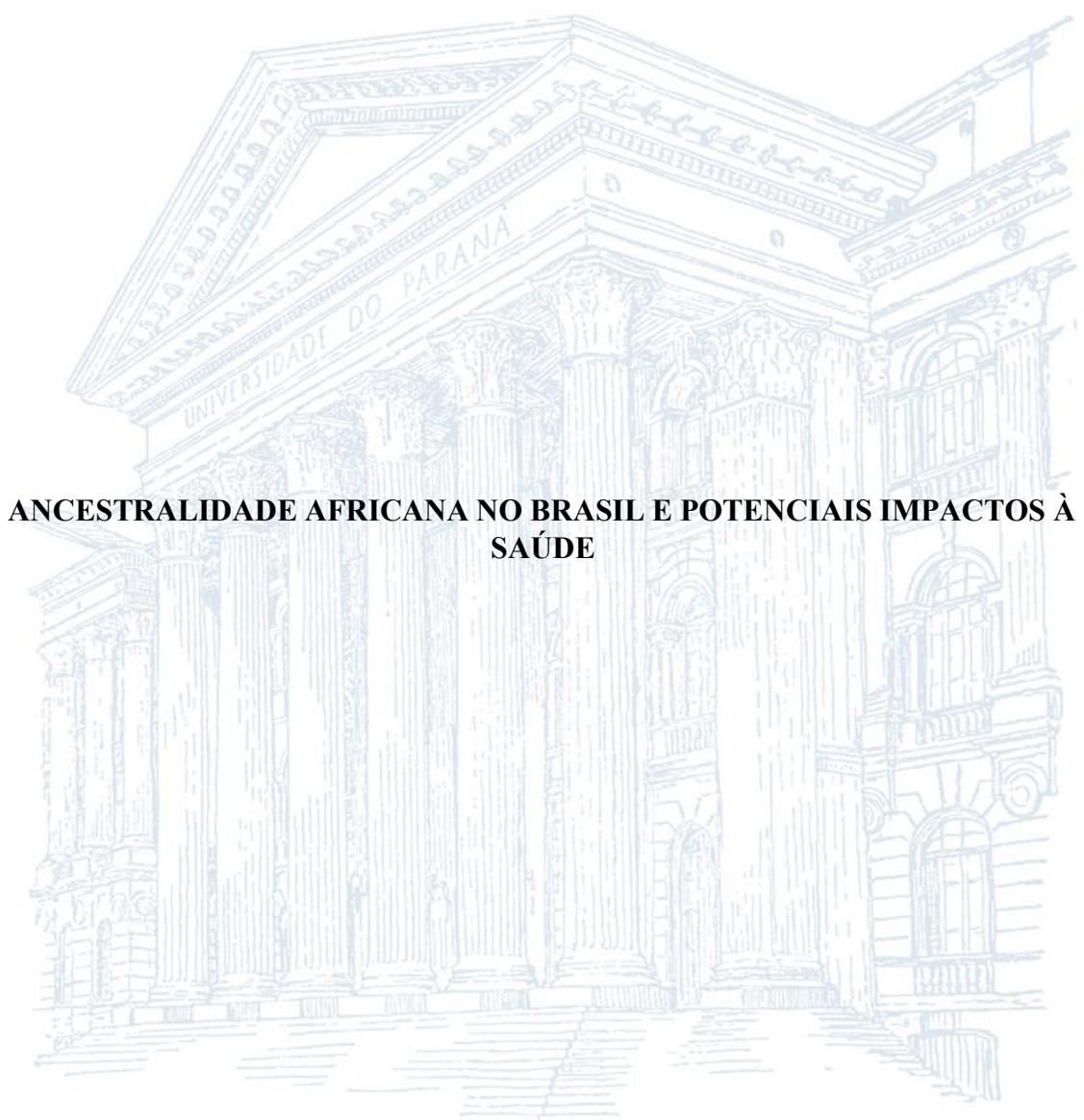


UNIVERSIDADE FEDERAL DO PARANÁ

IRIEL ARACELI JOERIN LUQUE



**ANCESTRALIDADE AFRICANA NO BRASIL E POTENCIAIS IMPACTOS À
SAÚDE**

CURITIBA

2023

UNIVERSIDADE FEDERAL DO PARANÁ

IRIEL ARACELI JOERIN LUQUE

**ANCESTRALIDADE AFRICANA NO BRASIL E POTENCIAIS IMPACTOS À
SAÚDE**

Tese de Doutorado apresentada como requisito
parcial à obtenção do grau de Doutora em
Genética ao Programa de Pós-Graduação em
Genética, Setor de Ciências Biológicas,
Universidade Federal do Paraná.

Orientadora: Prof^a. Dr^a. Marcia Holsbach Beltrame

Co-orientadora: Prof^a. Dr^a. Maria Luiza Petzl-Erler

CURITIBA

2023

AGRADECIMENTOS

Agradeço as minhas orientadoras, as Professoras Marcia Beltrame e Maria Luiza Petzl-Erler pelos ensinamentos ao longo destes anos de doutorado. Obrigada por me permitir trabalhar no Laboratório de Genética Molecular Humana e pela oportunidade de fazer parte de um projeto de pesquisa tão apaixonante.

Aos meus colegas e amigos do LGMH. Obrigada pela colaboração, pelas conversas, risadas, pelo tempo compartilhado e pelas ricas discussões durante as horas de trabalho.

Aos colegas e amigos do Programa de Pós-Graduação em Genética. Aos alunos da graduação que participaram nas diversas etapas deste trabalho.

Aos meus pais e ao meu irmão. Obrigada pelo apoio incondicional, pela paciência e pelo afeto. Por acompanhar este processo.

Ao meu namorado. Obrigada pelo afeto e apoio e por me ajudar a transitar este período com mais leveza. Obrigada pelas inúmeras colaborações com o projeto de pesquisa.

Aos meus avos pelo carinho mesmo de longe e por compreender as minhas limitações para visitá-los com a frequência que gostaria.

Às minhas amigas de Argentina pelo apoio, pelo carinho e pela colaboração.

À Universidade Federal do Paraná e aos amigos encontrados ao longo dessa caminhada.

A todos os participantes das comunidades quilombolas que fizeram possível a execução deste projeto.

Às agências de fomento CNPq, CAPES e Fundação Araucária pelo financiamento dos projetos de pesquisa. À CAPES pela bolsa de doutorado.

LISTA DE SIGLAS E ABREVIATURAS

ADN	Ácido desoxirribonucleico
ADNmt	Ácido desoxirribonucleico mitocondrial
AIM	Marcadores informativos de ancestralidade
ARN	Ácido ribonucleico
CA	Análise de correspondência
G6PD	Glicose-6-fosfato desidrogenase
Hg	Haplogrupo
HV	Região hipervariável
ka	Mil anos atrás
nt	Nucleotídeo
pb	Pares de bases
PCA	Análise de componentes principais
RC	Região controle
SNP	Polimorfismo de nucleotídeo único
STR	Repetição curta em tandem

LISTA DE FIGURAS

Figura 1 - Regiões geográficas do continente africano de acordo com o esquema da Organização das Nações Unidas.....	17
Figura 2 - Troncos linguísticos africanos e a sua distribuição geográfica no continente africano.....	19
Figura 3 - Mapa de miscigenação no continente africano quando considerados quatro componentes de ancestralidade.....	20
Figura 4 - Rotas da expansão Bantu no continente africano de acordo com a teoria de maior consenso científico.	22
Figura 5 - Rotas do tráfico transatlântico de escravizados nos anos 1500 até 1900.....	25
Figura 6 - Esquema do genoma mitocondrial humano mostrando as regiões hipervariáveis (HV).	34
Figura 7 - Esquema simplificado da filogenia do ADNmt humano.	37
Figura 8 - Mapa do Brasil mostrando o município da Lapa e a localização das três comunidades quilombolas ali presentes.....	47
Figura 9 - Esquema da região codificante do gene <i>G6PD</i> destacando a localização das variantes associadas à deficiência de G6PD.	52
Figura 10 - Esquema dos potenciais impactos em saúde da deficiência de G6PD.....	54

LISTA DE TABELAS

Tabela 1 - Sequência dos <i>primers</i> utilizados para a genotipagem de marcadores de linhagem materna e as condições de reação otimizadas para cada par alelo-específico..	60
Tabela 2 - Sequência dos <i>primers</i> utilizados para a genotipagem de marcadores de linhagem paterna e as condições de reação otimizadas para cada par alelo-específico...	62
Tabela 3 - Sequência dos <i>primers</i> utilizados na genotipagem de variantes do gene <i>G6PD</i>.	
.....	64
Tabela 4 - Combinações de <i>primers</i> utilizadas para a genotipagem de variantes do gene <i>G6PD</i> e condições de reação utilizadas.	65
Tabela 5 - Sequência dos <i>primers</i> utilizados como controles internos nas reações de genotipagem alelo-específicas.	66
Tabela 6 - Sequência dos <i>primers</i> utilizados para amplificação e sequenciamento das regiões hipervariáveis I, II e III do ADN mitocondrial (HVI, HVII e HVIII, respectivamente) e condições de reação utilizadas.	67

SUMÁRIO

1 INTRODUÇÃO	14
2 REVISÃO BIBLIOGRÁFICA	16
2.1 BREVE CONTEXTUALIZAÇÃO SOBRE A ESTRUTURA GENÉTICA HUMANA NA ÁFRICA	16
2.2 AS RAÍZES AFRICANAS DO BRASIL	24
2.3 MARCADORES GENÉTICOS E O RESTABELECIMENTO DE VÍNCULOS COM A ANCESTRALIDADE	27
2.3.1 MARCADORES UNIPARENTAIS E O ENFOQUE FILOGEOGRÁFICO	28
2.3.2 MARCADORES GENÉTICOS BIPARENTAIS	42
2.4 A POPULAÇÃO AFRO-BRASILEIRA E A RESISTÊNCIA À ESCRAVIDÃO: QUILOMBOS	44
2.4.1 A FORMAÇÃO DE QUILOMBOS NO MUNICÍPIO DA LAPA, PARANÁ	45
2.5 ASPECTOS GENÉTICOS E EVOLUTIVOS DA ANCESTRALIDADE AFRICANA NA INTERSECÇÃO COM A SAÚDE	49
2.5.1 POLIMORFISMOS NO GENE <i>G6PD</i> : ASPECTOS EVOLUTIVOS E IMPACTOS NA SAÚDE	49
3 ESTRUTURA DA TESE.....	55
4 OBJETIVOS	56
4.1 OBJETIVO GERAL	56
4.2 OBJETIVOS ESPECÍFICOS.....	56
5 MATERIAIS E MÉTODOS	57
5.1 COLETA E PROCESSAMENTO DE AMOSTRAS DE SANGUE	58
5.2 EXTRAÇÃO DE ADN.....	59
5.3 PCR ALELO-ESPECÍFICA	59
5.4 ELETROFORESE EM GEL DE AGAROSE	66
5.5 SEQUENCIAMENTO DE SANGER	67
6 RESULTADOS	70
6.1 CAPÍTULO I: Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil.....	71
6.2 CAPÍTULO II: Ancestry, diversity, and genetics of health-related traits in African-derived communities (<i>quilombos</i>) from Brazil	89
6.3 CAPÍTULO III: New clues on the formation of Afro-diasporic communities (<i>quilombos</i>) from Southern Brazil reveal African roots and a high degree of Amerindian admixture (in preparation)	104

6.4 CAPÍTULO IV: G6PD African variants prevalence correlates with ancestry and is higher in <i>Quilombos</i> from Southern Brazil (in preparation)	134
7 CONCLUSÕES.....	158
8 CONSIDERAÇÕES FINAIS.....	160
9 REFERÊNCIAS BIBLIOGRÁFICAS	161
ANEXO I - Projeto de pesquisa aprovado pelo Comité de Ética em Pesquisa e Ciências da Saúde.....	169
ANEXO II - Artigo de divulgação: Saúde de brasileiros negros em tempos de COVID-19	208

RESUMO

O povo brasileiro deriva da miscigenação de três grupos continentais principais: nativos americanos, europeus e africanos trazidos forçadamente através do tráfico transatlântico. Enquanto os ancestrais europeus foram bem documentados, as origens africanas são desconhecidas para a maioria da população, mesmo sendo o Brasil o país americano com a maior população afro-diaspórica. Africanos e os seus descendentes no Brasil contribuíram com o desenvolvimento de populações urbanas e núcleos de resistência à escravidão denominados quilombos. A fim de contribuir com o conhecimento da história populacional brasileira, assim como também das bases genéticas de doenças comuns na população negra, estudamos marcadores do cromossomo Y, do genoma mitocondrial e do cromossomo X em três regiões urbanas e dois quilombos. Estudamos a ancestralidade uniparental nas Regiões Sul (n = 43), Sudeste (n = 71) e Centro-Oeste do Brasil (n = 323). Também estudamos a ancestralidade uniparental de dois quilombos do estado de Paraná, Feijo (n = 117) e Restinga (n = 47), no sul do Brasil, região onde as raízes africanas foram historicamente invisibilizadas. Analisamos 2009 polimorfismos de nucleotídeo único (SNPs) no Y e 338 SNPs mitocondriais genotipados com microarranjo genômico nas amostras urbanas. Nos quilombos, 9 SNPs do Y e 6 SNPs mitocondriais foram genotipados por PCR-alelo específica. As amostras atribuídas ao clado mitocondrial L foram também sequenciadas nas regiões HVI-HVIII para a determinação do sub-haplogrupo (Hg) africano. Adicionalmente, estudamos a frequência de 27 variantes no gene *G6PD*, associadas à deficiência de glicose-6-fosfato desidrogenase, incluindo as variantes africanas *G6PD*A* e *G6PD*A-*, nas três regiões urbanas. As duas variantes africanas foram também estudadas pela primeira vez no Feijo e na Restinga. Revelamos miscigenação com viés de sexo em todas as populações estudadas, com a ancestralidade paterna europeia majoritária em todos os grupos e a materna mostrando proporções variáveis. No Centro-Oeste, Sudeste e no Feijo predominaram os Hgs mitocondriais ameríndios, enquanto no Sul e na Restinga predominaram os Hgs africanos. Pudemos estimar uma contribuição sub-continental materna africana predominantemente do Centro-Oeste Bantu no Centro-Oeste brasileiro. Nas regiões Sul e Sudeste a contribuição Niger-Congo não Bantu foi a mais expressiva. Inferimos a origem subcontinental africana das fundadoras do Feijo e da Restinga, detectando linhagens de provável origem no Centro-Oeste ou no Sudeste Africano Bantu no primeiro, e no Oeste Africano não Bantu no segundo. Os Hgs paternos nos quilombos foram majoritariamente do Centro-Leste ou Sul africanos. A fração ameríndia também mostrou contribuições variáveis, com frequências que sugerem a incorporação de linhagens andinas e do Brasil Central no Centro-Oeste e Sudeste brasileiros e de grupos ameríndios do sul do Brasil (Kaingang e Guarani) na região Sul e nos quilombos. Detectamos um gradiente na frequência das variantes *G6PD*A* e *G6PD*A-* da Região Centro-Oeste para o Sudeste e Sul do Brasil que contrastou fortemente com a maior frequência das mesmas nos quilombos. Nossos resultados contribuem com o conhecimento da história da diáspora africana no Brasil e trazem novos dados que podem auxiliar na compreensão das bases genéticas de doenças na população afro-brasileira. Descrevemos pela primeira vez a ancestralidade uniparental e as contribuições sub-continentais africanas e ameríndias da Região Centro-Oeste do Brasil e dos quilombos Feijo e Restinga. Ainda, nosso trabalho serve de base para que pesquisas futuras investiguem o efeito que variantes genéticas africanas possuem na saúde da população afro-brasileira.

Palavras chave: quilombos, ADNmt, Y, G6PD, diáspora africana.

RESUMEN

La población brasileña deriva del mestizaje de tres grupos continentales principales: pueblos originarios, europeos y africanos traídos en forma forzada por el tráfico transatlántico. Mientras que los ancestros europeos son bien documentados, la mayoría de la población desconoce sus orígenes africanos, aun siendo Brasil el país americano con mayor población afrodiásporica. Africanos y sus descendientes en Brasil contribuyeron con el desarrollo de poblaciones urbanas y núcleos de resistencia a la esclavitud denominados quilombos. A fin de contribuir con el conocimiento de la historia poblacional brasileña, así como también de las bases genéticas de enfermedades comunes de la población negra, estudiamos marcadores del cromosoma Y, del genoma mitocondrial y del cromosoma X en regiones urbanas y quilombos. Estudiamos la ancestría uniparental en el Sur ($n = 43$), Sureste ($n = 71$) y Centro-Oeste de Brasil ($n = 323$). También estudiamos la ancestría uniparental en dos quilombos del estado de Paraná, Feijo ($n = 117$) y Restinga ($n = 47$), en el Sur de Brasil, región donde las raíces africanas han sido históricamente invisibilizadas. Analizamos 338 polimorfismos de nucleótido simple (SNPs) mitocondriales y 2009 SNPs del Y, genotipados por *microarray* genómico, en las muestras urbanas. En los quilombos, genotipamos 6 SNPs mitocondriales y 9 SNPs del Y, y secuenciamos las regiones HVI-HVIII de las muestras asignadas al clado mitocondrial L para la determinación del sub-haplogrupo (Hg) africano. Además, estudiamos la frecuencia de 27 variantes del gen *G6PD* asociadas a la deficiencia de glucosa-6-fosfato deshidrogenasa, incluyendo las variantes africanas *G6PD*A* y *G6PD*A-*, en las tres regiones urbanas. También estudiamos, por primera vez, las variantes *G6PD*A* y *G6PD*A-* en Feijo y Restinga. Revelamos un mestizaje sexo-asimétrico en todas las poblaciones estudiadas, con la ancestría paterna europea mayoritaria en todos los grupos y la materna mostrando proporciones variables. En el Centro-Oeste, Sudeste y en Feijo predominaron los Hgs mitocondriales amerindios, mientras que en el Sur y en Restinga predominaron los Hgs africanos. Pudimos estimar una contribución sub-continental materna africana predominantemente del Centro-Oeste Bantu en el Centro-Oeste brasileño. En las regiones Sur y Sureste, la contribución Niger-Congo no-Bantu fue más significativa. Inferimos el origen sub-continental africano de las fundadoras de Feijo y Restinga, detectando linajes de probable origen en el Centro-Oeste o Sudeste Africano Bantu en el primero y en el Oeste africano no-Bantu en el segundo. Los Hgs paternos en los quilombos fueron principalmente oriundos del Centro-Este o Sur Africanos. La fracción amerindia también mostró contribuciones variables, con frecuencias que sugieren la incorporación de linajes andinos y del Brasil central en el Centro-Oeste y Sureste brasileños y de grupos amerindios del sur de Brasil (Kaingang y Guarani) en la región Sur y en los quilombos. Detectamos un gradiente en la frecuencia de las variantes *G6PD*A* y *G6PD*A-* de la región Centro-Oeste hacia el Sureste y Sur de Brasil que contrastó claramente con la mayor frecuencia de las mismas en los quilombos. Nuestros resultados contribuyen con el conocimiento de la historia de la diáspora africana en Brasil y traen nuevos datos que pueden ayudar a la comprensión de las bases genéticas de enfermedades en la población negra brasileña. Describimos por primera vez la ancestría uniparental y las contribuciones sub-continentales africanas y amerindias de la región Centro-Oeste de Brasil y de los quilombos Feijo y Restinga. Además, nuestro trabajo sirve de base para que trabajos futuros investiguen el efecto que variantes genéticas africanas ejercen en la salud de la población afro-brasileña.

Palabras clave: quilombos, ADNmt, Y, G6PD, diáspora africana.

ABSTRACT

The Brazilian population derives from the admixture between three main continental groups: Native Americans, Europeans, and Africans forcibly brought through the transatlantic traffic. While the European ancestors are well documented, African roots are unknown for most Brazilians, even with Brazil being the American country with the most expressive Afro-diasporic population. Africans and their descendants in Brazil contributed to developing urban populations and nuclei of resistance to slavery, called *quilombos*. To contribute to the knowledge of the Brazilian population's history and the genetic bases of diseases common in Black people, we studied Y chromosome, mitochondrial, and X chromosome markers in three urban regions and two *quilombos*. We studied uniparental ancestries in Southern ($n = 73$), Southeastern ($n = 41$), and Central-Western ($n = 323$) Brazil. We also examined the uniparental ancestry in two *quilombos* from the Paraná state, Feixo ($n = 117$) and Restinga ($n = 47$), in Southern Brazil. In this region, the African roots have been historically invisibilized. We analyzed 338 mitochondrial single nucleotide polymorphisms (SNPs) and 2009 Y chromosome SNPs genotyped through genomic microarray in the urban samples. In the *quilombos*, 6 mtDNA and 9 Y-SNPs were genotyped through allele-specific PCR and sequencing of the HV1-HVIII mitochondrial regions in samples assigned to clade L to resolve the African sub-haplogroup (Hg). We also studied the frequency of 27 variants in the *G6PD* gene associated with Glucose-6-phosphate dehydrogenase deficiency, including the African variants *G6PD*A* and *G6PD*A-*, in the three urban regions. Both African variants were studied for the first time in Feixo and Restinga. We reveal sex-biased admixture in all the populations studied, with paternal European ancestry predominant in all groups and maternal ancestry showing variable proportions. In Central-Western, Southeastern, and Feixo, the Amerindian maternal Hgs prevail, whereas in the South and Restinga, African Hgs were more expressive. We could estimate a maternal African sub-continental contribution mainly from Central-Western Bantu in Central-Western Brazil. The Western Niger-Congo non-Bantu contribution was more expressive in the South and Southeastern. We infer the sub-continental origin of the African women who formed Feixo and Restinga, detecting lineages most probably from Central-Western or Southeastern Bantu in the first and Western Niger-Congo non-Bantu in the second. The paternal Hgs in both *quilombos* were probably from Central-Eastern or Southern Africa. The Amerindian maternal fraction was also variable, with frequencies suggesting the incorporation of Andean lineages and from Central Brazil in the Central-Western and Southeastern regions of Brazil, and from Southern Brazilian Amerindians (Guarani and Kaingang) in the Southern region and *quilombos*. We detected a cline of *G6PD*A* and *G6PD*A-* variants frequencies between Central-Western, Southeastern, and Southern Brazil. Those frequencies contrasted sharply with the observed higher frequencies of the same variants in the *quilombos*. Our results contribute to the knowledge of the history of the African diaspora in Brazil and bring new insights that could help to understand the genetic bases of disease in the African Brazilian population. We described for the first time the uniparental ancestries and the sub-continental African and Amerindian contributions to Central-Western Brazil and the *quilombos* Feixo and Restinga. Furthermore, this work serves as a basis for future research that explores the effect that African variants have on the health of the African Brazilian population.

Keywords: *quilombos*, mtDNA, Y, G6PD, African diaspora.

1 INTRODUÇÃO

A população brasileira atual é o fruto de quase cinco séculos de miscigenação entre três grupos continentais: populações originárias americanas, imigrantes europeus e imigrantes africanos. Estima-se que aproximadamente 2,5 milhões de nativos americanos habitavam o Brasil no ano 1500, momento da chegada dos portugueses. A partir de então, aproximadamente 500.000 portugueses chegaram no Brasil (IBGE, 2007). Nos navios portugueses, cerca de 4 milhões de africanos foram trazidos forçadamente através do tráfico transatlântico de escravizados (IBGE, 2007). A partir de 1820, na tentativa de “embranquecer” a população, o governo brasileiro incentivou a migração de novas levas de europeus, permitindo também o ingresso de asiáticos (IBGE, 2007; PENA et al., 2009). Na atualidade, segundo dados do Instituto Brasileiro de Geografia e Estatística (2022), 55,9% da população brasileira se autodeclara negra. No estado do Paraná, essa porcentagem é de 33,5% da população (IBGE, 2022).

Numerosos estudos genéticos têm mostrado que não é possível predizer, ao nível individual, a ancestralidade genética de uma pessoa pela cor da sua pele ou outros traços fenotípicos comumente utilizados para esse fim, podendo existir uma ampla variação e sobreposição nas proporções dos componentes ancestrais ameríndio, africano e europeu entre pessoas com um mesmo fenótipo (PENA et al., 2009). Ainda, o uso da ancestralidade autodeclarada (denominada raça-cor no Brasil) como aproximação à ancestralidade genética de um indivíduo na medicina de precisão pode confundir o risco biológico com o risco associado ao ambiente, incluindo fatores socioeconômicos e racismo (TATONETTI; ELHADAD, 2021).

Nas populações da África encontra-se a maior diversidade genética do mundo, decorrente dos quase 200 mil anos de evolução dos humanos modernos neste continente e da sua adaptação aos seus múltiplos ambientes (GOMEZ; HIRBO; TISHKOFF, 2014). No entanto, os registros históricos existentes não nos permitem conhecer com exatidão a origem étnica ou geográfica dos africanos que formaram cada população brasileira. Isto não apenas reflete o tratamento desumano ao que estes eram submetidos durante o período colonial e gera uma dívida histórica com os descendentes dessas pessoas, traz também importantes consequências no presente, principalmente relacionadas à abordagem da saúde dessa fração majoritária da população brasileira. Hoje existe um número crescente de estudos que relacionam a ancestralidade com fenótipos complexos, como susceptibilidade a doenças e

resistência aos fármacos (CAMPBELL; TISHKOFF, 2008). Esses estudos, contudo, quando realizados em populações de um continente não podem ser extrapolados a outros (SIRUGO; WILLIAMS; TISHKOFF, 2019). Isto soma-se ao pronunciado viés existente nas pesquisas em saúde, as quais são voltadas majoritariamente à população europeia e nas quais africanos e outras minorias estão sub representados (PEPRAH et al., 2015). Estudos recentes demonstram inclusive que, devido à presença de estrutura genética sub-continental, existe uma limitação para extrapolar resultados nesse nível (BELBIN et al., 2021). Assim, a falta do marco teórico adequado somada ao racismo estrutural¹ faz com que brasileiros de ancestralidade africana não recebam uma adequada atenção à saúde (BATISTA; KALCKMANN, 2005). Por este motivo, a elucidação das populações ancestrais da diáspora africana no Brasil, assim como das suas particularidades genéticas e fenotípicas, torna-se essencial para que possam ser aplicadas as abordagens corretas de aconselhamento genético e atenção à saúde da população brasileira.

A presente tese de doutorado teve como objetivo investigar a diversidade genética da população afro-brasileira, a fim de rastrear suas origens às populações africanas específicas e contribuir para o conhecimento da sua história. Também, buscou-se estudar variantes população-específica com potenciais impactos na saúde dessa população.

¹ De acordo com o filósofo Silvio Almeida (2019), existem três níveis de racismo na sociedade: o racismo individual, que se reconhece por atos de preconceito contra o indivíduo, o racismo estrutural, que está fundamentado numa sociedade racista e tem um caráter político. O racismo estrutural é exercido através de práticas sociais e institucionais que atuam na imposição de regras e padrões sociais com a finalidade de atribuir privilégios a um grupo não racializado (os brancos, no caso brasileiro). Todas essas práticas se traduzem no terceiro tipo: o racismo institucional, originando desvantagens para o grupo discriminado (ALMEIDA, 2019).

2 REVISÃO BIBLIOGRÁFICA

2.1 BREVE CONTEXTUALIZAÇÃO SOBRE A ESTRUTURA GENÉTICA HUMANA NA ÁFRICA

O continente africano ocupa quase um quarto do território terrestre (GUTIÉRREZ, 2006; HALL, 2005; PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015; SLENES, 1992). Ele pode ser dividido geograficamente (de acordo com uma dentre várias classificações) em cinco grandes regiões: a África Setentrional ou Norte, Ocidental ou Oeste, Oriental ou Leste, Austral ou Sul e Central ou Centro-Oeste (Figura 1) (Organização das Nações Unidas, <https://unstats.un.org>). O continente abrange uma faixa de ambientes extremamente diversa que inclui desertos, savanas e florestas tropicais (CAMPBELL; TISHKOFF, 2010). As populações autóctones africanas são caracterizadas por altos níveis de diversidade genética dentro e entre populações, sendo a maior encontrada no mundo (CAMPBELL; TISHKOFF, 2010).

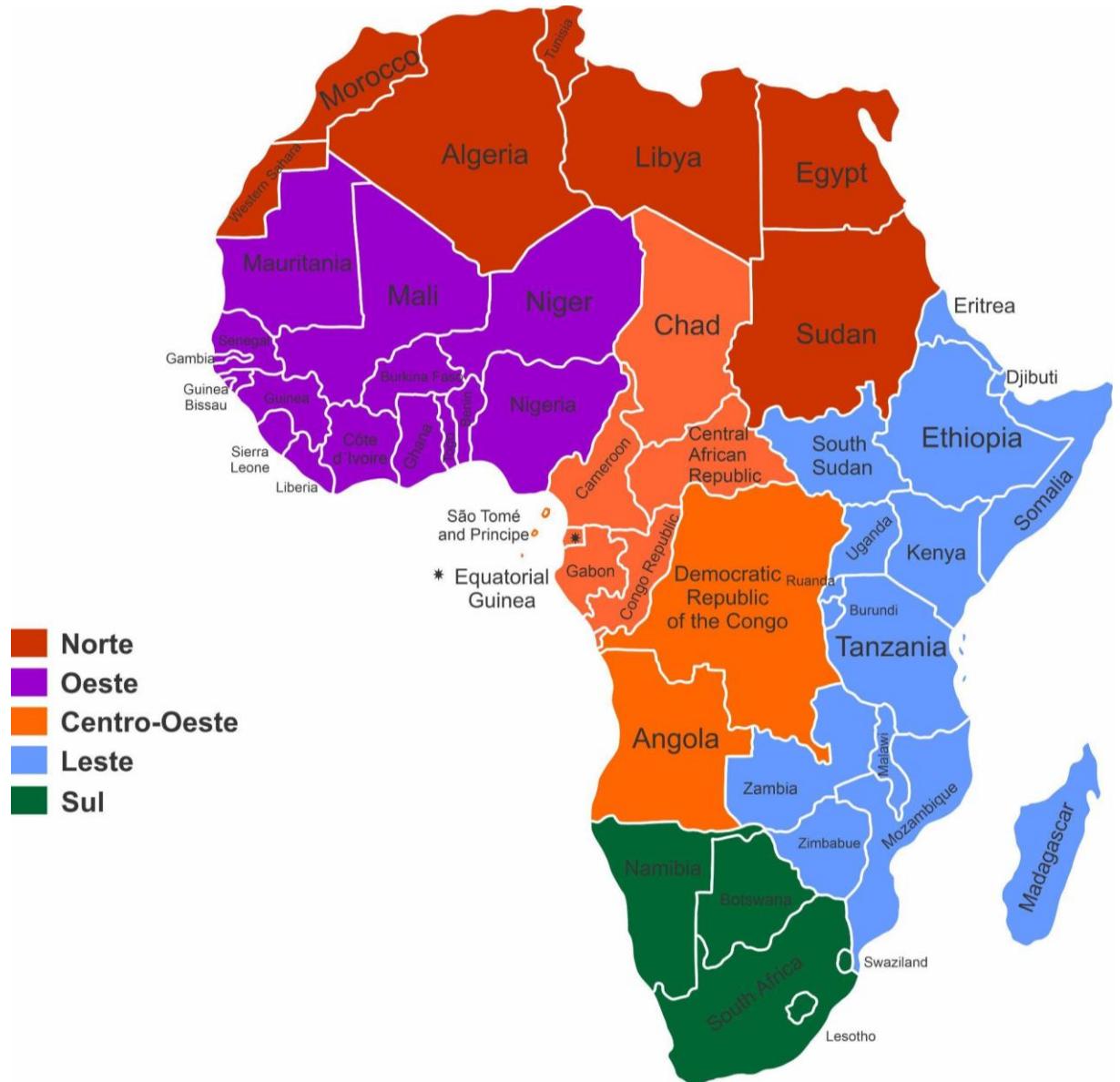


Figura 1 - Regiões geográficas do continente africano de acordo com o esquema da Organização das Nações Unidas. Fonte: A autora (2023)

Em estudos antropológicos, uma importante ferramenta para a diferenciação de populações são os estudos linguísticos (JOBLING et al., 2014). Na África, são faladas hoje mais de 2000 línguas, um indício da diversidade de povos abrigados neste vasto continente. Estas são classificadas em quatro grandes famílias linguísticas. A família Níger-Cordofaniana (ou Níger-Congo) é oriunda do Oeste Africano. É a mais diversa e amplamente distribuída, abrange línguas que são faladas predominantemente por populações agriculturalistas da África Subsaariana (DE FILIPPO et al., 2011). Deste grupo deriva o grande conjunto de línguas Bantu (que abrange mais de 500 línguas), hoje faladas por povos em sua maioria agriculturalistas, más

também por grupos pastoralistas e caçadores-coletores (BARBIERI et al., 2014b). A segunda grande família é a de línguas Nilo-Sarianas, faladas por grupos pastoralistas do Centro e do Leste Africanos. Já o conjunto de línguas Afro-Asiáticas é falado por populações agro-pastoralistas e pastoralistas no norte e no leste do continente. Por último, Khoisan (abreviatura dos denominadores de grupos étnicos Khoi-Khoi e San) denomina um agrupamento de línguas que não pertencem a outras famílias linguísticas e que compartilham o uso de consoantes de tipo “click” e outros fonemas, porém sem representar uma família linguística no sentido estrito (BARBIERI et al., 2014a). Línguas Khoisan são faladas por populações caçadoras-coletoras e pastoralistas do sul da África, as quais são caracterizadas por uma ampla diversidade linguística, cultural e genética (BARBIERI et al., 2014a). Outros grupos linguísticos, como o Indo-europeu ou o Malaio-polinésio, também falados no continente, não representam populações autóctones africanas e por isso não serão discutidas aqui (Figura 2).

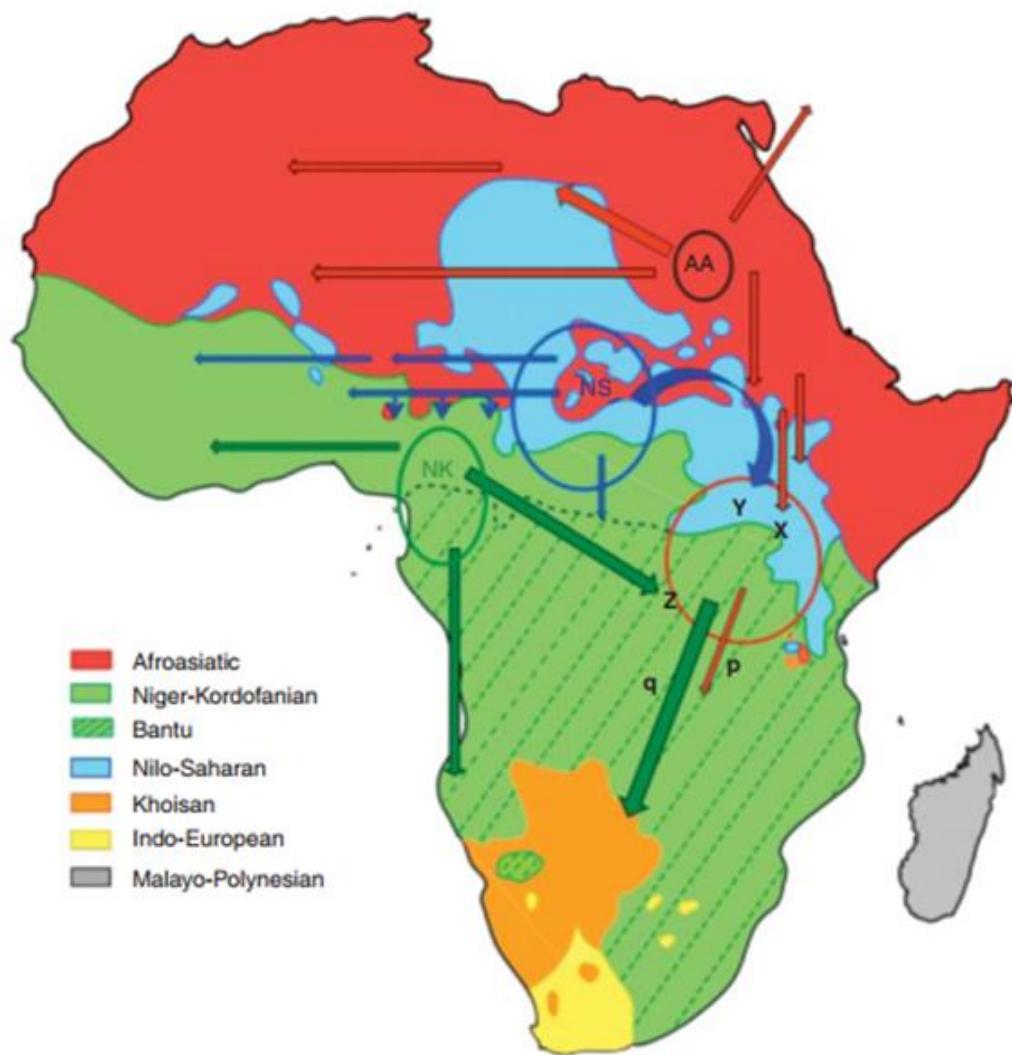


Figura 2 - Troncos linguísticos africanos e a sua distribuição geográfica no continente africano.
AA: Afro-Asiático. NS: Nilo-Saariano. NK: Níger-Congo. Fonte: Gomez et al. (2014).

Os estudos sugerem que muito antes da saída dos humanos modernos da África (cerca de 100.000 anos antes daquele evento), o continente africano já apresentava marcada estrutura genética (SILVA et al., 2015). A mesma foi determinada por oscilações climáticas e pelas marcadas barreiras geográficas presentes no continente, como o deserto do Saara - ao norte -, o mar Vermelho - no Leste -, as florestas tropicais centro-africanas e o deserto de Kalahari - ao sudoeste do continente, abrangendo a maior parte de Namíbia, Botswana e África do Sul (PFENNIG et al., 2023; SILVA et al., 2015). Assim, fazendo uma análise geral do continente, pode-se diferenciar quatro componentes de ancestralidade que se sobrepõem, parcialmente, com a distribuição de famílias linguísticas (Figura 3). Na região Norte da África há uma ancestralidade mais próxima à Euroasiática, que está relacionada à proximidade com o Oriente

Médio e a Europa, assim como também a povos falantes de línguas Afro-Asiáticas (PFENNIG et al., 2023). Na região Oeste observa-se um componente próprio daquela região, que representa muitos povos falantes de línguas Níger-Congo não Bantu (PFENNIG et al., 2023). Já a partir da região do Congo e na direção sudoeste do continente começa-se a perceber o componente Khoisan, que está restrito a essa região (PFENNIG et al., 2023). Por último, um componente Leste-Africano se estende desde o chifre da África (Somália, Etiópia, entre outros países), decrescendo até o noroeste do continente e na direção sul (PFENNIG et al., 2023).

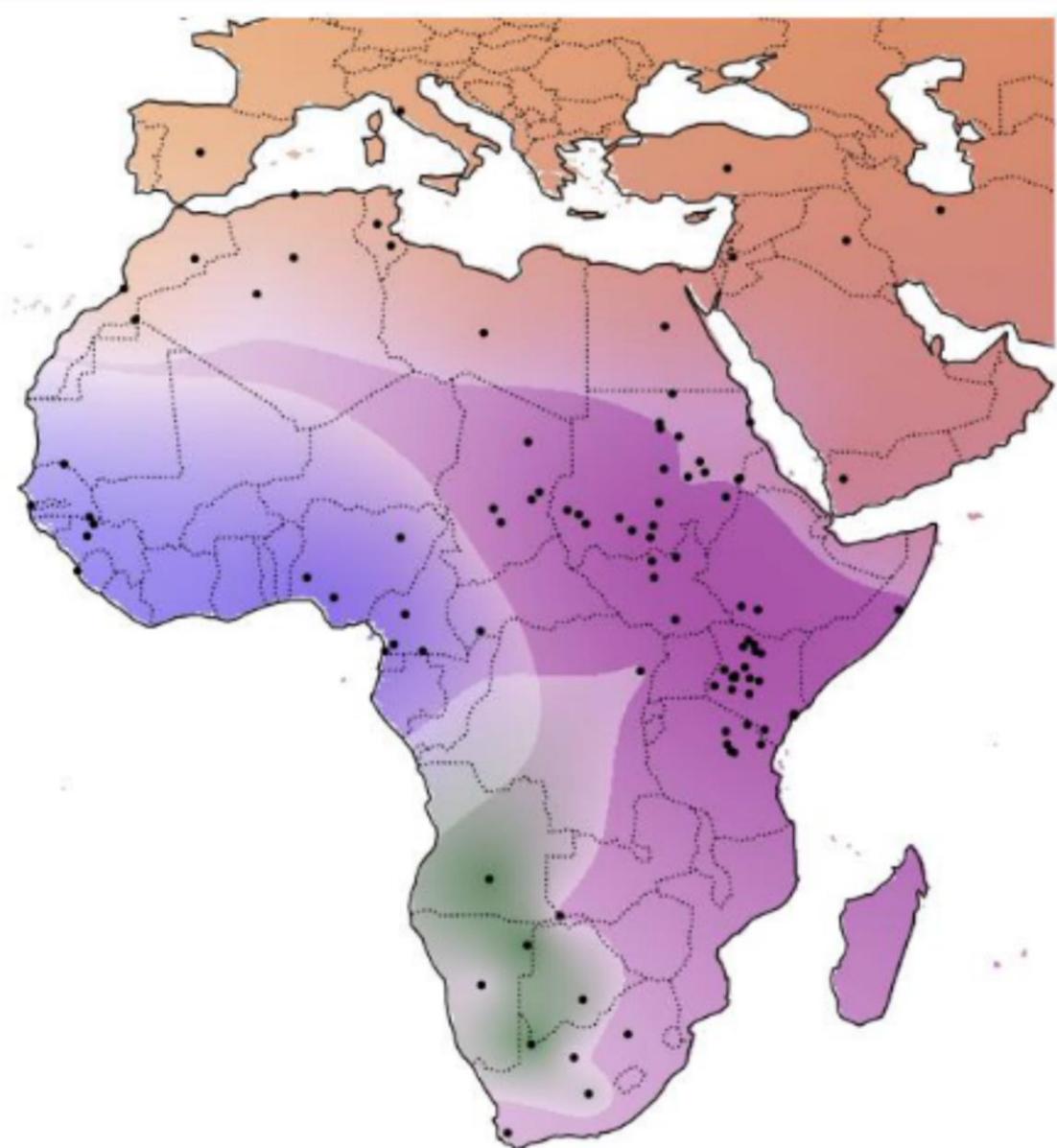


Figura 3 - Mapa de miscigenação no continente africano quando considerados quatro componentes de ancestralidade. Laranja: tipo Euroasiático, Azul: Oeste Africano, Roxo: Leste Africano, Verde: Khoisan. Fonte: Pfenning et al. (2023)

Um dos principais determinantes da estruturação genética africana atual, e de especial relevância para o nosso estudo, é a expansão de populações de língua Bantu, originárias do Centro-Oeste Africano, para o Leste e Sul (SILVA et al., 2015). As populações Bantu ocupam hoje a maior parte da África Subsaariana e falam mais de 500 línguas derivadas desse tronco linguístico (Figura 2).

Acredita-se que a migração dos Bantu tenha começado na região dos gramados localizados entre o sudeste da Nigéria e o oeste de Camarões, não mais de 5 ka (PEREIRA et al., 2001) (Figura 4) já que naquela região, o tronco Bantu apresenta a maior diversidade linguística (BARBIERI et al., 2014b). A expansão Bantu está associada com a inovação de práticas tecnológicas como a metalurgia e a agricultura, utilizadas por estes povos, que eram até então coletores e/ou pastoralistas, assim como também de um estilo de vida mais sedentário (DE FILIPPO et al., 2011; KOILE et al., 2022). As datas e rotas precisas da sua expansão, entretanto, são objeto de inúmeros debates e estudos (KOILE et al., 2022). Diferenças linguísticas entre as populações Bantu do Leste e do Oeste Africano suportam a ocorrência de duas rotas migratórias no passado (PFENNIG et al., 2023). Línguas Bantu do Leste se distribuem na região leste dos Grandes Lagos, desde arredores do Lago Victoria até a atual África do Sul (KOILE et al., 2022). As línguas Bantu do Oeste abrangem o território desde o Golfo de Guiné até o norte da atual Namíbia, ao sul do continente (KOILE et al., 2022).

Duas hipóteses principais têm sido propostas para explicar a expansão Bantu. Elas têm sido chamadas de hipóteses da separação inicial (“*early split*”) e da separação tardia (“*late split*”) (PFENNIG et al., 2023). A hipótese com maior consenso científico é a da separação tardia, que afirma que os Bantus teriam seguido uma rota de expansão inicial através da floresta tropical da África Central (região do Congo e arredores) até às atuais Angola e Zâmbia, para depois se dividir em dois grupos. O primeiro grupo teria continuado na direção sul (para as atuais Botswana e África do Sul), em torno de 3.500 anos atrás (3,5 ka), e outra na direção dos Grandes Lagos do Leste Africano em torno de 2,5 ka (GOMEZ; HIRBO; TISHKOFF, 2014; PATIN et al., 2017; PFENNIG et al., 2023) (Figura 4). Na região de Uganda, os migrantes teriam permanecido por um período de alguns milhares de anos até finalmente se deslocar na direção sul, chegando a Moçambique cerca de 1,8 ka (SALAS et al., 2002). Durante a travessia pelo Leste Africano, os grupos Bantu teriam se encontrado com as populações autóctones de lá, como os povos nilóticos e cuxitas (ambos do tronco Afro-asiático) (SILVA et al., 2015). Como consequência da expansão Bantu, povos dos grupos étnicos Khoisan mantiveram-se restritos ao Sudoeste Africano, na região do deserto de Kalahari, e conservaram a sua identidade

étnica e linguística (SALAS et al., 2002). Outra consequência foi a reclusão dos povos caçadores-coletores da África central nas profundidades das florestas tropicais daquela região (SALAS et al., 2002). Alguns destes grupos, entretanto, chegaram a se miscigenar com os Bantu, inclusive adotando suas línguas (SALAS et al., 2002). Contudo, essa miscigenação foi assimétrica, tanto em termos populacionais como sexuais, como o demonstra a maior proporção de ancestralidade Bantu (predominantemente paterna) em grupos caçadores-coletores centro-africanos do que ancestralidade destes em populações Bantu do Oeste Africano (PATIN; QUINTANA-MURCI, 2018; PFENNIG et al., 2023).

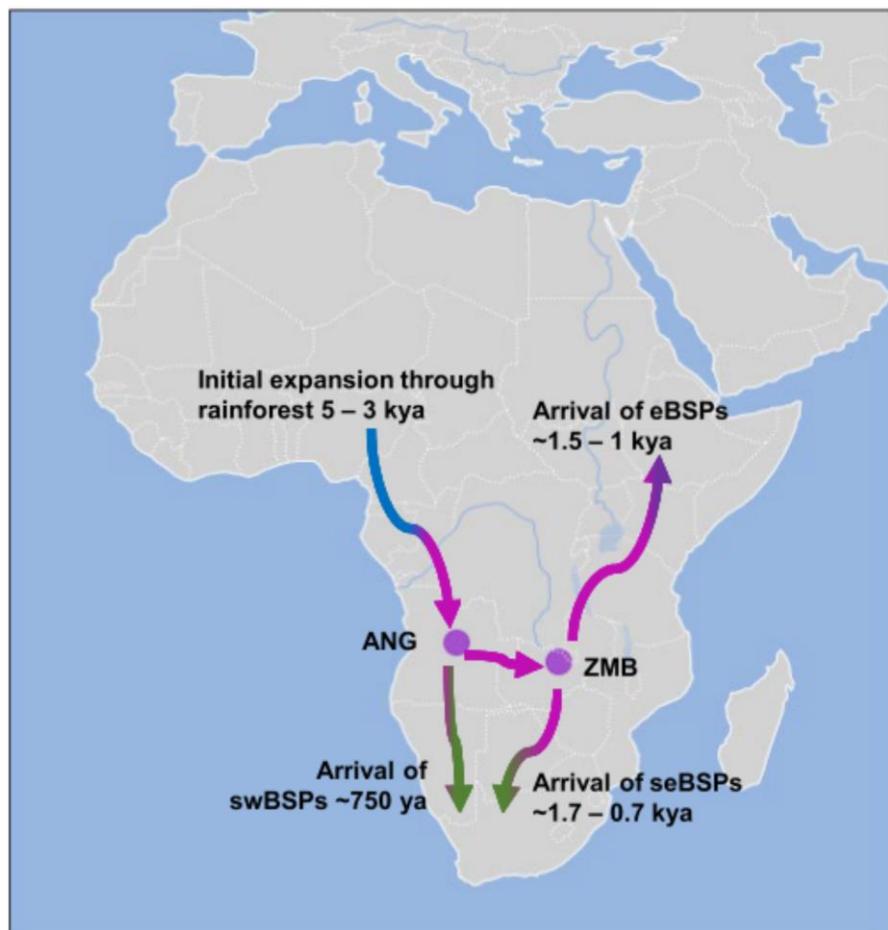


Figura 4 - Rotas da expansão Bantu no continente africano de acordo com a teoria de maior consenso científico. kya: mil anos atrás; ANG: Angola; swBSPs: populações Bantu do Sudoeste; eBSPs: populações Bantu do Leste. ZMB: Zâmbia; seBSPs: populações Bantu do Sudeste.

Fonte: Pfenning et al. (2023)

As populações Bantu contemporâneas na região Sul Africana possuem estilos de vida diversos, que compreendem pastoralismo, caça e coleta, e agricultura, e falam línguas derivadas tanto dos Bantu do Leste quanto do Oeste (BARBIERI et al., 2014b). Estudos recentes demonstraram que povos Bantu na África do Sul, onde cerca de 80% da população pertence ao tronco Leste-Bantu, possuem entre 1,5-20% de ancestralidade Khoisan, possivelmente derivada de múltiplos eventos de miscigenação (SENGUPTA et al., 2021). Já no Sudeste Africano, os grupos Bantu do Leste parecem ter substituído as populações caçadoras-coletoras, a julgar pela alta proporção de ancestralidade Bantu ($\geq 97\%$) observada em Malawi e Moçambique (PFENNIG et al., 2023).

Nos últimos 500 anos (séculos 15 a 19), outro evento impactou fortemente a estrutura genética do continente africano. O tráfico europeu de africanos causou alterações na estrutura populacional do continente, levando milhões de africanos para Europa e América mas também causando a relocalização de pessoas de regiões mais afastadas da costa africana para os arredores da costa ocidental e para ilhas do Atlântico (PRIMATIVO et al., 2017; RODNEY, 1975). Nesse período, o tráfico de pessoas entre as diversas regiões africanas também foi frequente, aumentando ainda mais a complexa estrutura populacional no continente e alterando as suas frequências alélicas e de linhagens uniparentais (PRIMATIVO et al., 2017).

2.2 AS RAÍZES AFRICANAS DO BRASIL

Os estados nacionais da América como a conhecemos hoje foram construídos pelas elites brancas e mestiças pertencentes às classes dominantes (NASCIMENTO, 2016; STAVENHAGEN, 2010). Durante esse processo, a exclusão proposital das populações originárias nativas americanas e dos africanos e os seus descendentes, mesmo estes constituindo as maiorias demográficas, tem resultado em sociedades estruturalmente racistas onde esses povos ocupam os lugares com os piores índices socioeconômicos (LEITE, 1999; STAVENHAGEN, 2010). Em muitas sociedades americanas, este processo tem resultado também na negação e no esquecimento das suas raízes étnicas (STAVENHAGEN, 2010).

Para poder nos aprofundar nos relatos históricos que podem trazer pistas sobre as origens dos africanos que formaram o Brasil precisamos conhecer todas as possíveis rotas de introdução de alelos africanos, tanto as diretas quanto as indiretas. Entre elas, o tráfico transatlântico de escravizados africanos foi a mais significativa. O comércio de africanos pelos portugueses teve início mais de um século antes dos europeus invadirem a América (HALL, 2005). Já em 1444, nas primeiras viagens portuguesas para a costa atlântica africana em busca de ouro, africanos eram atacados, sequestrados e levados para Portugal para a sua comercialização (HALL, 2005). Entretanto, a influência africana (tanto no nível genético quanto cultural) na região do Mediterrâneo e, em especial, na Península Ibérica e na Itália, era expressiva ainda antes do período colonial (CEREZO et al., 2012). Isto por causa das expansões históricas de muçulmanos africanos naquela região desde 1805 e durante quase 400 anos (GONZÁLEZ et al., 2003; SECHER et al., 2014). Inclusive, alguns estudos genéticos sugerem a ocorrência de fluxo gênico do Norte Africano ainda antes do período Neolítico (12-6,5 ka na África) (GONZÁLEZ et al., 2003; HALL, 2005; PFENNIG et al., 2023; SECHER et al., 2014).

Durante o período escravocrata, que, no Brasil, se estendeu entre os séculos XV e XIX, aproximadamente 13 milhões de pessoas foram forçadas a deixar a África e vir para a América (SALAS et al., 2004). Estimativas apontam que apenas 11 milhões delas sobreviveram à viagem transatlântica, e dentre elas, muitas faleceram durante os primeiros anos de escravidão (SALAS et al., 2004). Embora o Brasil seja o país onde o tráfico atingiu as maiores dimensões (cerca da metade do total) é onde ele foi menos documentado (BUCCIFERRO, 2013).

Os portugueses mantiveram uma posição permanente em Angola desde o início do século XVI (KLEIN, 2002). Assim, conservaram um fluxo constante de escravizados para o Brasil desde os portos de Luanda, Benguela e Cabinda, trazendo a maior parte dos escravizados daquela região, juntamente com a região do Congo, ambas no Centro-Oeste Africano (KLEIN, 2002). O Oeste Africano (principalmente a Baía de Benin, na região de Gana e Nigéria) teve um tráfico intenso durante o século XVI e início do século XVII, assim como também no início do século XVIII e entre 1770 e 1790 (KLEIN, 2002; SÁ JUNIOR, 2016). Com a proibição do tráfico pela Inglaterra, em 1845, o tráfico de Moçambique passou a ser mais expressivo devido ao menor alcance do controle inglês das águas ao sudeste do continente africano (REIS, 2007) (Figura 5). É por isso que, durante o século XIX, Angola e Moçambique passam a ser as regiões mais exploradas (REIS, 2007; SÁ JUNIOR, 2016). Contudo, é importante destacar que, embora cada período tenha tido uma região de maior protagonismo, o tráfico continuava acontecendo das demais regiões simultaneamente (SÁ JUNIOR, 2016).

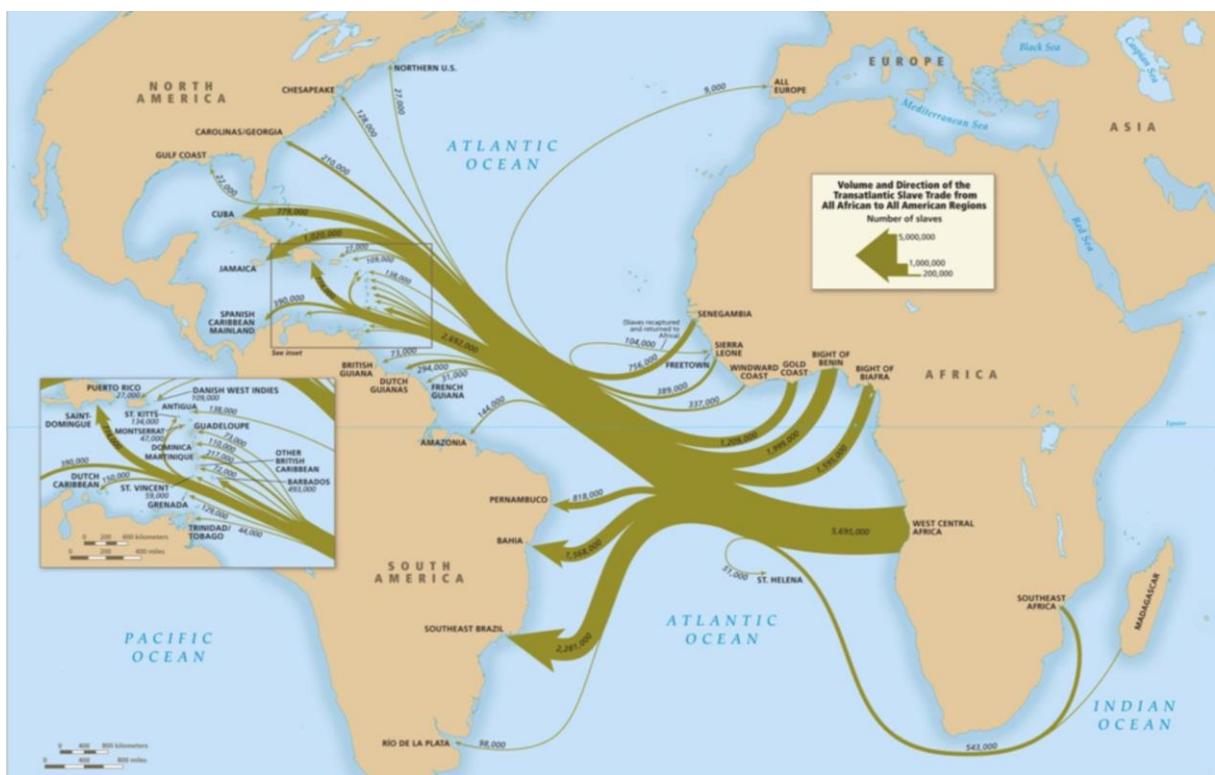


Figura 5 - Rotas do tráfico transatlântico de escravizados nos anos 1500 até 1900. Fonte: Eltis e Richardson (2010)

As redes que envolviam o tráfico de pessoas estavam organizadas de forma a minimizar e facilitar a viagem transatlântica para os europeus (HALL, 2005). Assim, o Sistema Atlântico

Norte conectava a região de Senegâmbia e região norte de Guiné com os Estados Unidos, o Caribe e o Nordeste brasileiro, enquanto o Sistema Atlântico Sul conectava a África Central, especialmente Angola e Moçambique, com o Sudeste brasileiro e o Rio de la Plata (HALL, 2005). A região de Senegâmbia e Serra Leoa, no extremo norte do Oeste-Africano, entretanto, teve um menor envolvimento com o tráfico de escravizados (de acordo com os registros históricos) (KLEIN, 2002).

Estimativas sugerem que do oeste da África vieram para a América aproximadamente 8 milhões de pessoas, do Centro-Oeste (Camarões até Angola) 4 milhões, e da região de Moçambique e Madagascar, no Sudeste, cerca de 1 milhão. Entretanto, a origem exata -quanto ao grupo étnico ou região geográfica- dessas pessoas é desconhecida para a grande maioria dos seus descendentes tanto no Brasil quanto no restante da América (FORTES-LIMA; VERDU, 2021; SLENES, 1992). Essa lacuna histórica é consequência, em parte, do trato desumanizante dado aos africanos, cujas identidades e origens não tiveram um registro detalhado (HALL, 2005). Quando existentes, esses registros mencionaram o trajeto realizado pelos navios negreiros, em adição a cifras e termos gerais de classificação dos africanos, sempre sob a perspectiva do colonizador europeu (HALL, 2005). A perspectiva do europeu, contudo, tendeu a apagar a multiplicidade de origens, culturas e línguas dos africanos nos seus registros, agrupando-os de acordo com a região de onde eles eram embarcados (GOMES, 2015; SLENES, 1992). Mas o tráfico tinha uma complexa estrutura dentro da África, portanto os locais de embarque nos navios negreiros não refletem necessariamente o local de origem das pessoas (SLENES, 1992). Estas muitas vezes eram trazidas de regiões muito afastadas da costa (SLENES, 1992). Além disso, após a abolição da escravidão no Brasil, em 1888, os limitados registros históricos existentes foram queimados na tentativa do governo brasileiro de evitar ações indenizatórias por parte dos escravocratas brasileiros (LACOMBE; SILVA; BARBOSA, 1988).

De acordo com estudos históricos, os africanos vindos para o Brasil teriam sido majoritariamente Bantus (GUTIÉRREZ, 2006; HALL, 2005; PINHEIRO; ASSUMPCÃO; FOLLMANN, 2015; SLENES, 1992). Essa classificação, entretanto, engloba milhares de grupos étnicos com distintas práticas de subsistência e estabelecidos em ambientes diversos por toda a África Subsaariana (GOMEZ; HIRBO; TISHKOFF, 2014). Já em termos geográficos, a principal origem seria a região de Angola e Congo, de onde os escravizados eram enviados para a América através do porto de Luanda (Angola), o mais movimentado da região durante todo o tráfico (HALL, 2005). É interessante destacar que, de acordo com alguns autores, as línguas e

práticas culturais dos povos naquela região eram muito próximos entre si (HALL, 2005). De acordo com alguns autores, dentre os africanos trazidos para o Brasil teria havido um predomínio do sexo masculino (3 de cada 4 africanos) (PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015; SLENES, 1992).

Considerando a estrutura política africana daquela época, cinco grandes reinos deram origem à diáspora africana no Brasil. São estes os três grandes reinos denominados “Sudaneses” pelos portugueses: o Reino de Gana (localizado na região da atual Mauritânia, sudeste do Mali e norte do Senegal), o Reino de Mali (localizado ao redor do reino de Gana e estendendo-se até o oceano Atlântico) e o Reino Songhai (atual Mali, Níger, Benim, Burkina Faso, Gâmbia, Guiné, Mauritânia e Senegal); e dois reinos Bantu: Kongo e Ndongo (PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015). A maioria dos Sudaneses vindos para o Brasil seriam falantes das línguas Iorubá e Hauçã, da Nigéria; os denominados Jeje, do Daomé (parte do atual Benim), e os Fanti e os Axanti da Costa do Ouro (atual República de Gana). A exceção da língua Hauçã, de origem Afro-Asiática, as demais pertencem à família Nígero-Cordofaniana (PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015). Dentre os reinos Bantu, o Reino de Kongo estava localizado na região norte da atual Angola, na República Democrática do Congo e na República do Congo, enquanto o Reino Ndongo situava-se na região que hoje ocupa Angola (GOMEZ; HIRBO; TISHKOFF, 2014; PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015; REECE, 2005). Dentre eles, o Reino de Ndongo teria sido o de maior contato com os traficantes portugueses (HALL, 2005).

2.3 MARCADORES GENÉTICOS E O RESTABELECIMENTO DE VÍNCULOS COM A ANCESTRALIDADE

Devido à sua variação interpopulacional, certos marcadores genéticos apresentam um grande potencial na reconstrução da história demográfica e evolutiva das populações humanas (UNDERHILL; KIVISILD, 2007). No contexto da formação da população brasileira, o uso de marcadores genéticos constitui uma ferramenta muito valiosa e com potencial para preencher, ao menos parcialmente, as lacunas históricas (TARAZONA SANTOS et al., 2016).

Os marcadores genéticos podem ser classificados de acordo com a sua taxa evolutiva (ou o número de mutações que ocorrem naquela região do genoma em uma certa quantidade de tempo). Com base nisso podemos descrever marcadores com velocidade de mutação lenta,

como são os polimorfismos de nucleotídeo único (SNPs, pelo nome em inglês *single nucleotide polymorphism*), inserções e deleções (ou indels) ou elementos retrotransponíveis, e marcadores de velocidade de mutação rápida, como as repetições em tandem de número variável (VNTRs, pelo nome em inglês *variable number of tandem repeats*) micro- e minissatélites (DE KNIJFF, 2000). Quando se avalia o uso de marcadores genéticos para o estudo de ancestralidade deve-se considerar aqueles polimorfismos cujas frequências nas populações ancestrais diferem consideravelmente (PFAFFELHUBER et al., 2020). Os marcadores de taxa evolutiva lenta são os mais adequados para análises de ancestralidade já que, apesar de apresentar menor diversidade alélica na população, possuem menor probabilidade de mutação recorrente e, portanto, maior potencial de apresentar diferença de frequência em populações diferentes (SANTOS; TYLER-SMITH, 1996). A presença de mutação recorrente dificulta a distinção entre alelos que são idênticos por descendência ou idênticos por estado, podendo levar a erros na inferência de ancestralidade (SANTOS; TYLER-SMITH, 1996).

2.3.1 MARCADORES UNIPARENTAIS E O ENFOQUE FILOGEOGRÁFICO

Os marcadores uniparentais são aqueles encontrados no genoma mitocondrial e na região específica masculina do cromossomo Y (MSY, do inglês, *Male Specific Y*). O seu modo de herança é exclusivamente uniparental e não sofrem recombinação intercromossómica, portanto a única variação que acumulam decorre de mutação (e recombinação intracromossómica no caso do cromossomo Y), sendo transmitidos praticamente intactos de geração para geração (UNDERHILL; KIVISILD, 2007). O seu estudo permite a inferência da história matri- e patrilinear, respectivamente, de uma população atual (UNDERHILL; KIVISILD, 2007). É importante destacar também que, pelo seu menor tamanho efetivo populacional (1/4 do tamanho efetivo de marcadores autossômicos, por exemplo) também são mais suscetíveis aos efeitos da deriva genética. Desta forma as suas linhagens (alelos) são mais facilmente perdidas pela população, o que acarreta a perda de parte da contribuição genética ancestral da população e acelera a diferenciação entre populações (JOBLING; TYLER-SMITH, 2003).

A combinação de estados alélicos encontrados em um determinado cromossomo é denominada haplótipo, sendo que um conjunto de haplótipos filogeneticamente relacionados formam um clado definido que é denominado haplogrupo (TORRONI et al., 1993). Excetuando eventos homoplásicos (idênticos por estado), o compartilhamento de alelos por dois haplótipos

indica que estes divergiram de um ancestral comum e, desde então, tem acumulado mutações diferentes (NOVELLETTO, 2007). Para cada haplogrupo, o motivo mutacional fundador (ancestral) - que pode incluir uma ou várias mutações - é aquele localizado no nó do ramo (JOBLING; TYLER-SMITH, 2017). O comprimento dos ramos numa filogenia é proporcional ao número de mutações e, desta forma, ao tempo evolutivo (JOBLING; TYLER-SMITH, 2017).

Em 1987, Avise e colaboradores estabeleceram as bases da filogeografia, a disciplina que estuda os princípios e processos que dirigem a distribuição geográfica das linhagens genealógicas (AVISE, 1987). A filogeografia proporcionou um ponto de vista filogenético aos estudos evolutivos intraespecíficos, conectando a filogenia e a genética de populações (AVISE, 1987). A filogeografia nasce a partir da observação de que muitas espécies mostram no ADNmt uma história filogenética geograficamente estruturada. Ela fundamenta-se no uso de marcadores genéticos ou sequências de alta taxa mutacional, na teoria da coalescência e no conceito de evolução neutra (AVISE, 1987). A teoria neutralista assume que a maior parte das mudanças no nível molecular são neutras e que, portanto, é a deriva genética a principal responsável por conduzir a evolução, e não a seleção natural (KIMURA, 1989). Por outro lado, a teoria da coalescência estabelece que, sob a neutralidade e numa população de tamanho constante, surgem permanentemente novos alelos por mutação enquanto outros são perdidos pela deriva genética (AVISE, 2009; VÁZQUEZ et al., 2009). Desta forma, todos os alelos ou as linhagens genéticas encontradas na atualidade derivam, ou coalescem, em um ancestral comum em algum momento da história evolutiva da população (AVISE, 2009; VÁZQUEZ et al., 2009). A teoria não implica que apenas o ancestral comum tenha existido na geração coalescente, mas que apenas as linhagens derivadas dele sobreviveram até a atualidade (AVISE, 2009).

Em filogeografia, as genealogias podem ser representadas em uma estrutura hierárquica através de cladogramas, aplicando os mesmos algoritmos (máxima parcimônia ou verossimilhança) utilizados para reconstruir árvores de táxons superiores (VÁZQUEZ et al., 2009). As árvores são usualmente associadas à distribuição geográfica a fim de compreender os processos de dispersão da espécie ou população estudada (VÁZQUEZ et al., 2009). Cada ramo da árvore constitui um haplótipo (VÁZQUEZ et al., 2009).

Os genomas mitocondrial e do cromossomo Y frequentemente apresentam homoplasia² (de substituições e indels, no ADNmt, e de STR no Y) (UNDERHILL; KIVISILD, 2007). Por esse motivo, a representação mais adequada é através de redes de haplótipos ou *networks*. As redes de haplótipos permitem a introdução de *loops* ou reticulações quando múltiplas reconstruções filogenéticas (descritas na rede através de *links* alternativos) são possíveis (BANDELT; FORSTER; RÖHL, 1999). As redes estabelecem uma relação de parentesco entre os haplótipos permitindo explicar os padrões filogeográficos com base na sua topologia (BANDELT; FORSTER; RÖHL, 1999). Diferente das árvores bifurcadas tradicionais, as redes permitem descrever relações evolutivas multibifurcantes, como aquelas em que um único haplótipo dá origem a muitos haplótipos (VÁZQUEZ et al., 2009). Nas redes, as conexões, ou *links*, entre os diferentes haplótipos representam eventos de coalescência (VÁZQUEZ et al., 2009). A interpretação das árvores é baseada nos seguintes pressupostos (VÁZQUEZ et al., 2009):

- Os haplótipos ancestrais são provavelmente os mais abundantes na atualidade.
- Os haplótipos ancestrais provavelmente possuem múltiplas conexões.
- Os haplótipos mais antigos ocupam um lugar central na rede, enquanto os haplótipos mais recentes ocupam posições externas.
- Os haplótipos ancestrais provavelmente possuem uma distribuição geográfica mais ampla (devido ao maior tempo de dispersão).
- Os haplótipos com uma única conexão estão unidos a haplótipos da mesma população, dado que possuem uma evolução recente e não tiveram tempo para se dispersar.

Existem dois algoritmos amplamente utilizados para a construção de redes de haplótipos não recombinantes, o *median-joining network* (MJ) e o *reduced median network* (RM) (BANDELT; FORSTER; RÖHL, 1999). Ambos se baseiam na construção de redes parcimoniosas de haplótipos conectados por vetores medianos (*median vectors*), embora o algoritmo MJ é capaz de manejar conjuntos maiores de dados e também dados com múltiplos estados - como sequências aminoacídicas. Os vetores medianos são sequências consenso que relacionam três haplótipos por vez selecionadas pelo algoritmo para a construção da árvore mais parcimoniosa (BANDELT; FORSTER; RÖHL, 1999). Eles podem ser interpretados biologicamente como haplótipos possivelmente existentes na população, mas que não foram amostrados ou como haplótipos ancestrais extintos (BANDELT; FORSTER; RÖHL, 1999).

² Homoplasia: ocorrência de mutações recorrentes na mesma posição da sequência.

2.3.1.1 O ADN MITOCONDRIAL

O genoma mitocondrial humano é uma molécula de ADN dupla fita circular que possui em torno de 16.569 pb (16,5 kbp) (ANDERSON et al., 1981). Ele foi sequenciado pela primeira vez em 1981 (ANDERSON et al., 1981). A sequência de referência resultante pertence ao haplogrupo H, sendo uma linhagem de origem europeia (PARSON et al., 2014). Atualmente essa sequência foi revisada e corrigida, encontrando-se disponível para comparação - sequência revisada de Cambridge (rCRS) (ANDREWS et al., 1999). Em 2012, Behar e colaboradores introduziram uma nova sequência de referência reconstruída a partir das sequências disponíveis de *Homo neanderthalensis* e com uma filogenia refinada com a informação obtida a partir de 8.216 mitogenomas modernos (BEHAR et al., 2012). A sequência (RSRS, por “*Reconstructed Sapiens Reference Sequence*”, ou Sequência Reconstruída Sapiens de Referência) foi construída numa tentativa de referenciar as análises mitocondriais humanas a uma linhagem ancestral, o que seria mais adequado em termos evolutivos (PARSON et al., 2014). A RSRS é situada próxima à raiz da filogenia mitocondrial humana, entre os haplogrupos L0 e L1'2'3'4'5'6 (PARSON et al., 2014). Porém, devido à maior antiguidade da sequência rCRS e o consenso estabelecido na nomenclatura dos estados ancestrais e derivados das variantes mitocondriais, ela continua a ser a mais utilizada, especialmente em estudos forenses e médicos (GASPARRE; PORCELLI, 2020; PARSON et al., 2014). Em estudos evolutivos, o uso de RSRS têm ganhado mais popularidade (GASPARRE; PORCELLI, 2020).

O uso do genoma mitocondrial para estudos genéticos fundamenta-se em um conjunto de características que este genoma apresenta. Entre elas, e especialmente útil em estudos forenses e de ADN antigo (onde outros cromossomos podem apresentar degradação), está o seu alto número de cópias (GASPARRE; PORCELLI, 2020). Cada mitocôndria humana possui de 2 a mais de 10 moléculas de ADNmt e, devido ao alto número de mitocôndrias nas células humanas, o número de moléculas de ADNmt por célula pode atingir centenas ou milhares de cópias (GASPARRE; PORCELLI, 2020).

O ADNmt possui uma taxa de mutação cerca de 5 vezes maior do que a do ADN nuclear devido à geração de espécies reativas de oxigênio no interior da mitocôndria, à ausência de histonas e à presença de um sistema de reparo menos eficiente quando comparado ao do ADN nuclear (GASPARRE; PORCELLI, 2020). A taxa de mutação, contudo, é variável ao longo da

molécula, podendo atingir taxas maiores nas regiões hipervariáveis e em posições específicas denominadas *hotspots* mutacionais (SOARES et al., 2009). O ADNmt, assim como os genomas bacterianos, está localizado numa região celular chamada nucleoide, onde encontra-se associado às proteínas necessárias para as suas funções biológicas (GASPARRE; PORCELLI, 2020).

Dentro do genoma mitocondrial encontramos uma região de aproximadamente 1121 pares de bases chamada Região Controle (RC), que se estende da posição 16024 até a 16569 e continua da posição 1 até a 576 (Figura 6) (JOBLING et al., 2014). A região recebe esse nome porque contém os promotores de transcrição e as origens de replicação do plasmídeo. Dentro da RC existe a região chamada D-loop (do inglês *displacement loop*) de aproximadamente 600 nt. Durante os estágios iniciais da replicação nessa região a cadeia recém sintetizada adquire um formato de alça (GASPARRE; PORCELLI, 2020). A RC possui uma taxa de mutação 5-10 vezes maior do que a do restante do genoma, sendo a mais utilizada em estudos filogenéticos por acumular a maior variabilidade interpessoal. A RC está dividida em três segmentos de acordo com o seu grau de variabilidade, conhecidos como regiões hipervariáveis 1, 2 e 3 (HVI, HVII e HVIII).

Por outro lado, devido à alta taxa mutacional, o genoma mitocondrial pode apresentar heteroplasmia - em contraposição ao estado de homoplasmia (QUINTERO FERRER; PARDO GOVEA; BORJAS FUENTES, 2019). Por este fenômeno, dentro de uma célula, mitocôndria ou até entre os distintos tecidos de um indivíduo, podem coexistir várias populações de genomas mitocondriais (AVISE, 1987). Habitualmente, a heteroplasmia consiste em uma única diferença entre genomas, sendo incomum a heteroplasmia em dois ou mais sítios (QUINTERO FERRER; PARDO GOVEA; BORJAS FUENTES, 2019). A heteroplasmia pode ser de comprimento, quando a diferença de sequências consiste na presença de um indel, ou de sequência, quando existe mais de uma base possível em uma determinada posição (QUINTERO FERRER; PARDO GOVEA; BORJAS FUENTES, 2019). A heteroplasmia de comprimento é frequente em regiões homopoliméricas do ADNmt, como as localizadas entre as posições 16183-16194, de HVI, e 303-315, de HVII (QUINTERO FERRER; PARDO GOVEA; BORJAS FUENTES, 2019). Sob a perspectiva da genética de populações, a presença de heteroplasmia é considerada derivada de um evento mutacional recente (BANDELT; MACAULAY; RICHARDS, 2006).

Outra das vantagens da utilização deste marcador é sua herança exclusivamente materna, assim como também a ausência de recombinação (JOBLING et al., 2014). Essas características fazem com que o ADNmt represente um único locus dentro do genoma humano

e tenha um menor tamanho efetivo populacional ($\frac{1}{4}$ do tamanho efetivo do genoma nuclear), sendo, portanto, mais suscetível à deriva genética (JOBLING et al., 2014).

Como ocorre usualmente com plasmídeos bacterianos e provavelmente decorrente da sua origem por endossimbionte, este genoma possui uma alta densidade gênica e ausência de íntrons. Apesar de contar com um número baixo de genes (37 no total), cerca de 90% da sua sequência é codificante (GASPARRE; PORCELLI, 2020). No ADNmt humano encontramos 13 genes que codificam proteínas (todas elas necessárias ao sistema de fosforilação oxidativa (OXPHOS), 2 genes que codificam ARNs ribossômicos e um conjunto de 22 ARNs transportadores necessários para a tradução dos ARNs mensageiros mitocondriais (GASPARRE; PORCELLI, 2020). As duas fitas do ADNmt possuem uma composição nucleotídica diferenciada que as identificam como fita pesada (ou H, pelo inglês *heavy*) que é rica em guaninas, e a fita leve (L, do inglês *light*), rica em citosinas (JOBLING et al., 2014) (Figura 6).

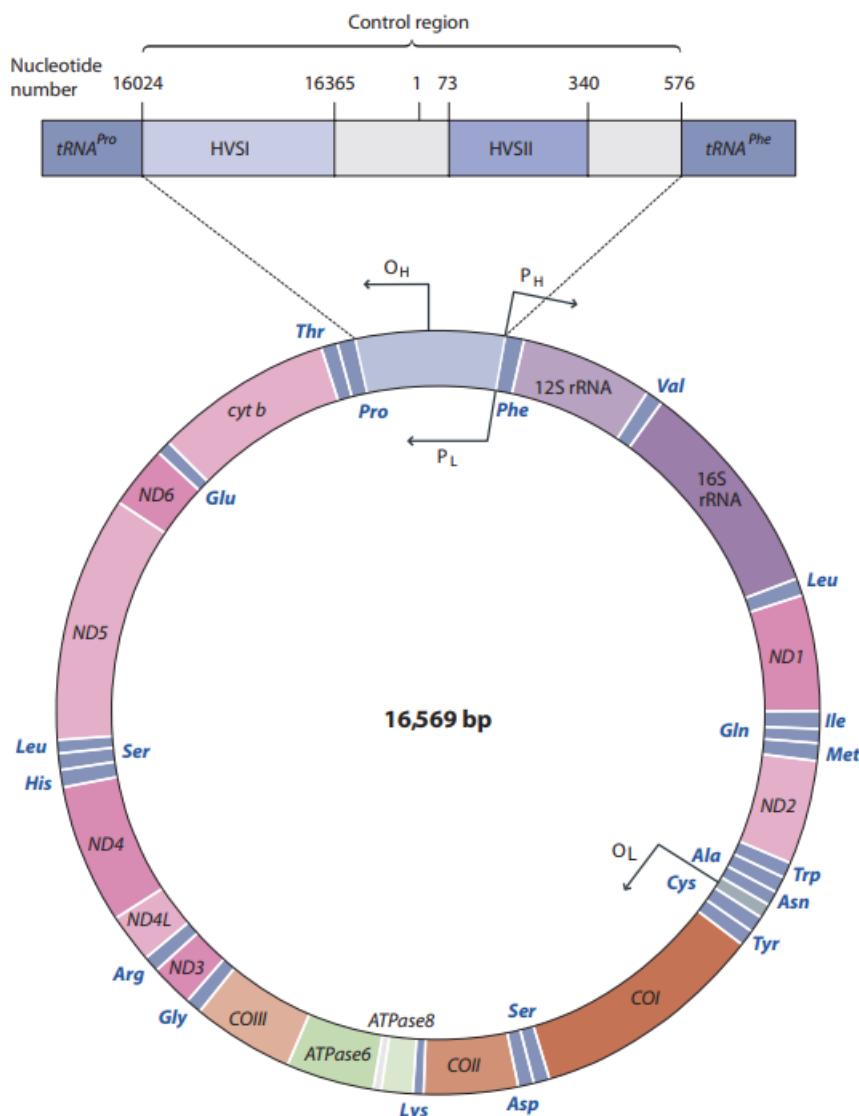


Figura 6 - Esquema do genoma mitocondrial humano mostrando as regiões hipervariáveis (HV).
 O: origem de replicação. P: promotor. H: fita pesada (heavy). L: fita leve (light). Fonte: Jobling (2014)

O tipo de polimorfismo mais comumente empregado em estudos com o ADNmt é o SNP. Embora indels também ocorram, estão presentes em uma taxa menor (JOBLING et al., 2014).

A nomenclatura e filogenia mundial consenso de haplogrupos do ADNmt é descrita pelo PhyloTree (<https://www.phylotree.org/>) (VAN OVEN; KAYSER, 2009). A última versão (Build 17; 18 de fevereiro de 2016) abrange 5400 haplogrupos e subhaplogrupos denominados com letras de A até Z, e números. Porém, atualizações recentes da mesma árvore têm atingido

o número de 6401 haplogrupos (DÜR; HUBER; PARSON, 2021). Cada um dos haplogrupos da árvore está incluído em um dos 33 clusters (A, B, L1, L2, etc.) (SCHÖNHERR et al., 2023).

2.3.1.1.1 BREVE INTRODUÇÃO À FILOGENIA DO ADNMT HUMANO

A análise de grandes conjuntos de mitogenomas completos tem permitido uma reconstrução ultra detalhada das relações filogenéticas humanas, sendo extremamente informativo para a datação de eventos migratórios (GASPARRE; PORCELLI, 2020). Desta maneira, a história das migrações humanas e o povoamento dos continentes foi fundamentalmente reconstruída através de estudos de marcadores uniparentais, dentre os quais aqueles baseados no ADNmt foram pioneiros (GASPARRE; PORCELLI, 2020).

Já nos primeiros estudos com populações atuais demonstrou-se que a maior diversidade genética se encontra na África, inferindo o ancestral comum mais recente aproximadamente 200.000 anos atrás e suportando a teoria da saída dos humanos modernos deste continente (CANN; STONEKING; WILSON, 1987). A filogenia do ADNmt humano pode ser colapsada em dois ramos irmãos: L0 e L1'6 (ou L1'2'3'4'5'6), com coalescência estimada em 200 ka (BEHAR et al., 2008; RITO et al., 2013) (Figura 7). O clado L0 constitui a ramificação mais antiga do ADNmt. Ele inclui linhagens restritas a populações Khoisan do Sul e do Leste da África (L0d e L0k) (BEHAR et al., 2008). Por outro lado, o clado L1'6 inclui todos os haplogrupos encontrados hoje no continente africano (L1-6) e no resto do mundo, já que M e N, macro-haplogrupos derivados do ramo L3, deram origem a todos os haplogrupos não africanos (BEHAR et al., 2008) (Figura 7). O clado L1 divergiu de L2'3'4'5'6 aproximadamente 150 ka e está representado predominantemente em populações do Leste e da África Central. Cerca de 130 ka, o clado L5 diferenciou-se no Leste Africano, onde ainda atinge as maiores frequências. L2'3'4'6 também se originou no Leste, enquanto L2 divergiu de L3'4'6 ~100 ka no Oeste Africano, onde hoje apresenta a sua maior diversidade e frequência. Cerca de 6-75 ka, o clado L3 emergiu no Leste Africano, expandindo-se rapidamente na região, assim como também para o Centro e Oeste Africanos e também para fora do continente através dos sub-haplogrupos M e N (BEHAR et al., 2012). Na mesma época, linhagens de L2 realizam o movimento recíproco, do Centro/Oeste Africanos para o Leste, favorecido pelo reverdecimento do Sahel (a região ao sul do deserto do Saara) no Holoceno (SOARES et al., 2012).

O modelo mais aceito de saída da África sugere uma única dispersão a partir do chifre africano (região que inclui Eritreia, Somália, Djibouti e Etiópia) ao redor de 60-70 ka.

Portadores do haplogrupo L3 teriam seguido uma rota através da costa sul, atingindo a Península Arábica e se dispersando em sentido leste e norte. Nessa região, o haplogrupo L3 teria divergido rapidamente nos super-haplogrupos M, N - e no seu derivado direto, R – enquanto o haplogrupo ancestral, L3, teria sido perdido (MACAULAY et al., 2005). É importante destacar que até hoje não foram encontradas linhagens pré-L3 no Oriente Médio nem linhagens L3 ancestrais na Europa ou na Ásia (MACAULAY et al., 2005).

O super-haplogrupo M encontra-se majoritariamente na Ásia, representado pelos haplogrupos C, D, G, E, Q, Z e numerosos clados de M (KIVISILD et al., 2002). O super-haplogrupo N apresenta a maior frequência na Eurásia, sendo representado pelos haplogrupos A, I, R (e os seus subclados B, J, T, U (e o seu derivado K) e H), S, W, X, Y e vários clados de N (HERNÁNDEZ, 2023). O haplogrupo R, derivado cedo de N, é considerado, junto com M e N, um dos fundadores da Eurásia (MACAULAY et al., 2005). R inclui numerosos subhaplogrupos e os haplogrupos R0, B, F, H, J, P, T, U e V. Dentre eles, o haplogrupo H se destaca por representar cerca de 50% da população europeia (TORRONI et al., 2006).

Devido aos inícios dos estudos antropológicos com ADNmt terem sido realizados com populações Nativas Americanas, as primeiras letras do alfabeto denominam os Hgs encontrados na América e cujos clados ancestrais foram posteriormente detectados na Sibéria e no Leste Asiático, suportando desta forma a origem asiática dos povos nativos americanos (TORRONI et al., 1992). Já nesses primeiros estudos foi demonstrado que os genomas mitocondriais nativos americanos exibiam uma diversidade relativamente baixa de linhagens, sendo todas elas compreendidas pelos haplogrupos A, B, C, D e X (GASPARRE; PORCELLI, 2020). No total, 16 linhagens fundadoras maternas de ancestralidade beríngia ou asiática foram identificadas em nativos americanos (GASPARRE; PORCELLI, 2020). Oito haplogrupos são definidos como panamericanos (A2, B2, C1b, C1c, C1d, D1 e D4h3a) e estavam presentes possivelmente na primeira onda migratória, desde a Beríngia até a América do Sul, ao longo da rota costeira, 16-18.000 anos atrás (BODNER et al., 2012). Outros haplogrupos (A2a, A2b, D2a e D3) são restritos às populações das regiões ártica e subártica Norte Americanas e possivelmente chegaram através de ondas migratórias mais recentes (BODNER et al., 2012; DE SAINT PIERRE et al., 2012). Outras linhagens como X2g ou D4e1 são extremamente raras e restritas à América do Norte (BODNER et al., 2012; DE SAINT PIERRE et al., 2012). Algumas linhagens como D1g, D1j, B2i2 e C1b13 são restritas ao extremo sul do continente (BODNER et al., 2012; DE SAINT PIERRE et al., 2012).

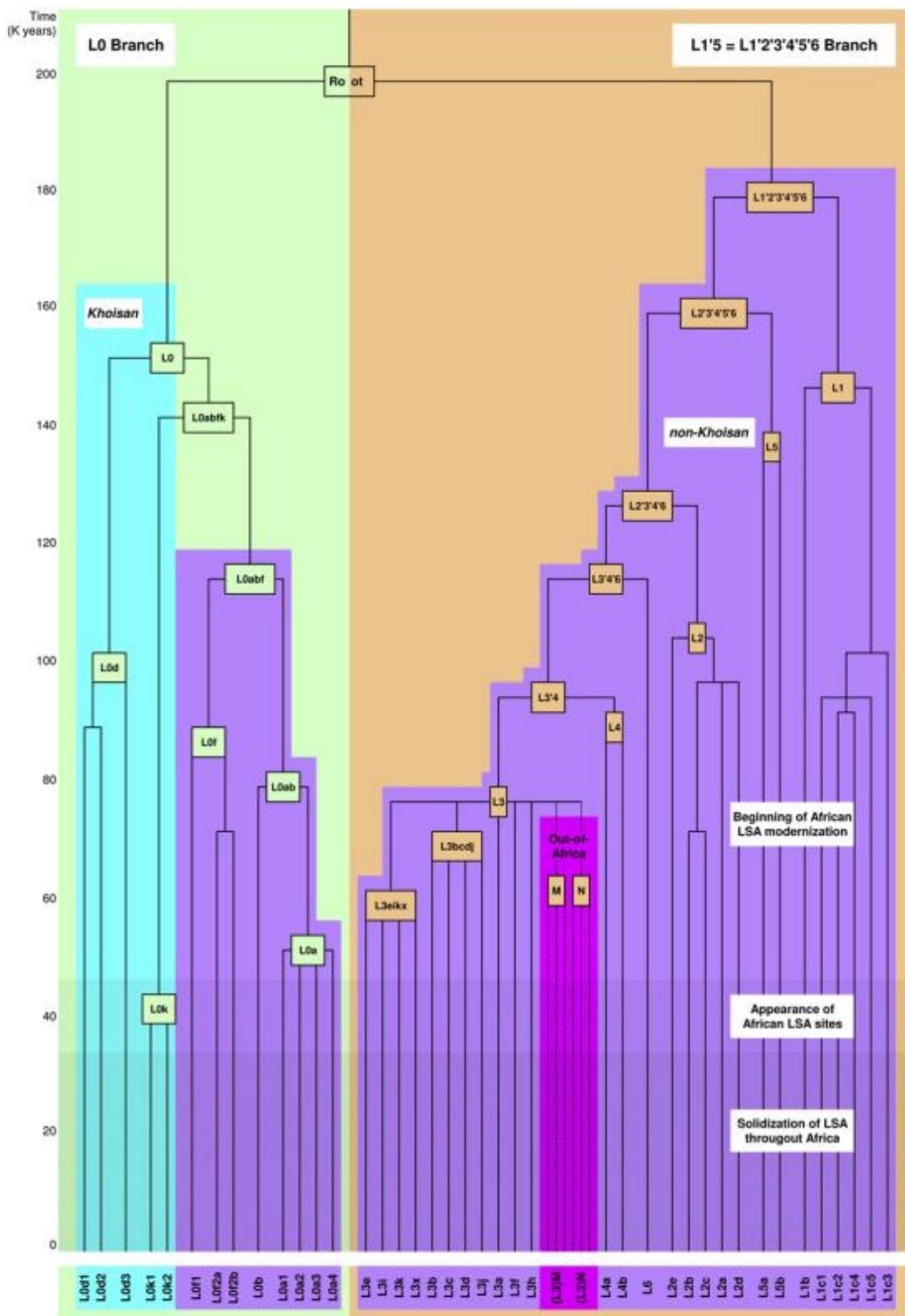


Figura 7 - Esquema simplificado da filogenia do ADNmt humano. Fonte: Behar et al. (2012)

2.3.1.2 O CROMOSSOMO Y COMO FERRAMENTA FILOGEOGRÁFICA

O cromossomo Y humano é um dos menores do genoma (possui aproximadamente 60 Mb). É constituído por heterocromatina (sequências altamente repetitivas) na região distal do braço longo (Yq) e por eucromatina no braço curto (Yp) e parte do braço longo (~23 Mb) (JOBLING; TYLER-SMITH, 2017). Na sua estrutura podem diferenciar-se três regiões: duas regiões pseudoautossômicas, de aproximadamente 3 Mb, que recombinam com o cromossomo X, e uma região central (MSY) que constitui 95% do cromossomo e não apresenta recombinação intercromossómica (JOBLING; TYLER-SMITH, 2017). A região utilizada para estudos filogeográficos é a região eucromática não recombinante, por possuir grande parte dos genes e polimorfismos. Contudo, a natureza altamente repetitiva do cromossomo Y determina que apenas cerca de 10 Mb da sua sequência (distribuída em segmentos descontínuos) seja de cópia única e passível de ser mapeada a uma sequência de referência (JOBLING; TYLER-SMITH, 2017). Devido à sua forma de herança exclusivamente de pais para filhos do sexo masculino e à ausência de recombinação intercromossómica, a MSY é altamente informativa para estudos evolutivos. O cromossomo Y apresenta polimorfismos binários do tipo SNP (substituições e indels) e de número de cópias, como os microssatélites. Ambos polimorfismos são utilizados nos estudos filogeográficos. Enquanto os SNPs definem os grupos monofiléticos (haplogrupos) devido a sua menor taxa evolutiva, os microssatélites definem sublinhagens dentro dos haplogrupos (ou haplótipos) (DE KNIJFF, 2000).

A árvore filogenética do cromossomo Y está descrita e permanentemente atualizada na *International Society of Genetic Genealogy* (ISOGG) (<https://isogg.org/tree/>). Possui, até o momento, vinte clados principais (A até T) e 10.263 haplogrupos, estando os seus nós definidos por mais de 70.000 SNPs (POZNIK et al., 2016; INTERNATIONAL SOCIETY OF GENETIC GENEALOGY, 2019).

Na árvore filogenética do cromossomo Y, os clados A e B correspondem aos mais antigos e são ambos exclusivamente africanos (POZNIK et al., 2016). A origem do clado E não possui um consenso científico, embora ele atinja a sua maior distribuição no continente africano (HABER et al., 2019). Já os demais clados (C até T) surgiram fora da África e distribuem-se quase exclusivamente fora da África (POZNIK et al., 2016).

O clado C é comum na Ásia e a sua origem sugerida se remonta a esse continente. C também se distribui amplamente na Oceania, e, em baixa frequência, na Europa (KARAFET et al., 2008). O Hg C3 é uma das linhagens fundadoras nativas americanas (PINOTTI et al., 2019).

O clado KLT tem uma origem sugerida no Oriente Médio, logo após a saída dos humanos modernos da África, e inclui aos Hgs LT, M, NO e P. Deste último derivam os Hgs Q e R. Q ocorre em frequências baixas na Europa, Ásia Oriental e no Oriente Médio e em frequências altas em populações da Sibéria e na América (KARAFET et al., 2008). A sua linhagem Q1b1a1a-M3 representa a maioria das populações nativas americanas estudadas até o momento, sendo considerada a linhagem fundadora nativa americana devido a que se encontra em frequência muito baixa na Sibéria (BATTAGLIA et al., 2013; BISSO-MACHADO et al., 2011). A linhagem ancestral de Q1b1a1a-M3, Q1b-M346 (xM3) também tem sido encontrada em nativos americanos, porém a sua frequência é baixa (BAILLIET et al., 2009; ROEWER et al., 2013).

A diversidade europeia está principalmente restrita aos clados E, G, I, J, N e R (JOBLING; TYLER-SMITH, 2003; CHIARONI et al., 2009). As linhagens derivadas do clado Hg R-M207 são as mais comuns na Europa, embora também são encontradas no centro, oeste e sul da Ásia, e sublinhagens específicas ocorrem na África (KARAFET et al., 2008; MYRES et al., 2011). Cerca de 50% da população do sexo masculino na Europa carrega linhagens derivadas de R-M207, sendo a maioria derivados de R1-M173 (MYRES et al., 2011). Enquanto linhagens derivadas de R1a-M420 são mais frequentes no Leste Europeu, linhagens derivadas de R1b-M343 são mais comuns no Oeste (MYRES et al., 2011). A única linhagem derivada de R1b com origem e distribuição exclusiva na África é R1b1b-V88, que ocorre no Norte Africano, no Sahel Central e no Centro-Oeste Africano (CRUCIANI et al., 2010). No Norte Africano, R1b1b-V88 ocorre em populações Berber, enquanto no Sahel Central ele ocorre em populações de línguas Niger-Congo, em particular no povo Hausa (CRUCIANI et al., 2010). Já no Centro-Oeste Africano, R1b1b-V88 tem sido observado em populações da Guiné Equatorial, Camarões, Gabon, e na bacia do Lago Chad (GONZÁLEZ et al., 2013). Embora a origem e rotas de dispersão de R1b1b-V88 na África sejam discutidas na literatura, alguns autores sugerem uma origem no Centro-Oeste Africano e posterior dispersão para o Norte Africano através de migrações pelo Saara (CRUCIANI et al., 2010; GONZÁLEZ et al., 2013). O clado ancestral, R1b1-L754³, é considerado de origem Europeia, sendo a sua alta frequência em populações do Centro-Oeste Africano explicada por uma migração reversa em épocas pré-

³ R1b1-L754 é atualmente definido pelos SNPs L754, PF6269 e YSC0000022 de acordo com ISOGG 2019-2020. Até 2016 ele era definido pelo SNP P25, que foi removido devido a sua alta taxa de mutação reversa (isogg.org).

históricas mediada por povos falantes de línguas Chádicas (Afro-Asiáticas) (GONZÁLEZ et al., 2007).

2.3.1.3 A ESTRUTURA POPULACIONAL AFRICANA E OS MARCADORES UNIPARENTAIS

A diversidade mitocondrial africana está representada nos Hgs L0-6, M1 e U6 (ROSA; BREHM, 2011). M1 e U6 ocorrem exclusivamente no Norte e no Leste Africanos, com uma baixa frequência de U6 no Oeste e no Centro-Oeste devido ao fluxo gênico (SALAS et al., 2002). Na África Subsaariana, cerca de 70% dos Hgs derivam dos clados L2 e L3, que, assim como o clado L1, encontram-se em todas as regiões do continente (ROSA; BREHM, 2011; SALAS et al., 2002). L2 atualmente é o clado mais complexo e amplamente distribuído no continente (SILVA et al., 2015). A disseminação deste clado e dos seus sub-Hgs para o Centro, Sul e Leste Africanos (principalmente os derivados de L2a e L2b) teria ocorrido durante a expansão Bantu (SILVA et al., 2015). Entretanto, o fluxo gênico oposto (de populações do Leste Africano para populações Bantu do Sul) não teria ocorrido de forma expressiva, a julgar pela baixa frequência das linhagens típicas do Leste Africano (como L4, L5, L6, e os clados de L3) em grupos Bantu sul-africanos (SILVA et al., 2015). A descoberta de linhagens de L2 em frequências altas (acima de 30%) no Sudão e na Etiópia sugeriram a introdução de L2 no Leste ainda antes da expansão Bantu, durante o início do Holoceno (ou Pleistoceno tardio), período em que as condições climáticas na África permitiram eventos migratórios em larga escala (SILVA et al., 2015). Já uma difusão de L2 para o Nordeste Africano teria acontecido graças ao tráfico Trans-Saariano (SALAS et al., 2002).

No cromossomo Y, a diversidade africana encontra-se restrita aos clados A, B, E e a sublinhagem R1b1b-V88 discutida na secção anterior. O clado A é encontrado principalmente entre populações Khoisan do sul da África, na Etiópia e no Sudão (DE FILIPPO et al., 2011). Entretanto, o clado B é comum entre caçadores-coletores centro-africanos de fenótipo pigmeu (com algumas linhagens restritas a esses grupos), povos Khoisan, e os seus vizinhos (CHIARONI et al., 2009).

Embora a origem do clado E é controversa, alguns autores sugerem que ele é oriundo da população leste africana que originou a migração fora da África e que a sua alta frequência neste continente decorre de uma migração regressiva (CHIARONI; UNDERHILL; CAVALLI-

SFORZA, 2009; HABER et al., 2019; POZNIK et al., 2016). Estudos mais recentes também suportam uma origem de E no continente africano (HABER et al., 2019). As suas linhagens atingem a maior diversidade e frequência na África subsaariana, constituindo cerca de 87% dos Hgs africanos. Sublinhagens específicas de E, entretanto, ocorrem em baixa frequência no Norte e Leste Africanos assim como também na Europa e no oeste da Ásia (CHIARONI; UNDERHILL; CAVALLI-SFORZA, 2009). Enquanto a sub-linhagem E1b1a1 (E-M2) está amplamente distribuída em grupos falantes de línguas do tronco Níger-Congo, tanto Bantu como não-Bantu, a sua linhagem derivada, E1b1a1a1c1a1a-U174, é considerada um marcador Bantu (DE FILIPPO et al., 2011).

Apesar das diferenças linguísticas entre grupos Bantu do Leste e do Oeste Africano, geneticamente esses povos apresentam homogeneidade, diferenciando-se entre si apenas pela inclusão de marcadores derivados do fluxo gênico com populações autóctones daquelas regiões durante a sua expansão (BARBIERI et al., 2014b). Por exemplo, no Sul Africano os grupos Bantu mostram frequências expressivas de linhagens marcadoras tais como os Hgs mitocondriais L0d e L0k, típicos de grupos Khoisan, e as linhagens do cromossomo Y A-M51, A-M23 e B-M112 (BARBIERI et al., 2013; COELHO et al., 2009; WOOD et al., 2005). Já no Sudeste Africano (Moçambique, leste da Zâmbia e Zimbabwe) a introgessão de marcadores Khoisan em grupos Bantu é menos expressiva (menos de 3%) (BARBIERI et al., 2014b). Linhagens A do cromossomo Y são típicas de grupos Khoisan e do Leste Africano, enquanto linhagens B são encontradas principalmente em populações caçadoras-coletoras da África Central, grupos Khoisan e populações vizinhas a eles (DE FILIPPO et al., 2011). Os estudos genéticos também demonstraram que o fluxo gênico com populações autóctones do Sul Africano ocorreu com um marcado viés sexual com tendência para a introgessão de linhagens maternas dos grupos autóctones nas populações Bantu e de linhagens Y no sentido oposto (BARBIERI et al., 2013; COELHO et al., 2009; WOOD et al., 2005). Contudo, observa-se homogeneidade genética entre os grupos falantes de línguas Bantu do Leste e do Oeste Africano, sugerindo a erosão de qualquer sinal da diversificação linguística original através do fluxo gênico posterior à mesma (BARBIERI et al., 2014b; DE FILIPPO et al., 2012).

Análises filogeográficas realizadas com haplogrupos uniparentais provém evidência de múltiplas ondas migratórias, com avanços e retrocessos, tanto ao longo da rota Leste quanto da Oeste, assim como também de uma última onda migratória oriental mais recente (ANSARI POUR; PLASTER; BRADMAN, 2013).

Assim como os estudos do genoma mitocondrial, estudos genômicos também suportam um componente genético comum entre a maioria dos grupos Bantu (GURDASANI et al., 2019). Um componente do Centro/Oeste Africano predomina entre grupos Bantu do Sul e do Leste. Enquanto os Bantu do Sul mostram um componente Khoisan, suportando o fluxo gênico entre os grupos, os Bantu do Leste mostram um componente leste africano não detectado nos Bantu do Sul (GURDASANI et al., 2019).

A população Luhya, do Quênia, é frequentemente incluída na comparação com amostras brasileiras por ser um grupo Bantu (KEHDY et al., 2015). Porém, as linhagens do ADNmt presentes nesse grupo demonstram uma ancestralidade tanto Bantu, como o sugere a alta frequência de L3b -típico do Oeste Africano-, quanto leste africana pela alta frequência de linhagens L0a, L4 e L5 (SILVA et al., 2015).

2.3.2 MARCADORES GENÉTICOS BIPARENTAIS

A ancestralidade genética se refere ao conjunto de ancestrais genealógicos dos quais um indivíduo ou população herdou o seu acervo genético (MATHIESON; SCALLY, 2020). É importante diferenciá-la do termo ancestralidade em um sentido amplo ou da genealogia genética, já que cada progenitor transmite apenas metade do seu ADN à progênie, então a ancestralidade genética de um indivíduo inclui apenas uma pequena fração de todos os seus ancestrais genealógicos (MATHIESON; SCALLY, 2020). A ancestralidade genética pode ser analisada de uma forma global, considerando a média da contribuição de cada população ancestral ao conjunto completo de cromossomos de um indivíduo, ou no nível local, considerando cada cromossomo separadamente ou ainda os distintos fragmentos de cada cromossomo (KORUNES; GOLDBERG, 2021). Desta forma, em populações miscigenadas, a proporção da contribuição de cada grupo ancestral pode variar em três níveis: 1) entre populações; 2) entre indivíduos da mesma população; 3) ao longo do genoma de um indivíduo (KORUNES; GOLDBERG, 2021).

Quanto ao número de marcadores utilizados na inferência de ancestralidade autossômica, existem duas estratégias possíveis. A primeira é baseada no estudo de um conjunto relativamente pequeno de marcadores informativos de ancestralidade (MIAs ou *AIMs*, do inglês *ancestry informative markers*) que maximizam a diferença de frequência alélica absoluta entre as populações ancestrais (KOSOY et al., 2009). O uso de MIAs oferece uma alternativa de

menor custo à segunda estratégia, baseada no estudo de microarranjos de alta densidade de SNPs, sequenciamento do genoma completo, entre outras (DE SOUZA et al., 2019).

Embora a análise de marcadores autossômicos seja considerada a abordagem mais completa para estimar a ancestralidade genômica de um indivíduo, ela possui algumas limitações. A estimativa de ancestralidade global é mais precisa quanto maior a cobertura das distintas partes do genoma de um indivíduo, isto é, quanto maior o número de marcadores analisados, sendo maximizada na análise do genoma completo (ESCHER et al., 2022). Por outro lado, uma estimativa correta depende de que o painel de populações de referência analisado seja representativo da diversidade encontrada nos grupos ancestrais (PFAFFELHUBER et al., 2020). Esta última condição geralmente não se cumpre para populações Ameríndias ou Africanas (entre outras) já que os painéis de referência atualmente disponíveis em bancos de dados públicos oferecem uma baixa representação de populações dessas origens (AUTON et al., 2015; CANN et al., 2002).

Por último, a análise de ancestralidade genética depende do modelo de miscigenação (*admixture model*, em inglês) utilizado, que por sua vez depende de quantas populações ancestrais sejam consideradas contribuintes ao acervo genético da população híbrida (ESCHER et al., 2022). No caso da população brasileira, dois modelos são utilizados, dependendo da história evolutiva da subpopulação estudada. O modelo tri-híbrido, que assume a contribuição predominante dos grupos europeu, africano e nativo-americano, e o modelo tetra-híbrido, onde a contribuição de pessoas do Leste Asiático é também considerada⁴ (ESCHER et al., 2022). No entanto, numa análise recente da performance de diferentes métodos de estimativa de ancestralidade genética em populações miscigenadas (inclusive a brasileira), foi demonstrado que a inclusão do componente leste-asiático introduz ruído na estimativa de ancestralidade nativo-americana e europeia em indivíduos sem ancestralidade leste-asiática (ESCHER et al., 2022). Este efeito é ainda mais acentuado quando se utilizam conjuntos pequenos de MIAs (ESCHER et al., 2022). Adicionalmente, estudos recentes realizados na população de São Paulo, Brasil, demonstraram que indivíduos oriundos do Leste Asiático possuem cerca de 100% dessa ancestralidade, indicando uma baixa miscigenação dessa população (NASLAVSKY et al., 2022). Este dado é coerente com o fato da migração do Leste Asiático no Brasil ser muito

⁴ A imigração leste-asiática para o Brasil teve o seu início no período colonial, com a chegada de imigrantes chineses. Porém, aumentou significativamente em 1908 com uma grande onda migratória de origem japonesa. A população leste-asiática no Brasil se concentra principalmente no estado de São Paulo, mas também em estados como Paraná, Mato Grosso, entre outros (IBGE, 2007).

mais recente que as demais (IBGE, 2007). Desta forma, o modelo tri-híbrido é o recomendado para populações brasileiras sem contribuição leste-asiática conhecida (ESCHER et al., 2022).

Pelas limitações acima mencionadas, qualquer valor de ancestralidade inferido para populações miscigenadas da América deve ser considerado apenas como uma estimativa, podendo a mesma variar em maior ou menor medida (SANTOS et al., 2016a).

Outro ponto importante a ser considerado é que cada população possui uma história evolutiva única, sendo necessário considerar as suas particularidades na hora de selecionar os marcadores de ancestralidade e o painel de populações de referência (ESCHER et al., 2022). Desta maneira, a estratégia utilizada para estimar a ancestralidade genética de uma população miscigenada não pode ser diretamente aplicada a outras populações (ESCHER et al., 2022).

2.4 A POPULAÇÃO AFRO-BRASILEIRA E A RESISTÊNCIA À ESCRAVIDÃO: QUILOMBOS

A população afro-brasileira foi propositalmente excluída da distribuição de terras no Brasil, o que não ocorreu com os migrantes europeus. A primeira Lei de Terras no Brasil (1850), projetada pela aristocracia brasileira, impediu aos africanos e seus descendentes se tornarem proprietários, por um lado excluindo-os da categoria de brasileiros e categorizando-os como “libertos”, e, por outro, estabelecendo um prazo para a compra e legalização das terras (LEITE, 1999). Dessa maneira, os negros libertos, que não possuíam recursos econômicos, foram sistematicamente expulsos dos seus territórios (LEITE, 1999). Mesmo quando as terras eram compradas ou herdadas dos senhores, os mecanismos do racismo estrutural brasileiro agiram para impedir ou, no mínimo, obstaculizar a obtenção do território (LEITE, 1999).

Apesar das barreiras impostas, múltiplas comunidades foram formadas ao longo do território brasileiro durante e depois de finalizado o período escravocrata (GOMES, 2015). Essas comunidades muitas vezes foram estabelecidas em locais isolados ou semi-isolados, de forma a preservar o direito à liberdade que lhes era negado pela sociedade escravista (GOMES, 2015). Outras, entretanto, se desenvolveram nas proximidades de centros urbanos após a abolição da escravidão (GOMES, 2015). Contudo, todas elas se constituíram como uma forma de resistência à escravidão e de preservação da sua ancestralidade (MOURA, 1993). As características de cada uma destas comunidades são únicas, e refletem a complexidade da sua história, o que inclui as interações com o seu entorno e populações vizinhas (GOMES, 2015).

Hoje, o Instituto Brasileiro de Geografia e Estatística contabiliza 5972 comunidades quilombolas no país (IBGE, <https://educa.ibge.gov.br>).

2.4.1 A FORMAÇÃO DE QUILOMBOS NO MUNICÍPIO DA LAPA, PARANÁ

A presença negra e a participação de negros e de indígenas no Paraná tem sido historicamente invisibilizadas por meio da construção de uma imagem branca e europeizada da região, onde os contingentes migratórios europeus são exaltados (RAGGIO; BLEY; TRAUCZYNSKI, 2018). Parte dessa imagem foi construída graças ao movimento Paranista, sendo um dos seus principais expoentes o historiador Alfredo Romário Martins (1874-1948), fortemente influenciado pelo pensamento racista eurocêntrico do final do século XIX e alinhado com as políticas de branqueamento da época (RAGGIO; BLEY; TRAUCZYNSKI, 2018). *História do Paraná*, de Martins, publicada em 1892, foi a obra legitimada como a história oficial do estado na época e reivindicada em 1953, com a sua reedição (RAGGIO; BLEY; TRAUCZYNSKI, 2018). Nela, a formação do Paraná se dá através da união entre portugueses e indígenas, havendo um esquecimento e uma diminuição proposital do componente africano (RAGGIO; BLEY; TRAUCZYNSKI, 2018). Pesquisas mais recentes desmistificaram essas ideias mostrando que as proporções reais de população negra nas principais cidades do estado, como Curitiba, Castro e a cidade portuária de Paranaguá, oscilaram entre 20 e 50% da população ao longo dos anos (RAGGIO; BLEY; TRAUCZYNSKI, 2018).

O certo é que a presença negra no Paraná está diretamente interligada ao desenvolvimento do Estado, desde a incorporação da mão de obra dos negros na mineração, nas lavouras da erva mate e café, na pecuária, na construção e nas tarefas domésticas (RAGGIO; BLEY; TRAUCZYNSKI, 2018; RIBAS, 2019). Já em 1640, ano da descoberta de ouro na Serra Negra, próxima ao litoral paranaense, originou-se a Vila de Paranaguá, levando à imigração de mineradores paulistas junto com seus escravizados africanos (RAGGIO; BLEY; TRAUCZYNSKI, 2018).

Um dado que chama a atenção é o número ligeiramente maior de escravizadas do sexo feminino relatado tanto na Lapa quanto no Paraná como um todo, dados que contrastam fortemente com os relatos do resto do Brasil (GUTIÉRREZ, 2006), onde é frequente o relato do ingresso de três homens africanos para cada mulher (PINHEIRO; ASSUMPÇÃO; FOLLMANN, 2015). De acordo com Gutierrez (2006), o equilíbrio entre os sexos foi a regra no Paraná (GUTIÉRREZ, 2006). Mulheres negras deviam ser para os seus escravizadores tão

eficientes quanto os homens, além de serem vítimas de estupro e utilizadas pelos “senhores” como reproduutoras da mão de obra escravizada (RIBAS, 2019). O estudo de Gutierrez (2006) também afirma que a população escravizada mencionada em registros de compra-venda no estado era predominantemente crioula (85%), ou seja, nascida nas vilas paranaenses (GUTIÉRREZ, 2006). Há também relatos recorrentes do ingresso de escravizados africanos a bordo de navios negreiros no litoral paranaense após a proibição do tráfico pela Lei Feijó⁵ (também chamada Lei de 7 de novembro de 1831) (WESTPHALEN, 1972). No tráfico estariam envolvidas as autoridades de Paranaguá, facilitando o tráfico clandestino, no mínimo, até 1850 (WESTPHALEN, 1972). Outros pontos de desembarque teriam sido a Ilha da Catinga, a Ilha do Mel e o litoral de Guaratuba (WESTPHALEN, 1972). Contudo, por se tratar de tráfico ilegal, não existe registro do número ou da procedência dos escravizados (WESTPHALEN, 1972). Uma exceção é o registro do ingresso de 2 africanos em navios vindos de Montevidéu e do Rio de Janeiro, em 1826 (WESTPHALEN, 1972).

Para o Paraná, africanos escravizados podem também ter vindo de São Paulo pela via dos tropeiros (e, portanto, ingressados ao Brasil através do porto do Rio de Janeiro), embora segundo as fontes históricas esses ingressos seriam ocasionais (GUTIÉRREZ, 2006). Quanto às etnias dos africanos relatadas nas vilas do Paraná (Antonina e Castro, 1804) estes seriam na sua maioria Bantus (em Antonina 88% sendo eles 27% Benguela, 55% Angola, 4% Congo, 2% Rebolo; em Castro 95,2% sendo eles 63,3% Benguela, 19% Angola, 6,8% Congo, 4,1% Rebolo, 1,4% Cassange e 0,7% Cabinda), enquanto uma porcentagem baixa seria da costa noroeste africana (em Antonina 12%, sendo eles 8% Mina e 4% Guiné; em Castro 4,8%, sendo eles 4,1% Mina e 0,7% Guiné) (GUTIÉRREZ, 2006). Proporções similares foram relatadas em Guaratuba e Paranaguá (GUTIÉRREZ, 2006).

Os territórios onde se organizaram quilombos no estado de Paraná foram diversos, desde terras ocupadas por ordens religiosas e logo cedidas a escravizados libertos, fazendas abandonadas e ocupadas por escravizados fugidos, fazendas cedidas por antigos escravizadores, pagamento por serviços prestados ao governo (como a participação na Guerra do Paraguai - 1864 a 1870-), e terras compradas por negros alforriados (PRIORI et al., 2012). Em muitos

⁵ A Lei Feijó proibia a importação de escravizados para o Brasil e obrigava o retorno dos africanos trazidos para solo brasileiro. A Lei, assim como as demais imposições antiescravistas do governo britânico na época obedeciam à intenção britânica de fazer tanto da África quanto da América um mercado de consumo para os seus produtos manufaturados, já que a potência tentava instalar o sistema capitalista e se beneficiar do mesmo. Para tal fim, precisava que a maior parte da população brasileira, então escravizada e sem dinheiro, fosse livre (Westphalen, 1988; Rodney, 1972).

casos, a doação de terras pelos escravistas esteve condicionada à permanência dos negros na prestação dos seus serviços à família do doador (RAGGIO; BLEY; TRAUCZYNSKI, 2018).

Entre os anos 2005-2010 o Grupo de Trabalho Clóvis Moura identificou a presença de 86 comunidades negras no território paranaense, ainda que não todas quilombolas. Dentre elas, 36 foram auto reconhecidas e certificadas pela Fundação Cultural Palmares (FCP) como Comunidades Remanescentes de Quilombo (CRQ) (GRUPO DE TRABALHO CLÓVIS MOURA, 2010).

No Município da Lapa, na região metropolitana de Curitiba, localizam-se três núcleos de quilombo: as comunidades Restinga, Feixo e a Vila Esperança de Mariental, a aproximadamente 10 km do centro da cidade da Lapa (Figura 8). As três comunidades foram certificadas enquanto Comunidades Remanescentes Quilombolas (CRQ) pela Fundação Cultural Palmares em 2006 (FUNDAÇÃO CULTURAL PALMARES, 2021). Nessa seção será abordada a história das comunidades Restinga e Feixo, cuja ancestralidade genética é o objeto de estudo desse trabalho.



Figura 8 - Mapa do Brasil mostrando o município da Lapa e a localização das três comunidades quilombolas ali presentes. Fonte: A autora (2023)

A adesão da região da Lapa ao modelo de economia escravista e a presença de escravizados africanos é registrada desde o início da urbanização (FAJARDO; CUNHA, 2021). A região foi formada por grandes fazendas originadas em sesmarias concedidas a portugueses ou seus descendentes (CIGOLINI; DA SILVA, 2018).

As comunidades Feixo e Restinga foram fundadas antes da promulgação da Lei Áurea do 13 de maio de 1888, por negros escravizados por famílias latifundiárias escravistas da oligarquia regional (RIBAS, 2019). A família Braga, que escravizou aos ancestrais dos quilombolas do Feixo, e a família Alves de Araújo, quem escravizou aos ancestrais dos quilombolas da Restinga, na Fazenda Santa Amélia (RIBAS, 2019). Os fundadores dos quilombos se assentaram nas terras onde já trabalhavam escravizados, recebendo-as como “doação” dos seus escravizadores (RIBAS, 2019). Em 2012, contabilizavam-se 343 pessoas morando no Feixo e 93 na Restinga (PRIORI et al., 2012). Porém, a comunidade do Feixo tem experimentado um aumento da migração de famílias externas em tempos recentes (RIBAS, 2019). Aproximadamente 50% do total de famílias destes quilombos residem no Feixo, que por sua vez encontra-se dividido nos subníveis Rincão (o maior), Maria Antonia, Campina e Pavão (PRIORI et al., 2012).

Nas comunidades existe uma agricultura orgânica e familiar, pela qual são responsáveis principalmente as mulheres, enquanto muitos quilombolas também prestam serviços para produtores agropecuários convencionais dentro e fora do quilombo, principalmente em épocas de colheitas (CIGOLINI; DA SILVA, 2018). Muitos quilombolas também trabalham fora da comunidade como empregados da indústria, comércio ou prestando serviço doméstico (RIBAS, 2019). A posse de terra é um problema sério para as comunidades que tem enfrentado a grilagem e a perda de terras por pagamentos muito abaixo do seu valor (PRIORI et al., 2012). O tamanho limitado dos terrenos e a necessidade de repartição dos mesmos entre os seus descendentes, fez com que a produção seja insuficiente para a subsistência das famílias e tenha aumentado a necessidade de busca de fontes externas de renda (CIGOLINI; DA SILVA, 2018).

Por outro lado, a criação de uma rodovia federal (a BR-476) em 1967 tem facilitado a locomoção e a comunicação com centros urbanos próximos, como a Lapa, Mariental e Curitiba, a capital do estado (CIGOLINI; DA SILVA, 2018). O centro urbano mais próximo às comunidades é a Colônia Mariental, formada a partir do assentamento de imigrantes russos e alemães (em 1978), onde muitos quilombolas residem hoje, trabalham ou frequentam a escola (CIGOLINI; DA SILVA, 2018; RIBAS, 2019). Dessa maneira, a interação com estes centros é intensa, não se tratando de quilombos isolados ou semi-isolados (CIGOLINI; DA SILVA,

2018). A partir da sua certificação como CRQ, os quilombolas contam com o auxílio de Agentes de Saúde Comunitária que devem realizar visitas regulares aos domicílios para verificar a saúde dos quilombolas (CIGOLINI; DA SILVA, 2018). Práticas alternativas de atenção à saúde, vinculadas ou não à tradição quilombola, como benzimentos, uso de plantas e bioenergia, também são frequentes entre os moradores.

2.5 ASPECTOS GENÉTICOS E EVOLUTIVOS DA ANCESTRALIDADE AFRICANA NA INTERSECÇÃO COM A SAÚDE

Numerosos estudos demonstraram que o uso de critérios baseados em raça/cor ou etnia autodeclarada como índices de ancestralidade genética são imprecisos e levam a confundir o risco biológico de doenças ou outros impactos em saúde com o seu risco devido a fatores ambientais (como fatores socioeconômicos ou o racismo, por exemplo) (TATONETTI; ELHADAD, 2021). Trabalhos recentes tem ainda demonstrado que não apenas os achados de estudos realizados em populações de um continente não podem ser extrapolados a outros continentes mas também que estes muitas vezes não se aplicam sequer a populações com ancestralidade subcontinental distinta (BELBIN et al., 2021).

Neste sentido, quando analisamos os impactos da ancestralidade genética na saúde de uma população é importante destacar que há uma necessidade tanto de estudos que permitam conhecer a ancestralidade genética da população (tanto no nível continental quanto subcontinental) como de estudos que forneçam um marco teórico adequado para o tratamento dessa população atendendo as suas demandas particulares (PEPRAH et al., 2015; SIRUGO; WILLIAMS; TISHKOFF, 2019; TATONETTI; ELHADAD, 2021).

Neste trabalho estudamos a ancestralidade autossômica e uniparental de populações brasileiras, incluindo as comunidades quilombolas Feijo e Restinga. Nessas populações, abordamos também o estudo de polimorfismos no gene *G6PD*, que possuem múltiplos impactos na saúde humana, como se descreve na secção seguinte.

2.5.1 POLIMORFISMOS NO GENE *G6PD*: ASPECTOS EVOLUTIVOS E IMPACTOS NA SAÚDE

A deficiência de glicose-6-fosfato desidrogenase (G6PDd) é uma enzimopatia ligada ao cromossomo X, recessiva, que apresenta maior prevalência em pretos e pardos do que em brancos, no Brasil (SANTOS et al., 2016b). Os sintomas da deficiência ocorrem principalmente em indivíduos do sexo masculino, os quais são hemizigotos para as variantes deletérias, e em indivíduos do sexo feminino homozigotos para as mesmas variantes (DALLOL et al., 2012). Sintomas também têm sido observados em pessoas do sexo feminino com variantes deletérias em heterozigose, devido aos efeitos de inativação do X (DALLOL et al., 2012). A G6PDd é causada por mutações no gene *G6PD*, localizado próximo da região telomérica no braço longo do cromossomo X (Xq28), que codifica a enzima glicose-6-fosfato desidrogenase (GÓMEZ-MANZO et al., 2016; OLIVEIRA et al., 2018) (Figura 9). Essa enzima desempenha uma função essencial no metabolismo energético a partir de glicose e na produção de NADPH nos eritrócitos e, por este motivo, possui um importante papel de proteção contra o dano oxidativo nessas células (SANTOS et al., 2016b). A G6PD catalisa a primeira reação da via das pentoses fosfato (PPP), que permite a conversão de glicose-6-fosfato a 6-fosfogluconolactona. Através dessa reação, nicotinamida adenina dinucleótido fosfato (NADP) é reduzido a NADPH (Figura 10). NADPH é um cofator fundamental na proteção celular contra níveis fisiologicamente elevados de substâncias oxidantes (MASON; BAUTISTA; GILSANZ, 2007). Nas hemácias maduras, a PPP é a única fonte de NADPH, portanto os danos oxidativos em indivíduos G6PD deficientes desencadeiam hemólise (MASON; BAUTISTA; GILSANZ, 2007). Pessoas G6PD deficientes não apresentam sintomas enquanto elas não consomem substâncias cujo processo de assimilação produz agentes oxidantes (como certos medicamentos, feijões fava, corantes, entre outros) assim como também não tenham processos infecciosos causados por determinados agentes (MASON; BAUTISTA; GILSANZ, 2007) (Figura 10). Os eventos antes mencionados podem desencadear três quadros clínicos principais: icterícia neonatal, anemia hemolítica crônica não esferocítica (CNSHA) e anemia hemolítica aguda (AHA) (HOWES et al., 2013) (Figura 10).

Mais de 1300 variantes já foram descritas no gene *G6PD* (GECK; POWELL; DUNHAM, 2023). Apenas 400 destas variantes possuem alguma interpretação funcional (GECK; POWELL; DUNHAM, 2023). Dentre elas, cerca de 200 são raras e estão associadas com as formas clinicamente mais severas da deficiência (GECK; POWELL; DUNHAM, 2023; OLIVEIRA et al., 2018) (Figura 9). A maioria das variantes associadas à G6PDd alteram a sequência aminoacídica da proteína, embora existem também algumas variantes regulatórias (LUZZATTO; ALLY; NOTARO, 2021).

Existe uma sobreposição da distribuição global da G6PDd com a extensão geográfica de espécies do gênero *Plasmodium* sp, atingindo frequências elevadas de G6PDd na África Central (9-20%) e apresentando frequências variáveis em outras regiões da África (DA ROCHA et al., 2021; GAMPIO GUEYE et al., 2019). A correlação entre ambos os eventos tem sido explicada com base numa hipótese de seleção natural, já que vários estudos associaram a G6PDd à proteção contra a malária (GAMPIO GUEYE et al., 2019; GOMEZ; HIRBO; TISHKOFF, 2014).

No Brasil, a área endêmica de malária corresponde à Região Amazônica (Acre, Amazonas, Amapá, Pará, Rondônia, Roraima, Tocantins e parte do Mato Grosso e do Maranhão) (PREFEITURA DE SÃO PAULO, 2018). Entretanto, existe transmissão residual de malária nos estados de Mato Grosso do Sul, Paraná, São Paulo, Minas Gerais, Rio de Janeiro e Espírito Santo e em regiões de Mata Atlântica (PREFEITURA DE SÃO PAULO, 2018). Desafortunadamente, a única droga antimalária que cura a infecção e previne a reinfecção por malária, a primaquina (entre outros medicamentos derivados de aminoquinolinas), pode produzir anemia hemolítica severa em indivíduos G6PD deficientes (DOMBROWSKI et al., 2017). No Brasil, a testagem da G6PD não é obrigatória antes do tratamento com primaquina, e, por este motivo, o estudo da prevalência das variantes que a causam é essencial, principalmente em indivíduos de ancestralidade africana (DOMBROWSKI et al., 2017). Além da primaquina, G6PDd é também um fator a ser considerado antes da administração de outros medicamentos antimaláricos também derivados do composto 8-aminoquinolina, como a cloroquina e a hidroxicloroquina (em diante CQ e HCQ, respectivamente). No Brasil, CQ e HCQ tem sido propostas como tratamento contra a doença pelo novo coronavírus (COVID-19), embora com estudos demonstrando que o seu uso não é eficaz contra a infecção (DA ROCHA et al., 2021; HOFFMANN et al., 2020). Desta maneira, o uso de CQ e HCQ para a prevenção e tratamento da COVID-19 tem sido amplamente fomentado por autoridades governamentais durante os primeiros anos da pandemia, levando a um aumento expressivo do seu uso pela população brasileira em 2020 (MELO et al., 2021). Ainda, estudos tem sugerido que o uso de CQ e HCQ podem estar associados com um incremento da mortalidade por COVID-19 (MELO et al., 2021).

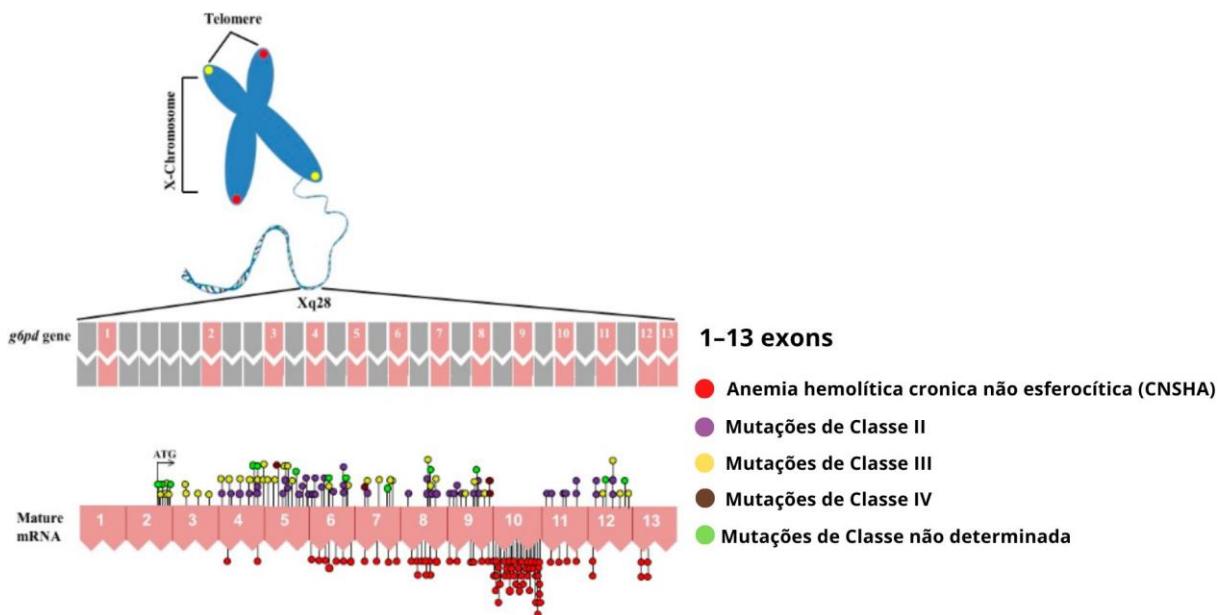


Figura 9 - Esquema da região codificante do gene *G6PD* destacando a localização das variantes associadas à deficiência de *G6PD*. Fonte: adaptado de GÓMEZ-MANZO et al. (2016)

A deficiência de G6PD também pode desencadear anemia hemolítica após o uso de outros medicamentos com potencial oxidativo amplamente utilizados, como aspirina, paracetamol ou dipirona, o que ressalta a necessidade da avaliação das variantes associadas a G6PDd na população brasileira (RANGEL; RANGEL, 2022).

Por outro lado, numerosos estudos sugerem que a G6PDd possui uma relação com a etiologia, desfechos e o diagnóstico do diabetes mellitus, doença crônica e multifatorial com maior incidência e mortalidade em negros (BRITO; LOPES; ARAÚJO, 2001). Vários estudos epidemiológicos apontam a deficiência de G6PD como um fator de risco para diabetes (QUARTEY et al., 2020; ZHANG et al., 2010). Tem sido sugerido que a deficiência de G6PD leva à disfunção e à morte das células β -pancreáticas, devido a uma alta sensibilidade das mesmas às espécies reativas de oxigênio (ZHANG et al., 2010). Em contrapartida, a hiperglicemias (aumento do nível de glicose no sangue) pode levar à diminuição da expressão da G6PD e, portanto, da sua atividade enzimática, levando a uma piora recíproca dessas duas condições (CARETTE et al., 2011a). Além disso, a ocorrência de hemólise em indivíduos G6PD deficientes pode levar ao subdiagnóstico da diabetes, já que um dos principais testes realizados para esse fim é a medida de hemoglobina glicada (HbA1c) (WHEELER et al., 2017).

O exame de HbA1c avalia o índice glicêmico do paciente quantificando a glicose ligada à hemoglobina presente nas hemácias, sendo potencialmente alterada pela hemólise (WHEELER et al., 2017). Outras interações da deficiência de G6PD com diabetes tem sido descobertas, como a ocorrência de episódios de hemólise decorrente de cetoacidose diabética e após o uso de medicamentos antidiabéticos como a glibenclamida (ANSARI et al., 2022; CARETTE et al., 2011b) (Figura 10).

Neste trabalho, foram analisados os SNPs rs1050829, rs1050828, rs5030868 e rs76723693 (entre outros – ver Capítulo IV -), que determinam as variantes G6PD*A- e G6PD*A, com origem e distribuição na África Central, e G6PD-Med (ou G6PD-B-), esta última de origem no Mediterrâneo com distribuição no oeste da Ásia (da Turquia, Arabia Saudita até a Índia) (HOWES et al., 2013). A variante G6PD*A- é comum na África subsaariana, afetando entre 15 e 20% da população (HOWES et al., 2013). Ela é determinada pela mutação rs1050829 em combinação com uma segunda mutação, sendo esta última mais frequentemente os SNPs rs1050828 ou rs76723693, assim como também outros não avaliados nesse trabalho (OLIVEIRA et al., 2018). G6PD*A- determina uma atividade enzimática correspondente a 10 até 60% da atividade normal (OLIVEIRA et al., 2018). A variante G6PD*A é originada pelo SNP rs1050829, determinando uma deficiência enzimática leve (atividade enzimática próxima da normal) e ocorre em cerca de 20 a 30% da população da África subsaariana. A variante G6PD-Med, causada pela mutação rs5030868, é responsável por um dos fenótipos mais severos, com menos de 1% da atividade enzimática (DOMBROWSKI et al., 2017).

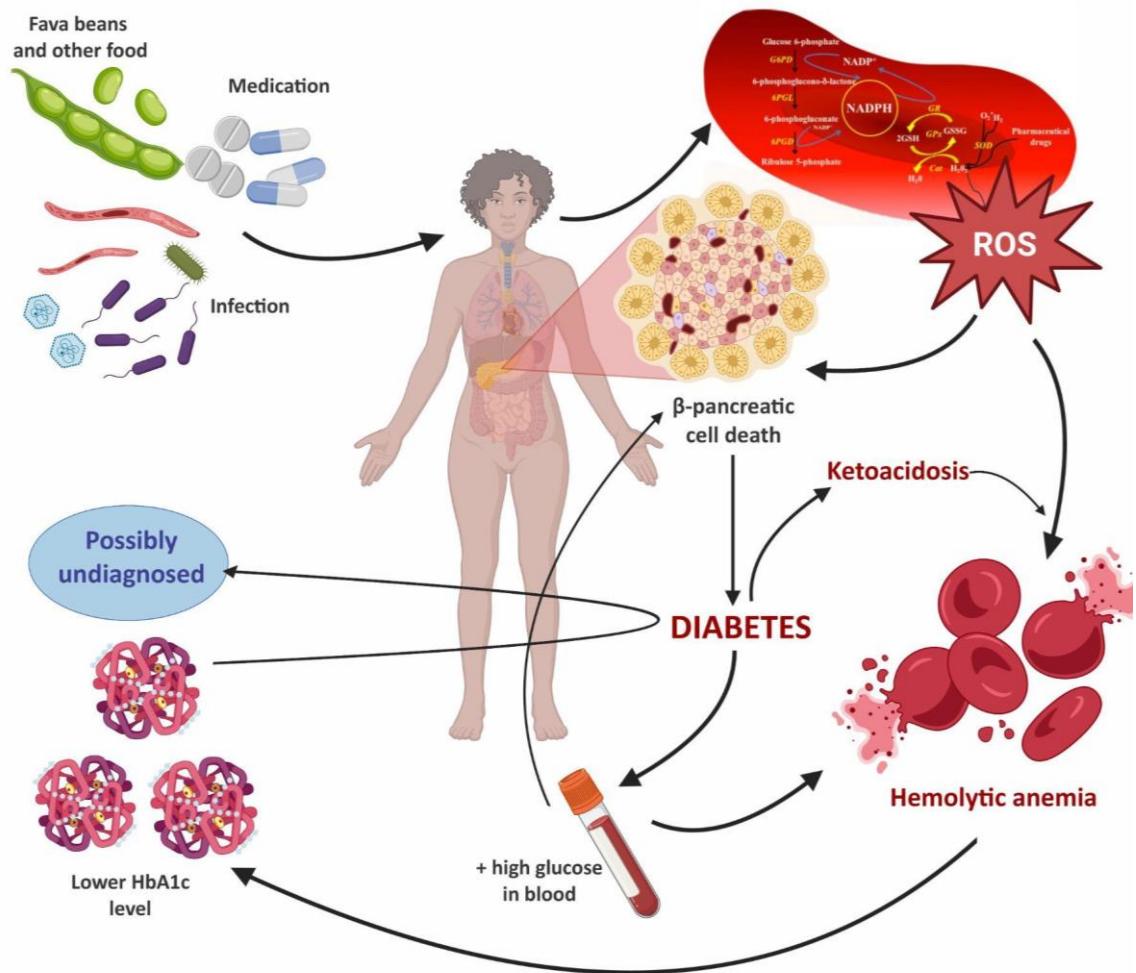


Figura 10 - Esquema dos potenciais impactos em saúde da deficiência de G6PD. Fonte: adaptado de GÓMEZ-MANZO et al. (2016). Criado com BioRender.

3 ESTRUTURA DA TESE

As seguintes secções da presente tese de doutorado descrevem os objetivos da mesma, assim como a metodologia empregada e os principais resultados obtidos. A secção de Materiais e Métodos descreve a estratégia geral utilizada para atingir os objetivos estabelecidos e as principais metodologias experimentais utilizadas. A secção de Resultados está dividida em quatro capítulos. Os capítulos 1 e 2 correspondem a artigos publicados em revistas internacionais. Os capítulos 3 e 4 são artigos de pesquisa científica ainda não publicados e que se encontram em preparação. A metodologia específica empregada para a obtenção dos resultados de cada capítulo encontra-se descrita no próprio artigo científico. No ANEXO I se encontra o projeto de pesquisa encaminhado ao Comité de Ética em Pesquisa do setor de Ciências da Saúde da UFPR. No ANEXO II se encontra o artigo de divulgação científica publicado no marco deste trabalho.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Determinar a estrutura populacional e as principais contribuições genéticas africanas de populações miscigenadas urbanas e afro-brasileiras (quilombos), através do uso de marcadores uniparentais e verificar a frequência de variantes africanas associadas à deficiência de glicose-6-fosfatodesidrogenase nas mesmas populações.

4.2 OBJETIVOS ESPECÍFICOS

1. Verificar o estado da arte dos estudos genéticos em populações quilombolas brasileiras gerando um artigo de revisão.
2. Determinar as principais contribuições continentais ao pool gênico uniparental de populações urbanas das Regiões Centro-Oeste, Sudeste e Sul do Brasil e dos quilombos Restinga e Feixo da Região Sul do Brasil e verificar se estas populações apresentam um padrão de miscigenação enviesado pelo sexo.
3. Determinar as principais contribuições genéticas maternas ameríndias e africanas, no nível subcontinental, ao pool gênico uniparental de populações urbanas das Regiões Centro-Oeste, Sudeste e Sul do Brasil e dos quilombos Restinga e Feixo da Região Sul do Brasil.
4. Determinar a frequência de polimorfismos de nucleotídeo único (SNPs) associados a deficiência de glicose-6-fosfatodesidrogenase (G6PD) em populações urbanas das Regiões Centro-Oeste, Sudeste e Sul do Brasil e nos quilombos Restinga e Feixo da Região Sul do Brasil.

5 MATERIAIS E MÉTODOS

Este trabalho foi dividido em duas etapas. Em uma etapa inicial foi realizado um estudo *in silico* de ancestralidade global a partir de um conjunto de dados genotípicos de uma amostra miscigenada predominantemente euro-brasileira das Regiões Centro-Oeste (n = 321), Sudeste (n = 42) e Sul (n = 74) do Brasil, e de uma amostra de 24 indivíduos indígenas dos grupos étnicos Guaraní Kaiowá (n = 7), Guaraní M'byá (n = 7) e Kaingang (n = 10), genotipados por meio do microarranjo Infinium® CoreExome-24 v1.1 BeadChip da illumina®. Estes dados foram previamente obtidos em um projeto do Laboratório de Genética Molecular Humana (AUGUSTO et al., 2021).

Inicialmente foi inferida a ancestralidade uniparental dos cromossomos mitocondrial e Y e a ancestralidade global do cromossomo X, as quais foram comparadas com ancestralidade autossômica previamente estudada para o mesmo conjunto de amostras (CALONGA-SOLÍS et al., 2020). Adicionalmente, no mesmo conjunto de dados, foram investigados SNPs associados à deficiência de glicose-6-fosfato desidrogenase (G6PD) quanto a sua frequência em negros e em brancos brasileiros. Para esta etapa, e por se tratar de dados previamente obtidos, a classificação étnico-racial dos participantes utilizada foi a heteroclasseificação (determinada pelo pesquisador no momento da coleta da amostra e baseado principalmente em observação de cor da pele e traços fenotípicos) ou a autodeclaração no caso de participantes autodeclarados indígenas.

Em uma segunda etapa foi realizado o estudo de duas populações afro-brasileiras do estado de Paraná: as comunidades quilombolas Restinga e Feixo, ambas do município da Lapa. Neste estudo foi realizada:

- Elaboração e aprovação do projeto de pesquisa pelo Comitê de Ética em Pesquisa em seres humanos, Setor de Ciências da Saúde (CEP/SD, CAAE número 42761521.9.0000.0102)
- Aplicação de questionários para levantamento de informação etnográfica, estilo de vida e dados epidemiológicos nas comunidades quilombolas Restinga e Feixo. Os questionários

foram baseados nos questionários oficiais do governo brasileiro utilizados na Pesquisa Nacional em Saúde (PNS, <https://www.gov.br/>) e podem ser consultados no ANEXO I.

- Coleta de amostras de sangue periférico, medidas antropométricas, teste glicêmico e aferição de pressão.
- Extração do ADN de amostras de sangue periférico e análises laboratoriais (perfil lipídico, glicemia, entre outros).
- Genotipagem de 6 SNPs do genoma mitocondrial (ADNm_t) e 9 SNPs do cromossomo Y das amostras de ADN através de PCR alelo específica e eletroforese em gel de agarose em 169 amostras dos quilombos Restinga e Feixo.
- Estudo de ancestralidade materna subcontinental africana nos quilombos Restinga e Feijo mediante o sequenciamento das regiões hipervariáveis I, II e III do genoma mitocondrial (HV-I, -II e -III) dos indivíduos previamente classificados como portadores de linhagem materna africana através da genotipagem dos 6 SNPs mitocondriais (76 indivíduos).
- Análise bioinformática de ancestralidade uniparental nas duas comunidades quilombolas.

Nesta segunda etapa, a determinação do grupo étnico-racial foi baseada na autodeclaração dos participantes e de acordo as categorias utilizadas pelo Instituto Brasileiro de Geografia e Estatística. O Projeto aprovado pelo Comitê de Ética em Pesquisa, Ciências da Saúde - UFPR, encontra-se no ANEXO I.

A continuação se detalham as principais metodologias experimentais empregadas neste trabalho.

5.1 COLETA E PROCESSAMENTO DE AMOSTRAS DE SANGUE

A coleta de sangue dos participantes foi realizada em dois tubos de 4ml com anticoagulante (EDTA) para a extração de ADN e separação de plasma. Foi realizada a homogeneização de todos os tubos com anticoagulante imediatamente após a coleta e até o seu

processamento. Dois tubos de 4ml com anticoagulante (EDTA) foram também coletados e utilizados para realização de hemograma e teste de hemoglobina glicada. Foi também coletado um tubo de 8 ml sem anticoagulante para a coleta de soro de cada participante. As amostras foram centrifugadas frescas. Para a separação de plasma, ADN, e soro os tubos foram centrifugados durante 15 minutos a temperatura ambiente (20°C) a 2500 rpm. Para a extração de ADN (ver próxima secção) a camada leucocitária foi separada e armazenada em microtubo estéril de 2 ml e congelado a -80°C até a extração. Os tubos reservados para a realização de hemograma não foram centrifugados e foram armazenados a 4°C para posterior análise. As amostras de plasma e soro foram aliquotadas em microtubos estéreis de 2 ml e congeladas a -80°C para posterior análise.

5.2 EXTRAÇÃO DE ADN

O ADN foi extraído da camada leucocitária isolada do sangue total, usando o kit comercial Wizard® Genomic DNA Purification Kit (Promega), seguindo as recomendações do fabricante. Após a extração, as amostras foram armazenadas na sua concentração original a -80°C. Posteriormente, as amostras foram diluídas a 20 ng/ul em água ultrapura e armazenadas a -20°C até a realização dos experimentos.

5.3 PCR ALELO-ESPECÍFICA

As genotipagens de polimorfismos de nucleotídeo único dos cromossomos mitocondrial, Y e X foram realizadas através de reação em cadeia da polimerase (PCR, do inglês *Polymerase Chain Reaction*) alelo-específica, utilizando reações separadas para cada alelo. As reações consistiram no uso de um oligonucleotídeo iniciador (*primer*) alelo-específico (*forward* ou *reverse*) e de um segundo *primer*, não alelo-específico, comum para ambas as reações (*forward* ou *reverse*), exceto no caso das variantes do gene *G6PD*, onde ambos os *primers forward* e *reverse* foram alelo-específicos. Os *primers* utilizados se detalham nas tabelas 1-4, junto com as condições experimentais otimizadas de cada reação.

Para a genotipagem de linhagens maternas foram avaliados os estados alélicos dos SNPs mitocondriais 10400, 10873, 1736, 827, 15487 e 5178 (rs28358278, rs2857284, rs28358574,

rs28358569, rs28357370 e rs28357984, respectivamente). Os programas de ciclagem utilizados consistiram em: uma etapa de desnaturação inicial a 94°C por 3 minutos, 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento de 30 segundos (a Tm é descrita na Tabela 1 para cada par de *primers*) e uma etapa de extensão a 72°C por 1 minuto. Por último, utilizou-se uma etapa de extensão final a 72°C por 1 minuto.

Tabela 1 - Sequência dos *primers* utilizados para a genotipagem de marcadores de linhagem materna e as condições de reação otimizadas para cada par alelo-específico. Hg: haplogrupo. **Tm:** temperatura de anelamento. **CI:** controle interno da PCR. Concentrações ([]) referem-se às finais otimizadas para cada par de *primers*. As condições de reação descritas foram otimizadas com a Taq polimerase GoTaq®Flexi DNA Polymerase (Promega). Com exceção do primer R2, cuja sequência foi retirada da referência CARDOSO e col., 2013, todos os demais *primers* foram retirados da referência SCHWAB, 2019. Fonte: A autora (2023)

Hg	SNP	Nome do primer	Sequência (5'-3')	Amplicon (pb)	Tm (°C)	[Mg ⁺⁺] (mM)	CI [CI] (μM)	[Taq] (U/μL)	[ADN] (ng/μL)
M	10400	F- 10400*T	taattataaaaaaggattagactgtct	149 (M)	58,6	1,5 [0,25]	HGH	0,04	3
		F- 10400*C	tacaaaaaggattagacagaacc	142 141	58,6 [0,25]	2	HGH	0,04	3
		R-10400	gaagtggatggataatgttag	-	-	-	-	-	-
N	10873	F- 10873*T	aatataccaaattttatcatcaac	136 (N)	58,6	1,5 [0,25]	HGH	0,04	3
		F- 10873*C	cctaattttatgcattcccc	131 130	58,6 [0,25]	1,5	HGH	0,04	3
		R-10873	tggtagtgatcgaggtagttag	-	-	-	-	-	-
A	1736	F-1736*A	taaagccaaaccattttccaa	410	53	1,5 [0,1]	HVI	0,04	3
		F-1736*G	agccaaaccatttatccag	407	53 [0,1]	1,5	HVI	0,04	3

		R2-1736	tgtccaaagagctgtcctct	-	-	-	-	-	-
B	827	F-827*G	cgggaaacagcagtgatttag	85 (B)	60	1	HGH [0,25]	0,04	3
		F-827*A	ttaacggaaacagcagtgaata a	89	60	1,5	HGH [0,25]	0,04	6
		R-827	tggcacgaaattgaccaac	-	-	-	-	-	-
M8 (C)	15487	R- 15487*T	gtcttggtegccttaggaagtca	70 (M8)	55	2	WSF1 [0,1]	0,02	3
		R- 15487*A	tggtcgcctaggagttct	66	54	1	HGH [0,2]	0,04	3
		F-15487	cttacttcttccctctc	-	-	-	-	-	-
D	5178	F-5178*A	tgatcaacgcacctaacaaga	107 (D)	55	1	HVI [0,2]	0,02	3
		F-5178*C	gtgcacactgaagcaagc	102	55	1	HVI [0,2]	0,02	3
		R-5178	cccatttgagcaaaaagcc	-	-	-	-	-	-

Para a genotipagem de linhagens paternas foram avaliados os estados alélicos dos seguintes SNPs do cromossomo Y: M242, M207, M89, M9, P152, M242, M3, M346, M2, M35 (rs8179021, rs2032658, rs2032652, rs3900, rs9786634, rs8179021, rs3894, rs568022901, rs9785941 e rs375228668, respectivamente). Os programas de ciclagem utilizados consistiram em: uma etapa de desnaturação inicial a 94°C por 3 minutos, 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento de 30 segundos (a Tm é descrita na Tabela 2 para cada par de primers) e uma etapa de extensão a 72°C por 1 minuto. Por último, utilizou-se uma etapa de extensão final a 72°C por 1 minuto.

Tabela 2 - Sequência dos primers utilizados para a genotipagem de marcadores de linhagem paterna e as condições de reação otimizadas para cada par alelo-específico. Hg: haplogrupo. Tm: temperatura de anelamento. CI: controle interno da PCR. Concentrações ([]) referem-se às finais otimizadas para cada par de primers. As condições de reação descritas foram otimizadas com a Taq polimerase GoTaq®Flexi DNA Polymerase (Promega). Primers com (*) foram retirados da referência MEDINA e col., 2014, primers com (**) foram desenhados neste trabalho. Fonte: A autora (2023)

Hg	SNP	Nome do primer	Sequência (5'- 3')	Amplicon (pb)	Tm (°C)	[Mg ⁺⁺] (mM)	CI/[CI] (μM)	[Taq] (U/μL)	[ADN] (ng/μL)
Q*	M24 2	M242F1C	aaagggtgaccaagggtgctc	131	55	1,5	HVII/II I [0,25]	0,02	3
		M242F2T	aacaataaaggttaccaagg tgctt	137 (Q)	55	1,5	HVII/II I [0,25]	0,02	3
		M242R	tcagcataatacccttacctag aacaac	-	-	-	-	-	-
R*	M20 7	M207F1G	cgagtcaagcaagcaattta g	124 (R)	55	1	HVI [0,15]	0,02	3
		M207F2A	gttcactaagtcaagcaaga aatctaa	118	53	2	HVI [0,15]	0,04	3
		M207R	aaaagctgaaggaaaagtg ga	-	-	-	-	-	-
F (XK,Q,R) *	M89	M89R1	cgcaggcaaagtgagagat a	85 (F)	55	1,5	HVI [0,15]	0,02	3
		M89R2	tttaatcaggcaaagtgaga gatg	89	55	1,5	HVI [0,15]	0,02	3
		M89F	cacagaaggatgtgtca	-	-	-	-	-	-
K(x Q,R)*	M9	M9F1G	ggcctaagatgggtggatg	70 (K)	55	1,5	HVI [0,15]	0,02	3

	M9F2C	tttggacctaagatggttgaa tc	74	55	1,5	HVI	0,02	3	
	M9R	tgtaagacattgaacgttgg a	-	-	-	-	-	-	
E*	P152	P152F1G	gccccatcggtttctgaag	141	55	1	HVI	0,02	3
	P152F2C	ataaaatgctcatccgtttctg aac	147 (E)	55	1,5	HVI	0,02	3	
	P152R	tctcagggttgtgcagattt	-	-	-	-	-	-	
Q1b1a1a*	M3	M3R1T	cgtaccagactttcccaatta (Q1b1a1a)	115	55	1,5	HVI	0,02	3
	M3R2C	atttttaagtaccagacttct aatttg	121	55	1,5	HVI	0,02	3	
	M3F	aaggcacattcgccggata	-	-	-	-	-	-	
Q1b*	M34 6	M346R1G	cgcacattacaacgacaagc 160 (Q1b)	55	1,5	HVI	0,02	3	
	M346R2C	aataatgtcaatttacaaag acaagg	167	55	1,5	HVI	0,02	3	
	M346F	aaggcctgaaaatgtggaa a	-	-	-	-	-	-	
E1b1a1**	M2	M2_G	cctttatecctccacagatctc ac	153 (E1b1a1)	60	1	HGH	0,02	3
	M2_A	cctttatcctccacagatctc at	153	60	1	HGH	0,02	3	
	M2_F	aagtccagacccaggaagg t	-	-	-	-	-	-	

E1b1b1**	M35	M35_G	ggagtcctgcctgtgtcc	713	55	1,5	HGH	0,02	3
						[0,15]			
		M35_C	ggagtcctgcctgtgtcg	713	55	1,5	HGH	0,02	3
						[0,15]			
		M35_F	ctgggacatccgttgcatt	-	-	-	-	-	-

Para a genotipagem de variantes do gene *G6PD* foram utilizados os *primers* apresentados na Tabela 3 e na Tabela 4. Os programas de ciclagem utilizados consistiram em: uma etapa de desnaturação inicial a 94°C por 3 minutos, 35 ciclos de desnaturação a 94°C por 30 segundos, etapa de anelamento de 30 segundos (a Tm é descrita na - para cada par de *primers*) e uma etapa de extensão a 72°C por 1 minuto. Por último, utilizou-se uma etapa de extensão final a 72°C por 1 minuto.

Tabela 3 - Sequência dos *primers* utilizados na genotipagem de variantes do gene *G6PD*. Os *primers* foram desenhados neste trabalho. Fonte: A autora (2023)

Gene	Nome do <i>primer</i>	Sequência (5' - 3')
<i>G6PD</i>	G6PD_isoA_T_F	ccccaggtggaggcatt
	G6PD_isoA_C_F	ccccaggtggaggcattc
	G6PD_isoB_G_R	tgcggaaaacaccttcatcg
	G6PD_isoB_A_R	tgcggaaaacaccttcatca

Tabela 4 - Combinações de primers utilizadas para a genotipagem de variantes do gene G6PD e condições de reação utilizadas. Tm: temperatura de anelamento otimizada para a reação de PCR alelo-específica. CI: controle interno da PCR. Concentrações ([]) referem-se às condições de reação finais otimizadas para cada par de primers com a Taq polimerase GoTaq®Flexi DNA Polymerase (Promega). Fonte: A autora (2023)

Variante	Nome do primer	Amplicon (pb)	Tm (°C)	[Mg ⁺⁺] (mM)	CI/[CI] (μM)	[Taq] (U/μL)	[ADN] (ng/μL)
G6PD*B (selvagem)	<u>G6PD_isoA_T_F</u>	763	55	1,2	HGH [0,15]	0,02	3
	G6PD_isoB_G_R						
G6PD*A	<u>G6PD_isoB_G_R</u>	763	55	1	HGH [0,15]	0,02	3
	G6PD_isoA_C_F						
G6PD*A-	<u>G6PD_isoA_C_F</u>	763	55	1	HGH [0,15]	0,02	3
	G6PD_isoB_A_R						

A fim de evitar falsos negativos por falha do sistema de reação, as reações alelo-específicas foram realizadas em multiplex junto com um controle interno. O mesmo consistiu em um par de iniciadores adicional com amplificação invariável entre indivíduos e tamanho de amplicon diferenciável do fragmento amplificado pelo par de iniciadores alelo-espécifico. As características dos controles internos utilizados se detalham na Tabela 5. No intuito de evitar a ocorrência de falsos positivos e negativos na genotipagem foram também utilizados sempre controles positivos e negativos (amostras de genótipo conhecido) para cada SNP avaliado.

Tabela 5 - Sequência dos *primers* utilizados como controles internos nas reações de genotipagem alelo-específicas. Fonte: A autora (2023)

Gene	Nome do <i>primer</i>	Sequência (5'- 3')	Amplicon (pb)
<i>HGH</i>	HGH_F	tgccttcccaaccattccctta	434
	HGH_R	ccactcacggattctgttgtttc	
<i>WSF1</i>	WSF1_Ex8_rs71532874Gr	ggcgccatgacgggtgc	189
	WSF1_Ex8_rs35932623Cf	agcgtgctgctcagcctgc	

Todas as reações foram executadas nos termocicladores Mastercycler EP Gradient S (Eppendorf) e QIAamplifier 96 (Quiagen).

5.4 ELETROFORESE EM GEL DE AGAROSE

A verificação do tamanho dos fragmentos amplificados foi realizada através de eletroforese em gel de agarose. As condições da corrida eletroforética foram ajustadas de acordo com o fragmento amplificado. Para fragmentos abaixo de 100 pb foi utilizada uma concentração de agarose de 2,5% e os géis foram corridos a 80 V durante 1h30. Fragmentos acima de 100pb e até 200pb foram corridos em gel de agarose 2,5% a 80 V durante 1h. Fragmentos acima de 200pb foram corridos em gel de agarose 0,8-1% a 80 V durante 45 minutos. O volume aplicado foi de 3 uL de produto amplificado, exceto para fragmentos de tamanho inferior a 200pb, quando foi aplicado o volume total do produto de reação (~10uL). Para tincão do ADN no gel foi utilizado o corante GreenSafe DNA Gel Stain seguindo as recomendações do fabricante. Juntamente às amostras foram aplicados 3 uL do marcador de peso molecular Thermo Scientific™ GeneRuler 50 bp DNA Ladder.

5.5 SEQUENCIAMENTO DE SANGER

A amplificação das regiões hipervaráveis do ADNmt foi realizada por PCR utilizando os *primers* e as condições detalhados na Tabela 6. O programa do termociclador utilizado consistiu em uma etapa de desnaturação inicial a 95°C por 15 minutos, 35 ciclos de desnaturação a 95°C por 30 segundos, anelamento a 55°C por 1 minuto e 50 segundos, extensão a 72°C por 1 minuto e uma etapa de extensão final a 72°C por 10 minutos. Todas as reações foram executadas nos termocicladores Mastercycler EP Gradient S (Eppendorf) e QIAampiler 96 (Quiagen).

Tabela 6 - Sequência dos *primers* utilizados para amplificação e sequenciamento das regiões hipervariáveis I, II e III do ADN mitocondrial (HVI, HVII e HVIII, respectivamente) e condições de reação utilizadas. Tm: temperatura de anelamento. CI: controle interno da PCR. Concentrações ([]) referem-se às condições de reação finais otimizadas para cada par de *primers* com a Taq polimerase GoTaq®Flexi DNA Polymerase (Promega). As sequências dos *primers* foram retiradas da referência SCHAAAN et al., 2017. Fonte: A autora (2023)

Região	Posição	Nome do <i>primer</i>	Sequência (5'-3')	Amplico n (pb)	Tm (°C)	[Mg ⁺⁺] (mM)	[Taq] (U/μL)	[ADN] (ng/μL)
HVI	16024-	L15997	caccattagcacccaaag ct	628	55	1,5	0,02	3
	16365	H017	cccgtagtggttaatag ggt					
HVII/III	73-576	L034	ccatgcatttgtatttcg	607	55	2	0,02	3
		H629	tttgttatgggtgatgtg a					

Posteriormente, os produtos de amplificação foram purificados pelo método enzimático. Para isso foram utilizadas as enzimas exonuclease I (para a degradação de ADN de fita simples -iniciadores) e fosfatase alcalina (para a eliminação de grupos fosfato e inativação dos nucleotídeos da reação anterior). A cada 5 µl de produto de PCR foram adicionados 2 µl de solução de purificação contendo 4 U de exonuclease I (Fermentas) e 0,7 U de fosfatase alcalina (Thermo Scientific). Esta solução foi levada ao termociclador com as seguintes condições de ciclagem: ativação enzimática a 37 °C por 60 minutos e inativação enzimática a 80 °C por 20 minutos para degradação das enzimas de purificação.

O sequenciamento dos produtos de amplificação foi realizado pelo método de terminação da cadeia de Sanger. Para um volume final de 10 µl foi utilizado: 0,3 µl de Big Dye Terminator Cycle Sequencing Standart v3.1 (Life Technologies), 0,2 µl de oligonucleotídeos a uma concentração inicial de 10 pmol/µl, 1,9 µl de Big Dye Terminator v.3.1 5X Sequencing Buffer e 4 µl do produto de PCR purificado. As reações foram realizadas em placas de reação de 96 poços Microamp Optical (Life Technologies) utilizando os mesmos iniciadores da etapa de amplificação inicial. As condições de ciclagem da reação de sequenciamento consistiram em uma etapa de desnaturação inicial a 95 °C por 1 minuto; seguida por 25 ciclos de desnaturação a 95 °C por 10 segundos, anelamento a 50 °C por 5 segundos e extensão a 60 °C por 4 minutos.

Os produtos da reação de sequenciamento foram purificados a fim de evitar a interferência dos reagentes na eletroforese capilar. Para isso, em cada amostra foram adicionados 2,5 uL de EDTA 125 mM (Merck) e 30 uL de etanol absoluto (Merck). Seguidamente as amostras foram incubadas por 15 minutos a temperatura ambiente e centrifugadas a 3.220 xg por 45 minutos a 18 °C. O sobrenadante foi descartado por inversão e a placa foi centrifugada invertida até 300 xg (apenas um *spin*) e coberta por papel toalha para eliminar o excesso de álcool. Posteriormente, foram adicionados 30 uL de etanol 70% a temperatura ambiente e foi feita uma centrifugação a 3.220 xg por 5 minutos a 25 °C. O sobrenadante foi novamente descartado por inversão e foi feito um novo *spin* com a placa invertida e coberta por papel toalha. Seguidamente a placa foi coberta com papel alumínio, a fim de evitar a degradação dos fluoróforos, e deixada a temperatura ambiente durante 20 minutos para permitir a evaporação do álcool restante. Em seguida, a placa foi selada e armazenada a -20 °C. Antes da injeção das amostras no sequenciador, as mesmas foram ressuspensas em 10 uL de Hi-Di Formamida (Life Technologies). A eletroforese capilar foi realizada no sequenciador 3500xl Genetic Analyzer (Life Technologies) utilizando os seguintes reagentes: POP-7 Polymer for 3500/3500xl Genetic Analyzers, Anode Buffer Container (ABC)

3500 Series, Cathode Buffer Container (CBC) 3500 Series e Conditioning Reagent 3500 Series (Life Technologies).

6 RESULTADOS

6.1 CAPÍTULO I: Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil



Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil

Iriel A. Joerin-Luque¹ · Danillo G. Augusto^{1,2} · Verónica Calonga-Solís¹ · Rodrigo Coutinho de Almeida⁴ · Claudemira Vieira Gusmão Lopes³ · Maria Luiza Petzl-Erler¹ · Marcia Holsbach Beltrame¹

Received: 17 February 2021 / Accepted: 4 January 2022
 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The Brazilian population is a product of asymmetric admixture among European men and Amerindian and African women. However, Brazilian subcontinental ancestry is scarcely documented, especially regarding its African roots. Here, we aimed to unveil the uniparental continental and subcontinental contributions from distinct Brazilian regions, including South ($n=43$), Southeast ($n=71$), the poorly genetically characterized Central-Western region ($n=323$), and a subset of unique Brazilian Amerindians ($n=24$), in the context of their genome-wide ancestral contributions. The overwhelming majority of European Y haplogroups (85%) contrast sharply with the predominant African and Amerindian mtDNA haplogroups (73.2%) in admixed populations, whereas in Amerindians, non-Native haplogroups could only be detected through the paternal line. Our in-depth investigation of uniparental markers showed signals of an Andean and Central-Brazilian Amerindian maternal contribution to Southeastern and Central-Western Brazil ($83.1 \pm 2.1\%$ and $56.9 \pm 0.2\%$, respectively), the last having the highest paternal Amerindian ancestry yet described for an admixed Brazilian region (9.7%) and contrasting with higher Southern-Brazilian Amerindian contribution to Southern Brazil ($59.6 \pm 1\%$). Unlike the higher African Bantu contribution previously reported for the South and Southeast, a relevant Western African non-Bantu contribution was detected in those regions ($85.7 \pm 5\%$ and $71.8 \pm 10.8\%$ respectively). In contrast, a higher Bantu contribution was described for the first time in the Central-West ($64.8 \pm 1.3\%$ maternal and $86.9 \pm 9.6\%$ paternal). We observed sex-biased signatures consistent with the historically recorded Brazilian colonization and added new insights in the subcontinental maternal ancestry of Brazilians from regions never studied at this level.

Keywords Mitochondrial lineages · Y chromosome · Haplotype · Sex-biased gene flow · African diaspora · Population structure

Introduction

A plethora of studies has previously addressed the Brazilian population's phylogeographic origins, focusing on the continental admixture proportions. In this way, the African, Amerindian, and European ancestral contributions to the population of different geographic areas in Brazil have been well established (Alves-Silva et al. 2000; Guerreiro-Junior et al. 2009; Schaan et al. 2017, 2020). The study of the subcontinental ancestry, on the other hand, can reveal how different regions within a continent have contributed to an admixed population and shed light on its fine-scale structure. Because Brazilian population's history contains many gaps regarding its subcontinental ancestry contributions, here we intended to recall them through the analysis of uniparental markers. In addition, we investigated the usefulness of both

Iriel A. Joerin-Luque
 irieljoerin@gmail.com

¹ Programa de Pós-Graduação em Genética, Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal do Paraná (UFPR), Centro Politécnico, Jardim das Américas, Curitiba, Paraná 81531-980, Brazil

² University of California, San Francisco, CA, USA

³ Câmara de Educação do Campo, UFPR, Setor Litoral, Matinhos, Paraná 83260-000, Brazil

⁴ Present Address: Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

uniparental markers and phenotype of admixed Brazilians as proxies for autosomal ancestry which is especially relevant in biomedical research.

Genetic ancestry studies become more powerful and interesting when interpreted jointly with the historical context. According to historical and genetic studies, multiple Amerindian ethnic groups originated in South America, descending from migrations from Northeast Asia more than 15 thousand years ago (Pinotti et al. 2019). Although there is no certainty on the number of Amerindians that inhabited Brazil in 1500, estimations reach over 3 million individuals distributed in more than a thousand ethnic groups (Oliveira and Freire 2006). According to the Brazilian Institute of Geography and Statistics (IBGE; <https://www.ibge.gov.br>), currently, these populations have been reduced to about 800,000 people. The second migratory wave in Brazil came from Portugal, between 1500 and 1850, a period in which approximately 500,000 Portuguese arrived in Brazil (IBGE 2007). From 1500 until 1855, the Portuguese and their descendants in Brazil forced about 4 million enslaved Africans to move to Brazil. This human traffic became illegal in 1845, and the Brazilian abolition of slavery only occurred in 1888. Notwithstanding, it continued for at least 20 additional years (Reis 2007).

Western Africa was the first region to be victimized by the Portuguese traffic of human beings, in 1445 (Green 2012). At the earliest stages, the first enslaved Africans were taken to Europe (Green 2012). On the other hand, the trans-Atlantic slave trade to Brazil started from Central-Western Africa in the XVI century and that region was the most numerically exploited (Pinheiro et al. 2015). The Southeast of Africa (especially Mozambique) was victimized after 1845 (Pinheiro et al. 2015). Later, between 1820 and 1975, the Brazilian governing elite forced the incorporation of the country into the capitalist system promoted by England and the replacement of the African workforce by a "whiter" one, based on a racist ideology (Nascimento 2016). In this context, the migration of about 6 million people from Europe and Asia was then encouraged (IBGE 2007). In decreasing order, Italy, Portugal, Spain, Japan, Germany, Lebanon, and Syria were the new contributors to this migration wave (IBGE 2007).

There is a scarcity of public records that leads to uncertainty on the specific subcontinental African contributions to Brazil. The African populations are characterized by great genetic and cultural diversity, often correlated with geographical regions (Gomez et al. 2014). The correlation between geography and culture is often identified by shared languages, which are classified into four linguistic families: Niger-Congo, mainly spoken by the agriculturalist populations from Western to Eastern and Southern Africa; Nilotic, from Central and Eastern pastoralists; Afroasiatic, from Northern and Eastern; and the broad group of Khoesan

languages, spoken in Southern and Eastern Africa by populations with diverse ways of living (Salas et al. 2002; Gomez et al. 2014).

Genetic studies conducted on urban and rural regions of Brazil showed signatures of asymmetric mating during the colonial period (Bortolini et al. 1999; Pena et al. 2009; Schaan et al. 2017), with a higher proportion of European Y haplogroups and African and Amerindian mitochondrial haplogroups than that expected based on historical migratory and population records (Bortolini et al. 1999; Hüne-meyer et al. 2007; Schaan et al. 2017, 2020). Interestingly, approximately only one out of four Africans brought to Brazil was a female (Pinheiro et al. 2015), but the mating of enslaved men was not allowed, especially with White women (Aidoo 2018). In this context, the genetic scenario of sex-biased admixture was favored by the asymmetric nature of the sexual relations between Portuguese men with African and Amerindian women, systematically occurring by violent domination or coercion (Nascimento 2016; Dias 2019). There is extensive documentation of sexual abuse against enslaved women (Walsh 1830; Aidoo 2018), which occupied the least privileged and more vulnerable position in Brazil (Mattoso 1991). However, that pattern of admixture varied among the Brazilian regions (Bortolini et al. 1999; Ongaro et al. 2019), highlighting the particularities of a local complex demographic process.

Here, we aimed to characterize the paternal and maternal ancestry of admixed populations from three Brazilian regions (Southern, Southeastern and Central-Western) and an Amerindian sample by studying the haplogroups (Hgs) of the non-recombining portion of the Y chromosome (NRY) and mitochondrial DNA (mtDNA) and characterizing the strata at each continental level. A particular attention was given to the Central-Western region which have not been characterized at that level before. We specially focused on the African and Amerindian uniparental contributions, whose study could help to fill historical gaps in Brazilian populations' history. Compared to the autosomal continental ancestry of the same individuals using ancestry-informative markers (AIMs), we expand our knowledge on Amerindian and African Brazilian origins and show a sex-biased admixture pattern consistent with the historically documented sexual exploitation of Black and Amerindian women during the Brazilian colonization.

Subjects and methods

Human subjects

This study was approved by the national ethics committee for research with human subjects (CONEP—Comissão Nacional de Ética em Pesquisa), under the protocol

number 02727412.4.0000.0096, according to Brazilian Federal Laws. Blood samples from 437 admixed individuals from Central-Western ($n=323$), Southeastern ($n=41$), and Southern ($n=73$) regions of Brazil were previously collected and described (Farias et al. 2018; Augusto et al. 2021). In addition to the samples described in the mentioned references, 31 additional samples from the same dataset were included in this work (Central-Western ($n=26$), Southeastern ($n=4$), and Southern ($n=1$)). The individuals were subdivided into Afro-Brazilians, Euro-Brazilians, and Amerindians based on self-classification, phenotypic characteristics, and a careful assessment of individual family history.

Amerindian samples from Kaingang and Guaraní Mbyá ethnic groups ($n=10$ and $n=7$, respectively) were previously collected (Petzl-Erler et al. 1993; Tsuneto et al. 2003) in the reservations of Rio das Cobras ($25^{\circ}18' S$, $52^{\circ}32' O$) and Ivaí ($24^{\circ}30' S$, $51^{\circ}40' O$), from the state of Paraná (Southern region). Amerindian samples from Guaraní Kaiowá ethnic group ($n=7$) were collected in the villages Amambáí, and Limão Verde from Mato Grosso do Sul ($23^{\circ}06' S$; $55^{\circ}12' O$ and $23^{\circ}12' S$; $55^{\circ}06' O$, respectively) (Fig. S1) (Petzl-Erler et al. 1993; Tsuneto et al. 2003).

Genotyping and haplogroup classification

Genotyping was performed through microarray hybridization on the Illumina platform CoreExome-24 v1.1 (Illumina, San Diego, CA). A standard quality control procedure was performed as described elsewhere (Anderson et al. 2010). Filtration of Y chromosome and mtDNA datasets was performed with PLINK v1.9.0 (Chang et al. 2015), and the complete list of SNPs analyzed is shown in Tables S6 and S7. The Y chromosome and mtDNA haplogroup classification were performed through yHaploTM and HaplotypeGrep2 v2.2.9 (Poznik 2016; Weissensteiner et al. 2016) considering the haplogroup defining mutations described in the International Society of Genetic Genealogy 2019 (ISOGG; <http://www.isogg.org/tree/>) and Phylotree build 17, respectively (van Oven and Kayser 2009; van Oven 2015). For disambiguation purposes in this work, the term clade will be used to design major lineages (L0, L3, M, N, etc.), and the term haplogroup (Hg) will indicate their derivatives (L0a, M1, etc.). Uniparental ancestry proportions were estimated directly based on the biogeographic specificity of each haplogroup. Confidence intervals of ancestry proportions were calculated according to Sison and Glaz, 1995 using the R package DescTools (Sison and Glaz 1995). Chi-square and Fisher's exact tests were performed using the R package stats (R Core Team; <https://www.R-project.org/>).

Population differentiation and correspondence analysis

To understand the subcontinental genetic contributions to Brazil, we compared the maternal Hg frequencies found here to that reported in the literature for ancestral populations. For this purpose, mitochondrial sequences or published haplogroups from African and Amerindian populations from the regions known to have contributed to the formation of Brazil were compiled (Klein 2002). Only sequences mapped against the revised Cambridge Reference Sequence (rCRS) (Anderson et al. 1981; Andrews et al. 1999) were used and were classified as described in the previous section. Published Hgs classified based on Phylotree Build 17 (van Oven and Kayser 2009; van Oven 2015) were included directly, as well as Amerindian clades. A hierarchical approach was used to allow the comparison of samples genotyped by different methods. We also pre-examined the genetic distance of our sample to 40 African populations from Western, Central-Western, Southern, and Southeastern Africa through the pairwise computation of the *Fst* index and excluded from correspondence analysis (CA) those populations with significant differentiation and *Fst* index above 30% (Fig. S2, Table S4). Population's comparison and pairwise genetic differentiation (*Fst*) were computed using Arlequin software v3.5.2.2 and visualized with R (Fig. S2) (Excoffier and Lischer 2010). Haplotype diversity indices and analysis of molecular variance (AMOVA) were calculated with the R packages pegas (Paradis 2010) and poppr (Kamvar et al. 2014), respectively. Correspondence analysis was performed using the function CA of FactoMineR package and visualized with the R package factoextra (Lê et al. 2008). Analysis of the relative subcontinental genetic contribution of African populations based on Hg frequencies was performed using the following formula developed by Bernstein in 1931 as described in Chakraborty (1986):

$$m_{1i} = (p_{hi} - p_{2i}) / (p_{1i} - p_{2i}),$$

where p_h , p_{1i} , and p_{2i} are the frequencies of the i th haplogroup in the Brazilian and ancestral regions 1 and 2, respectively, and m_{1i} is the contribution of one African ancestral region. The reciprocal contribution for this haplogroup can be then estimated as

$$m_{2i} = 1 - m_{1i}.$$

The variance error and the total contribution of each region can be estimated with the equation of Cavalli-Sforza and Bodmer (1971), and the extension for multiple alleles developed by Korey (1978), respectively (Chakraborty 1986)

$$\text{Var}(\bar{m}_1) = 1 \left/ \left[\sum_{i=1}^L (1/V_i) \right] \right.$$

$$\bar{m}_1 = \frac{\sum_{i=1}^L \frac{1}{r_i} \sum_{j=1}^{r_i} m_{1(j)} V^{-1}(m_{1(j)})}{\sum_{i=1}^L \frac{1}{r_i} \sum_{j=1}^{r_i} V^{-1}(m_{1(j)})},$$

where $m_{1(j)}$ is the estimated m_j from the j th allele at the i th locus, $V^{-1}(m_{1(j)})$ is the inverse of its variance, and r_i is the number of alleles at the i th locus.

Ancestry informative markers' selection and analysis

The ancestry proportions of individuals and populations were estimated as described in a previous study of our group (Calonga-Solís et al. 2020). Estimation of ancestry in the autosomal genome was performed with the software ADMIXTURE (Alexander et al. 2009) using the unsupervised model and testing for $K=2-11$. We used genotyped samples publicly available from the HGDP-CEPH (Human Genome Diversity Cell Line Panel Database), from: Sub-Saharan African ($n=120$)—Biaka Pygmy, Mbuti Pygmy, Mandenka, Yoruba, San, Bantu; Amerindians ($n=83$)—Colombian, Surui, Karitiana, Pima; and European ($n=118$)—French, French Basque, North Italian, Orcadian, Sardinian, and Tuscan (Cann et al. 2002). We analyzed 71 ancestry-informative markers (AIMs), which differ in their allelic frequencies across continental groups ($Fst > 0.25$, $8 > 0.40$) (Calonga-Solís et al. 2020).

Principal component analysis (PCA) was performed with the snpgdsPCA function of the SNPRelate (Zheng et al. 2012) package implemented in R. Kruskal–Wallis and Pairwise Wilcoxon tests with Benjamini and Hochberg p -adjustment for multiple comparisons were applied with the R package stats (R Core Team).

Results

Paternal lineages in Brazil are predominantly of European origin

A total of 33 haplogroups classified in 8 major clades (R, E, I, Q, J, G, T, and L, in decreasing order of frequency) were detected in the whole admixed sample ($n=205$) (Table S2, Fig. 1). Hgs from clades R and E were the most common (together accounting for $> 70\%$). Over 80% of the Y chromosomes were of European origin, with Hgs from other continents detected in the Southern region (10.3% African and 3.4% Amerindian) and the Central-Western region (9.7%

Amerindian, 6.5% African, and 0.6% Asian) but not in the Southeast.

The majority (49%) of Y chromosomes in our admixed sample belong to the European R1 branch, with 66% of the R1 subHgs belonging to the R-S116 sublineage. Other R sublineages detected in our sample were R-U106, R-M529, and R-U152, as well as other European Hgs such as G-M201 and G2-L43 (together accounting for 5%). Hgs I1 and I2 accounted for 8%, and Hg J represented 7% of the samples.

Hg E is the most frequent Hg of Northern and sub-Saharan Africa, but some of its lineages reach maximum frequencies in Europe and the Middle East (Underhill and Kivisild 2007). The European E lineages observed in our sample were E-V13, E-M123, and E-M183 (altogether accounting for 15% of the total). On the other hand, we detected several low-frequency sub-Saharan African Hgs such as E-M180, its sublineages E-U174 and E-M4254, and Hg E2b1-M54 (6.3% altogether). We also detected low frequencies of Hg T and the Asian L1-M22 (1% and 0.5%, respectively). Finally, we detected the Amerindian autochthonous lineage Q1b1a1a-M3 (7.8%), which was found in the Central-Western region at an elevated frequency (9.7%).

In sharp contrast, Amerindian and African maternal lineages are predominant in Brazil

We found 21 clades and 138 mtDNA haplogroups in our admixed sample (Table S1, Fig. 1), reflecting the high level of admixture of the Brazilian population. The most frequent clades were B, A, L3, C, H, L2, and L1. Amerindian and African Hgs were the majority (42.7 and 30.5%, respectively) in the whole admixed sample, followed by 24.3% of European Hgs and 2.5% of Asian or Eurasian. These proportions exhibited some variation at the regional level, with African predominant in the Southeast (39%), and similar Amerindian and European proportions in the South (38.4% each). Clades A–D are of Amerindian origin, and together, corresponded to 42.6% of the maternal lineages in our admixed sample. Hgs L0, L1, L2, L3, and U6, of African origin, represent 30.4% of the total, while H represents the most frequent European Hg (9.8%) (Table S1).

Hg B, found in our admixed sample at frequency of 17.6%, was the most frequent Amerindian Hg in the Central-Western and Southeastern regions (19.2% and 17.6% of the total Hgs for each region). Hg A and C were found in our sample at frequencies of 11.7% and 10.5%, respectively, being the most frequent Amerindian Hgs in the Southern region (12.3% each). Hg D was found in our total sample at a low frequency (2.7%) and was completely absent from the Southeastern region.

Regarding African lineages, we found Hgs from clades L1, L2, and L3 (7.8, 9.2, and 6.6%, respectively). Clade L0 was the less frequent (2.3%); L0a1 and L0a2 (1.1% each)

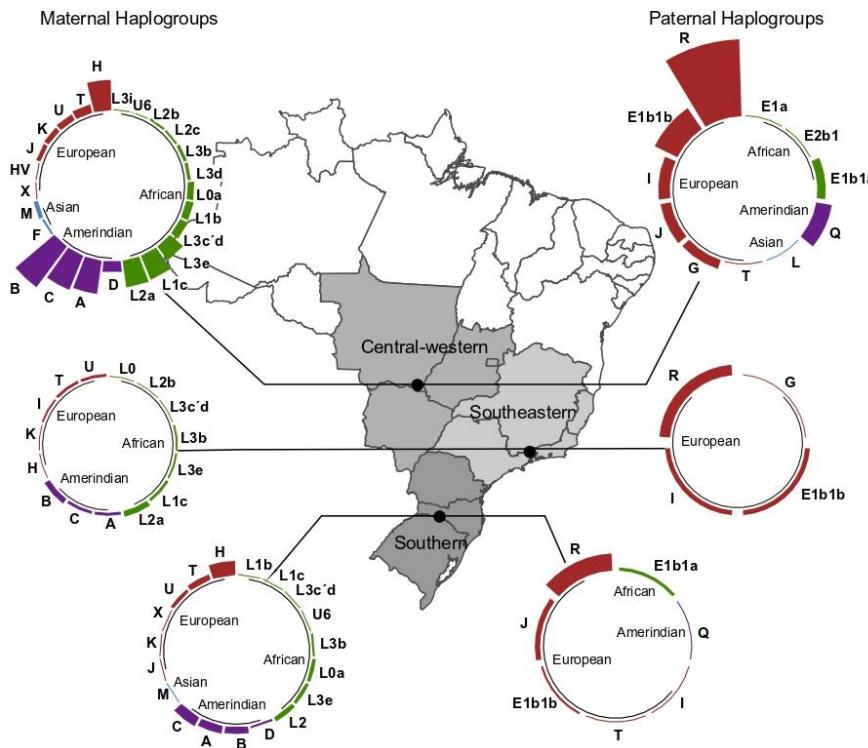


Fig. 1 Maternal ancestry in Brazilian populations is predominantly Amerindian and African, while paternal ancestry is European. The mitochondrial phylogenetic tree has been regularly updated on the PhyloTree website and counts today with more than 5000 nodes defined by over 4000 mutations (van Oven 2015). Worldwide maternal variation can be divided into continental-specific clades (Underhill and Kivisild 2007). Mitochondrial deepest branches emerge in Africa represented by clades L0–L6. The clades M and N, which derive from L3, are thought to have differentiated in Asia, representing the out-of-Africa migration, whose subclades are spread worldwide (Underhill and Kivisild 2007). Thus, clades A–G and O, P, Q, S, Y, and Z are mainly concentrated in Asia, whereas clades H, HV, I, J, JT, K, R, T, U, V, W, X occur mainly in Europe (Underhill and Kivisild 2007). Lineages derived from clades A–D, in turn, represent the trans-continental migration from Asia to America, from which all known Amerindian founder Hgs derive (Underhill and Kivisild 2007). The Y chromosome phylogenetic tree, described and regularly

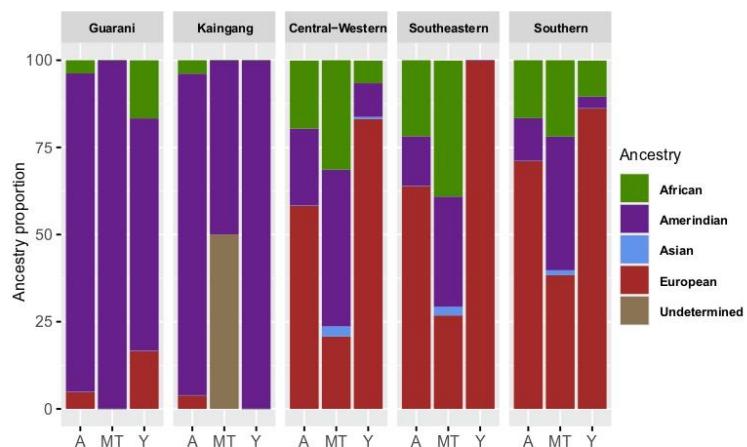
updated in the International Society of Genetic Genealogy (ISOGG), has currently 20 major clades (namely, A–T). Among these clades, A and B are the deepest branches of the Y phylogeny and exclusively African, mainly found among Khoisan and Pygmy populations, respectively (Chiaroni et al. 2009). Clade E, more recent, probably descends from the east African population that generated the out of Africa migration and its lineages are the most common in sub-Saharan Africa (Chiaroni et al. 2009). Whereas, Hgs from clades E to T are thought to have evolved out of Africa and dispersed worldwide (Underhill and Kivisild 2007). Differently, clades C and Q are common in Asia and its derived Hgs, C3 and Q-M242, are considered Native American founder lineages (Underhill and Kivisild 2007). European variation, on the other hand, is mainly restricted to clades E, G, I, J, N, L, K, and R (Underhill and Kivisild 2007). Haplotype (Hg) absolute frequencies in the urban population of the three regions studied are shown as circular bar plots. Biogeographic origin of each Hg is indicated as different colors

were the only Hgs present in our sample. We found the L1 Hgs L1b and L1c at frequencies of 1.8 and 5.9%, respectively. From the clade L2, we detected L2a-c at frequencies of 9.2, 0.7, and 0.5%, respectively. On the other hand, lineages L3b, -d, -e, and -f were found at 1.6, 3.2, 4.4, and 1.4%, respectively. Finally, we found a low frequency of the Northern African lineage U6 (0.5%).

The sex-biased ancestry proportions are evident in admixed and Amerindian populations

In the admixture analysis based on the 71 selected AIMs, $K=3$ presented the lowest cross-validation values (Salzano and Sans 2014). The European autosomal ancestry predominated in the three studied regions ($>58\%$) (Fig. 2 and Fig.

Fig. 2 Ancestry proportions of Brazilian sample show evidence of uniparental sex-biased ancestry. Uniparental ancestry was plotted as percentage and global ancestry as the average for each population. A: autosomal ancestry, MT: maternal ancestry, Y: paternal ancestry



S3). The second ancestry was Amerindian for the Central-West (22.1%) and African for the South and Southeast (16.6 and 21.8%, respectively) (Table S9). An asymmetric landscape emerged, however, when the uniparental ancestry patterns were taken into consideration.

The paternal ancestry proportions were predominantly European and homogeneous in the three regions (Fisher's exact test, $p=0.4818$) (Fig. 2 and Table S9). For the maternal lineages, on the other side, the ancestry proportions showed significant differences only among Central-Western and Southern regions (Pearson Chi-squared $p=0.007029$). The highest proportions were African in the Southeast and Amerindian in Central-West, followed by Amerindian in the Southeastern and African in the Central-Western region. The European ancestry was 26.8% and 20.0%, while Asian ancestry was 2.4% and 2.8% for Southeastern and Central-Western regions, respectively. The Southern region presented similar proportions of Amerindian and European maternal ancestry, followed by African and Asian. For both uniparental genomes, the AMOVA did not show significant molecular variance among regions (Table S13).

For Amerindian communities, the degree of non-native admixture varied with sex, as previously observed (Bortolini et al. 2003; Marrero et al. 2007). In Guarani, we observed 100% Amerindian ancestry according to the mitochondrial genome but 66.6% according to the NRY (Figs. 2 and 3c). Among the Kaingang, five individuals were classified to mitochondrial haplogroup N. An Asian origin has been suggested for Hg N and its basal branches are found today in that continent. However, Hg N has also been detected at very low frequencies (<1%) in Northern Africa and Portugal (Fernandes et al. 2012; Marques et al. 2015; Bekada et al. 2015). The haplogroups A, B, and X (Amerindian), S, O, Y, F and P (Asian), and W, I, H, V, HV, K, U, J, and T (European)

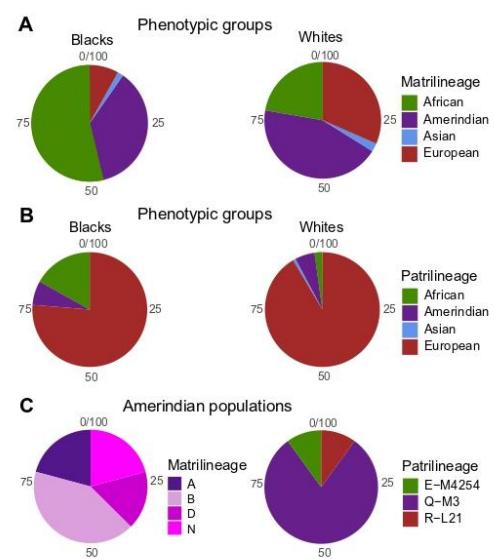


Fig. 3 Uniparental ancestry proportions according to phenotypic classification based on skin color. Proportions are represented in percentage. **a** Maternal ancestry in Black and White urban individuals. **b** Paternal ancestry in Black and White urban individuals. **c** Maternal and paternal haplogroups in Amerindian partialities. Tones of violet in matrilineages of Amerindian populations indicate their Amerindian Hg. Continental ancestry of patrilineages in Amerindian populations is indicated by colors as in **a** and **b**

derive from the clade N, but the alleles that define these N-derived Hgs were not detected in the five Kaingang individuals (van Oven and Kayser 2009). We have not genotyped

a defining SNP for the Asian Hg S, but we have genotyped SNPs defining its derived branches. Nevertheless, the recent introduction of Asian lineages in Amerindian populations is not expected as there are no reports of East Asian settling in Amerindian communities (IBGE 2007). African and European lineages, on the other hand, have been reported before in another Kaingang population (Marrero et al. 2007). Alternatively, as the non-Amerindian autosomal ancestry in those individuals was very low (0–5%) we could also suggest a reversion of the A Hg defining SNP, 1736, the only SNP we used to assign A haplogroup membership. Thus, to correctly determine the ancestry of those individuals, the genotyping of additional variants from the hypervariable region I is required. All the Kaingang Y chromosomes were Amerindian. We detected one European Y chromosome belonging to Hg R-L21 in an individual of the ethnic group Guarani Kaiowá as well as an African chromosome belonging to Hg E-M4254 in a Guarani individual from Rio das Cobras (Fig. 3c). The average Amerindian autosomal ancestry for this sample was 80%, with 18% European, followed by African ancestry (1.3%). Nei's haplotype diversity index (H) and number of haplotypes (N_{hap}) were high for Amerindian maternal haplotypes both in the Amerindian and in the admixed sample ($H=0.998$ and $H=0.956$, respectively). We found 16 different mtDNA haplotypes among 24 Amerindian individuals, whereas 167 different mtDNA haplotypes were detected among the 186 individuals of the admixed sample ($N_{hap}/N=0.897$ in the admixed sample and 0.666 in the Amerindian sample). Both H and N_{hap}/N were slightly higher in the admixed sample (Table S12).

Only in individuals from Amerindian populations and in the individuals with the highest Amerindian autosomal ancestry from the admixed sample, the presence of maternal and paternal Amerindian Hgs reflected their autosomal ancestry. For the other Brazilian populations, there was no clustering of uniparental ancestries with autosomal ancestry (Fig. 4).

Skin “color” categories in Brazil do not reflect autosomal or uniparental ancestries

The Brazilian Institute of Geography and Statistics (IBGE) distinguishes phenotypic categories based predominantly on skin color, classifying individuals as Amerindian, White, Black, and Admixed. In our analysis, we merged both Admixed and Blacks in the category Blacks, as was implemented in the official IBGE classification system in response to the demands of the Brazilian Black movement (Daflon et al. 2017).

When we analyzed the autosomal ancestry components, both Black and White samples presented on average predominantly European ancestry (48% and 68.1%, respectively), being it higher in Whites ($p < 2.2e-16$). On the other hand,

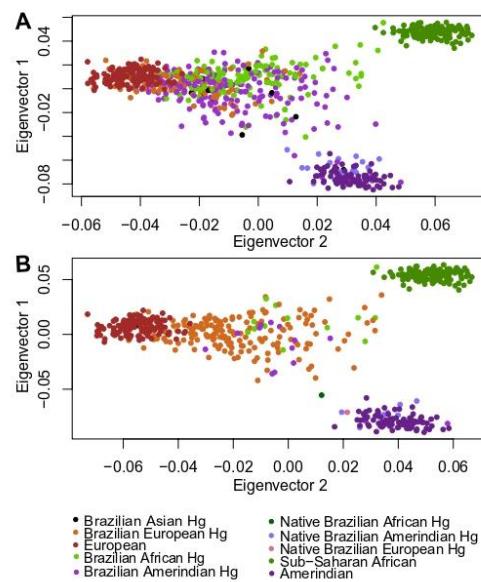


Fig. 4 Uniparental ancestry does not reflect autosomal ancestry in admixed Brazilians. In the picture, plots of Principal Component Analysis (PCA) based on ancestry-informative markers (AIMs) genotypes from all samples studied are shown. Sample color indicates uniparental **a** maternal and **b** paternal ancestry. Hg: haplogroup

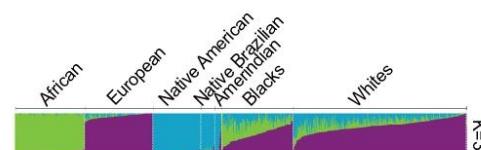


Fig. 5 Phenotypic categories do not reflect autosomal ancestry of Brazilians. Admixture analysis based on AIMs genotypes in ancestral populations (African, European, and Native American), Native Brazilian partialities, urban self-declared Amerindian individuals, and heteroclassified Black and White individuals

African ancestry was higher in Blacks (33.6% in Blacks and 14.1% in Whites, $p < 2.2e-16$). Similar low proportions of Amerindian ancestry were observed in both groups (18.4% and 17.8%, respectively, and $p = 0.3102$) (Fig. 5 and Tables S10 and S11). The AMOVA showed that a low molecular variance exists among Blacks and Whites, while the maximum variance occurs within the groups (among groups variance = 2.50, $p = 0.001$, Table S13). Those who were self-declared Amerindians presented a larger proportion of Native component (80.7%) and differentiated from the other

two groups on the three components ($p < 0.01$, Tables S10, S11 and Fig. 5). To further explore the concordance of phenotypic categories with autosomal ancestry, we plotted the individuals count on each category against the computed ancestry components. Although self-declared Amerindian individuals showed a clear pattern of differentiation on the three ancestry components compared with Blacks and Whites, we could not observe the same clear division for the others. We observed a considerable overlapping at the individual level among Blacks and Whites when considering the African and European components (Figs. 5 and 6a and b).

We also observed that phenotypic categories did not reflect either uniparental ancestries, as the total sample presented mainly Amerindian maternal contribution and European Y Hgs (Fig. 3 and Table S10). The Black sample showed mainly African maternal contribution, whereas the White sample presented mainly Amerindian contribution (Fig. 3 and Table S10). Only 2.7% of our total sample self-declared as members of Amerindian ethnic groups. For them, both maternal and paternal ancestries were mainly Amerindian (>80 and 100%, respectively), with just a small proportion of European maternal ancestry (<10% each).

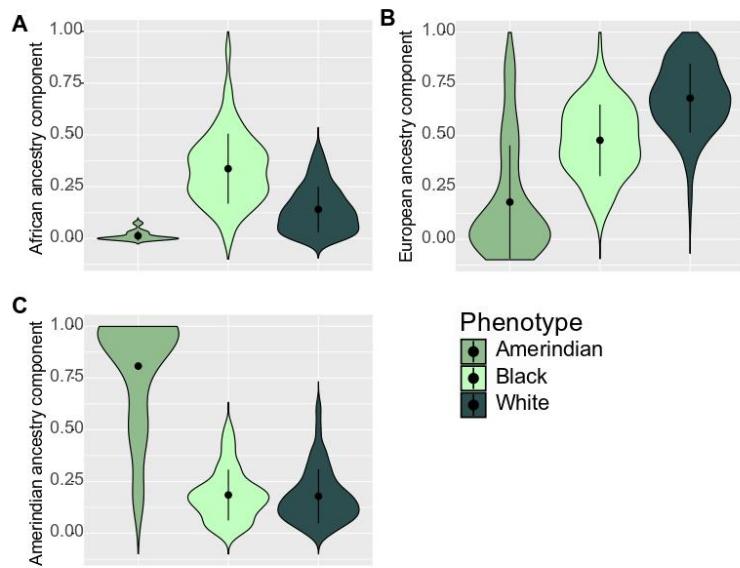
Analysis of the subcontinental African and Amerindian maternal contributions shows regional variation

For the African paternal fraction, admixture analysis indicated a predominantly Central-African Bantu contribution (86.8% against 13.1% Western African contribution, $\pm 9.6\%$).

In the correspondence analysis (CA) of mitochondrial African Hgs (Fig. 7a), as also observed in the *Fst* analysis (Fig. S2), the populations more similar to the Brazilians were Western Africans, especially from Benin and Nigeria, which is particularly evident in the close cluster with Southern Brazil. In a north–south direction, the Brazilian samples clustered with Northern populations (closer to Central-Western Gabon and Cameroon than to Angola, Mozambique, and Zambia) due to the absence of Hg L0d. Consistently, the admixture analysis indicated a higher Western African contribution for Southern (85.7% Western and 14.3% Central-Western $\pm 5\%$), and Southeastern Brazil (71.8% Western and 28.2% Central-Western $\pm 10.4\%$). The Central-Western region of Brazil showed higher Central-Western Bantu and also expressive non-Bantu contribution (64.8% Central-Western and 35.2% Western $\pm 1.3\%$) (Fig. S4).

The CA of mitochondrial Amerindian Hgs showed a clear clustering of Amerindian populations according to language and region (Fig. 7b). The Andean populations (Aymara and Quechua-speaking) clustered with Central-Brazilian, Paraguayan, and Western Argentinian Amerindians (Xavante, Aché and Chiriguano groups), due to the high frequency of

Fig. 6 Autosomal ancestry components in urban individuals according to phenotypic categories. **a** African, **b** European, and **c** Amerindian ancestry



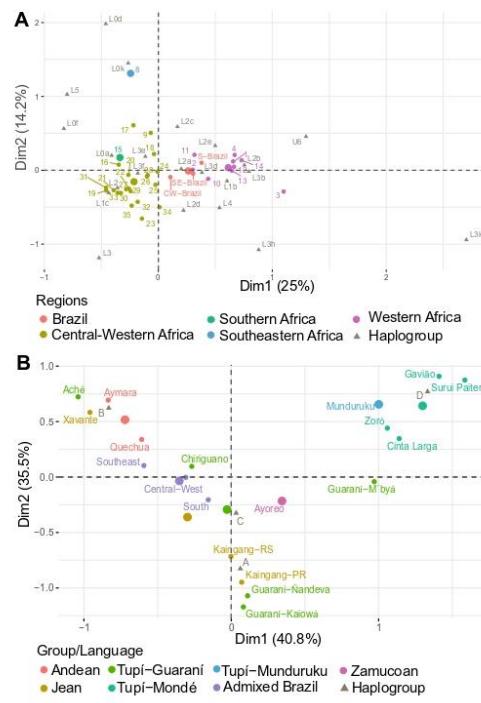


Fig. 7 The subcontinental African and Amerindian ancestries vary among Brazilian regions. In the figure, Correspondence Analyses of the African (a) and Amerindian (b) maternal haplogroup pools in urban Brazilian populations and reference African and Amerindian populations, respectively, are shown. In A, 35 populations from five African regions are numbered as follows: 1: Yoruba in Ibadan, Nigeria; 2: Esan in Nigeria; 3: Mandenka in Western Division of Gambia; 4: Mende in Sierra Leone; 8: Mozambique; 9: Angola; 10: Bariba in Benin; 11: Dendi in Benin; 12: Fon in Benin; 13: Guinea-Bissau; 14: Burkina Faso; 15: Zambia; 16: Gangela in Angola; 17: Nyenka in Angola; 18: Ovimbundu in Angola; 19: Akele in Gabon; 20: Ateke in Gabon; 21: Benga in Gabon; 22: Duma in Gabon; 23: Eshira in Gabon; 24: Eviya in Gabon; 25: Fang in Gabon; 26: Galoa in Gabon; 27: Mitsogo in Gabon; 28: Ndumu in Gabon; 29: Nzebi in Gabon; 30: Obama in Gabon; 31: Orungu in Gabon; 32: Punu in Gabon; 33: Shake in Gabon; 34: Fang in Cameroon; 35: Ngumba in Cameroon. CW: Central-Western; S: Southern; SE: Southeastern; RS: Rio Grande do Sul; PR: Paraná. Reference populations were taken from the literature (Ward et al. 1996; Bonatto and Salzano 1997; Schmitt et al. 2004; Dornelles et al. 2004, 2005; Kivisild et al. 2006; Marrero et al. 2007; Quintana-Murci et al. 2008; Behar et al. 2008; Coelho et al. 2009; Sala et al. 2010; Gayà-Vidal et al. 2011; Barbieri et al. 2014; Auton et al. 2015; Larmuseau et al. 2015; Pramitivo et al. 2017)

Hg B. Southeastern and Central-Western Brazilian clustered closer to that group. On the other hand, Central and Southern-Brazilian Amerindians of the Tupí-Guaraní branch (such

as Mbyá, Kaiowá, and Nandeva), Macro-Jé (Kaingang), and the Paraguayan Zamucoan speakers (Ayoreo), are clustered by their expressive frequency of Hgs C and A. The South of Brazil is positioned between the Andean-Central-Brazil cluster and the Southern-Brazilian Amerindian populations.

We also calculated the Amerindian admixture proportions for the three regions. The Central-Western Brazil presented 56.9% contribution from the Andean-Central-Brazil cluster with the reciprocal percentage from Southern-Brazilian Amerindians ($43.1\% \pm 0.2\%$); the Southeastern region presented 83.1% ($\pm 2.1\%$) of contribution from that same cluster (and $16.9\% \pm 2.1\%$ Southern Amerindian contribution), and the Southern region presented mainly Southern Amerindian contribution (59.6% and 40.4% Andean-Central-Brazil contribution $\pm 1\%$) (Fig. S4).

Discussion

We reveal new insights into the subcontinental African and Amerindian ancestry by leveraging the data for mitochondrial and Y chromosome haplogroups in Brazilian regions never previously studied for both uniparental markers. We show that a sex-biased admixture pattern is evident in those regions, consistent with the historical data.

The genetic footprint of the Africans and Amerindians in the current Brazilian population

Historical records suggest that a higher proportion of Bantu (mainly from Central-Western and, secondarily, Southeastern Africa) than of Niger-Congo non-Bantu people (Western Africa) were brought to Brazil over the colonial period (Klein 2002). These proportions, however, fluctuated over time and space, and diverse patterns of African ancestry have been detected for different Brazilian regions in genetic and anthropological studies (Bortolini et al. 1999; Klein 2002; Silva et al. 2006; Hünemeier et al. 2007; Gonçalves et al. 2008; Gouveia et al. 2020).

In our urban sample, the African paternal background was represented by Hgs E-M180, E-U174, E-M4254, and E2b1-M54, which are the main Hgs in Bantu populations (Cruciani et al. 2002; Gomes et al. 2010; de Filippo et al. 2011; Oliveira et al. 2018). Hg E-U174, however, is also frequent in Niger-Congo Yoruba-speaking populations from Western Africa (de Filippo et al. 2011; Montano et al. 2011; Auton et al. 2015; Oliveira et al. 2018). The maternal African background, on the other hand, was mostly represented by Hgs from clade L2, with L2a as the most frequent. L2a is broadly distributed over Sub-Saharan Africa (Salas et al. 2004), being modal in Benin (Pramitivo et al. 2017). L2b and -c are common in Western Africa and are associated with the Western fringe of the Bantu expansion, being found

at lower frequencies in West-Central Africa (Beleza et al. 2005). An exception to this distribution is Hg L2a5 (found here at a low frequency in Central-Western Brazil, 0.3%), restricted to Southeastern Bantu populations and Eastern Africa (Silva et al. 2015). The second most common clade, L1, also has a broad sub-Saharan African distribution. The subclades found here, L1b and L1c, frequently occur in West-Central and Southeastern Africa, with a higher frequency of L1c among Bantu speakers from West-Central and of L1b in coastal Western and West-Central Africa (Vigilant et al. 1991; Salas et al. 2002, 2004; Plaza et al. 2004; Rito et al. 2013). From clade L3, we observed L3b, -d, -e, and -f lineages, which distribute mainly in West-Central Africa (Rosa and Brehm 2011; Soares et al. 2012). L3e is the most widely distributed L3 lineage, with frequency peaks in Southeastern and West-Central Africa (Rosa and Brehm 2011). Haplotype L3f is from East Africa, but the lineage found here, L3f1, mainly spreads into Western Africa (Salas et al. 2002, 2004). We also found a very low frequency of the autochthonous Northern African Berber Hg U6 (0.5%), which occurs at a low frequency in Western Africa (Salas et al. 2002). The less represented clade in our sample was L0, whose derived Hgs L0a1 (typical of Eastern and Southeastern Africa) and L0a2 (an east–south Bantu expansion marker) were detected here (Rosa and Brehm 2011; Soodyall 2013).

As we observed a higher maternal than paternal African contribution in the admixed samples (only 6.3% paternal but 30.5% African maternal lineages), we only estimated the overall sample African admixture proportions in the Y chromosome, which indicated a predominantly Central-Western African contribution (86.9%). These results, however, should be considered with caution due to our limited male sample size.

Regarding the mtDNA, we could detect a regional variation. We observed a high proximity of our sample to the populations of Western Africa, especially the Bight of Benin, and, secondarily, to Central-Western Africa. This pattern was especially pronounced for Southern and Southeastern Brazil, whereas the Central-Western region was predominantly Central-Western Bantu with also an expressive Western African contribution.

From the Bight of Benin, Africans highly specialized in mining technologies were kidnapped by the Portuguese and shipped to Brazil, especially to the North and Northeastern (Fausto 2015; Sá Junior 2016; Primitivo et al. 2017). It is known that during the mining cycle (1701–1828), a high contingent of Africans was relocated from those regions to Central-Western and Southeastern Brazil (Fausto 2015; Sá Junior 2016). Similarly, the state of Paraná (> 90% of our Southern sample) was urbanized after the Portuguese gold discoveries, in 1648 (Raggio et al. 2018). That could explain the high Western African

contribution to our regional samples, which also agrees with higher Western African ancestry in Northeastern Brazil, detected by autosomal markers (Silva et al. 2010; Kehdy et al. 2015).

We could not estimate subcontinental Amerindian paternal ancestry due to the finding of only one Hg in our sample (Q1b1a1a-M3). Q1b1a1a-M3 derives from the major founder lineage of Amerindian populations, Q-L54, and is the most common in South Amerindians (Zegura 2003; Karafet et al. 2008; Bailliet et al. 2009; Battaglia et al. 2013). The Hg Q-M346, a less frequent ancestral Amerindian lineage, was not identified in our sample (Bailliet et al. 2009; Resque et al. 2016). Maternal Hgs, on the other hand, reflected regional particularities. Hg B, the most frequent in Central-Western and Southeastern Brazil, is highly frequent (84–100%) in the Amerindian ethnic groups Xavante and Aché from Central Brazil and Paraguay, respectively, but is also a typical Andean haplogroup (Corella et al. 2007; Marrero et al. 2007; Gayá-Vidal et al. 2011). In agreement, Central-Western and Southeastern Brazilian samples were genetically closer to Andean populations (Quechua and Aymara-speakers) from Bolivia, the Argentinian Chiriguano (which are related to the Terena from Central Brazil and the Amazonian Guarani), the Central-Brazilians Xavante, and the Paraguayan Aché (Sala et al. 2018). Thus, the high frequency of Hg B in the admixed samples could be related to the historical records that describe the Portuguese practice of enslaving Amerindians, even brought from neighbor colonies, for mine work (Chamorro and Combès 2015). Furthermore, the high Amerindian paternal ancestry detected in Central-Western Brazil could be related to the geographical proximity to Bolivia, where it reaches proportions as high as 28% (Cárdenas et al. 2015). Alternatively, the high frequency of Hg B could be reflecting pre-Columbian gene flow among Andean Amerindians and the Amerindian populations that were involved in the formation of the admixed Central-Western and Southeastern Brazilian regions. The existence of gene flow among Andean and Amazonian Amerindian populations has already been suggested (Corella et al. 2007). Hg A and C, the most frequent in Southern Brazil, are highly frequent in populations from the Tupí-Guarani branch (such as Nandeva, Kaiowá, and Mbya) which, according to the Brazilian Socioenvironmental Institute (ISA; <https://pib.socioambiental.org>), are spread over Central to Southern Brazil, Argentina, and Paraguay. Consistently, those populations, the Paraguayan Ayoreo, and the Kaingang were the closest to Southern Brazil. Hg D has a variable frequency (0–87%) in Native Brazilian populations (Marrero et al. 2007; Tavares et al. 2019), and was found at a low frequency in our sample. These results, however, are limited by the scarcity of genetic studies in Southern Amerindian populations.

Uniparental markers' patterns are consistent with sex-biased admixture that occurred during the colonization of Brazil

The settling of colonial society was mediated by the systematic abuse of Amerindian and Black women, a process that has been widely documented in anthropological studies (Dias 2019). Our results of the maternal, paternal, and autosomal ancestries detected genetic signatures of sex-biased admixture in the current admixed populations. We observed that the great majority of mitochondrial Hgs in all regions were Amerindian and African, whereas the Y Hgs were European, mainly Portuguese, as previously reported for other Brazilian populations (Guerreiro-Junior et al. 2009; Chiaroni et al. 2009; Carvalho et al. 2010; Myres et al. 2011; Resque et al. 2016; Schaan et al. 2020). Y Hgs typical of other European regions were found at lower frequencies, probably reflecting more recent European migration (IBGE 2007). A similar pattern was observed in Amerindian partialities as European and African admixture was only observed at the paternal line among the Guarani and, to a less extent, in autosomal ancestry. Their average high autosomal Amerindian ancestry (80%), with more European than African admixture (18 against 1.3%), agrees that none of the Amerindian participants declared a recent admixture. Thus, the admixture in Amerindian partialities can potentially be related to the first centuries of the colonial time when sex exploitation of Amerindian women by European men, and gene flow between *quilombos* (traditional rural African Brazilian communities) and Native tribes were frequent, as described in historical and anthropological records (Ribeiro 1995; Kimura et al. 2013; Nascimento 2016; Aidoo 2018).

The paternal ancestry in the Southern region agreed with that reported by others (Ferreira et al. 2006; Resque et al. 2016). The finding of only European Y Hgs in the South-eastern region was expected, as our small male sample size ($n=23$) probably limited the detection of lower frequency Amerindian and African lineages detected by others (Ferreira et al. 2006; Resque et al. 2016; Braganholi et al. 2017). The maternal ancestry in those regions was also similar to that reported in the literature (Alves-Silva et al. 2000; Cardena et al. 2013; Ambrosio et al. 2015; Braganholi et al. 2017; Poletto et al. 2019).

The low proportion of Asian lineages (<3%) found in all regions is consistent with the migration of East Asian people in the XX century, as reported by others (Alves-Silva et al. 2000; Braganholi et al. 2017).

Skin "Color" categories are linked to racism in Brazil, not genomic ancestry

The structural racism in Brazil originated in the colonial time and accompanied the transition to capitalism (Santos

et al. 2009; Nascimento 2016; Raggio et al. 2018). Until today, racial discrimination in Brazil relies on skin color and other phenotypic characteristics rather than on genomic ancestry and sustains social inequalities (Wade et al. 2014). These phenotypic traits are roughly differentiated in IBGE phenotypic categories which, as reported by others, do not correlate directly with genomic ancestry (Santos et al. 2009; Pena et al. 2009).

In our study, the phenotype of the individuals sampled did not correlate directly with uniparental or autosomal ancestry, except for self-declared Amerindians, who had a high autosomal Amerindian ancestry (over 80%). Despite that African and Amerindian paternal ancestries were higher in Blacks than in Whites, these proportions were low in both groups, with an overwhelming majority of European Y lineages. A similar pattern was observed for autosomal ancestry, confirmed by the strong overlapping of Blacks and Whites with similar African and European ancestry components and a low and similar Amerindian ancestry. In agreement, the AMOVA confirmed that the highest molecular variance occurs within phenotypic groups, with only a small variance detected among them. Mitochondrial ancestry contributions were more uniform, mainly Amerindian in Whites, and African in Blacks, as observed by others (Cardena et al. 2013; Poletto et al. 2019). This lack of correlation among "color" categories and genomic ancestry was expected, as an extensive variation exists in skin pigmentation in Amerindian and African populations (Schmitt et al. 2004; Crawford et al. 2017; Adhikari et al. 2019). Besides, to date, a limited number of genomic regions have been linked to this trait (Crawford et al. 2017; Adhikari et al. 2019), which reflects only a small fraction of an individual's genome.

As ancestry classification is a common approach to control for sample stratification in epidemiological studies, we further investigated the usefulness of Hgs as proxies to autosomal ancestry of admixed Brazilians through the performance of a PCA of AIM genotypes. We observed individuals clustered by autosomal ancestry but not by shared uniparental ancestry, except self-declared Amerindian individuals or from Amerindian communities. Overall, these results demonstrate that both phenotypic categories and uniparental markers are not good estimators for assessing the autosomal ancestry of admixed Brazilians, as it was already established for other worldwide populations and in South-eastern Brazil (Alves-Silva et al. 2000; Cardena et al. 2013; Emery et al. 2015). The highly admixed nature of the Brazilian population, and the very low correlation of phenotype and genomic ancestry, only highlight the fact that racism in this country is solely based on phenotype and not the genetic origin, proving human race as a social construct with no biological foundation (Wade et al. 2014; Lima-Costa et al. 2015). Since human race is not biological, the genetic admixture of our population cannot be used to invalidate

racism and the need to improve public policies for Black people in Brazil, including health policies (Ribeiro 1995; Lima-Costa et al. 2015).

There is a regional variation in the autosomal ancestry in Brazil

Several authors have shown a gradient of autosomal ancestry in Brazil, with higher Amerindian and African components in the North and Northeastern regions, respectively, and higher European in Southern and Southeastern, with an intermediate proportion of the three components in the Central-Western region (Souza et al. 2019). Our results agree with that pattern and showed that the African component was similar in Central-Western and Southeastern regions and significantly lower in the South (Fig. S3). Among phenotypic groups, however, this pattern was only evident in Whites. That could be reflecting the Brazilian policies of XIX–XX centuries, which favor the migration of European families predominantly to Southern Brazil, possibly reducing the admixture bias (IBGE 2007).

The Central-Western Brazil has a high percentage of Amerindian ancestry in admixed populations

For the first time, we report the uniparental and autosomal admixture proportions of Brazil's Central-Western region, representing the largest proportion (73.9%, $n = 323$) of our urban sample. We included samples from its four states and have found the highest Amerindian paternal ancestry yet described for a Brazilian region, only comparable with that found in Northern Brazil (>8%) (Carvalho-Silva et al. 2001; Abe-Sandes et al. 2004; Marrero et al. 2005; Barcelos et al. 2006; Ferreira et al. 2006; Hüneimer et al. 2007; Gonçalves et al. 2008; Carvalho et al. 2010; Resque et al. 2016; Figueiredo et al. 2016; Braganholi et al. 2017). Southeastern and Southern regions presented even lower Amerindian paternal ancestry (<5%), as already reported (Carvalho-Silva et al. 2001; Marrero et al. 2005; Ferreira et al. 2006; Hüneimer et al. 2007; Gonçalves et al. 2008; Resque et al. 2016; Figueiredo et al. 2016; Braganholi et al. 2017).

In agreement with the high Amerindian paternal ancestry observed in the Central-West, we detected a higher Amerindian autosomal ancestry than other studies (Manta et al. 2013) (22.1%). Consistently, the Amerindian maternal ancestry in the Central-West was the highest among the studied regions (45%), only comparable to that found by others in the Northeast (43.5%) (Schaan et al. 2017). This was expected as the State of Mato Grosso do Sul is home to the second-highest indigenous population in the country (Chamorro and Combès 2015), and whose urbanization is associated with the mine discoveries in the XVIII century (Ribeiro 1995; Fausto 2015; Sá Junior 2016). That event attracted

migrants from other regions of the country, as well as Portuguese, and caused the relocation of enslaved African and Amerindians from North and Northeastern Brazil but also from neighboring countries such as Bolivia to compulsory work in the mines and associated activities (Fausto 2015; Chamorro and Combès 2015; Sá Junior 2016). Thus, autochthonous and relocated Amerindians were deeply involved in urban life and supported the development of the region (Ribeiro 1995; Fausto 2015; Chamorro and Combès 2015).

Brazilian urban populations carry relatively low Amerindian ancestry but are relevant reservoirs of Amerindian maternal haplotypes

According to the historical records, the invasion of America was a violent process of imposition of the European belief systems and economic interests on Native American and African populations (Ribeiro 1995; Nascimento 2016). It caused the extinction of entire Amerindian ethnic groups and the development of diverse forms of indigenous and African resistance (Ribeiro 1995; Oliveira and Freire 2006). In our study, this reduction of Amerindian diversity could only be related to the low autosomal and paternal Amerindian ancestry detected in admixed samples, as well as on the finding of only one Amerindian paternal Hg. In agreement, a slightly higher number of mitochondrial haplotypes and haplotype diversity indices were found in urban than in Native populations (those corresponding to 14 Hgs against 3). In comparison to other populations from Central and Southern Brazil, and from the Amazon region, however, mitochondrial haplotype diversity was high in the Amerindian sample studied (Ward et al. 1996; Bonatto and Salzano 1997; Schmitt et al. 2004; Dornelles et al. 2005; Marrero et al. 2007; Arias et al. 2018; Tavares et al. 2019). Despite the small sample size, this pattern was similar to that observed for urban and Amerindian populations from the South of Brazil, reinforcing the role of admixed populations as reservoirs of Amerindian haplotypes (Tavares et al. 2019).

The inclusion of more Black participants in research could reveal novel information about subcontinental African ancestry

It is important to highlight that our sample included mainly Whites (69.4% Whites and 27.9% Blacks), proportions that are close to that of the Southern region but could bias for European ancestry in the Central-Western and Southeastern regions. In those regions, the proportion of Blacks in the population is higher (Central-Western Blacks = 56%, Amerindian = 0.9% and Whites = 41%; Southeastern Blacks = 43.8%, Amerindian = 0.1% and Whites = 54.9%) (IBGE). Thus, the ancestry proportions registered for the Central-Western and Southeastern regions (at both the

uniparental and autosomal levels) could be over-representing the European components and, in this way, hiding African ancestry and Hg diversity. This is especially relevant in the study of the paternal ancestry, where the African contribution is low.

Limitations

Limitations of our study include the lack of whole mitochondrial and Y chromosome sequences which could increase the knowledge of the subcontinental African and Amerindian contributions to Brazil and increase our knowledge of Amerindian and African genetic diversity. Besides, the small Amerindian sample size restricted our possibilities of inter-population comparisons. Additionally, the European bias of our admixed sample could have limited the African and Amerindian Hg diversity found and over-estimated the overall Brazilian European ancestry. It is important to highlight the need for additional studies on African-Brazilians to help identify the African subcontinental origins of the Brazilian population.

Conclusions

We observed an overwhelming predominance of maternal African and Amerindian haplogroups and European paternal haplogroups in the study populations, consistent with the well-documented historically violent colonization of Brazil. It is critical to highlight that we do not claim that all sex bias in the study populations results from violence against women or that our study has the power to demonstrate that violence occurred. However, it is plausible that our results reflect, at least partially, the consequences of the colonization-associated historical records of sexual violence against enslaved women (Walsh 1830; Nascimento 2016; Aidoo 2018). Interestingly, we also observed the same pattern in two populations from Brazil never previously studied for both mitochondrial and Y chromosome markers (from the Central-Western region and the State of Paraná). For the first time, we have described the variation in the maternal subcontinental Amerindian ancestry for those regions, showing a higher contribution of Central-Brazilian and Andean-related populations in the Central-Western and Southeastern regions of Brazil and of Southern-Brazilian Amerindians in the Southern region. Finally, we have detected a higher non-Bantu Western African maternal contribution than previous studies (Silva et al. 2006; Hüneimer et al. 2007; Gonçalves et al. 2008) in the Brazilian Southeast and provided the unprecedented characterization of the higher Western African and Bantu ancestries in Paraná and Central-Western Brazil, respectively. Additional studies on African-Brazilians are required to better understand and identify the African

subcontinental origins of the Brazilian population. Our results also reinforce the importance of autosomal markers investigation for accessing ancestry of admixed Brazilians in biomedical studies as it cannot be inferred through the analysis of uniparental markers or individual phenotype.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00438-022-01857-7>.

Author contributions MHB contributed to the study's conception and supervision. Genotyping and data generation were performed by DGA. Data analysis was performed by IAJL, VCS, and RCDA. MHB, MLP-E, and DGA contributed with reagents and/or DNA samples. The first draft of the manuscript was written by IAJL and MHB. All authors commented on previous versions of the manuscript. All authors revised and approved the final manuscript.

Funding We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PROAP—Finance Code 001) for financial support and for the scholarships provided to Iriel Araceli Joerin Luque and Verônica Calonge Solis, and for financial support through the Young Talent Attraction, Science without Borders Program (CAPES—88881.067970/2014-01) provided to Danillo G. Augusto. We also thank the Programa de Apoio a Núcleos de Excelência—Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Paraná (PRONEX-FA—Convênio 116/2018—Protocol 50530, Grant recipient: Maria Luiza Petzl-Erler).

Availability of data and materials The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical approval This study was approved by the Brazilian National Ethics Committee (CONEP, approval number 2970200) under Brazilian Federal laws and the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication Not applicable.

References

- Abe-Sandes K, Silva WA, Zago MA (2004) Heterogeneity of the Y chromosome in Afro-Brazilian populations. *Hum Biol* 76:77–86
- Adhikari K, Mendoza-Revilla J, Sohail A et al (2019) A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nat Commun* 10:358
- Aidoo L (2018) Slavery unseen: sex, power, and violence in Brazilian history. Duke University Press, Durham
- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19:1655–1664

- Alves-Silva J, da Silva SM, Guimarães PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Prado VF (2000) The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet* 67:444–461
- Ambrosio IB, Braganholi DF, Polverari FS, Cicarelli RMB (2015) Mitochondrial DNA 30-SNP data confirm high prevalence of African lineages in the population of Espírito Santo, Brazil. *Forensic Sci Int Genet Suppl Ser* 5:e346–e347
- Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT (2010) Data quality control in genetic case-control association studies. *Nat Protoc* 5:1564–1573
- Anderson S, Bankier AT, Barrell BG, de Brujin MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465
- Andrews RM, Kubacka I, Chinnery PF, Lightowler RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147–147
- Arias L, Barbieri C, Barreto G, Stoneking M, Pakendorf B (2018) High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions, and population mobility in Northwestern Amazonia. *Am J Phys Anthropol* 165:238–255
- Augusto DG, de Almeida RC, Farias TDJ, Magalhães WCS, Malheiros D, Lima-Costa MF, Barreto ML, Horta BL, Kumar V, Wittig M, Franke A, Busch H, Schmidt E, Roselino AM, Tarazona-Santos E, Boldt ABW, Petzl-Erler ML (2021) Unsuspected associations with variants within the genes NOTCH4 and STEAP2-ASI uncovered by a genome-wide association study in endemic pemphigus foliaceus. *J Invest Dermatol.* <https://doi.org/10.1016/j.jid.2021.04.017>
- Auton A, Abecasis GR, Altshuler DM et al (2015) A global reference for human genetic variation. *Nature* 526:68–74
- Bailliet G, Ramallo V, Muzzio M, García A, Santos MR, Alfaro EL, Dipierri JE, Salceda S, Carnese FR, Bravi CM, Bianchi NO, Demarchi DA (2009) Brief communication: restricted geographic distribution for Y-Q^{*} paragroup in South America. *Am J Phys Anthropol* 140:578–582
- Barbieri C, Vicente M, Oliveira S, Bostoen K, Rocha J, Stoneking M, Pakendorf B (2014) Migration and interaction in a contact zone: mtDNA variation among Bantu-speakers in Southern Africa. *PLoS ONE* 9:e99117
- Barcelos R, Ribeiro G, Silva W, Abe-Sandes K, Godinho N, Marinho-Neto F, Gigonzac M, Klautau-Guimarães M, Oliveira S (2006) Male contribution in the constitution of the Brazilian Centro-Oeste population estimated by Y-chromosome binary markers. *Int Congr Ser* 1288:228–230
- Battaglia V, Grugni V, Pergo UA, Angerhofer N, Gomez-Palmieri JE, Woodward SR, Achilli A, Myres N, Torroni A, Semino O (2013) The first peopling of South America: new evidence from Y-chromosome haplogroup Q. *PLoS ONE* 8:e71390
- Behar DM, Villemans R, Soodyall H, Blue-Smith J, Pereira L, Metspalv E, Sczzari R, Makkhan H, Tzur S, Comas D, Bertranpetti J, Quintana-Murci L, Tyler-Smith C, Wells RS, Rosset S (2008) The dawn of human matrilineal diversity. *Am J Hum Genet* 82:1130–1140
- Bekada A, Arauna LR, Deba T, Calafell F, Benhamamouch S, Comas D (2015) Genetic heterogeneity in Algerian human populations. *PLoS ONE* 10:e0138453
- Beleza S, Gusmão L, Amorim A, Carracedo A, Salas A (2005) The genetic legacy of western Bantu migrations. *Hum Genet* 117:366–375
- Bonatto SL, Salzano FM (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci* 94:1866–1871
- Bortolini M-C, Salzano FM, Thomas MG et al (2003) Y-Chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet* 73:524–539
- Bortolini MC, Da Silva WA, De Guerra DC, Remonatto G, Mirandola R, Hutz MH, Weimer TA, Silva MCBO, Zago MA, Salzano FM (1999) African-derived South American populations: a history of symmetrical and asymmetrical matings according to sex revealed by bi- and uni-parental genetic markers. *Am J Hum Biol* 11:551–563
- Braganholi DF, Ambrosio IB, Polverari FS, Martinez J, Cicarelli RMB (2017) 500 years later: understanding the genetic ancestry of the southeast region, Brazil. *Forensic Sci Int Genet Suppl Ser* 6:e370–e371
- Calonga-Solís V, Amorim LM, Farias TDJ, Petzl-Erler ML, Malheiros D, Augusto DG (2020) Variation in genes implicated in B-cell development and antibody production affects susceptibility to pemphigus. *Immunology* 162:58–67
- Cann HM, De Toma C, Cazes L et al (2002) A human genome diversity cell line panel. *Science* 296:261b–2262
- Cardena MMSGSG, Ribeiro-dos-Santos Á, Santos S, Mansur AJ, Pereira AC, Fridman C (2013) Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. *PLoS ONE* 8:e62005
- Cárdenas JM, Heinz T, Pardo-Seco J, Álvarez-Iglesias V, Taboada-Echalar P, Sánchez-Díz P, Carracedo A, Salas A (2015) The multiethnic ancestry of Bolivians as revealed by the analysis of Y-chromosome markers. *Forensic Sci Int Genet* 14:210–218
- Carvalho-Silva DR, Santos FR, Rocha J, Pena SDJ (2001) The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet* 68:281–286
- Carvalho M, Brito P, Lopes V, Andrade L, Anjos MJ, Real FC, Gusmão L (2010) Analysis of paternal lineages in Brazilian and African populations. *Genet Mol Biol* 33:422–427
- Cavalli-Sforza LL, Bodmer WF (1971) The genetics of human populations. Freeman, San Francisco
- Chakraborty R (1986) Gene admixture in human populations: models and predictions. *Am J Phys Anthropol* 29:1–43
- Chamorro G, Combès I (2015) Povos indígenas em Mato Grosso do Sul: História, cultura e transformações sociais. UFGD, Dourados
- Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:1–16
- Chiaroni J, Underhill PA, Cavalli-Sforza LL (2009) Y chromosome diversity, human expansion, drift, and cultural evolution. *Proc Natl Acad Sci* 106:20174–20179
- Coelho M, Sequeira F, Luiselli D, Beleza S, Rocha J (2009) On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evol Biol* 9:80
- Corella A, Bert F, Pérez-Pérez A, Gené M, Turbón D (2007) Mitochondrial DNA diversity of the Amerindian populations living in the Andean Piedmont of Bolivia: Chimane, Moseten, Aymara and Quechua. *Ann Hum Biol* 34:34–55
- Crawford NG, Kelly DE, Hansen MEB et al (2017) Loci associated with skin pigmentation identified in African populations. *Science* 358:eaan8433
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Sczzari R, Underhill PA (2002) A back migration from Asia to Sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* 70:1197–1214
- Dafion VT, Carvalhaes F, Feres Júnior J (2017) Sentido na Pele: Percepções de Discriminação Cotidiana de Pretos e Pardos no Brasil. *Dados* 60:293–330

- de Filippo C, Barbieri C, Whitten M, Mpoloka SW, Gunnarsdóttir ED, Bostoen K, Nyambe T, Beyer K, Schreiber H, De Knijff P, Luiselli D, Stoneking M, Pakendorf B (2011) Y-chromosomal variation in sub-Saharan Africa: Insights into the history of Niger-Congo groups. *Mol Biol Evol* 28:1255–1269
- de Souza AM, Resende SS, de Sousa TN, de Brito CFA (2019) A systematic scoping review of the genetic ancestry of the Brazilian population. *Genet Mol Biol* 42:495–508
- Dias L (2019) Colonialism to Carnival : Tracking centuries of racialized imagery of Brazilian woman. Dissertation, University of Missouri-St. Louis
- Dornelles CL, Battilana J, Fagundes NJR, Freitas LB, Bonatto SL, Salzano FM (2004) Mitochondrial DNA and Alu insertions in a genetically peculiar population: The Ayoreo Indians of Bolivia and Paraguay. *Am J Hum Biol* 16:479–488
- Dornelles CL, Bonatto SL, De Freitas LB, Salzano FM (2005) Is haplogroup X present in extant South American Indians? *Am J Phys Anthropol* 127:439–448
- Emery LS, Magnaye KM, Bigham AW, Akey JM, Bamshad MJ (2015) Estimates of continental ancestry vary widely among individuals with the same mtDNA haplogroup. *Am J Hum Genet* 96:183–193
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567
- Farias TDJ, Augusto DG, De ARC, Malheiros D (2018) Screening the full leucocyte receptor complex genomic region revealed associations with pemphigus that might be explained by gene regulation. *Immunology* 156:86–93
- Fausto B (2015) História concisa do Brasil. Edusp, São Paulo
- Fernandes V, Alshamali F, Alves M, Costa MD, Pereira JB, Silva NM, Cherni L, Harich N, Cerny V, Soares P, Richards MB, Pereira L (2012) The Arabian Cradle: mitochondrial relicts of the first steps along the southern route out of Africa. *Am J Hum Genet* 90:347–355
- Ferreira LB, Mendes CT, Wiezel CEV, Luizone MR, Simões AL (2006) Y-STR diversity and ethnic admixture in White and Mulatto Brazilian population samples. *Genet Mol Biol* 29:605–607
- Figueiredo R, Ambrosio IB, Braganholi DF, Chemale G, Martins JA, Gomes V, Gusmão L, Ciccarelli RMB (2016) Male-specific contributions to the Brazilian population of Espírito Santo. *Int J Legal Med* 130:679–681
- Gayá-Vidal M, Moral P, Saenz-Ruales N, Gerbault P, Tonasso L, Vilena M, Vasquez R, Bravi CM, Dugoujon J-M (2011) mtDNA and Y-chromosome diversity in Aymaras and Quechuas from Bolivia: different stories and special genetic traits of the Andean Altiplano populations. *Am J Phys Anthropol* 145:215–230
- Gomes V, Sánchez-Díz P, Amorim A, Carracedo Á, Gusmão L (2010) Digging deeper into East African human y chromosome lineages. *Hum Genet* 127:603–613
- Gomez F, Hirbo J, Tishkoff SA (2014) Genetic variation and adaptation in Africa: implications for human evolution and disease. *Cold Spring Harb Perspect Biol* 6:6-a008524
- Gonçalves VF, Carvalho CMB, Bortolini MC, Bydlowski SP, Pena SDJ (2008) The Phylogeography of African Brazilians. *Hum Hered* 65:23–32
- Gouveia MH, Borda V, Leal TP et al (2020) Origins, admixture dynamics, and homogenization of the African gene pool in the Americas. *Mol Biol Evol* 37:1647–1656
- Green T (2012) The rise of the trans-Atlantic slave trade in western Africa, 1300–1589. Cambridge University Press, New York
- Guerreiro-Junior V, Bisso-Machado R, Marrero A, Hünemeier T, Salzano FM, Bortolini MC (2009) Genetic signatures of parental contribution in black and white populations in Brazil. *Genet Mol Biol* 32:1–11
- Hünemeier T, Carvalho C, Marrero AR, Salzano FM, Junho Pena SD, Bortolini MC (2007) Niger-Congo speaking populations and the formation of the Brazilian gene pool: mtDNA and Y-chromosome data. *Am J Phys Anthropol* 133:854–867
- IBGE (2007) Brasil: 500 anos de Povoamento. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830–838
- Kehdy FSG, Gouveia MH, Machado M et al (2015) Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proc Natl Acad Sci* 112:8696–8701
- Kimura L, Ribeiro-Rodrigues EM, De Mello Auricchio MTB, Vicente JP, Batista-Santos SE, Mingroni-Netto RC (2013) Genomic ancestry of rural African-derived populations from Southeastern Brazil. *Am J Hum Biol* 25:35–41
- Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A, Scozzari R, Modiano D, Coppa A, de Knijff P, Feldman M, Cavalli-Sforza LL, Oefner PJ (2006) The role of selection in the evolution of human mitochondrial genomes. *Genetics* 172:373–387
- Klein HS (2002) As origens africanas dos escravos brasileiros. In: Pena SDJ (ed) Homo Brasilis: Aspectos genéticos, linguísticos, históricos e socioantropológicos da formação do povo brasileiro, 1st edn. Ribeirão Preto, FUNPEC, pp 93–112
- Korey KA (1978) A critical appraisal of methods for measuring admixture. *Human Biology* 50:343–360
- Larmuseau MHD, Vessi A, Jobling MA, Van Geystelen A, Primitivo G, Biondi G, Martínez-Labarga C, Ottoni C, Decorte R, Rickards O (2015) The paternal landscape along the Bight of Benin - Testing regional representativeness of West-African population samples using Y-chromosomal markers. *PLoS ONE* 10:1–19
- Lê S, Josse J, Husson F (2008) FactoMineR : An R package for multivariate analysis. *J Stat Softw* 25:253–258
- Lima-Costa MF, Rodrigues LC, Barreto ML, Gouveia M, Horta BL, Mambrini J, Kehdy FSG, Pereira A, Rodrigues-Soares F, Victora CG, Tarazona-Santos E (2015) Genomic ancestry and ethnoracial self-classification based on 5,871 community-dwelling Brazilians (The Epigen Initiative). *Sci Rep* 5:9812
- Manta FSN, Pereira R, Vianna R, Beuttenmüller R, de Araújo A, Leite Góes Gitá D, Aparecida da Silva D, de Vargas WE, da Mota PI, Ivan Aguiar J, Ozório Moraes M, Fagundes de Carvalho E, Gusmão L (2013) Revisiting the genetic ancestry of Brazilians using autosomal AIM-Indels. *PLoS ONE* 8:e75145
- Marques SL, Goios A, Rocha AM, Prata MJ, Amorim A, Gusmão L, Alves C, Alvarez L (2015) Portuguese mitochondrial DNA genetic diversity—an update and a phylogenetic revision. *Forensic Sci Int Genet* 15:27–32
- Marrero AR, das Leite FPN, de Carvalho BA, Peres LM, Kimmers TC, da Cruz IM, Salzano FM, Ruiz-Linares A, da Silva WA, Bortolini MC (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
- Marrero AR, Silva-Junior WA, Bravi CM, Hutz MH, Petzl-Erler ML, Ruiz-Linares A, Salzano FM, Bortolini MC (2007) Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *Am J Phys Anthropol* 132:301–310
- Mattoso KMQ (1991) To be a slave in Brazil: 1550–1888. Rutgers University Press, New Brunswick
- Montano V, Ferri G, Marcari V, Batini C, Anyaele O, Destro-Bisol G, Comas D (2011) The Bantu expansion revisited: a new analysis

- of Y chromosome variation in Central Western Africa. *Mol Ecol* 20:2693–2708
- Myres NM, Rootsi S, Lin AA et al (2011) A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet* 19:95–101
- Nascimento A (2016) O genocídio do Negro brasileiro: Processo de um racismo mascarado. Perspectiva, São Paulo
- Oliveira JP, Freire CAR (2006) A presença indígena na formação do Brasil. MEC/UNESCO, Brasília
- Oliveira S, Fehn AM, Aço T, Lages F, Gayá-Vidal M, Pakendorf B, Stoneking M, Rocha J (2018) Matrilineal shape populations: Insights from the Angolan Namib Desert into the maternal genetic history of southern Africa. *Am J Phys Anthropol* 165:518–535
- Ongaro L, Sciliar MO, Flores R et al (2019) The genomic impact of European colonization of the Americas. *Curr Biol* 29:3974–3986. e4
- Paradis E (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420
- Pena SDJ, Bastos-Rodrigues L, Pimenta JR, Bydlowski SP (2009) DNA tests probe the genomic ancestry of Brazilians. *Brazilian J Med Biol Res* 42:870–876
- Petzl-Erler LM, Luz R, Sotomaior VS (1993) The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. *Tissue Antigens* 41:227–237
- Pinheiro AA, Assumpção JE, Follmann JI (2015) África e afrodescendentes no sul do Brasil: história, religião e educação. Casa Leiria, São Leopoldo
- Pinotti T, Bergström A, Geppert M et al (2019) Y chromosome sequences reveal a short beringian standstill, rapid expansion, and early population structure of native american founders. *Curr Biol* 29:149–157.e3
- Plaza S, Salas A, Calafell F, Corte-Real F, Bertranpetti I, Carracedo A, Comas D (2004) Insights into the western Bantu dispersal: mtDNA lineage analysis in Angola. *Hum Genet* 115:439–447
- Poletto MM, Malagolini M, Silva JS, Bicalho MG, Braun-Prado K (2019) Mitochondrial DNA control region diversity in a population from Paraná state—increasing the Brazilian forensic database. *Int J Legal Med* 133:347–351
- Poznik GD (2016) yHaplotype* Identifying Y-chromosome haplogroups in arbitrarily large samples of sequenced or genotyped men. bioRxiv. <https://doi.org/10.1101/088716>
- Primitivo G, Ottoni C, Biondi G, Serafino S, Martínez-Labarga C, Larmuseau MHD, Scardi M, Decorti R, Rickards O (2017) Bight of Benin: a maternal perspective of four Beninese populations and their genetic implications on the American populations of African ancestry. *Ann Hum Genet* 81:78–90
- Quintana-Murci L, Quach H, Harmant C et al (2008) Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *Proc Natl Acad Sci* 105:1596–1601
- Raggio AZ, Bley RB, Trauczynski SC (2018) Abordagem histórica sobre a população negra no Estado de Paraná. SEJU, Curitiba
- Reis JJ (2007) Presença negra: conflitos e encontros. In: IBGE (ed) Brasil: 500 anos de povoamento, 1st edn. IBGE, Rio de Janeiro, pp 80–99
- Resque R, Gusmão L, Geppert M, Roewer L, Palha T, Alvarez L, Ribeiro-Dos-santos A, Santos S (2016) Male lineages in Brazil: Intercontinental admixture and stratification of the European background. *PLoS ONE* 11:1–17
- Ribeiro D (1995) O povo brasileiro: A formação e o sentido do Brasil. Companhia das Letras, São Paulo
- Rito T, Richards MB, Fernandes V, Alshamali F, Cerny V, Pereira L, Soares P (2013) The first modern human dispersals across Africa. *PLoS ONE* 8:e80031
- Rosa A, Brehm A (2011) African human mtDNA phylogeography at-a-glance. *J Anthropol Sci* 89:25–58
- Sá Junior MT (2016) Não é feitiçaria! É tecnologia! Africanos e descendentes na mineração do Mato Grosso setecentista. *Revista Eletrônica História Em Reflexão* 10:81–98
- Sala A, Argüelles CF, Marino ME, Bobillo C, Fenocchio A, Corach D (2010) Genetic analysis of six communities of Mbyá-Guaraní inhabiting Northeastern Argentina by means of nuclear and mitochondrial polymorphic markers. *Hum Biol* 82:433–456
- Sala A, Caputo M, Gimart S, Theiler G, Parolin ML, Carnese RF, Fainboim L, Corach D (2018) Historical records under the genetic evidence: “Chiriguano” tribe genesis as a test case. *Mol Biol Rep* 45:987–1000
- Salas A, Richards M, De la Fe T, Lareu M-V, Sobrino B, Sánchez-Díaz P, Macaulay V, Carracedo Á (2002) The making of the African mtDNA landscape. *Am J Hum Genet* 71:1082–1111
- Salas A, Richards M, Lareu M-V, Scozzari R, Coppa A, Torroni A, Macaulay V, Carracedo Á (2004) The African diaspora: mitochondrial DNA and the Atlantic Slave Trade. *Am J Hum Genet* 74:454–465
- Salzano FM, Sans M (2014) Interethnic admixture and the evolution of Latin American populations. *Genet Mol Biol* 37:151–170
- Santos RV, Fry PH, Monteiro S, Maio MC, Rodrigues JC, Bastos-Rodrigues L, Pena SDJ (2009) Color, race, and genomic ancestry in Brazil. *Curr Anthropol* 50:787–819
- Schaan AP, Costa L, Santos D, Modesto A, Amador M, Lopes C, Rabenhorst SH, Montenegro R, Souza BDA, Lopes T, Yoshioka FK, Pinto G, Silbiger V, Ribeiro-dos-Santos Á (2017) mtDNA structure: the women who formed the Brazilian Northeast. *BMC Evol Biol* 17:185
- Schaan AP, Gusmão L, Januzzi J, Modesto A, Amador M, Marques D, Rabenhorst SH, Montenegro R, Lopes T, Yoshioka FK, Pinto G, Santos S, Costa L, Silbiger V, Ribeiro-Dos-Santos Á (2020) New insights on intercontinental origins of paternal lineages in Northeast Brazil. *BMC Evol Biol* 20:1–9
- Schmitt R, Bonatto SL, Freitas LB, Muschner VC, Hill K, Hurtado AM, Salzano FM (2004) Short Report: Extremely limited mitochondrial DNA variability among the Aché Natives of Paraguay. *Ann Hum Biol* 31:87–94
- Silva M, Alshamali F, Silva P, Carriollo C, Mandlak F, Jesus Trovoada M, Černý V, Pereira L, Soares P (2015) 60,000 years of interactions between Central and Eastern Africa documented by major African mitochondrial haplogroup L2. *Sci Rep* 5:12526
- Silva S, Klautau-guimarães MDN, Grisolia CK (2010) b-globin haplotypes in normal and hemoglobinopathic individuals from Recôncavo Baiano, State of Bahia, Brazil. *Genet Mol Biol* 33:411–417
- Silva WA, Bortolini MC, Schneider MPC, Marreca A, Elion J, Krishnamoorthy R, Zago MA (2006) mtDNA haplogroup analysis of Black Brazilian and Sub-Saharan populations: implications for the Atlantic Slave Trade. *Hum Biol* 78:29–41
- Sison CP, Glaz J (1995) Simultaneous confidence intervals and sample size determination for multinomial proportions. *J Am Stat Assoc* 90:366–369
- Soares P, Alshamali F, Pereira JB, Fernandes V, Silva NM, Afonso C, Costa MD, Musilova E, Macaulay V, Richards MB, Cerny V, Pereira L (2012) The expansion of mtDNA haplogroup L3 within and out of Africa. *Mol Biol Evol* 29:915–927
- Soddyall H (2013) Lemba origins revisited: Tracing the ancestry of Y chromosomes in South African and Zimbabwean Lemba. *S Afr Med J* 103:1009
- Tavares GM, Reales G, Bortolini MC, Fagundes NJR (2019) Measuring the impact of European colonization on Native American populations in Southern Brazil and Uruguay: evidence from mtDNA. *Am J Hum Biol* 31:1–16

- Tsuneto LT, Probst CM, Hutz MH, Salzano FM, Rodriguez-Delfin LA, Zago MA, Hill K, Hurtado AM, Ribeiro-Dos-Santos AKC, Petzl-Erler ML (2003) HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. *Tissue Antigens* 62:512–526
- Underhill PA, Kivisild T (2007) Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet* 41:539–564
- van Oven M (2015) PhyloTree Build 17: growing the human mitochondrial DNA tree. *Forensic Sci Int Genet Suppl Ser* 5:e392–e394
- van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30:386–394
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson A (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253:1503–1507
- Wade P, López Beltrán C, Restrepo E, Ventura Santos R (2014) *Mestizo genomics: race mixture, nation, and science in Latin America*. Duke University Press, Durham
- Walsh R (1830) Notices of Brazil in 1828 and 1829, vol I. Frederick Westley and A.H Davis, London
- Ward RH, Salzano FM, Bonatto SL, Hutz MH, Coimbra CEA, Santos RV (1996) Mitochondrial DNA polymorphism in three Brazilian Indian tribes. *Am J Hum Biol* 8:317–323
- Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt HJ, Kronenberg F, Salas A, Schönherz S (2016) HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res* 44:W58–W63
- Zegura SL (2003) High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol Biol Evol* 21:164–175
- Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28:3326–3328

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

6.2 CAPÍTULO II: Ancestry, diversity, and genetics of health-related traits in African-derived communities (*quilombos*) from Brazil



Ancestry, diversity, and genetics of health-related traits in African-derived communities (*quilombos*) from Brazil

Iriel A. Joerin-Luque¹ · Natalie Mary Sukow¹ · Isabela Dall'Oglio Bucco² · Joana Gehlen Tessaro² · Claudemira Vieira Gusmão Lopes³ · Ana Angélica Leal Barbosa⁴ · Marcia H. Beltrame¹

Received: 27 December 2022 / Revised: 22 February 2023 / Accepted: 23 February 2023
 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Brazilian *quilombos* are communities formed by enslaved Africans and their descendants all over the country during slavery and shortly after its abolition. *Quilombos* harbor a great fraction of the largely unknown genetic diversity of the African diaspora in Brazil. Thus, genetic studies in *quilombos* have the potential to provide important insights not only into the African roots of the Brazilian population but also into the genetic bases of complex traits and human adaptation to diverse environments. This review summarizes the main results of genetic studies performed on *quilombos* so far. Here, we analyzed the patterns of African, Amerindian, European, and subcontinental ancestry (within Africa) of *quilombos* from the five different geographic regions of Brazil. In addition, uniparental markers (from the mtDNA and the Y chromosome) studies are analyzed together to reveal demographic processes and sex-biased admixture that occurred during the formation of these unique populations. Lastly, the prevalence of known malaria-adaptive African mutations and other African-specific variants discovered in *quilombos*, as well as the genetic bases of health-related traits, are discussed here, together with their implication for the health of populations of African descent.

Keywords Black Brazilians · Hemoglobinopathies · G6PD · Genetic ancestry · Sex bias · Complex disease

Introduction

Brazil has the highest proportion of Afro-diasporic population in the world, and more than 50% of Brazilians are self-declared Blacks (IBGE, <https://www.ibge.gov.br/>). Brazil

was the last country to declare slavery illegal in 1888 and was the destination of about 40% of the Africans forced to move to America, with estimations that range from 4 to 15 million people (Moura 1993). During the approximately four hundred years that slavery lasted, Africans and their descendants settled over almost the entire Brazilian territory (FCP, <http://www.palmares.gov.br/>). Due to the inhumanity of the traffic, however, there are no written records of the identity and geographic origins of enslaved Africans (Hall 2005). The only written records of the slave trade in the country were burned by the Brazilian authorities after the abolition to avoid compensation claims by slavers, remaining only rough estimations based on anthropological studies and navy files (Lacombe et al. 1988). Due to this lack of historical records, the ethnicities and/or original populations of the Africans who formed the diverse geographic regions in Brazil as well as specific African-derived populations, like *quilombos*, are mostly unknown (Hall 2005).

Africans and their descendants developed several ways of resistance to exploitation during slavery. Among them, the formation of communities where African traditional ways of living transcended the oppression of slavery was the most

✉ Iriel A. Joerin-Luque
 irieljoerin@gmail.com

¹ Programa de Pós-Graduação Em Genética, Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal Do Paraná (UFPR), Centro Politécnico, Jardim das Américas, Curitiba, Paraná 81531-980, Brazil

² Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal Do Paraná (UFPR), Centro Politécnico, Jardim das Américas, Curitiba, Paraná 81531-980, Brazil

³ Câmara de Educação Do Campo, Universidade Federal Do Paraná (UFPR), Setor Litoral, Matinhos, Paraná 83260-000, Brazil

⁴ Laboratório de Biologia E Genética Humana, Departamento de Ciências Biológicas, Universidade Estadual Do Sudoeste da Bahia (UESB), Campus de Jequié, Bahia, Brazil

important (Moura 1993). Popularly known as Black rural communities in Central-Western and Southern Brazil, and as Black territories or *mocambos* in the North and the Northeast, these settlements have more recently acquired the legal name of *quilombos*, as recalled by the National Coordination of Joint of the Rural Black Quilombolas Communities (CONAQ, <http://conaq.org.br/>). The word *quilombo* derives from Central African Bantu languages, and in Africa, it defines several kinds of human settling (Leite 1999). Over the years in America, however, it has been used with different meanings, being in the last decades vindicated by the Brazilian Black movements as the legal word to identify Afro-derived communities in the context of their fight for land rights and respect for their ancestral traditions (Matos and Eugenio 2019). The legal definition of *quilombo* has changed over the years, given the impossibility to universalize and encompass the diversity of histories and the heterogeneity of their characteristics (Matos and Eugenio 2019). Their common denominator, however, has been their link to African ancestry and the community use of land for subsistence and culture preservation. In most cases, these populations have traditionally subsisted on agriculture or are directly dependent on the land's natural resources (Gomes 2015). The latter make *quilombos*, like other traditional populations in Brazil, crucial for the conservation and sustainable usage of biodiversity (Shiraishi Neto 2007). In addition, these characteristics give *quilombos* their identification as an ethnic group in Brazil (Gomes 2015). Today, there are 5972 Black rural communities throughout Brazil (IBGE). Even so, only 2807 of them have already been legally certified as *quilombos* and have their land right recognized by the official laws (FCP).

Each *quilombo* has a particular history, which has shaped its demographic processes and interaction with the environment and populations around it (Matos and Eugenio 2019). Given that mainly Africans and their descendants participated in their formation, it is expected that African genetic ancestry in *quilombos* should be predominant (Nunes et al. 2020). That fact has important implications for the study of health-related variants in populations of African descent, currently underrepresented in biomedical studies (Peprah et al. 2015; Tawfik et al. 2023). Africa is known to host the highest human genetic and phenotypic diversity in the world, given the origin of modern humans in this continent approximately 200 to 500 thousand years ago and their adaptation to its extremely diverse environments (Pereira et al. 2021). Thus, genetic studies in *quilombos* could help to unveil an important portion of the human genetic variation.

Here we review 57 genetic studies carried out in Brazilian *quilombos* (Fig. 1, Table S1). The data obtained from these studies has great importance in the knowledge of Brazilian, and especially African-Brazilian, history, but also has implications for the health of the entire Brazilian population and

other populations of the African diaspora. We expect that this work will serve as a reference and motivation for the continuity of research in the field.

Genetic ancestry and the evolutionary history of *quilombos*

Altogether, we reviewed 15 studies on autosomal genomic ancestry: 14 regarding continental and only one about subcontinental ancestry (Table S2). The studies included 30 communities from all five Brazilian geopolitical regions: North (N, $n=8$), Northeast (NE, $n=8$), West-Central (WC, $n=1$), Southeast (SE, $n=11$), and South (S, $n=2$) (Fig. 1, Table S2). We also reviewed 12 studies on uniparental ancestry (determined through the analysis of the non-recombining region of the Y chromosome and the mitochondrial DNA—mtDNA). These encompass Northern ($n=9$), Northeastern ($n=11$), West-Central ($n=1$), and the Southeastern region ($n=9$) (Fig. 1, Table S3). Only one study (of mtDNA) included a Southern *quilombo* (Bortolini et al. 1997) (Table S3), but since it did not estimate ancestry proportions, it was not included in our discussion.

Slow (protein loci, single-nucleotide polymorphisms—SNPs, and indels), and fast-evolving genetic markers (microsatellites) were used to calculate continental ancestry proportions, with more than one type of marker for the same community in some cases. Due to the heterogeneity of markers, the proportions of the African, Amerindian, and European ancestries obtained differ considerably between communities and between the types of markers used (Fig. 1, Table S2 and S3). However, with a few exceptions, such variations did not change the predominance of each ancestry (Table S2 and S3). The use of different markers along with the use of different sample sets could explain, in part, the divergences between the proportions obtained in the studies. Fast-evolving markers provide information on smaller time scales, so the resulting data is more informative about subcontinental ancestry and recent demographic processes. In contrast, slow-evolving markers furnish information about larger time scales, being more informative about ancestry at the continental level (de Knijff 2000). For the following section of the discussion, we selected the most representative study for each *quilombo*, prioritizing slow over fast-evolving markers, the higher numbers of markers, and the higher sample sizes, when possible (Fig. 1, Table S2 and S3). It should be noted that the most important feature behind the differences in ancestry proportions between *quilombos* is the particularity of each community's demographic processes throughout their development.

Regarding autosomal ancestry, as expected, most of the evaluated communities have predominant African ancestry, ranging from 38.2% in Pacoval (Northern Brazil) to

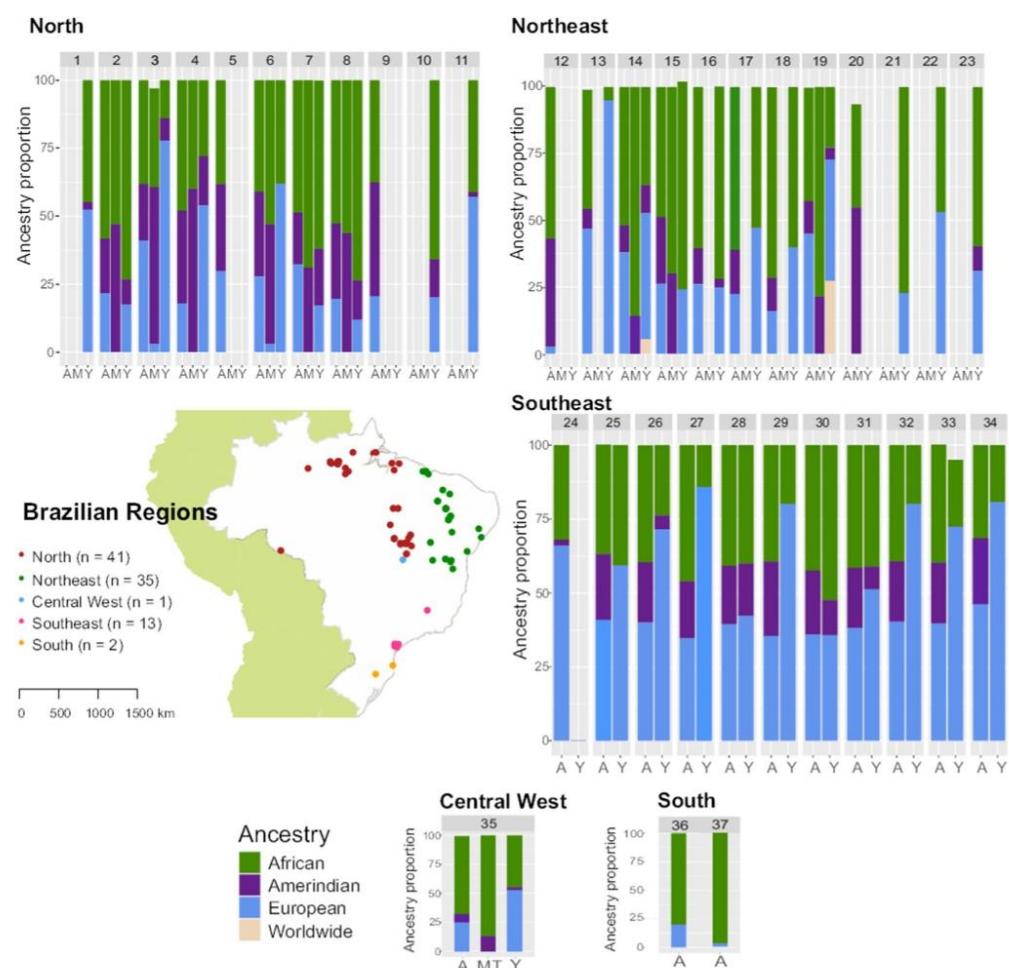


Fig. 1 Autosomal and uniparental ancestry and geographic distribution of Brazilian *quilombos* included in genetic studies so far. A, autosomal; M, mitochondrial; Y, Y-chromosome ancestries. 1—Mazagão Velho; 2—Curiaú; 3—Mazagão; 4—Cametá; 5—Pacoval; 6—Marajó; 7—Pitimandeuá; 8—Trombetas; 9—Santo Antônio do Guaporé; 10—Itacoá; 11—Saracura; 12—Santiago do Iguaçú; 13—Riacho de Sacutíaba and Sacutíaba; 14—Rio das Rãs; 15—Cajueiro; 16—Pontal; 17—Mimbó; 18—Sítio Velho; 19—Mocambo; 20—Tamaúari; 21—Barra; 22—São Gonçalo; 23—Gauçinha; 24—Marinhos; 25—Abobral; 26—André Lopes; 27—Galvão; 28—Ivapó-runduva; 29—Maria Rosa; 30—Nhunguara; 31—Pedro Cubas; 32—Pilões; 33—São Pedro; 34—Sapatu; 35—Kalunga; 36—Paredão; 37—Sertão do Valongo (Schneider et al. 1987; Bortolini et al. 1995, 1997, 1998, 1999; Arpini-Sampaio et al. 1999; Guerreiro et al. 1999; Ribeiro-Dos-Santos et al. 2002; Cayres Vallinoto et al. 2003; Abesandes et al. 2004; de Souza and Culpí 2005; Barbosa et al. 2006; Silva et al. 2006; Carvalho et al. 2008; Sciliar et al. 2009; Ribeiro et al. 2009, 2011; Amorim et al. 2011; Maciel et al. 2011; Palha et al. 2011; Kimura et al. 2013, 2017; Wiezel et al. 2013; Gontijo et al. 2014)

97.33% in Sertão do Valongo (Southern Brazil) (Da Silva et al. 1999; De Oliveira et al. 2001; de Souza and Culpí 2005; Maciel et al. 2011) (Fig. 1, Tables 1 and S2). Variability in ancestry proportions is probably associated with

the founder effect (as many *quilombos* are known to have been founded by few individuals) as well as with differences in the level of admixture over the generations. Thus, greater proportions of African ancestry are observed in

Table 1 Autosomal ancestry measures in *quilombos* from the five Brazilian regions. AFR, African; EUR, European; AMER, Amerindian ancestry (Schneider et al. 1987; Bortolini et al. 1995, 1998; Arpini-Sampaio et al. 1999; Guerreiro et al. 1999; Da Silva et al. 1999; Cayres Vallinot et al. 2003; de Souza and Culpí 2005; Barbosa et al. 2006; Sclar et al. 2009; Amorim et al. 2011; Maciel et al. 2011; Kimura et al. 2013; Gontijo et al. 2014, 2018)

Geographic region	Ancestry average, median, and standard deviation (%)								
	AFR			EUR			AMER		
Central West	67.3	67.3	0	24.9	24.9	0	7.2	7.2	0
North	46.5	44.6	9.5	27.2	27.4	7.4	26.3	26	10.1
Northeast	55.9	56.8	10.3	28.3	32.6	15.2	15.6	13.4	11.4
South	88.8	88.8	12.1	11.2	11.2	12.1	0	0	0
Southeast	42.2	39.8	9	38.5	39.4	3.7	19.3	20.4	6.6

traditionally more isolated communities—such as Sertão do Valongo, Paredão, and Bananal (Maciel et al. 2011).

Only in communities from the North and the Northeast, there are *quilombos* with predominant Amerindian ancestry or the second highest, agreeing with several historical reports which mention that the admixture between Native Americans and *quilombolas* from the Amazon region was quite intense and higher than that observed in other regions (Gomes 2015). This is also consistent with the larger Native American population that inhabits the Brazilian Amazon (Brazilian Socioenvironmental Institute; <https://pib.socioambiental.org>). It is common for *quilombola* communities and indigenous settlements to be geographically close in Brazil's regions where the Amerindian presence is high, such as the Amazon region. That proximity has favored the gene flow between these groups—despite not always resulting in a harmonious coexistence (Gomes 2015).

The Santo Antônio do Guaporé community, located in Rondônia (N), is the only *quilombo* with predominantly Amerindian ancestry, followed by African and European ancestries (Gontijo et al. 2014). In the Guaporé Valley (*Vale do Guaporé*, located on the Bolivian-Brazilian border), the region where the community is established, several Amerindian groups have traditionally been settled, which favored the relationship with the Afro-Brazilians when they founded *quilombos* in this area in the late nineteenth century (Teixeira and Xavier 2018).

The Amerindian presence in the *quilombos* was frequently associated with voluntary entry, given that the *quilombos* may have provided refuge to Amerindians and other social classes marginalized in the colonial society, including prostitutes and military service deserters (Moura 1992; Gomes 2015). However, it is important to emphasize that there are several reports of abduction of Amerindian women by *quilombolas*, highlighting that the female Amerindian presence in *quilombos* could also have been forced (Gomes 2015). Since both populations were widely enslaved, some authors suggest that the first resistance communities of enslaved people were formed by Africans, their

descendants, and Amerindians (Gomes 2015). Accordingly, although Santo Antônio do Guaporé was the only community in which Amerindian ancestry exceeded the others, considerable proportions of this ancestry (> 20%) are also identified in other *quilombos* such as Santiago do Iguape (Bahia, NE), Pacoval, and Cametá (Pará, N), and the communities of the Ribeira River Valley (*Vale do Ribeira*, São Paulo, SE) (Bortolini et al. 1995; Maciel et al. 2011; Gontijo et al. 2014).

According to Brazilian demographic history, high proportions of European ancestry would be expected in Southern communities due to the larger contingent of European migrants received in that region during the nineteenth and twentieth centuries (IBGE, 2007). In the Southern *quilombos* studied so far (Sertão do Valongo and Paredão), nevertheless, the proportions of European ancestry described do not exceed 20% (Bortolini et al. 1995, 1999; de Souza and Culpí 2005) (Fig. 1, Table S2). Notably, the proportion reported for Sertão do Valongo (2.67%) was the second smallest yet reported in Brazilian *quilombos*, except for the Santiago do Iguape community (NE, 2.5%) (de Souza and Culpí 2005; Gontijo et al. 2014). It should be noted that the low contribution of European ancestry in both communities mentioned contrasts sharply with that reported for other Brazilian *quilombos*, in which the estimates for such ancestry exceed at least 10%. This particularity possibly comes from the semi-isolation of the communities, resulting from their remote location and the historically recorded segregation generated by European-derived neighboring populations (de Souza and Culpí 2005; Gontijo et al. 2014).

Predominant European ancestry, followed by African and Amerindian, was found in communities from the North, Northeast, and Southeast of Brazil (Fig. 1, Table S2). In addition to these, most of the evaluated communities have significant estimates of European ancestry (> 20%)—especially those located in the Ribeira River Valley (SE), in which the estimates exceed 30% (Table S2). The European contribution to *quilombos*, possibly stems from founders of the communities already having European ancestry but also

from the more recent migration of people with this ancestry (Amorim et al. 2011; Kimura et al. 2013).

The data reviewed here demonstrated that, differently from the gradient of ancestry proportions that has already been shown for the Brazilian urban populations (with the highest European ancestry at Southern longitudes and higher Amerindian and African ancestry in the North and the Northeast, respectively), it is not possible to identify a pattern associating the geographic location of *quilombos* to their ancestry proportions (Fig. 1 and Tables 1, S2) (Pena et al. 2011; de Souza et al. 2019). Interestingly, however, we observed that, on average, *quilombos* from all five Brazilian geographic regions had predominant African ancestry and the highest Amerindian contribution is observed in the North of Brazil (Table 1).

Complementing autosomal ancestry studies, the analysis of uniparental ancestry patterns provides a fine-scale picture of demographic processes in human populations (Underhill and Kivisild 2007). Despite the fact that both mtDNA and the Y chromosome were not studied simultaneously for all the populations, we could observe some recurring patterns. In the first place, the European gene flow was exclusively male-mediated in 89.5% of the *quilombos*, being the predominant paternal ancestry in more than half of the *quilombos* assessed, representing all the regions studied (Fig. 1 and Table S2). Indeed, European ancestry was observed in Y chromosomes from all the *quilombos* studied, whereas its introduction by the maternal line was only observed in a few communities from the Northern region (less than 30% of the populations). Even though the African paternal contribution was observed in all *quilombos*, different from the expectations, it was the highest only in less than 40% of them. Even more striking, the African paternal contribution was less than 10% in some Northeastern communities. In sharp contrast, the African maternal contribution was detected in all *quilombos*, being African the predominant maternal ancestry in more than 70% of them. In the only communities where African maternal ancestry was lower than 40% (Mazagão and Tamaúari, in the Amazon region), the reciprocal contribution was almost exclusively from Amerindian women (Fig. 1, Table S3).

That evident sex-biased admixture pattern observed in the *quilombos* studied so far agrees with that already reported for other admixed populations from Brazil and populations throughout America, as well as from other European colonized countries in the world (Trovadoa et al. 2007; Ongaro et al. 2019; Korunes et al. 2020; Martínez et al. 2020). This sex-biased gene flow can be related to the extensive historical records of sexual exploitation of enslaved African and Amerindian women by European men, common during the colonial period (Walsh 1830; Nascimento 2016; Aidoo 2018). The continued abuse of enslaved women by their slavers

frequently resulted in children not only unrecognized by their male progenitors but also enslaved (Aidoo 2018). In contrast, despite the fact that the majority of Africans brought to America were male, their relatively short life expectancy due to maltreatment and compulsory work (10 years, on average) can potentially be reflected in the relatively low paternal African contribution (Carvalho et al. 2008; Myscofski 2013).

Regarding the Amerindian ancestry, the paternal contribution was less expressive, being detected in 60% of the populations. None of the communities showed more than 21% of Amerindian paternal ancestry. In comparison, the Amerindian maternal contribution was observed in all populations in higher proportions (10–60%), being predominant in some populations of the Amazon region (Tamaúari, Mazagão, and Cametá) (Fig. 1, Table S3).

Quilombos were not always isolated

The persecution and recurrent destruction of *quilombos* by colonial slavers led several communities to settle in remote locations (de Andrade 1995). As expected from their relatively higher isolation in comparison to urban Brazilian populations, the levels of inbreeding of *quilombos* are also higher (Lemes et al. 2014). However, and as expected, they varied greatly according to the *quilombo* studied, reflecting the diverse demographic processes involved in their formation (Lemes et al. 2014; Cruz et al. 2020). High levels of inbreeding have been seen in the more isolated *quilombos* such as Sertão do Valongo and Tucum (from Southern and Northeastern Brazil, respectively). In particular, Sertão do Valongo was founded in the 1880s by four couples including seven freed enslaved Africans and a white man (de Souza and Culp 2005). Since then, the community has experienced population growth with low exogamy, which explains the high levels of inbreeding and high African ancestry observed in that community (de Souza and Culp 2005). On the other hand, a significant amount of gene flow and shared ancestry has been detected in the *quilombos* of the Ribeira River Valley (Southeastern Brazil), suggesting a high level of interaction among nearby *quilombos* (Lemes et al. 2014). The interaction between different African-derived communities, as well as with the urban and Native populations around them, has been extensively recorded as the means of articulation and subsistence of *quilombos* (de Andrade 1995; Gomes 2015). Bearing in mind that the definition of *quilombo* in Brazil is critical for preserving their land rights and affirmative actions benefits, it is crucial to consider that not only one historical trajectory or demographic process can be attributed to *quilombos* in general (Guimarães Paiva et al. 2020).

Insights on the within-Africa roots of African-Brazilians

From the limited extant historical records on the origin of enslaved Africans (mainly transatlantic slave trade voyage documents), it is known that, roughly speaking, two African coasts were the main points of departure of the people brought to Brazil (Hall 2005; Pinheiro et al. 2015). From West Africa (mainly from what is today Ghana, Nigeria, Benin, and Togo but also from the Greater Senegambia region – the region between Senegal and Sierra Leona rivers) they came speakers of Niger-Congo non-Bantu languages, such as Yoruba, Fon, Monde, etc. From the Portuguese colonies of Angola, Congo, and Mozambique, in West-Central and Southeastern Africa, respectively, were brought people who spoke Bantu languages (like Kikongo, Kimbundu, and Mbundu) (Hall 2005). There is no information, however, about the specific ethnicities of these people nor if they came from regions on the continent more distant from the coasts (Hall 2005). Despite uniparental markers showing biogeographic specificity, and being useful to estimate the relative contribution from specific regions within the African continent, no study has inferred subcontinental African contributions to the formation of *quilombos* (Table S3). However, a few studies have discussed the possible origin of the uniparental haplotypes detected. For example, Y chromosomes belonging to haplogroup (Hg) E1b1a, with haplotypes typical from West-Central Africa, were detected in *quilombos* from Southeastern Brazil (Kimura et al. 2017).

In most *quilombos*, the mitochondrial haplogroups (Hgs) detected belong to the African clades L1, L2, and L3, which have a broad distribution over sub-Saharan Africa (Table S3). Similarly, all African Y Hgs were shown to belong to clade E, which is also the most common Hg in sub-Saharan Africa (Table S3). No study detected the Y Hgs typical from West-Central African hunter-gatherer populations nor Southern or Eastern African can populations, such as the lineages derivatives from clades A or B (Table S3) (Knight et al. 2003). Also, no Southern African Khoisan mitochondrial Hg (such as L0d and L0k) was detected in the *quilombos* assessed so far (Table S3) (Salas et al. 2002; Knight et al. 2003). Nevertheless, both mtDNA Hg L0d and Y chromosome clade B derivatives have been detected in low frequencies in urban admixed Southern Brazilian populations (Hünemeier et al. 2007). As the studies reviewed here have sampled an average of 30 individuals per study per *quilombo*, there is a possibility that some of the less frequent lineages could have remained undetected or lost by genetic drift. Thus, despite these results being consistent with the absence of an expressive involvement of West-Central African hunter-gatherer populations or Southern and Eastern African populations in the trans-Atlantic

traffic of human beings, a more in-depth investigation of subcontinental ancestry in *quilombos* and a broader sampling could be helpful to fill the historical gaps (Salas et al. 2004; Hall 2005).

We identified only one study that estimated the subcontinental African contributions in *quilombos* from Northern and Northeastern Brazil. Those contributions were inferred based on the frequency of the beta-globin S (HbS) haplotypes (Pante-De-Sousa et al. 1999; Silva et al. 2010). Five haplotypes of the beta-globin gene, in chromosome 11, have been described and are identified based on the geographic locations where they occur more frequently. These are the Arab-Indian, Benin, Cameroon, Bantu (also called the Central African Republic (CAR) haplotype), and Senegal haplotypes (Cruz et al. 2019). The *quilombos* Santiago do Iguapé (Northeast), Curiau, Pacoval, and Trombetas (North) were studied for these loci. In Santiago de Iguapé, 52.9% of the haplotypes were from Benin, whereas 32.5% were Bantu, with an absence of the Senegal haplotype (Silva et al. 2010). In the *quilombos* from the Amazonian region (Curiau, Pacoval, and Trombetas), the Bantu haplotype was more frequent, followed by the Senegal and Benin haplotypes (60, 30, and 10%, respectively) (Pante-De-Sousa et al. 1999). The frequencies from Santiago do Iguapé, however, were computed together with other admixed populations from the state of Bahia. Hence, they are not informative about the subcontinental ancestry of the *quilombo* itself. The prevalence of the Bantu haplotype in the Amazonian *quilombos* agrees with that observed in other admixed Brazilian populations. In contrast, the Senegal haplotype, which proved to be relatively common in the assessed *quilombos*, has been proportionally less observed in other admixed populations from Brazil (Silva et al. 2010). Caution must be taken because, as the Senegal HbS haplotype has been associated with a more benign clinical course, hence less hospitalization rate, it could be underrepresented in studies including only sickle cell patients from health centers (Silva et al. 2010). Besides, another two elements should be considered when the contribution of Africans from Senegambia is concerned. In the first place, this contribution is expected in the Northern region, due to the known migration of enslaved Africans from French and Dutch Guyana to that region (Ribeiro-Dos-Santos et al. 2002). And, secondarily, there is a known underestimation of the numbers of Africans brought from the Greater Senegambia region (the region between the Senegal and Sierra Leona rivers) to America in historical records (Hall 2005). Thus, for the Amazonian *quilombos*, these results both corroborate the West-Central and Western origin of the Africans brought to the region between the sixteenth and the nineteenth centuries and highlight the significant contribution from Senegambia (Pante-De-Sousa et al. 1999).

African adaptations in the Brazilian environment

Genomic studies have shown that the main drivers of recent human evolution are those related to skin pigmentation, metabolism, and infectious disease resistance (Grossman et al. 2013). Among them, the exposure to *Plasmodium* species (including *P. falciparum*, *P. ovale*, *P. malariae*, and *P. vivax*), transmitted by female *Anopheles* mosquitoes, and resulting in the parasitic disease malaria in susceptible individuals, is considered one of the strongest selective pressures and has been linked to a plethora of adaptations (Pereira et al. 2021). The growth of the mosquito population, and thus the spreading of malaria, is associated with the introduction of agriculture in African populations (Relethford 2012). In Africa, malaria spreads in the equatorial belt (the Sub-Saharan region) and diminishes its incidence southward after the Kalahari Desert. Several variants that confer resistance to infection originated in that region and have reached high frequencies due to selection (Pereira et al. 2021). Common African variants associated with malaria resistance include those in the genes *HBB*, *G6PD*, and *DARC* (Pereira et al. 2021). Some of these variants have been studied in *quilombos*, among them, those that cause hemoglobinopathies (hemoglobin S and C, and α- and β-thalassemia) and glucose-6-phosphate dehydrogenase (G6PD) deficiency (Table S4) (Gomez et al. 2013).

Hemoglobinopathies are hemoglobin hereditary disorders that can affect the structure of hemoglobin itself, or alter its production, as occurs in thalassemias (Gomez et al. 2013; Nussbaum et al. 2016).

Hemoglobin S (HbS) mutation is the result of a nucleotide change in the sixth codon of the β-globin gene (rs334(T)), replacing an adenine with thymine (GAG → GTG), causing the substitution of a single amino acid, with valine replacing glutamic acid ($\beta^6\text{Glu} \rightarrow \text{Val}$). That changes the round-shaped hemoglobin to a crescent or “sickle” shape, which can block blood flow to the rest of the body. Hemoglobin C (HbC) mutation also changes the sixth amino acid of the β-globin gene (rs33930165(A)), replacing it with a lysine ($\beta^6\text{Glu} \rightarrow \text{Lys}$) (Nussbaum et al. 2016).

These variants confer resistance to malaria in the heterozygous state whereas, in homozygosity, constitute severe genetic disorders, being only maintained in the population by balancing selection (Gomez et al. 2013). The presence of two hemoglobin S alleles causes sickle cell anemia, being the most common type of sickle cell disease (SCD). The association between HbS with any other hemoglobin variant (HbC, HbD, HbE, or β-thalassemia), also causes sickle cell disease, having an autosomal recessive pattern of inheritance. Approximately two-thirds of the SCD-affected newborns occur in Africa (Piel et al. 2013a).

Today, there is evidence for a single African origin of the HbS variant, at least 7000 years ago in the Sahara or

in the rainforests of Central Africa (vicinity of present-day Cameroon) (Shriner and Rotimi 2018; Esoh and Wonkam 2021). However, HbS occurs throughout Sub-Saharan Africa surrounded by at least four different haplotype backgrounds with well-defined geographic distributions (discussed in the previous section) (Ngo Bitoungui et al. 2015). HbC, on the other hand, is assumed to have emerged only once in Western Africa, reaching its highest frequency around Burkina Faso and the Bight of Benin (Piel et al. 2013a).

The introduction of the *Plasmodium* sp. (and, in this way, malaria) in Brazil is thought to have occurred through the navies of the Transatlantic Slave Trade, during colonial times. Since then, the Brazilian Amazon has become malaria-endemic, whereas other Brazilian regions, such as the West-Central, Southeastern, and Southern regions, maintain a residual transmission (Prefeitura de São Paulo 2018). For this reason, the knowledge of the prevalence of malaria resistance variants in populations of African descent is of public health concern in Brazil. Here, we reviewed 17 studies encompassing 60 *quilombos* from the five Brazilian regions (Table S4).

The distribution of the HbS and HbC alleles in *quilombos* is heterogeneous, and their frequency varies widely, even within the same region (Fig. 2 and Table S4). The allele frequencies for hemoglobin S (HbS) ranged from 0 to 13% in the *quilombos* assessed, with the highest number found in the *quilombo* Riacho de Sacutiába (Northeast). This variation is in agreement with the one observed in African populations (Piel et al. 2013b). The allele for hemoglobin C (HbC) was present in 18 out of 51 populations analyzed and its frequency ranged from 0.20% (Queimada Nova, Amapá, and Paulistana, NE) to 12.5% (Lajeado, N). Only 3 *quilombos* (from the Northern region) did not present the S and C hemoglobin mutations.

Due to the African origin of these variants, it should be reasonable to expect that their frequency increases with African ancestry. In the *quilombos* assessed for the HbC and HbS variants frequencies, the average of African ancestry was 51.6 ± 16.2 (Table S2 and S4). However, the frequency of HbC and HbS variants did not show a significant correlation with the percentage of African ancestry (Pearson's correlation p -value > 0.05 , Figure S1 a, b, and c). This lack of correlation could be explained not only by the founder effect in the formation of *quilombos*, but also by the variable frequency of these variants in the African populations that contributed to the Brazilian gene pool (Piel et al. 2013b; Ngo Bitoungui et al. 2015). Furthermore, we observed that the frequency of the HbC variant showed significant variation among *quilombos* from different geographic regions, with a higher frequency of HbC in the North and Northeast when compared to the Southeastern region (Fisher's exact test p -value < 0.05 and Table S4). This result agrees with the historical records showing that the greatest contribution

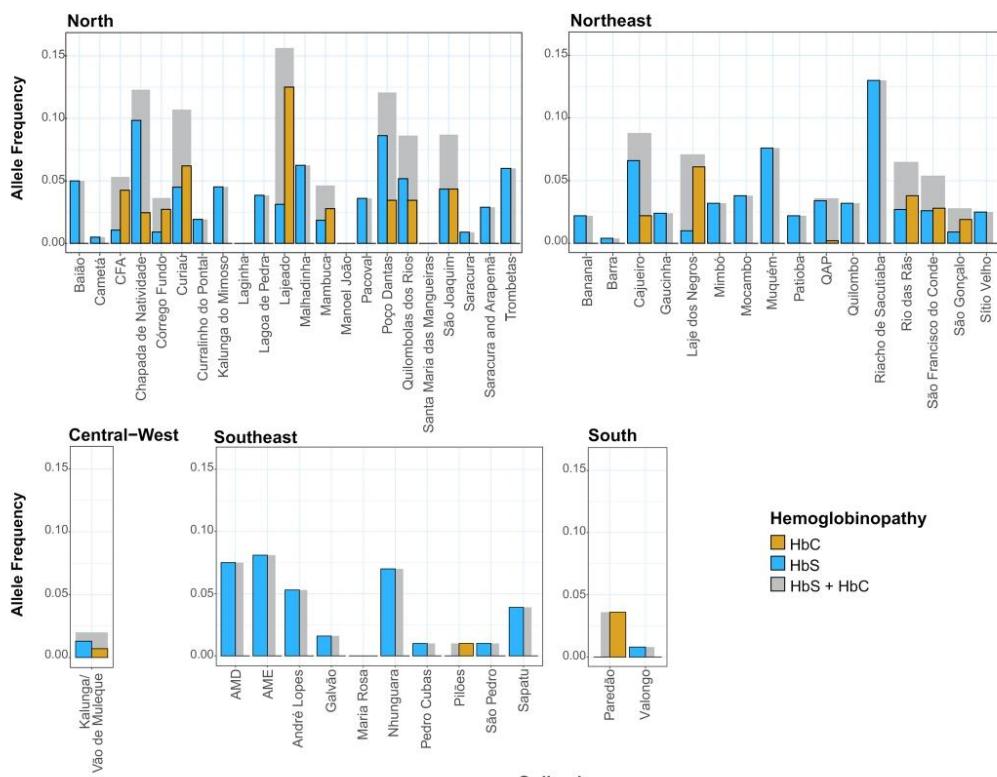


Fig. 2 HbS and HbC allele frequencies and distribution in Brazilian *quilombos*. CFA, Carrapato, Formiga and Ambrósio; CN, Chapada de Natividade; CP, Curralinho do Pontal; KM, Kalunga do Mimoso; QR, Quilombolas dos Rios; SMM, Santa Maria das Mangueiras; SA, Saracura and Arapemá; QAP, Queimada Nova, Amarante, and Pau-listana; RS, Riacho de Sacutíabá; SFC, São Francisco do Conde; VAO, Vão de Muleque. AMD, Abobral Margem Direita; AME, Abobral Margem Esquerda (Schneider et al. 1987; Bortolini et al. 1992, 1998; Arpini-Sampaio et al. 1999; Guerreiro et al. 1999; Oliveira et al. 2002; De Mello Auricchio et al. 2007; Cardoso et al. 2012; de Souza et al. 2013; De Assis et al. 2015; Dantas et al. 2016; Santiago et al. 2017; Teles et al. 2017)

from Western Africans, where this variant predominates, was for the North and Northeast regions of Brazil (Piel et al. 2013b).

On the other hand, α - and β -thalassemias have been studied only in the *quilombo* of Saracura (N) where it was reported a frequency of 9.5% for the 3.7-kb deletion for α -thalassemia ($-\alpha^{3.7}$) and 1.3% of the Mediterranean deletion. β -thalassemias were detected at 8.9%, in a spectrum of seven different mutations, being only 2.1% African, and the remaining forms of Mediterranean origin (Cardoso et al. 2012). The frequency of $-\alpha^{3.7}$ is higher than the 3.3% frequency found in the urban area of Santarém, PA (Cardoso et al. 2012), where the *quilombo* is located; however, it is similar to that observed in African-Brazilians from other

urban Brazilian regions (Adorno et al. 2008; Wagner et al. 2010). It is remarkable, however, that the highest proportion of thalassemias detected in this *quilombo* were most possibly introduced by European gene flow (Cardoso et al. 2012).

Besides hemoglobinopathies, another two red blood cell traits have been studied (despite that less comprehensively), in *quilombos* from the Brazilian Amazon. In the first place, the deficiency of the G6PD enzyme (G6PDd) and, secondarily, the Duffy antigen (Fy).

G6PDd is an X chromosome-linked enzymopathy determined by missense single nucleotide variants (SNVs) in the *G6PD* gene (Howes et al. 2013). Two known G6PD variants of African origin have clinical importance: *G6PD*A* (determined by the SNP rs1050829, a c.376 T > C mutation), and

*G6PD**A- (determined by rs1050829 and a second mutation: rs1050828 (c.202C>T), rs137852328 (c.680G>T), or rs76723693 (c.968 T>C)). Besides being associated with protection against malaria infection, G6PDd is of great public health importance in malaria-endemic regions, such as the Brazilian Amazon, for causing hemolytic crises induced by antimalarial drugs—such as primaquine—in G6PD deficient individuals. Besides that, other anti-malarial drugs, such as chloroquine and hydroxychloroquine, have been widely promoted by Brazilian official authorities and health professionals for treatment and prevention of coronavirus disease (COVID-19) (even with existent studies reporting on the ineffectiveness and adverse events of these drugs) further compromising the health of G6PDd individuals (Brito-Sousa et al. 2019; Kuipers et al. 2020; Melo et al. 2021; Singh et al. 2021; da Rocha et al. 2021). As screening G6PD activity is not mandatory before treatment with amquinolines in Brazil, monitoring the prevalence of G6PDd is crucial, especially in vulnerable populations with a higher predisposition to G6PDd, such as African-Brazilians. Notwithstanding its relevance, only one study has evaluated G6PDd prevalence in Brazilian *quilombos* so far. Oliveira et al., (2018) evaluated *G6PD**A- and *G6PD**A variants in nine *quilombos* from the Amazon region (the *quilombos* Arancuan, Tapagem, Abuí, and Cachoeira Porteira, along the Trombetas River, and Serrinha, Jauari, Araçá de Fora, Jarauacá and Boa Vista do Cuminá, along the Erepecuru and Cuminá Rivers). The allele frequencies reported for *G6PD**A- and *G6PD**A were 6.1% and 10.4%, respectively, substantially lower than those found in Sub-Saharan African populations (15–20%) (Oliveira et al. 2018). In these populations, 1.9% of males and 1.5 of females were G6PD deficient (hemizygous and homozygous, respectively).

The *DARC* gene encodes a membrane-bound chemokine receptor necessary for the infection of red blood cells by *P. vivax* and *P. knowlesi*. It has two codominant alleles, *FY**A and *FY**B, determining three serological phenotypes: Fy(a-b+), Fy(a+b-), and Fy(a+b+). A fourth phenotype, the Duffy-negative blood group (Fy(a-b-)), is determined by the disruption of its erythroid expression by a substitution in the gene's promoter (*FY**B^{ES}) (Howes et al. 2011). As it generates resistance against infection, the Duffy-negative blood group is fixed or near fixation in most Sub-Saharan African populations and other malaria-endemic regions, being rare outside these regions (Howes et al. 2011). In the *quilombos* around the Trombetas, Erepecuru, and Cuminá rivers, 20.4% of the individuals were Duffy negative, whereas 41.3% were heterozygous for the *FY**B^{ES} allele (Oliveira et al. 2018). The *FY**B^{ES} allele was the most frequent in these communities (41%) (Oliveira et al. 2018). On the other hand, the *quilombos* of Curiaú and Mazagão Velho (Amapá, N), and Pitimandeua (Pará, N) were studied in a different work, demonstrating also that the *FY**B^{ES} allele

was the most prevalent (Perna et al. 2007). The frequencies of the Duffy-negative blood group in these communities were 32.3%, 50%, and 58% in Mazagão Velho, Pitimandeua, and Curiaú, respectively (Perna et al. 2007). The lowest frequency of the Duffy-negative blood group in Mazagão Velho is consistent with the history of this *quilombo* and its foundation by descendants of enslaved North Africans from Morocco, a region where Duffy-negative group frequency is lower than in Sub-Saharan African populations (> 50%) (Fernández-Santander et al. 1999).

Despite that Hemoglobinopathies have been the most studied traits in *quilombos* it is important to highlight that 83% of the communities studied for these traits were from the North and Northeast regions. Also, the prevalence of G6PDd and the different Duffy antigens in *quilombos* from regions outside of the Amazonian regions remain to be elucidated.

African-specific variants occur in *quilombos*

Despite that African-derived populations are expected to harbor higher genetic diversity, studies of genetic variability not linked to known African adaptations in Brazilian *quilombos* are scarce. During this research, we have found only two studies that discovered new variants in *quilombos*. One of them evaluated the haplotypic profile of *HLA-A*, *HLA-B*, *HLA-C*, and *HLA-DRB1* in a sample of 144 individuals from *quilombos* of the Ribeira Valley, in the Brazilian Southeast. The researchers detected a new null *HLA-C* allele shared by three individuals. The new variant was found to occur in a haplotype common in populations of African descent and was not detected in African-Brazilians from nearby urban centers (Nunes et al. 2016). On the other side, a rare and possibly African allele from the FRAXAC1 locus has been detected in *quilombos* from the Brazilian Northeast (Mingroni-Netto et al. 2002). Locus FRAXAC1 flanks the fragile X mental-retardation-1 gene (FMR-1) and its alleles have been correlated with the copy number variation that causes X fragile syndrome (Mingroni-Netto et al. 2002).

The genetics of health-related traits in *quilombos* and the implications for the health of populations of African descent

Black Brazilians have lower socio-economical and health indices when compared to White Brazilians (IBGE). This higher vulnerability worsens the outcome of both infectious-parasitic as well as chronic non-transmissible diseases. The prevalence of chronic non-transmissible diseases has increased due in part to the epidemiologic transition taking place in Brazil, which has led to increased consumption of industrialized food and an increase in population densities (Silva et al. 2016; Nunes et al. 2020). In particular,

the restrictions on hunting, fishing, and the reduction of their lands by predatory entrepreneurship have been forcing changes in the lifestyle of the *quilombolas*, who were traditionally involved in subsistence agriculture (Shiraishi Neto 2007; Nunes et al. 2016; Silva et al. 2016). It is worth noting that a higher level of urbanization, linked to a shift to a capitalist economy, is related to more sedentary lifestyles, higher stress, and a decrease in dietary quality, associated with increased chronic disease risk (Wells 2012; Ghimire et al. 2022; Kidokoro et al. 2022). Brazilian *quilombos* are not exempt from this pattern, with more urbanized *quilombos* showing higher chronic disease-related risk factors than their urban counterparts (Paiva et al. 2022).

Despite that some public policies have been promoted to reduce the mentioned health disparities, they have not been efficiently established yet (Voss kominek and Vanali 2018). In part, that responds to the structural racism present in Brazil, which makes the living conditions and access to health services difficult for the Black population (Voss kominek and Vanali 2018), but also there is a scarcity of studies aimed to improve the knowledge of the ancestry specific parameters and genetic bases of diseases, that are crucial for the proper health care of this population. Black Brazilians present a higher prevalence of diabetes, hypertension, and related conditions (such as obesity and cardiovascular disease) than Whites (Ministério da Saúde 2001).

Some *quilombos* are semi-isolated and, for that reason, represent optimal populations for the study of the genetics of complex traits. They present relatively higher inbreeding and founder effect, which can empower genetic association studies by drifting up the frequency of rare variants (Zeggini 2014).

So far, a few studies have assessed the genetics of multifactorial diseases in *quilombos* (Table S6). Angeli et al., (2011) studied the association between multiple candidate loci and obesity in the *quilombos* from the Ribeira Valley (SE). They detected a significant interaction between *LEPR* (Arg16Gly) and *ADRB2* polymorphisms (Gln223Arg), suggesting a role of the interplay between the Leptin-signaling and the catecholaminergic pathways for the susceptibility to obesity-related traits (Angeli et al. 2011).

Regarding hypertension, Kimura and collaborators detected a putative multilocus effect of the *GNB3* variants, rs5441, and rs5443, on diastolic blood pressure levels (DBP) in a family-based study in the *quilombos* from the Ribeira Valley (SE). The same authors also showed an interaction between *NOS3-GRK4* variants rs1799983 and rs1801058, respectively, in an unrelated case-control study in the same populations, concerning DBP (Kimura et al. 2012). On the other hand, Neto et al., (2021) also detected an association between the *NOS3* SNP rs1799983 and the *IGFBP3* SNP rs11977526 with arterial hypertension in a sample constituted only by women from *quilombos* of Alagoas (NE) (Neto

et al. 2021). The authors also highlight the need to test for multilocus effects in the study of complex traits, as the individual polymorphisms did not show significant associations (Yeh et al. 2008; Angeli et al. 2011; Kimura et al. 2012).

On the other hand, infectious diseases like acquired immunodeficiency syndrome (AIDS) are more prevalent and mortal for Black Brazilians (Brazilian Ministry of Health 2017). Although social factors are pointed out as the major determining factor of susceptibility, genetic factors have also been studied. A 32 bp deletion in the *CCR5* gene, which codes for the β-chemokine receptor 5, results in a truncated protein associated with resistance to HIV-1 infection (Samson et al. 1996). The mutated allele (*CCR5D32*) is considered an indicator of European ancestry, as it is more frequent in that continent, being virtually absent in Asian, African, and Amerindian populations. The presence of the *CCR5D32* allele has been studied in the *quilombos* Mocambo, Rio das Rãs, and São Gonçalo (NE). It was detected exclusively in the heterozygous state, both in native and immigrant inhabitants of those communities. Its presence has been interpreted as introduced by European gene flow, possibly due to recent admixture but also to the founder effect (Carvalho et al. 2004). The last is potentially the case of the *quilombo* Mocambo, which showed the highest frequency of the *CCR5D32* allele (5.6%) (Carvalho et al. 2004).

Final considerations and perspectives

A total of 57 articles addressing genetics in Brazilian *quilombos* were analyzed to describe the ancestry and genetics of health-related phenotypes in these populations. Only approximately 93 out of 5972 extant African-derived communities have been included in genetic studies so far (Table S5). We also observed a strong bias in the regional distribution of the *quilombos* studied throughout Brazil, with the great majority of them concentrated in populations from the Northern and Northeastern regions (> 80%), and underrepresentation of Southern, Southeastern, and Central Western *quilombos*.

The variability in the African, Amerindian, and European admixture proportions observed in Brazilian *quilombos* emphasizes the well-known singularity of their history and the great influence of the surrounding populations in their formation, stressed by the effect of genetic drift on these historically small populations. Some examples of these particular histories are the strong effect of the segregation caused by Euro-descendant populations to Sertão do Valongo or the high Amerindian genetic contribution (predominantly maternal) to Amazonian *quilombos*. The uniqueness of each *quilombo* certainly contributed to the absence of a pattern of ancestry referring to the geographic region in which the community is located, in contrast with Brazilian urban populations. The uniparental ancestry patterns, on the other hand,

in general, reflects sex-biased gene flow and are consistent with the historically recorded sexual exploitation of African and Amerindian women by European men. It is remarkable the scarcity of subcontinental ancestry studies—only one was developed until the publication of this work. Even so, the study of different molecular markers allowed the demonstration of the heterogeneous within-Africa roots of *quilombolas*, with high contribution from Central-Western African Bantus (from Angola and Kongo), and Western Africans (from Senegambia and the Bight of Benin) but also from more remote regions such as Morocco, in Northern Africa. Altogether, the higher prevalence of HbC and the HbS Senegal and Benin haplotypes in Northern and Northeastern *quilombos* agrees with historical estimations of a higher number of Western Africans brought to these regions through the transatlantic slave trade.

The study of malaria adaptive variants is of public health concern in Brazil, a country where *Plasmodium vivax* is prevalent and where the use of anti-malarial drugs can cause hemolytic crises in G6PDd individuals. This review reflects a paucity of studies of red blood cell traits and malaria adaptive variants in *quilombos* from the Central-Western, Southeastern, and Southern regions of Brazil and reinforces the need to assess their prevalence.

Despite the scarcity of studies detected here, we reinforce the potential of the study of *quilombos* for the detection of African-specific variants and for a better understanding of the genetic basis of complex diseases such as diabetes, hypertension, and obesity that affect, to a greater extent, populations of African descent.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10142-023-00999-0>.

Author contribution The idea for the article belongs to IAJL. IAJL, NMS, IDOB, and JGT performed the literature search, data analysis, and drafted the work. CVGL, AALB, and MHB critically revised the work.

Funding We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PROAP—Finance Code 001) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support and for the scholarship provided to Iriel Araceli Joerin-Luque and Natalie Mary Sukow.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

- Abe-Sandes K, Silva WA, Zago MA (2004) Heterogeneity of the Y chromosome in Afro-Brazilian populations. *Hum Biol* 76:77–86. <https://doi.org/10.1353/hub.2004.0014>
- Adorno EV, Zanette Á, Lyra I et al (2008) Clinical and molecular characteristics of sickle cell anemia in the northeast of Brazil. *Genet Mol Biol* 31:621–625. [https://doi.org/10.1590/S1415-4757200800040003](https://doi.org/10.1590/S1415-47572008000400003)
- Aidoo L (2018) Slavery unseen: sex, power, and violence in Brazilian history. Duke University Press, Durham
- Amorim CEG, Carvalho GONTIJO, Falcão-Alencar G et al (2011) Migration in Afro-Brazilian rural communities: crossing demographic and genetic data. *Hum Biol* 83:509–521. <https://doi.org/10.3378/027.083.0405>
- Angeli CB, Kimura L, Auricchio MT et al (2011) Multilocus analyses of seven candidate genes suggest interacting pathways for obesity-related traits in Brazilian populations. *Obesity* 19:1244–1251. <https://doi.org/10.1038/oby.2010.325>
- Arpini-Sampaio Z, Costa MCB, Melo AA et al (1999) Genetic polymorphisms and ethnic admixture in African-derived black communities of northeastern Brazil. *Hum Biol* 71:69–85
- Barbosa AAL, Sousa SMB, Abé-Sandes K et al (2006) Microsatellite studies on an isolated population of African descent in the Brazilian state of Bahia. *Genet Mol Biol* 29:23–30. <https://doi.org/10.1590/S1415-47572006000100004>
- Bortolini MC, Weimer TA, Franco MH et al (1992) Genetic studies in three South American black populations. *Gene Geogr* 6:1–16
- Bortolini MC, de Weimer TA, Salzano FM et al (1995) Evolutionary relationships between black South American and African populations. *Hum Biol* 67:547–559
- Bortolini MC, Zago MA, Salzano FM et al (1997) Evolutionary and anthropological implications of mitochondrial DNA variation in African Brazilian populations. *Hum Biol* 69:141–159
- Bortolini MC, Silva-Junior TAW, Zago MA et al (1998) Protein and hypervariable tandem repeat diversity in eight African-derived South American populations: Inferred relationships do not coincide. *Hum Biol* 70:443
- Bortolini MC, Da Silva WA, De Guerra DC et al (1999) African-derived South American populations: a history of symmetrical and asymmetrical matings according to sex revealed by bin and uni-parental genetic markers. *Am J Hum Biol* 11:551–563. [https://doi.org/10.1002/\(SICI\)1520-6300\(1999\)11:4%3c551::AID-AJHB15%3e3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1520-6300(1999)11:4%3c551::AID-AJHB15%3e3.0.CO;2-Z)
- Brazilian Ministry of Health (2017) Comprehensive health care national policy of Negro population: an unified health system policy (SUS – Brazil). 44
- Brito-Sousa JD, Santos TC, Avalos S et al (2019) Clinical spectrum of primaquine-induced hemolysis in glucose-6-phosphate dehydrogenase deficiency: a 9-year hospitalization-based study from the Brazilian Amazon. *Clin Infect Dis* 69:1440–1442. <https://doi.org/10.1093/cid/ciz122>
- Cardoso GL, Takanashi SYL, Guerreiro JF (2012) Inherited hemoglobin disorders in an Afro-Amazonian community: saracura. *Genet Mol Biol* 35:553–556. <https://doi.org/10.1590/S1415-47572012005000041>
- Carvalho MWP, Leboute APM, Oliveira SF et al (2004) CCR5D32 mutation in three Brazilian populations of predominantly Sub-Saharan African ancestry. *Genet Mol Biol* 27:321–325. <https://doi.org/10.1590/S1415-47572004000300002>
- Carvalho BM, Bortolini MC, dos Santos SEB, Ribeiro-dos-Santos ÁKC (2008) Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations. *Genet Mol Biol* 31:12–22. <https://doi.org/10.1590/S1415-47572008000100002>

- Cayres Vallinoto IMV, Rosário Vallinoto AC, Duarte Valente CM, Farias Guerreiro J (2003) Allele frequency distributions of six hypervariable loci (D1S80, APOB, D4S43, vW1, F13A, and DYS19) in two African-Brazilian communities from the Amazon region. *Genet Mol Biol* 26:235–240. <https://doi.org/10.1590/s1415-47572003000300003>
- Cruz PRS, Ananina G, Gil-da-Silva-Lopes VL et al (2019) Genetic comparison of sickle cell anaemia cohorts from Brazil and the United States reveals high levels of divergence. *Sci Rep* 9:1–9. <https://doi.org/10.1038/s41598-019-47313-2>
- da Rocha JEB, Othman H, Tiemessen CT et al (2021) G6PD distribution in sub-Saharan Africa and potential risks of using chloroquine/hydroxychloroquine based treatments for COVID-19. *Pharmacogenomics J*. <https://doi.org/10.1038/s41397-021-00242-8>
- Da Silva WA, Bortolini MC, Meyer D et al (1999) Genetic diversity of two African and sixteen South American populations determined on the basis of six hypervariable loci. *Am J Phys Anthropol* 109:425–437. [https://doi.org/10.1002/\(SICI\)1096-8644\(199904\)109:4%3c425::AID-AJPA1%3e3.0.CO;2-D](https://doi.org/10.1002/(SICI)1096-8644(199904)109:4%3c425::AID-AJPA1%3e3.0.CO;2-D)
- Dantas CR, Silva RAA da, Group S et al (2016) Sickle cell anemia in the Boqueirão Kilombo community Bahia Brazil. *Scientia Plena* 12:1–10. <https://doi.org/10.14808/sciplena.2016.091002>
- de Andrade LMM (1995) Os Quilombos da Bacia do Rio Trombetas: Breve Histórico. *Rev Antropol* 38:79–99
- de Knijff P (2000) Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *Am J Human Gen* 67:1055–1061. [https://doi.org/10.1016/S0002-9297\(07\)62935-8](https://doi.org/10.1016/S0002-9297(07)62935-8)
- de Souza IR, Culpí L (2005) Valongo, genetic studies on an isolated Afro-Brazilian community. *Genet Mol Biol* 28:402–406
- de Cruz JO, Lima PSP, Sousa SMB (2020) Population structure of a remnant quilombo community: an analysis of the inbreeding and ethnic admixture. *Oalib* 07:1–10. <https://doi.org/10.4236/oalib.1106207>
- de Oliveira HSS, da Silva ANLM, Andrade GB et al (2018) Molecular genotyping of G6PD mutations and Duffy blood group in Afro-descendant communities from Brazilian Amazon. *Genet Mol Biol* 41:758–765. <https://doi.org/10.1590/1678-4685-gmb-2017-0253>
- de Souza LO, Teles AF, de Oliveira RJ et al (2013) Triagem das hemoglobinas S e C e a influência das condições sociais na sua distribuição: um estudo em quatro comunidades quilombolas do Estado do Tocantins. *Saúde e Sociedade* 22:1236–1246. <https://doi.org/10.1590/S0104-12902013000400024>
- de Souza AM, Resende SS, de Sousa TN, de Brito CFA (2019) A systematic scoping review of the genetic ancestry of the Brazilian population. *Genet Mol Biol* 42:495–508. <https://doi.org/10.1590/1678-4685-gmb-2018-0076>
- De Assis ES, Araújo JA, De Rezende MIRC et al (2015) Prevalence of variant hemoglobins and thalassemias in a maroon community in Sergipe, Brazil. *Acta Scientiarum - Health Sciences* 37:211–216. <https://doi.org/10.4025/actascihealthsci.v37i2.23444>
- De Mello Auricchio MTB, Vicente JP, Meyer D, Mingroni Netto RC (2007) Frequency and origins of hemoglobin S mutation in African-derived Brazilian populations. *Hum Biol* 79:667–677. <https://doi.org/10.1353/hub.2008.0017>
- De Oliveira SF, Dos Santos EB, De Souza Mendonça PJ et al (2001) Group-specific component (GC) in Curiaú and Pacoval, two African-derived Brazilian populations. *Am J Hum Biol* 13:718–720. <https://doi.org/10.1002/ajhb.1116>
- Fernández-Santander A, Kandil M, Luna F et al (1999) Genetic relationships between southeastern Spain and Morocco: new data on ABO, RH, MNS, and DUFFY polymorphisms. *Am J Hum Biol* 11:745–752. [https://doi.org/10.1002/\(SICI\)1520-6300\(199911/12\)11:6%3c745::AID-AJHB4%3e3.0.CO;2-W](https://doi.org/10.1002/(SICI)1520-6300(199911/12)11:6%3c745::AID-AJHB4%3e3.0.CO;2-W)
- Ghimire U, Yasmine S, Chand S et al (2022) Cardiovascular disease risk factors distribution and clustering across different geographic levels in Nepal. *Am J Hum Biol* 34:1–11. <https://doi.org/10.1002/ajhb.23787>
- Gomes FDS (2015) Mocambos e Quilombos: Uma história do campesinato negro no Brasil, 1st edn. Claro Enigma, São Paulo
- Gomez F, Ko WY, Davis A, Tishkoff SA (2013) Impact of natural selection due to malarial disease on human genetic variation. In: Brinkworth J, Pechenkina E (eds) *Primates, pathogens and evolution*. Springer, New York, pp 117–160
- Gontijo CC, Guerra Amorim CE, Godinho NMO et al (2014) Brazilian quilombos: a repository of Amerindian alleles. *Am J Hum Biol* 26:142–150. <https://doi.org/10.1002/ajhb.22501>
- Gontijo CC, Mendes FM, Santos CA et al (2018) Ancestry analysis in rural Brazilian populations of African descent. *Forensic Sci Int Genet* 36:160–166. <https://doi.org/10.1016/j.fsigen.2018.06.018>
- Grossman SR, Andersen KG, Shlyakhter I et al (2013) Identifying recent adaptations in large-scale genomic data. *Cell* 152:703–713. <https://doi.org/10.1016/j.cell.2013.01.035>
- Guerreiro JF, Ribeiro-dos-Santos ÁKC, dos Santos EJM et al (1999) Genetical-demographic data from two amazonian populations composed of descendants of african slaves: Pacoval and Curiaú. *Genet Mol Biol* 22:163–167. <https://doi.org/10.1590/S1415-4757199900200004>
- Guimarães Paiva S, Carvalho Gontijo C, Carvalho Gontijo R, et al (2020) Migration in Brazilian Afro-descendants communities: a new approach to illustrate the meaning of contemporary quilombo. *Revista da Associação Brasileira de Pesquisadores(as) Negros(as) – ABPN* 12:188–208. <https://doi.org/10.31418/2177-2770.2020.v12.n.32.p188-208>
- Hall GM (2005) Slavery and African ethnicities in the Americas: restoring the links. The University of North Carolina Press, North Carolina
- Howes RE, Patil AP, Piel FB, et al (2011) The global distribution of the Duffy blood group. *Nature Communications* 2: <https://doi.org/10.1038/ncomms1265>
- Howes RE, Battle KE, Satyagraha AW, et al (2013) G6PD deficiency. In: *Advances in parasitology*. Elsevier 133–201
- Hünemeier T, Carvalho C, Marrero AR et al (2007) Niger-Congo speaking populations and the formation of the Brazilian gene pool: mtDNA and Y-chromosome data. *Am J Phys Anthropol* 133:854–867. <https://doi.org/10.1002/ajpa.20604>
- Kidokoro T, Tian X, Fukui N et al (2022) Segmented physical activity and sedentary behavior patterns among children in Maasai village and Nairobi city. *Am J Hum Biol* 34:1–15. <https://doi.org/10.1002/ajhb.23649>
- Kimura L, Angeli CB, Auricchio MTBM et al (2012) Multilocus family-based association analysis of seven candidate polymorphisms with essential hypertension in an african-derived semi-isolated Brazilian population. *Int J Hypertens* 2012:1–8. <https://doi.org/10.1155/2012/859219>
- Kimura L, Ribeiro-Rodrigues EM, De Mello Auricchio MTB et al (2013) Genomic ancestry of rural African-derived populations from Southeastern Brazil. *Am J Hum Biol* 25:35–41. <https://doi.org/10.1002/ajhb.22335>
- Kimura L, Nunes K, Macedo-Souza LI et al (2017) Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes. *Am J Hum Biol* 29:1–11. <https://doi.org/10.1002/ajhb.22930>
- Knight A, Underhill PA, Mortensen HM et al (2003) African Y Chromosome and mtDNA divergence provides insight into the history of click languages. *Curr Biol* 13:464–473. [https://doi.org/10.1016/S0960-9822\(03\)00130-1](https://doi.org/10.1016/S0960-9822(03)00130-1)
- Korunes KL, Soares-Souza GB, Bobrek K, et al (2020) Sex-biased admixture and assortative mating shape genetic variation and influence demographic inference in admixed Cabo Verdeans. *bioRxiv* 0–3. <https://doi.org/10.1101/2020.12.14.422766>

- Kuijpers MT, Zwieten R, Heijmans J et al (2020) Glucose-6-phosphate dehydrogenase deficiency-associated hemolysis and methemoglobinemia in a COVID-19 patient treated with chloroquine. *Am J Hematol* 95:E194–E196. <https://doi.org/10.1002/ajh.25862>
- Lacombe AJ, Silva E, de Barbosa FA (1988) Rui Barbosa e a queima dos arquivos, 1st edn. Fundação Casa de Rui Barbosa, Rio de Janeiro
- Leite IB (1999) Quilombos e quilombolas: cidadania ou folclorização? *Horizontes Antropológicos* 5:123–149. <https://doi.org/10.1590/s1044-7183199900010006>
- Lemes RB, Nunes K, Meyer D et al (2014) Estimation of inbreeding and substructure levels in African-derived Brazilian quilombo populations. *Human Biology* 86:276. <https://doi.org/10.13110/humanbiology.86.4.0276>
- Maciel LGL, Rodrigues EM, dos Santos NPC et al (2011) Afro-derived Amazonian populations: inferring continental ancestry and population substructure. *Hum Biol* 83:627–636. <https://doi.org/10.3378/027.083.0504>
- Martínez B, Simão F, Gomes V et al (2020) Searching for the roots of the first free African American community. *Sci Rep* 10:20634. <https://doi.org/10.1038/s41598-020-77608-8>
- Matos WSD, Eugenio BG (2019) Comunidades quilombolas elementos conceituais para sua compreensão. PRACS: Revista Eletrônica de Humanidades do Curso de Ciências Sociais da UNIFAP 11:141. <https://doi.org/10.18468/pracs.2018v1n2.p141-153>
- Melo JRR, Duarte EC, de Moraes MV, et al (2021) Automedicação e uso indiscriminado de medicamentos durante a pandemia da COVID-19. *Cadernos de Saúde Pública* 37. <https://doi.org/10.1590/0102-311x00053221>
- Mingroni-Netto RC, Angeli CB, Auricchio MTBM et al (2002) Distribution of CGG repeats and FRAXAC1/DXSS548 alleles in South American populations. *Am J Med Genet* 111:243–252. <https://doi.org/10.1002/ajmg.10572>
- Ministério da Saúde (2001) Manual de Doenças Mais Importantes, por Razões Étnicas, na População Brasileira Afro-Descendente, 1st edn. Ministério da Saúde, Brasília
- Moura C (1992) História do Negro Brasileiro, 2nd edn. Ática, São Paulo
- Moura C (1993) Quilombos: Resistência ao escravismo, 3rd edn. Ática, São Paulo
- Myscofski CA (2013) Amazons, wives, nuns, and witches: women and the Catholic church in colonial Brazil, 1500–1822, 1st edn. University of Texas Press, Austin
- Nascimento A (2016) O genocídio do Negro brasileiro: Processo de um racismo mascarado, 2nd edn. Perspectiva, São Paulo
- Neto ABL, Vasconcelos NBR, dos Santos TR et al (2021) Prevalence of IGFBP3, NOS3 and TCF7L2 polymorphisms and their association with hypertension: a population-based study with Brazilian women of African descent. *BMC Res Notes* 14:186. <https://doi.org/10.1186/s13104-021-05598-5>
- Ngo Bitoungui VJ, Pule GD, Hanchard N et al (2015) Beta-globin gene haplotypes among Cameroonian and review of the global distribution: is there a case for a single sickle mutation origin in Africa? *OMICS* 19:171–179. <https://doi.org/10.1089/omi.2014.0134>
- Nunes K, Piovezan B, Torres MA et al (2016) Population variation of HLA genes in rural communities in Brazil, the Quilombos from the Vale do Ribeira, São Paulo – Brazil. *Hum Immunol* 77:447–448. <https://doi.org/10.1016/j.humimm.2016.04.007>
- Nunes K, Kimura L, Gontijo CC et al (2020) Estudos de dinâmica populacional, ancestralidade genética e saúde em comunidades quilombolas Relato de uma experiência. *Tessituras Revista de Antropologia e Arqueologia* 8:219–251. <https://doi.org/10.15210/tes.v8i2.18396>
- Nussbaum RL, McInnes RR, Willard HF (2016) Thompson and Thompson Genética Médica, 8th edn. Elsevier, Rio de Janeiro
- Oliveira SF, Pedrosa MAF, Sousa SMB et al (2002) Heterogeneous distribution of HbS and HbC alleles in Afro-derived Brazilian populations. *Int J Hum Genet* 2:153–159. <https://doi.org/10.1080/09723757.2002.11885799>
- Ongaro L, Scliai MO, Flores R et al (2019) The genomic impact of European colonization of the Americas. *Curr Biol* 29:3974–3986.e4. <https://doi.org/10.1016/j.cub.2019.09.076>
- Paiva SG, Rivara AC, de Castro Nóbrega M, et al (2022) Cardiovascular risk factors across different levels of urbanization in Brazilian Afro-derived communities (quilombos). *American Journal of Human Biology* 1–19. <https://doi.org/10.1002/ajhb.23839>
- Palha TDJB, Ribeiro-Rodrigues EM, Ribeiro-dos-Santos Â et al (2011) Male ancestry structure and interethnic admixture in African-descent communities from the Amazon as revealed by Y-chromosome STRs. *Am J Phys Anthropol* 144:471–478. <https://doi.org/10.1002/ajpa.21436>
- Pante De-Sousa G, De Cassia M-R, Melo Dos Santos EJ, Guerreiro JF (1999) β-globin haplotypes analysis in Afro-Brazilians from the Amazon region: evidence for a significant gene flow from Atlantic West Africa. *Ann Hum Biol* 26:365–373. <https://doi.org/10.1080/030144699282679>
- Peña SDJJ, Di Pietro G, Fuchshuber-Moraes M et al (2011) The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS ONE* 6:e17063. <https://doi.org/10.1371/journal.pone.0017063>
- Peprah E, Xu H, Tekola-Ayele F, Royal CD (2015) Genome-wide association studies in Africans and African Americans: expanding the framework of the genomics of human traits and disease. *Public Health Genomics* 18:40–51. <https://doi.org/10.1159/000367962>
- Pereira L, Mutesa L, Tindana P, Ramsay M (2021) African genetic diversity and adaptation inform a precision medicine agenda. *Nat Rev Genet* 22:284–306. <https://doi.org/10.1038/s41576-020-00306-8>
- Perna SJQ, Cardoso GL, Guerreiro JF (2007) Duffy blood group genotypes among African-Brazilian communities of the Amazon region. *Gen Mol Res: GMR* 6:166–172
- Piel FB, Howes RE, Patil AP et al (2013a) The distribution of haemoglobin C and its prevalence in newborns in Africa. *Sci Rep* 3:1671. <https://doi.org/10.1038/srep01671>
- Piel FB, Patil AP, Howes RE et al (2013b) Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet* 381:142–151. [https://doi.org/10.1016/S0140-6736\(12\)61229-X](https://doi.org/10.1016/S0140-6736(12)61229-X)
- Pinheiro AA, Assumpção JE, Follmann JI (2015) África e afrodescendentes no sul do Brasil: história, religião e educação. Casa Leiria, São Leopoldo
- Prefeitura de São Paulo (2018) Informe Técnico sobre Malária. Cidade de São Paulo Comunicação. [https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/Informe%20Tecnico_Malaria\(1\).pdf](https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/Informe%20Tecnico_Malaria(1).pdf). Accessed 22 May 2022
- Esoh K, Wonkam A (2021) Evolutionary history of sickle-cell mutation: Implications for global genetic medicine. *Hum Mol Genet* 30:R119–R128. <https://doi.org/10.1093/hmg/ddab004>
- Relethford JH (2012) Human Population Genetics, 1st edn. Wiley-Blackwell, New Jersey
- Ribeiro GGBL, De Lima RR, Wiegel CEV et al (2009) Afro-derived Brazilian populations: male genetic constitution estimated by Y-chromosomes STRs and Alu YAP element polymorphisms. *Am J Hum Biol* 21:354–356. <https://doi.org/10.1002/ajhb.20875>
- Ribeiro GGBL, Abe-Sandes K, Barcelos RDSS et al (2011) Who were the male founders of rural Brazilian Afro-derived communities? A proposal based on three populations. *Ann Hum Biol* 38:237–240. <https://doi.org/10.3109/03014460.2010.500471>

- Ribeiro-Dos-Santos ÁKC, Pereira JM, Lobato MRF et al (2002) Dis-similarities in the process of formation of Curiaú, a semi-isolated Afro-Brazilian population of the Amazon region. *Am J Hum Biol* 14:440–447. <https://doi.org/10.1002/ajhb.10059>
- Salas A, Richards M, De la Fe T et al (2002) The making of the African mtDNA landscape. *Am J Human Genet* 71:1082–1111. <https://doi.org/10.1086/344348>
- Salas A, Richards M, Lareu MV et al (2004) The African Diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Human Genet* 74:454–465. <https://doi.org/10.1086/382194>
- Samson M, Libert F, Doranz BJ et al (1996) Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382:722–725. <https://doi.org/10.1038/382722a0>
- Santiago RP, Oliveira RM, Soares LF et al (2017) Hemoglobin variant profiles among Brazilian quilombola communities. *Hemoglobin* 41:83–88. <https://doi.org/10.1080/03630269.2017.1321014>
- Schneider H, Guerreiro JF, Santos SEB et al (1987) Isolate breakdown in Amazonia - the blacks of the Trombetas River. *Rev Brasil Genet* 10:565–574
- Scliar MO, Vaintraub MT, Vaintraub PMV, Fonseca CG (2009) Brief communication: admixture analysis with forensic microsatellites in Minas Gerais, Brazil: The ongoing evolution of the capital and of an African-derived community. *Am J Phys Anthropol* 139:591–595. <https://doi.org/10.1002/ajpa.21046>
- Shiraishi Neto J (Org) (2007) Direito dos povos e das comunidades tradicionais no Brasil: declarações, convenções internacionais e dispositivos jurídicos definidores de uma política nacional, 1st edn. UEA, Manaus
- Shriner D, Rotimi CN (2018) Whole-genome-sequence-based haplotypes reveal single origin of the sickle allele during the holocene wet phase. *Am J Hum Genet* 102:547–556. <https://doi.org/10.116/j.ajhg.2018.02.003>
- Silva WA, Bortolini MC, Schneider MPC et al (2006) mtDNA haplogroup analysis of Black Brazilian and Sub-Saharan populations: implications for the Atlantic slave trade. *Hum Biol* 78:29–41. <https://doi.org/10.1353/hub.2006.0028>
- Silva S, Klautau-guimarães MDN, Grisolia CK (2010) b-globin haplotypes in normal and hemoglobinopathic individuals from Recôncavo Baiano, State of Bahia, Brazil. *Genet Mol Biol* 33:411–417
- Silva HP, Padez C, Moura EAF, Filgueiras LA (2016) Obesity, hypertension, social determinants of health and the epidemiologic transition among traditional Amazonian populations. *Ann Hum Biol* 43:371–381. <https://doi.org/10.1080/03014460.2016.1197967>
- Singh B, Ryan H, Kredo T, et al (2021) Chloroquine or hydroxychloroquine for prevention and treatment of COVID-19. Cochrane Database of Systematic Reviews 2021: <https://doi.org/10.1002/14651858.CD013587.pub2>
- Tawfik SM, Elhosseiny AA, Galal AA et al (2023) Health inequity in genomic personalized medicine in underrepresented populations: a look at the current evidence. *Funct Integr Genomics* 23:54. <https://doi.org/10.1007/s10142-023-00979-4>
- Teixeira MAD, Xavier DFB (2018) Santo Antônio do Guaporé: direitos humanos, conflitos e resistência socioambiental. *Revista Direito e Práxis* 9:351–371. <https://doi.org/10.1590/2179-8966/2018/32718>
- Teles AF, da Silva LC, da Silva AC et al (2017) Hemoglobinas de origem africana em comunidades quilombolas do estado do Tocantins, Brasil. *Revista Pan-Amazonônica De Saúde* 8:39–46. <https://doi.org/10.5123/S2176-62232017000100006>
- Trovoadas MJ, Tavares L, Gusmão L et al (2007) Dissecting the genetic history of São Tomé e Príncipe: a new window from Y-chromosome biallelic markers. *Ann Hum Genet* 71:77–85. <https://doi.org/10.1111/j.1469-1809.2006.00309.x>
- Underhill PA, Kivisild T (2007) Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet* 41:539–564. <https://doi.org/10.1146/annurev.genet.41.110306.130407>
- Voss kominek AM, Vanali AC (2018) Roteiros temáticos da Diáspora: Caminhos para o enfrentamento ao racismo no Brasil, 1st edn. Editora Fi, Porto Alegre
- Wagner SC, de Castro SM, Gonzalez TP et al (2010) Prevalence of common α -thalassemia determinants in south Brazil: importance for the diagnosis of microcytic anemia. *Genet Mol Biol* 33:641–645. <https://doi.org/10.1590/S1415-47572010005000086>
- Walsh R (1830) Notices of Brazil in 1828 and 1829, vol I. Frederick Westley and A.H Davis, London
- Wells JCK (2012) Obesity as malnutrition: The role of capitalism in the obesity global epidemic. *Am J Hum Biol* 24:261–276. <https://doi.org/10.1002/ajhb.22253>
- Wiezel CEV, Luizão MR, Sousa SMB et al (2013) Y-Linked microsatellites in Amazonian Amerindians applied to ancestry estimates in Brazilian Afro-derived populations. *Am J Hum Biol* 25:313–317. <https://doi.org/10.1002/ajhb.22361>
- Yeh E, Kimura L, Errera FIV et al (2008) Association of polymorphisms at the ADIPOR1 regulatory region with type 2 diabetes and body mass index in a Brazilian population with European or African ancestry. *Braz J Med Biol Res* 41:468–472. <https://doi.org/10.1590/S0100-879X2008000600005>
- Zeggini E (2014) Using genetically isolated populations to understand the genomic basis of disease. *Genome Medicine* 6:83. <https://doi.org/10.1186/s13073-014-0083-5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

6.3 CAPÍTULO III: New clues on the formation of Afro-diasporic communities (*quilombos*) from Southern Brazil reveal African roots and a high degree of Amerindian admixture (in preparation)

Iriel A. Joerin Luque^{1,2*}, Isadora Baldon Blaczyk², Priscila Ianzen dos Santos^{2,3}, Ana Cecília Guimarães Alves^{1,2}, Natalie Mary Sukow^{1,2}, Ana Carolina Malanczyn de Oliveira², Angela Rodrigues do Amaral Bispo², Aymee Fernanda Gros², Maria Letícia Saatkamp², Victor Dobis Barros², Joana Gehlen Tessaro², Maria Eduarda da Silveira Costa², Luana Leonardo Garcia², Isabela Dall Ogglio Bucco², Denise Raquel de Moura Bones^{1,2}, Sarah Elisabeth Cupertino², Letícia Boslooper Gonçalves^{1,2}, Adriana Inês de Paula⁴, Claudemira Vieira Gusmão Lopes⁵, Marcia Holsbach Beltrame^{1,2}

¹Programa de Pós-graduação em Genética, Departamento de Genética, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ²Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ³Programa de Pós-graduação em Medicina Interna, Hospital de Clínicas (HC), Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ⁴Departamento de Educação Física, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ⁵Setor Litoral, Universidade Federal do Paraná (UFPR), Matinhos, Paraná, Brasil.

*Correspondence E-mail: irieljoerin@gmail.com

Abstract

In Brazil, African-derived communities (*quilombos*) were founded by ex-enslaved people during and shortly after the abolition of slavery. The state of Paraná, in Southern Brazil, concentrates 86 *quilombos*, but little is known about their history or the African origins of their founders. Here we analyzed 9 Y chromosome single nucleotide polymorphisms (SNPs), 6 mitochondrial SNPs, and D-loop sequencing to contribute to unveil the history of two Southern Brazilian *quilombos*, Feixo (N = 117) and Restinga (N = 47). We detected a previously unknown predominant Amerindian maternal ancestry in Feixo (59.8%) and a haplogroup (Hg) composition that suggests the participation of Kaingang people in its formation. In Restinga, the Amerindian fraction suggests admixture with Guarani populations. However, most maternal Hgs were African in Restinga (78.7%). Analyzing the African Hg frequency and through a phylogeographic approach, we observed that Feixo presents mtDNA haplotypes congruent with a contribution of African Bantu populations, mainly from Central-Western/Southeastern Africa (possibly Angola, Congo, and Mozambique). In comparison, Restinga showed a Hg composition more typical of Western Africa. A sex-biased picture emerged by analyzing Y chromosome haplotypes that were

predominantly European in both communities (57.8%). The African paternal fraction (35.6%) included only Hgs from clade E, with an outstanding contribution of E-M35, suggesting an origin in Central, Eastern or Southern Africa. Our results bring new insights into the history of the African diaspora in a Brazilian region not yet studied, revealing a high degree of Amerindian admixture and new insights into their within-Africa roots.

Keywords: uniparental markers, sex bias, sub-continental ancestry, Maroon.

Introduction

The trans-Atlantic traffic of human beings triggered an unprecedented diaspora in the history of humanity. It has been estimated that it forcibly relocated more than 12 million Africans to America, Europe, and the Atlantic islands (FORTES-LIMA; VERDU, 2021). Africans and their descendants formed systems of resistance all over the continent. Some of them translated to developing communities known as *quilombos* in Brazil and *cumbes*, *palenques*, or *maroons* in other Latin American countries (GOMES, 2015). Despite receiving nearly half of the Africans brought to America, Brazil has the least trans-Atlantic traffic data coverage (BUCCIFERRO, 2013).

In Brazil, Africans and African descendants formed *quilombos* during and shortly after the abolition of slavery, and those are spread along almost the entire country (GOMES, 2015). *Quilombos* developed as a form of resistance to the oppression of slavery in colonial society (MOURA, 1993). It often involved other ethnic groups, such as neighboring Amerindian populations or marginalized people (MOURA, 1992, 1993). Some of these communities have been settled in isolated or rugged lands to maintain their freedom (GOMES, 2015). However, as each *quilombo* developed in unique environments and was founded by persons originating from potentially diverse ethnic backgrounds, they all have a particular history that together can help recall the history of the largest diaspora in human history (JOERIN-LUQUE et al., 2023; MOURA, 1993).

Feixo and Restinga are two *Quilombos* located in the region of Lapa, Paraná, in Southern Brazil (Figure 1). Based on a longstanding eugenic ideology, the white oligarchy of the state constructed an

official history that hid the involvement of the African workforce during the region's development (RAGGIO; BLEY; TRAUCZYNSKI, 2018). It was not until recent years (2005-2010) that the Clóvis Moura Working Group (*Grupo de Trabalho Clóvis Moura*), a governmental organization, made visible the 86 African-derived rural communities that are distributed all over the state territory, especially in the regions which most benefit economically from slavery (Grupo de Trabalho Clóvis Moura, 2010). In agreement with those historical narratives, the last Brazilian demographic census (2022) demonstrated that the cities of Palmas and Lapa comprise the most significant proportion of *quilombola* population, which constitutes 0.06% of the total inhabitants of the state (representing 7,113 people) (Governo do Estado de Paraná, <https://www.aen.pr.gov.br/>).

The *quilombos* Feixo and Restinga were formed before the abolition of slavery, which occurred in 1888, by Blacks enslaved in two plantations of the region (RIBAS, 2019). Their founders received part of the lands where they worked, and the *quilombos* were developed there (RIBAS, 2019).

Different from other more isolated or semi-isolated *quilombos* from Brazil, Feijo and Restinga maintained, especially in more recent times, intense interaction with neighboring populations, in particular with the closest Colony of Mariental, but also with the city of Lapa and the capital of the state (Curitiba) (CIGOLINI; DA SILVA, 2018; RIBAS, 2019).

In Brazil, African ancestry ranges from 10.9-28.8% on average in urban regions and 40-90% on average in the *quilombos* studied so far (DE SOUZA et al., 2019; JOERIN-LUQUE et al., 2023). However, most of the population has not yet discovered the specific origin of Brazilians' African ancestors (FORTES-LIMA; VERDU, 2021). As the settlement of the Lapa region was related to the *tropeirismo* (a kind of stockbreeding that involved the transport of cattle between the South and the Southeast of Brazil), it is thought that several enslaved Africans could have entered the region through these movements, either from São Paulo or from Rio de Janeiro (FAJARDO; CUNHA, 2021). From São Paulo, it is expected that most Africans had come from the port of Rio de Janeiro, which mainly received Africans from Angola, Congo, and, after 1845, Mozambique (IBGE, 2007). Speculations apart, there are no specific records on the African ethnic origins of the enslaved population of Lapa (GUTIÉRREZ, 2006).

Here, we leveraged the genetic data obtained through the analysis of uniparental markers (mitochondrial DNA and Y chromosome SNPs) and historical records to restore, for the first time, the phylogenetic links of the *Quilombos* Feixo and Restinga to Africa. Through this work, we contribute to the knowledge of the history of Paraná and shed a new light on the origins of the African diaspora.

Subjects and methods

Sample collection and DNA extraction

This project obtained ethics approval from the local institutional review board, Comitê de Ética em Pesquisa em seres humanos, Setor de Ciências da Saúde, CEP/SD (CAAE number 42761521.9.0000.0102). Written informed consent was obtained from all participants for sample collection and interviews. Blood samples were collected along with ethnographic information (name, sex, age, place of birth, and self-declared ancestry). We also collected data on family history (self-declared ancestry and parents' place of birth). DNA was extracted from the buffy coat of blood samples using a commercial kit following the manufacturer's instructions (Wizard® Genomic DNA Purification Kit, Promega). All analyses were performed anonymously.

Genotyping

Overall, 164 samples (Feijo n = 117, Restinga n = 47) were genotyped for 6 mtDNA coding region single nucleotide polymorphisms (SNPs) - 827, 1736, 5178, 10873, 10400, and 15487 - through allele-specific PCR and the product size confirmed by agarose gel electrophoresis (Schwab 2019 and C. M. Bravi, personal communication). After the initial screening, samples with the ancestral state for positions 10873 and 10400 were selected for sequencing of the control region of the mitochondrial genome - D-loop - (see Sanger sequencing section). Male samples (n = 45, Feijo n = 33, Restinga n =

12) were genotyped for 5 Y chromosome SNPs - M242, M207, M89, M9, and P152- through allele-specific PCR with the primers described by Medina and colleagues (MEDINA et al., 2014). PCR product sample sizes were confirmed by agarose gel electrophoresis. Samples positive for the derived state of M242 were also tested for M3 and M346 mutations to check the sub-lineage of Native American haplogroup Q (MEDINA et al., 2014). Samples positive for the derived state of P152 were additionally tested for SNPs M35 and M2. PCR primers for allele-specific PCR are described in Table S1.

Sanger sequencing

MtDNA control hypervariable regions 1 to 3 (HVRI, HVRII, and HVRIII) were sequenced using Sanger's Method with the primers described in Schaan and colleagues 2017 (SCHAAN et al., 2017). HVRI was PCR-amplified with primers L15997 and H017, whereas HVRII-III was amplified with primers L034 and H629. The PCR products were purified with the enzymes exonuclease I (Fermentas, United States) and alkaline phosphatase (Thermo Fisher Scientific, United States) on a Mastercycler EP Gradient S® (Eppendorf) at 37°C for 1 h and 80°C for 15 min. The purified products were then amplified using BigDye® Terminator Cycle Sequencing Standard v3.1 (Life Technologies, United States), according to the manufacturer's instructions, and using the same primers. The sequencing thermocycler program consisted of a first step at 95°C for 1 min, followed by 25 cycles of 95°C for 10 s, 50°C for 5 s, and 60°C for 4 min. The products were purified using ethanol (Merck, Germany), resuspended in Hi-Di Formamide (Life Technologies), and, finally, submitted to capillary electrophoresis in a 3500xl Genetic Analyzer Sequencer (Life Technologies).

Bioinformatic analysis

MtDNA control region electropherograms were aligned to the rCRS (NC_012920.1) (ANDERSON et al., 1981; ANDREWS et al., 1999). Polymorphic positions were analyzed using Mutation Surveyor R © 3.30 (SoftGenetics, United States). MtDNA fasta sequences retrieved from

GenBank for comparisons were aligned to the rCRS using the software Mega (TAMURA; STECHER; KUMAR, 2021). Haplogroup assignation was done using EMPOP (<https://empop.online/>, database version v4/R13) and HaploGrep3 (SCHÖNHERR et al., 2023). Haplogroup frequencies were calculated by direct counting. Ethnographic information available - kinship until first grade - was used to double-check the haplotypes observed. Positions with missing data and InDels (especially from poly-C regions from positions 303-315 and 16183-16194) were removed from all analyses. For the construction of haplotype networks, we retrieve from the literature the GenBank codes of mitochondrial sequences matching two criteria: 1- Be assigned to one of the haplogroups detected in our sample; 2- Have been collected in populations from the African regions known to be involved with the slave trade to Brazil. Despite the absence of records of enslaved Africans brought to Brazil from Madagascar, sequences from that location were also used in the comparison as proxies to the Bantu lineages found in Mozambique, known to be a contributor to Madagascar and Brazilian gene pools (HALL, 2005; PEREIRA et al., 2001; PIERRON et al., 2017). In the same line, samples from Morocco were used as proxies to lineages from unsampled populations from West Africa. Haplotype parsimony networks and indices of haplotype (H) and nucleotide diversity (π) were calculated with the R package *pegas* (PARADIS, 2010). Analysis of molecular variance (AMOVA) was performed using the R package *poppr* (KAMVAR; TABIMA; GRÜNWALD, 2014). Correspondence analysis was performed using the function CA of the R package *FactoMineR* and visualized with the package *factoextra* (LÊ; JOSSE; HUSSON, 2008).

Results and discussion

The mtDNA and Y chromosome haplotypes detected in Feijo and Restinga are described in Table S2, along with the corresponding haplogroup classifications and the mtDNA control region range analyzed. The EMPOP algorithm and HaploGrep v.3 classified the haplotypes identically (Table S2).

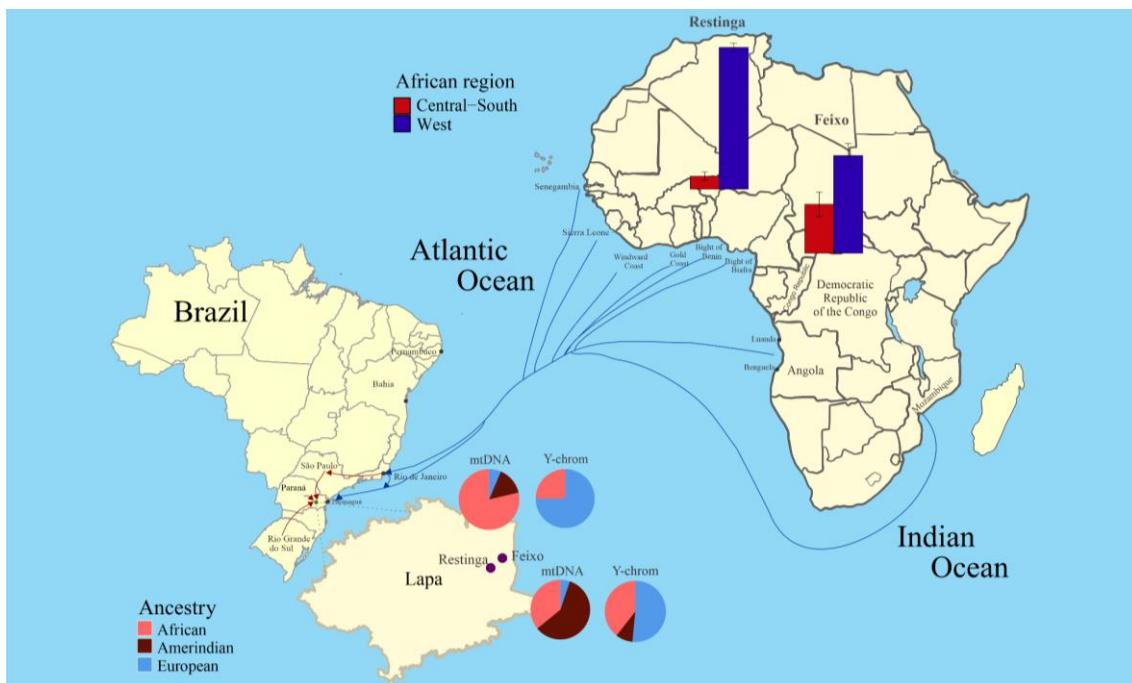


Figure 1 Geographical location of the African-derived populations Feixo and Restinga and uniparental ancestry proportions observed. The region of Lapa, in the state of Paraná, where the *quilombos* Feixo and Restinga are located is enlarged in the map. Uniparental ancestry proportions observed in each community are showed in pie charts. Blue arrows indicate possible slave vessel routes from the main coast of departure during the trans-Atlantic traffic and to the *quilombos*, based on historical records (ELTIS; RICHARDSON, 2010; WESTPHALEN, 1972). Red arrows outline possible land routes for the entry of Africans enslaved to the *quilombos* based on historical records (GUTIÉRREZ, 2006; RAGGIO; BLEY; TRAUCZYNSKI, 2018). Bar charts show the estimated proportions of sub-continental African maternal ancestry estimated according to the absolute number of haplotypes from each African region observed in both *quilombos*.

The female component

We genotyped 164 samples (Feixo n = 117, Restinga n = 47) for 6 mtDNA SNPs (10400, 10873, 1736, 827, 15487, and 5178), which define the macro-haplogroups M and N, and the clades A, B4b'd'e'j, M8 (ancestral to clade C), and D, respectively. Clades A, B4b'd'e'j, M8, and D are ancestral to all Native American founder Hgs (A2, B2, C1b, C1d, C1c, C4c, D1, D2a, D3, D4h3a, D4e1c, X2a, and X2g) (ÁLVAREZ-IGLESIAS et al., 2007; MOTTI et al., 2017). We considered all mitochondrial chromosomes positive to A, B4b'd'e'j, M8, and D markers to be Native American (from A, B, C, and D clades, respectively) as the intermediate Hgs between those ancestral clades, and the Amerindian ones only occur in Asian populations, which are not reported to be involved in the formation of *quilombos*.

(GOMES, 2015). For the same reason, as all the mitochondrial lineage diversity from out of Africa derives from macro-haplogroups M and N, haplotypes attributed to them (but not to A, B, C, and D clades) were considered of European origin (UNDERHILL; KIVISILD, 2007). MtDNA chromosomes with the ancestral states for 10873 (N) and 10400 (M) were attributed to the African macro-haplogroup L.

Considering both communities, most haplogroups were African or Amerindian (African = 47.6%, Amerindian = 47%). However, when we observed each community separately, a striking majority of African lineages was observed in Restinga (78.7%), whereas Feixo presented mostly Amerindian Hgs (59%) (Figure 1). The European fraction was less expressive in both *quilombos* (European = 5.5%). Looking at the whole picture of maternal ancestry in Brazil, this pattern is perfectly consistent, as the sum of Amerindian and African ancestries, or one out of those two, is predominant in most Brazilian populations (JOERIN-LUQUE et al., 2022, 2023; POLETTO et al., 2019; SCHAAN et al., 2017). It is interesting to note that the tremendous Amerindian contribution of maternal lineages in Feixo agrees with that detected in other studies of urban populations from the state of Paraná, where that ancestry ranged from 38.4 to 49.2% (JOERIN-LUQUE et al., 2022; POLETTO et al., 2019).

In the same line, Amerindian maternal ancestry ranges from 10-60% in the Brazilian *quilombos* studied so far (JOERIN-LUQUE et al., 2023). However, the elevated proportion observed in Feixo is only comparable with that observed in *quilombos* from the Amazon region (BORTOLINI et al., 1999; CARVALHO et al., 2008; SILVA et al., 2006). In the Amazon, a high degree of Amerindian admixture in *quilombos* was favored because the region concentrates the greatest density and diversity of Indigenous populations from Brazil (Brazilian Socioenvironmental Institute; <https://pib.socioambiental.org>). There are still no reports on the maternal ancestry of other *quilombos* from southern Brazil (JOERIN-LUQUE et al., 2023). However, an average of 4.7% Amerindian paternal ancestry (with most communities showing null Amerindian paternal contribution) and an average of 19.3% autosomal Amerindian ancestry observed in *quilombos* from southeastern Brazil (mainly from the Ribeira Valley, a region geographically close to Paraná), suggests a predominantly female mediated Amerindian gene-flow in those populations (KIMURA et al., 2013, 2017).

Haplogroup	<i>Quilombos</i>		Feixo		Restinga	
mtDNA	n	%	n	%	n	%
A	16	9.8	12	10.3	4	8.5
B	6	3.7	5	4.3	1	2.1
C	55	33.5	53	45.3	2	4.3
D	0	0	0	0	0	0
Total Amerindian	77	47	70	59.8	7	14.9
M	1	0.6	1	0.9	0	0
N	8	4.9	5	4.3	3	6.4
Total European	9	5.5	6	5.1	3	6.4
L	78	47.6	41	35	37	78.7
L0a2a2	1	0.6	1	0.9	0	0
L1b1a+189	4	2.4	4	3.4	0	0
L1b1a10	4	2.4	3	2.6	1	2.1
L1c1b	1	0.6	0	0	1	2.1
L1c1b1	1	0.6	0	0	1	2.1
L2a1+143+16189+(16192)	2	1.2	2	1.7	0	0
L2a1+16189+(16192)	18	11	18	15.4	0	0
L2a1c2a	1	0.6	0	0	1	2.1
L2a11	32	19.5	0	0	32	68.1
L2c2b1b	7	4.3	6	5.1	1	2.1
L3d3a	5	3	5	4.3	0	0
Total African	78	47.6	41	35	37	78.7
Y chromosome						
Q	3	6.7	3	9.1	0	0

Q-M3	3	6.7	3	9.1	0	0
Total Amerindian	3	6.7	3	9.1	0	0
R	19	42.2	13	39.4	6	50
F	7	15.6	4	12.1	3	25
Total European	26	57.8	17	51.5	9	75
E(xM35, M2)	5	11.1	5	15.2	0	0
E-M35	8	17.8	6	18.2	2	16.7
E-M2	3	6.7	2	6.1	1	8.3
Total African	16	35.6	13	39.4	3	25

Table 1 Frequency of mtDNA and Y chromosome haplogroups detected in the two *quilombos* studied.

Have different Amerindian ethnic groups contributed to the formation of Feixo and Restinga?

Only Amerindian clades A, B, and C were detected in Feixo and Restinga, and a closer look at their Hg composition highlights significant differences between them (Fisher exact test *p*-value = 0.02214). In Feixo, the striking majority of lineages belonged to Hg C (45.30% of the total Hgs), whereas in Restinga, the highest proportion corresponded to A Hg (8.5% of the total Hgs) (Table 1).

Among the original settlers of the region where the *quilombos* are located (Guarani Mbya and Guarani Nandeva from the Tupian linguistic branch; Kaingang and Xocleng from the Macro-Je linguistic branch), genetic data is only available from Guarani and Kaingang ethnic groups (BISSO-MACHADO; CÁTIRA BORTOLINI; SALZANO, 2012; MARRERO et al., 2007). The absence of Hg D in both *quilombos* strongly suggests the incorporation of Kaingang women, as D has never been detected in that ethnic group, whereas it has been found in most Guarani populations studied so far (BISSO-MACHADO; CÁTIRA BORTOLINI; SALZANO, 2012; TAVARES et al., 2019). Furthermore, the highest Hg C prevalence in Feixo and the Kaingang also points in that direction

(BISSO-MACHADO; CÁTIRA BORTOLINI; SALZANO, 2012). On the other hand, the majority of Hg A in Restinga could suggest the incorporation of Guarani women, as that Hg prevails in most Guarani populations studied so far (TAVARES et al., 2019). The frequency of Hg D varies greatly in Guarani populations (0-50%) as well as the frequency of B (0-17%), which was the less frequent Hg in both *quilombos* (TAVARES et al., 2019).

Indigenous women's involvement in *quilombos'* formation is an interesting anthropological question. On the one hand, the political articulation and cooperation of African descendants with Amerindians, as well as with other minorities discriminated against in colonial society, has been repeatedly cited (GOMES, 2015; MOURA, 1993). On the other hand, there are reports of Amerindian women kidnapping by *quilombolas* (GOMES, 2015; MOURA, 1993). In this particular case, the *quilombo* Feixo is inserted in a region with a great maternal contribution of Amerindian women where, interestingly, haplogroup C prevails (JOERIN-LUQUE et al., 2022; POLETTI et al., 2019). Thus, this pattern could also reflect a high degree of gene flow from neighboring urban populations, as previously reported for that community (RIBAS, 2019).

Feixo and Restinga could have had distinct African genetic contributors

Overall, 78 samples were assigned to mitochondrial African clades (L-derived) according to the genotyping of SNPs 10400 and 10873. For those samples, we sequenced the mitochondrial D-loop to establish the specific lineage to which they belong (Feijo n = 39, Restinga n = 37) (Table S2). Two samples failed the sequencing reaction and, were excluded from downstream analysis. To help unveil the origin of the African contributors to the maternal lineage pool of Feijo and Restinga, we used a phylogeographic approach to compare their haplotypes with those described in published literature.

Only sub-branches of clades L0-L3 were detected in both communities, representing 11 haplogroups (Table 1). The predominant haplogroups were lineages derived from L2 (78.9%) and L1 (13.2%) (Table 1 and Figure 2). From clade L1, only lineages derived from L1b were detected in Feijo, whereas lineages derived from L1b and L1c were present in Restinga (Table 1 and Figure 2). L1b and

L1c have different distribution patterns in Africa. Whereas the sub-branches of L1b are more enriched and diverse in West Africa, those of L1c occur both in Western and West-Central Africa (ROSA; BREHM, 2011). L1c-derived branches are found in Huther-gatherers and Bantu populations from Central Africa, possibly derived from a common ancestral population (QUINTANA-MURCI et al., 2008).

In Feixo, we detected two lineages derived from Hg L1b1a: L1b1a+189 and L1b1a10. When compared with the haplotypes from African reference populations, the samples L1b1a10 found in Feixo matched haplotypes from Bantu and Niger-Congo non-Bantu populations from Angola, Nigeria, and Burkina Faso (Figure S1). Haplotypes from Hg L1b1a+189 did not match any haplotypes from the populations evaluated but were similar to one haplotype found in Bantus from Senegal (Figure S1).

The lineages L1c1b and L1c1b1 were only detected in Restinga (Table 1 and Figure 2). Those lineages have been proposed as originating from Bantu populations (BATINI et al., 2007). In agreement, one of the haplotypes detected in our sample was identical to that described in Bantu populations from Angola and Zambia. The haplotype L1c1b1 found in Restinga did not match any of the sequences evaluated. Still, it was similar to one ancestral L1c1b from Zambian Bantus, suggesting a Southern Bantu origin for those lineages in Restinga (Figure S2).

Clade L2 was the most frequent in both communities (66.7 in Feijo and 91.9% in Restinga), but only two out of the five main L2-derived clades (L2a, -b, -c, -d, and -e) were detected in the *quilombos*. Lineages derived from haplogroups L2a1 and L2c2b1b were observed in the *quilombos* (Table 1 and Figure 2).

Clade L2a1 originated in West Africa and is the most complex L2a sub-branch (SILVA et al., 2015). It englobes lineages from all African and non-African regions (SILVA et al., 2015). Lineages derived from L2a1 detected in Feijo were L2a1+143+16189+(16192) and L2a1+16189+(16192), the last at a high frequency (46.2%) (Table 1). None of the haplotypes found in Feijo matched one of those observed in the African reference populations evaluated (Figure S3).

The *quilombo* Restinga presented a surprisingly elevated frequency of Hg L2a11, constituting most of the haplotypes detected (86.5%). Interestingly, all L2a11 chromosomes were represented by only

two haplotypes that differ by the presence of the hotspot mutation 16519T (Figure 2) (BANDELT; MACAULAY; RICHARDS, 2006; SOARES et al., 2009). This finding suggests a unique founding haplotype for this Hg in Restinga (haplotype II in Figure S4, of higher frequency), where the transition 16519T emerged (Figure 2). Haplogroup L2a11 has been proposed to be originally from Western Africa. It occurs today in Western and Central Africa, with subclades suggesting its involvement in recent Bantu migrations southwards (SILVA et al., 2015). As mutations outside the control region define most sub-branches of L2a11, it is not straightforward to infer the origin of that Hg in Restinga (VAN OVEN, 2015). The comparison with L2a11 haplotypes from African populations suggests a source of this lineage in West/North Africa, as the haplotypes found in Restinga matched haplotypes from Burkina Faso and Morocco (Figure S4). As North Africa was not a source of the traffic of human beings to America during the Atlantic Slave Trade, and as, due to geographic proximity, several haplotypes from the North also occur in West Africa, the last region emerges as a more probable source of that lineage to Restinga (HALL, 2005).

Other minor haplogroups were L2a1c2a and L2c2b1b. L2a1c shows a pattern similar to L2a11, with origin and distribution in Western and Central Africa (SILVA et al., 2015). However, the haplotype found in Restinga did not match any of the haplotypes found in reference populations evaluated here (Figure S5). L2c2b1b was the only haplogroup shared between both communities (Figure 2). It is associated with Southern African Bantu populations and has no representatives in Eastern Africa (SILVA et al., 2015). The L2c2b1b haplotypes observed in Feijo and Restinga only differ from those found in Angola, South Africa, Zambia, and Madagascar by the presence of the hotspot mutation 16519C (Figure S6), suggesting a Central/Southern African Bantu origin for that lineage.

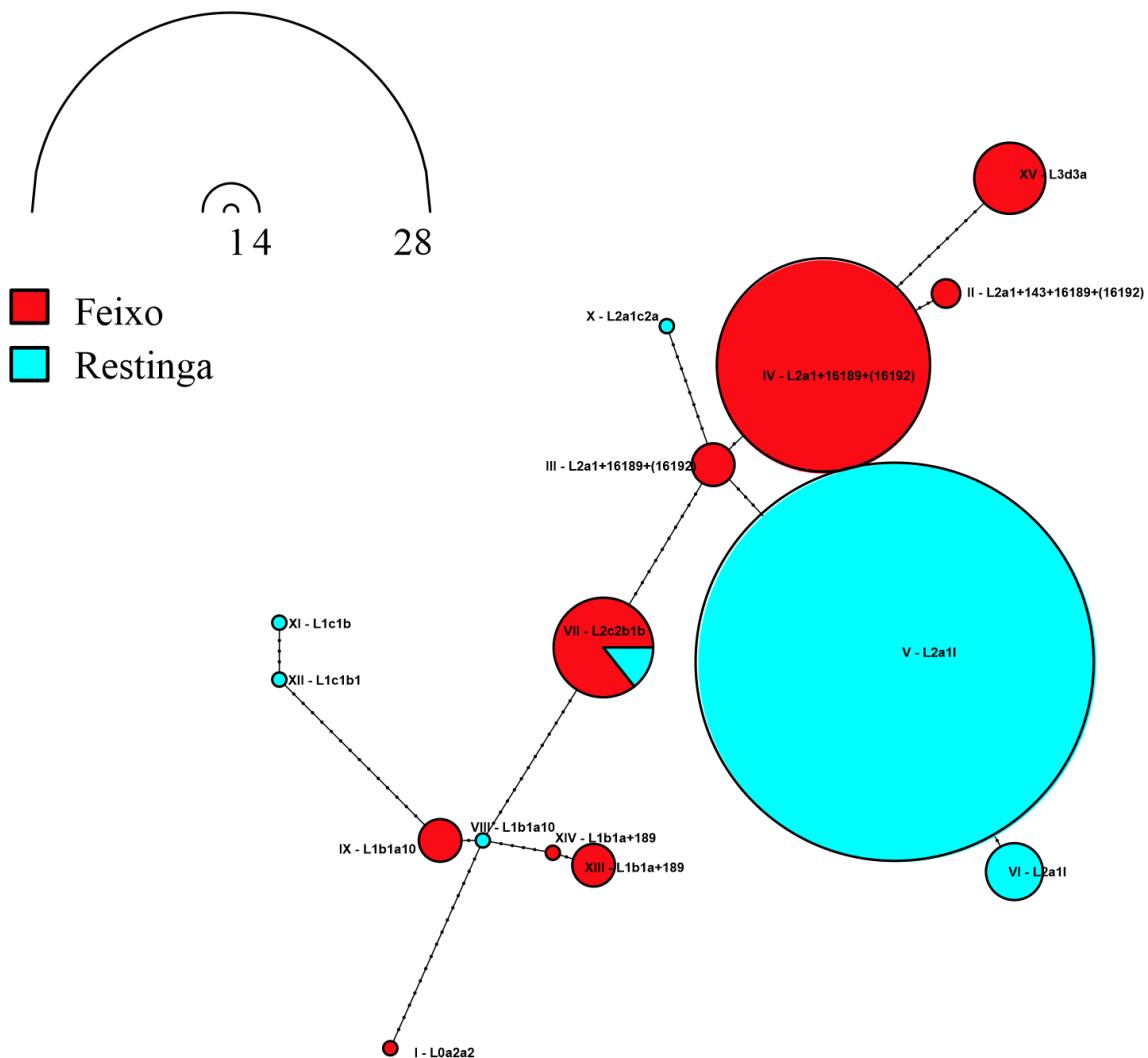


Figure 2 Haplotype network of African mitochondrial haplotypes detected in Feixo and Restinga. Haplotypes are identified with Roman numerals. Black circles on the network links represent mutational steps between the haplotypes.

The only L0 lineage detected in these *quilombos* was L0a2a2, seen at a low frequency (2.6%) in Feixo (Table 1 and Figure 2). Most L0a-derived lineages (75%) show a Central African origin despite the Eastern origin of their ancestral clade, L0a (SILVA et al., 2015). Its common lineages in Africa are derived ones, such as L0a2a2a and its subclades (SILVA et al., 2015). However, one exception to the Central African origin of L0a-derived lineages is the frequent branch L0a2a2a, which may represent an Eastern African lineage assimilated by Bantu migrating groups and disseminated southwards during the Bantu expansion (SILVA et al., 2015). As L0a2a2a is defined by the mutation C13116T, which is outside of the control region, not analyzed here, it is still possible that the chromosome found in Feixo

could belong to that branch (VAN OVEN, 2015). In either case, the haplotype found in Feijo is likely derived from Central or Southern African Bantu groups.

Finally, the lineage L3d3a was detected in Feijo at 12.8% (Table 1 and Figure 2). The ancestral clade of this Hg, L3d, is found across Africa at low frequency among Bantu and Khoisan populations (BARBIERI et al., 2014b). It reaches a maximum in populations from the Southwest of Africa (Namibia) (BARBIERI et al., 2014b). In that region, lineages mainly derived from its sub-brach L3d3a1 are unusually frequent in the pastoralist Bantu ethnic groups Herero and Himba (47% and 38%, respectively) and their neighbors, the Khoisan-speaking group Damara (63%), which are foragers and small stock-herders (BARBIERI et al., 2014b). Furthermore, L3d3a has been suggested to be carried by Khoe-Kwadi-speaking pastoralist groups and incorporated by Bantu populations through gene flow (BARBIERI et al., 2014b). As mutations outside the control region define L3d3a-derived lineages, we cannot discard that the chromosomes found in Feijo could belong to these sub-clades (VAN OVEN, 2015). Also, the haplotype network with Southern African samples showed a match of the haplotypes found in Feijo with the several available L3d3a sequences of Bantu populations from Angola, Namibia, Zambia, South Africa, and Botswana (Figure S7). In addition, as Khoisan-speaking people are not thought to have been brought to Brazil during the Trans-Atlantic slave Trade, a Central/Southern African Bantu origin is more reasonable for this haplotype in Feijo (SALAS et al., 2004).

Different patterns of sub-continental ancestry are observed in both *quilombos*

To explore the general patterns of sub-continental African ancestry in each *quilombo*, we considered the frequency of the Hgs detected and compared it with that of African populations from the regions known to be sources of the African diaspora in America through a Correspondence Analysis (CA). Due to the high number of mtDNA Hgs described to date, we used a hierarchical approach to reduce this number (reference populations and their Hg frequencies are described in Table S3).

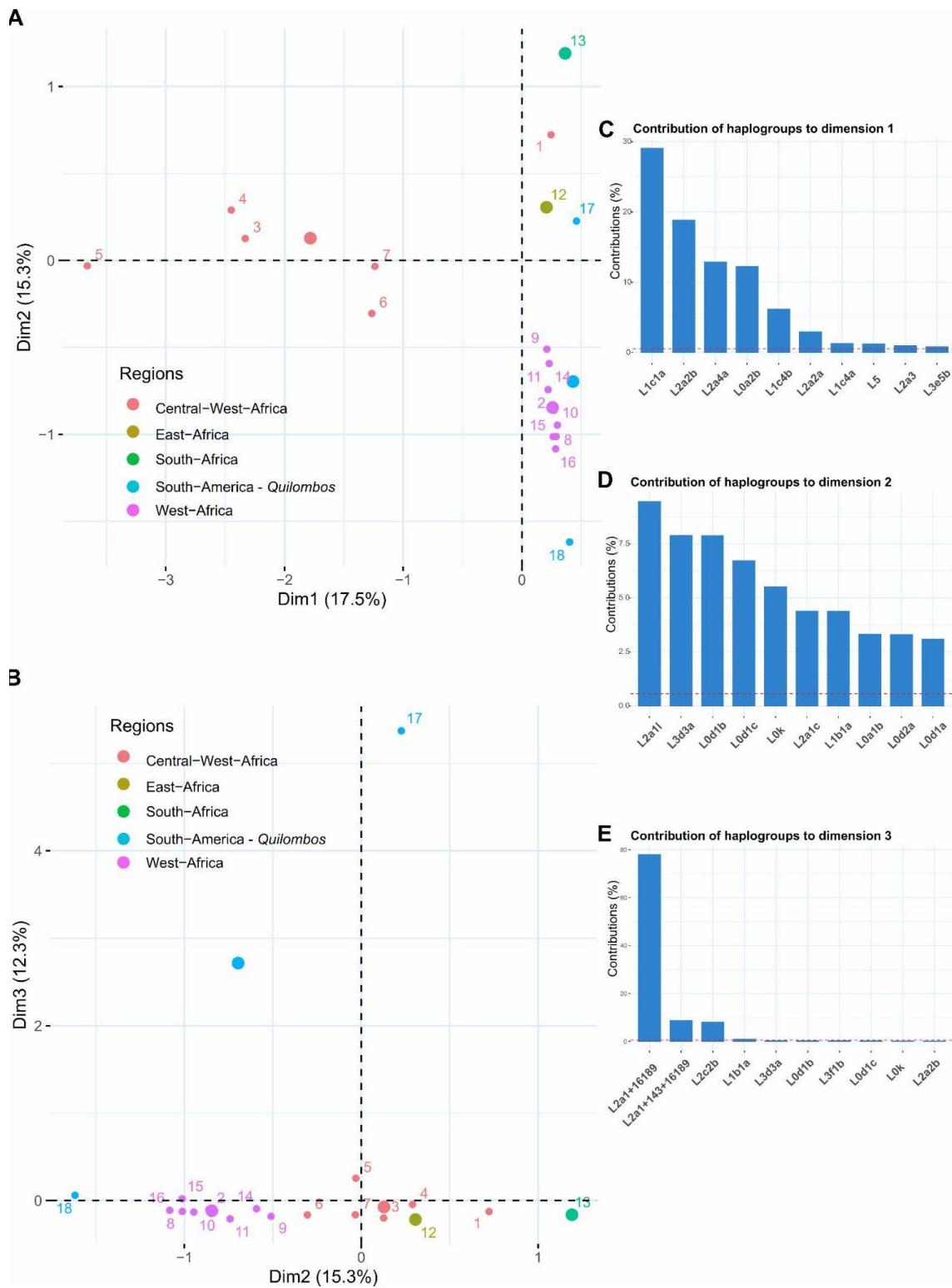


Figure 3 Correspondence Analysis of mtDNA haplogroup composition in Feixo, Restinga, and African reference populations. A) Dimensions 1 and 2. B) Dimension 2 and 3. C), D), and E) Contribution of haplogroups to each dimension. 1: Angola, 2: Burkina Faso, 3: Cameroon, 4: Central African Republic, 5: Congo, 6: Equatorial Guinea, 7: Gabon, 8: Gambia, 9: Ghana, 10: Guinea-Bissau, 11: Ivory Coast, 12: Mozambique, 13: Namibia, 14: Nigeria, 15: Senegal, 16: Sierra Leone, 17: Feixo, 18: Restinga.

As expected, reference African populations clustered according to the geographical distribution of mitochondrial Hgs, from North to South along the second dimension, which captures 15.3% of the variance (Figure 3, A and B). The first dimension accounted for 17.5% of the variance and mainly separated Central African populations in a gradient from west to east (Figure 3, A and B). In this layout, Feixo and Restinga adopted very divergent positions. Whereas Feixo was closer to the Central/Southern/Southeastern populations, Restinga clustered closely to the Western African ones. It is important to note that, as *quilombos* are populations usually formed by few individuals, they are subjected to strong genetic drift, particularly the founder effect (GOMES, 2015). A certain level of isolation, especially in the first generations after the formation of the *quilombo*, could potentially increase endogamy in these populations (LEMES et al., 2014). Thus, the patterns observed are heavily influenced by the enrichment in specific lineages. In this line, when we look at the haplogroups with the highest contributions to the CA dimensions, we see that the arrangement of the populations along the second dimension is mainly influenced by the frequencies of Hgs L2a11, which approximates Restinga to Western populations and by L3d3a, which approximates Feixo to Southern populations (Figure 3A, B, D, and Figure S). L0-derived lineages, enriched in Southern Africa, also contribute to the North/South separation of the populations and their clustering with Feixo (Figure 3A, B, D, and Figure S8).

When an additional third dimension is considered (Figure 3B, and E), we observed that lineages L2a1+143+16189 and L2a1+16189 push Feixo away from the other populations. Even though the enrichment of these lineages in Feixo could be the result of genetic drift, we cannot discard that the presence of the homoplasic positions 143, 16189, 16192, and 16309 can make the correct Hg assignation of those samples difficult (SILVA et al., 2015).

In addition, when the African maternal fraction was compared to other *quilombos* and urban Brazilian regions with available data, there was no clustering according to geographical area (Figure S1 and 2). Regardless of their geographical location in Brazil, *quilombos* show distinct clade composition, although restricted to clades L0a, L1, L2, and L3, the commonest in Sub-Saharan Africa. Feixo and Restinga clustered near the Northern *quilombo* of Marajo but far from Southern Brazilian urban samples.

The lasts, in contrast, cluster with Central-Western and Southeastern urban samples (Figure S1 and 2). In this sense, the clustering of the big urban centers could reflect a broader picture of the mitochondrial lineages introduced by the trans-Atlantic slave trade in Central-Southern Brazil, as the three regions analyzed, according to historical records, received enslaved Africans mainly through the port of Rio de Janeiro. On the other side, the position of *quilombos* could have been more influenced by the effect of genetic drift in their formation. However, it is essential to highlight that this approach is limited by the need for more data from most Brazilian *quilombos* and due to the different markers used in each study (coding SNPs, HV1, or complete control region).

Not surprisingly, the male lineages in both *quilombos* were mainly European

We genotyped 45 samples (Feixo $n = 33$, Restinga $n = 12$) for 9 Y SNPs. Most Y chromosomes were allocated to clade R-M207 (42.2%). Clade R-M207 is the most common in the urban admixed populations of Brazil and predominant in Western European countries such as Portugal, from which Brazil received the most migrants (NAVARRO-LÓPEZ et al., 2021; RIBEIRO, 1995). To a lesser extent, other Y chromosomes belonged to clade F-M89 (15.6%), from which other common European Hgs (such as I and J, that have been frequently described in other *quilombos*) derive (JOERIN-LUQUE et al., 2023).

The paternal ancestry of Feijo and Restinga follows a pattern previously observed in other *quilombos* from Brazil (JOERIN-LUQUE et al., 2023). In Brazil, as in other European colonized populations, the admixture process was largely biased by a predominantly male-mediated European gene flow, contrasting with Amerindian and African contributions being frequently female-mediated (BORTOLINI et al., 1999; ONGARO et al., 2019). The relative proportion of the last two ancestries usually varies more according to the formation of the population (JOERIN-LUQUE et al., 2022, 2023). This sex bias has been longly associated with the countless reports of rape of Amerindian and African women by European men and their descendants, as well as with the asymmetric relation of power between the sexes in the American colonies (AIDOO, 2018; NASCIMENTO, 2016; WALSH, 1830).

The same pattern has been frequently observed in other Brazilian *quilombos*, especially in Southeastern Brazil, as well as in other African-derived communities formed by ex-enslaved persons in America (D'ATANASIO et al., 2020; JOERIN-LUQUE et al., 2023; KIMURA et al., 2017; MARTÍNEZ et al., 2020). In particular, two scenarios could have determined the excess of European Y chromosomes in Feijo and Restinga. First, that could result from the sexual exploitation of African enslaved women, possibly used as reproducers of the slave labor force (DAVIS, 1944). That hypothesis is reinforced by the historical records reporting that most enslaved African descendants in Paraná were born in Brazil (GUTIÉRREZ, 2006). Secondarily, it can be caused to some degree by more recent gene flow from surrounding urban regions (CIGOLINI; DA SILVA, 2018).

Only 6.7% of the chromosomes (all of them detected in Feijo) belonged to Hg Q-M242, a Pan-American Q lineage from which most Amerindian lineages derive (GRUGNI et al., 2019). All Q-M242 chromosomes were also positive for Q1b-M346 and its derived lineage Q1b1a1a-M3. Q1b1a1a-M3 is an autochthonous Amerindian Hg and the most frequently observed in Natives from South America (BAILLIET et al., 2009).

The African male fraction is enriched in Eastern African haplogroups

Most African Y haplogroup diversity is constituted by lineages from the clades A, B, and E. Clade A is considered the most ancient and prevails in eastern and southern Africa (COELHO et al., 2009; UNDERHILL; KIVISILD, 2007). B lineages have a broader distribution over Sub-Saharan Africa but with the highest prevalence among Central African hunter-gatherers (COELHO et al., 2009). Most African Y chromosome diversity is constituted by sub-branches of Hg E (JOBLING; TYLER-SMITH, 2003).

Hgs A and B were not detected in Feijo and Restinga, as demonstrated by all Y chromosomes presenting derived states for at least one of the SNPs tested.

Overall, 35.6% of the Y chromosomes belonged to clade E-P152 (Table 1). P152 defines the basal clade E (also defined by the SNPs M96 and M40) of African origin and from which most African

haplogroups derive (UNDERHILL; KIVISILD, 2007). Of them, 17.8% (18.2% in Feijo and 16.7% in Restinga) were also positive for the derived allele of SNP M35, which defines Hg E1b1b1 (Table 1). Only three chromosomes (6.7% of the total) presented the derived state of M2, the mutation that defines Hg E1b1a1 (Table 1).

Haplogroup E1b1b1-M35 has an eastern African origin and is one of the most broadly distributed E clades (TROMBETTA et al., 2015). Its lineages reach high frequencies in a broad area from northern and eastern Africa to Europe and western Asia (TROMBETTA et al., 2015). Even though specific E1b1b1-derived lineages (V6, M78, M81, M123, and M293) have been reported at high frequencies in European populations, the same lineages are also present in northern, eastern, and southern Africa (CRUCIANI et al., 2004; NAVARRO-LÓPEZ et al., 2021; TROMBETTA et al., 2015). Some of its terminal branches are restricted to Eastern African pastoralists and Cushitic populations from the Horn of Africa (TROMBETTA et al., 2015). Furthermore, its paraphyletic branch, E1b1b1-M35*(xV6, M78, M81, M123, M293), is present at low frequencies in Western Africa and is common in the East and the South of the continent both in Bantu, as in Nilo-Saharan, and Afroasiatic populations, but is very rare in Europe (CRUCIANI et al., 2004; ROWOLD et al., 2014; TROMBETTA et al., 2015).

It is important to highlight that all the participants carrying Hg E1b1b1-M35 reported that their fathers were also born in the *quilombo*. Thus, despite its phylogeographic ambiguity, we classified all E chromosomes, including the E1b1b1-M35 lineages, as African, based on the strong links of *quilombos* with that ancestry. E1b1a1-M2 and its derived lineages are the most common in Niger-Congo populations, such as Bantu-speaking, among other Sub-Saharan African groups, such as the non-Bantu Khoisan (Southern Africa) (DE FILIPPO et al., 2011).

Looking closely at the proportion of lineages in these *quilombos* raises an intriguing picture. The higher prevalence of Hg E1b1b1-M35 and E(xM35, M2) suggests a main male contribution from Eastern/Southern Africa. It is worth noting that populations from regions known as sources of Africans brought to Brazil, such as the Hema (Nilotic) from Northeastern Congo, have exceptionally high frequencies of that Hg (CRUCIANI et al., 2004; HENN et al., 2008).

Based on uniparental markers, Feixo and Restinga are genetically similar

We run a molecular variance analysis (AMOVA) based on the polymorphic sites studied in this work to evaluate the degree of differentiation of the two *quilombos*. Interestingly, we did not detect a significant variance between populations nor in the mtDNAs (Monte-Carlo test of significance p -value = 0.487) or in the Y chromosomes (Monte-Carlo test of significance p -value = 0.99), as most variance is condensed between the haplogroups.

In addition, we assessed the genetic diversity parameters of the African maternal fraction from these two communities. We observed that, despite its minor African maternal contribution when compared to Restinga (Table 2), Feixo presented higher values for all the parameters evaluated, which could be congruent with a higher number of African female founders from different ethnic backgrounds or a stronger effect of the genetic drift in Restinga. Compared to the urban centers of Paraná, the haplotype diversity indices in the *quilombos* were smaller (H : 0.997) (POLETTO et al., 2019).

Population	N	NHg	Nha	Nhap/N	H	Variance	π	Variance
				p				
<i>Quilombos</i>	76	16	16	0.211	-	-	-	-
Feixo	39	7	10	0.256	0.812	0.002	0.131	0.005
Restinga	37	6	6	0.162	0.390	0.010	0.049	0.001

Table 2 Population diversity indices for the two *quilombos* studied. H: Nei's haplotype diversity

index. π : Nucleotide diversity index. Hg: haplogroups. Hap: haplotypes.

Self-declared color is associated with maternal but not paternal ancestry

In Brazil, a large part of the anti-racist social policies is based on self-declaration of ancestry, officially denominated “color or race”, here relaying in the social dimension of the last concept (KENT; WADE, 2015). Thus, government surveys and censuses use the phenotypic categories the Brazilian Institute of Geography and Statistics (IBGE) defines. These categories are Preto, Pardo, Branco, Amarelo, and Indigena, translated here as Black, Admixed, White, Asian, and Amerindian. It has been demonstrated that a complex interplay between culture, genealogy, appearance, and shared history of slavery or discrimination plays a role in self-identification in Brazil (KENT; WADE, 2015; LIMA-COSTA et al., 2015; RUIZ-LINARES et al., 2014).

Previous studies have evaluated the association of self-perceived ethno-racial categories with autosomal and uniparental ancestries in urban populations (CARDENA et al., 2013; LIMA-COSTA et al., 2015; RUIZ-LINARES et al., 2014). These studies demonstrated that self-declared perception of ancestry depends heavily on non-biological cultural factors which can vary according to the population studied (LIMA-COSTA et al., 2015). Here, we evaluated the association of self-declared color with uniparental ancestry in *quilombos*.

Both communities differed in the proportion of self-declared ancestries, with most people declaring to be Admixed in Feijo (63.5%) and Black in Restinga (57.4%) (Fisher exact test *p*-value = 0.0008185). This finding is in agreement with the differences observed in the proportion of maternal lineages in both *quilombos*, with Restinga presenting the highest proportion of African lineages (Figure 1, Table 1). A minority of people declared to be white in both *quilombos* (7% in Feijo and 8.5% in Restinga). These proportions are also highly divergent from that reported in the last census for the general population of Southern Brazil where 78.2% people self-declared White, 20.9% Admixed, and 5.4% Black (Pearson's Chi-squared *p*-value = < 2.2e-16) (IBGE, 2022).

When the uniparental ancestry of the participants was compared to their self-declared color, the maternal ancestry was associated to self-declared ancestry in both communities (Fisher exact test *p*-value = 0.006899 in Feijo and *p*-value = 0.005867 in Restinga). People who self-declare White or

Admixed had predominantly Amerindian maternal ancestry (Whites 50% Amerindian, 33.3% European, and 16.7% African; Admixed 52.8% Amerindian, 44.9% African, and 2.2% European ancestry). Whereas people who self-declared Black had predominantly African ancestry (Blacks 55.7% African, 39.3% Amerindian, and 4.9% European ancestry) (Figure 4).

On the other hand, paternal ancestry did not show a significant association to self-declared color in both communities (Fisher exact test p -value = 0.47). It is interesting to note that the European paternal ancestry was predominant in most color categories (100% in Whites and 62.5% in Admixed). Only in Blacks the European was equivalent to African paternal ancestry (being both 44.4%).

Similar associations between mtDNA ancestry and self-declared color have been previously observed for Brazilian urban populations (CARDENA et al., 2013). As European paternal ancestry predominates in most Brazilian populations (JOERIN-LUQUE et al., 2022, 2023; RESQUE et al., 2016; SCHAAAN et al., 2020) we could assume that mitochondrial lineages capture at a greater extent the population substructure of the Brazilian population. The last hypothesis could be especially suitable in communities which are more aware of their ancestry due to political reasons, such as *quilombos*, which need to self-declare their ethnic background in order to be legally recognized and to preserve their land rights (RIBAS, 2019).

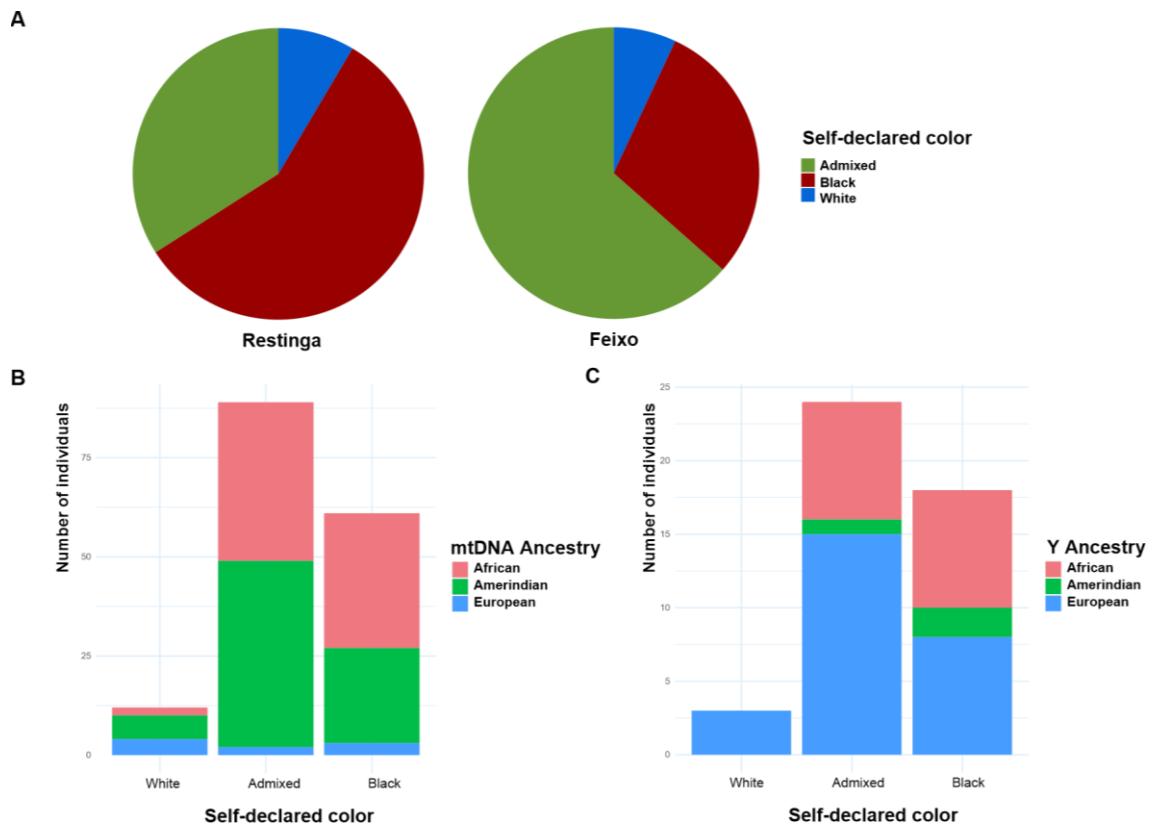


Figure 4 Self-declared color proportions in *quilombos* from Southern Brazil are associated to maternal but not to paternal ancestry. A) Self-declared color proportions in the *quilombos* studied B) Mitochondrial ancestry proportions according to self-declared color C) Y chromosome ancestry proportions according to self-declared color.

Conclusions

Winds and shorter distances minimized the travel between the ports of West Central and Southeastern Africa (mainly Angola and Mozambique) and Rio de Janeiro in the Southeast of Brazil (HALL, 2005). Rio de Janeiro is expected to be the main entrance port for the Africans brought to Paraná state. That is because it is thought that most Africans came to Paraná enslaved along the livestock-associated movements between the states of Rio Grande do Sul and São Paulo. Also, they could have been brought by the incoming migrants from São Paulo, representing a great fraction of the immigration

during the early formation of the urban centers of Paraná (FAJARDO; CUNHA, 2021). Furthermore, some historical records narrate the entrance, by the port of Paranaguá, of navies from Angola and Mozambique (WESTPHALEN, 1972). The last records represent events after the English prohibition of enslavement of African people (1831), which caused the traffic of Africans to continue illegally through minor and less monitored ports from Brazil, such as the port of Paranaguá (WESTPHALEN, 1972). At that time, the traffic from Mozambique intensified (HALL, 2005).

Considering both historical records and the genetic data obtained through this work, we can propose that more probably the Southern, Central, and/or Southeastern African regions (possibly from Angola, Congo, and Mozambique, respectively) were the sources of most maternal haplotypes detected in Feixo. On the other hand, a West African region seems to be the source of most mitochondrial lineages detected in Restinga. On the paternal side, we can suggest that Central-Eastern (with Congo as a candidate) or Southern Africa could have contributed most African Hgs (from clade E-M35). Genetic drift could have substantially modified the frequencies of certain lineages increasing expressively the estimated proportion of Western haplotypes in both *quilombos*.

Especially for Feixo, the proportion of mitochondrial Amerindian Hgs detected suggests a great participation of Kaingang women in forming the *quilombo*. Restinga showed a smaller Amerindian maternal fraction, but with Hg proportions suggesting a contribution from Guarani ethnic groups. However, the lack of genetic data from other ethnic groups that lived in the state of Paraná limited the analysis.

Both populations present smaller indices of genetic diversity when compared to the urban populations of Paraná, as expected as populations subjected to strong founder effects. Although these two *quilombos* did not present a significant molecular differentiation, the smaller diversity indices estimated for Restinga suggest a stronger effect of genetic drift in that population.

This work represents the first approach to the allelic architecture and demographic history of Afro-diasporic communities from the region of Paraná, in Southern Brazil, a region where the African roots have been longly hidden. We also estimated for the first time the sub-continental African and Amerindian contributions to Brazilian *quilombos* based on uniparental markers. Furthermore, it adds

new clues to the big picture of the trans-Atlantic traffic of human beings digging deeper into the history of a previously unexplored Brazilian slave region.

Final considerations

Here we investigate for the first time the genetic ancestry of two *quilombos* from Paraná state and perform an in-depth analysis of their uniparental African contributions. In addition, we carried out, for the first time for a Brazilian *quilombo*, a thorough examination of the maternal Amerindian lineages together with the historical data in order to outline the interaction with indigenous groups.

It is essential to highlight that these results do not represent a certainty of the origins of the founders of the two *quilombos* studied. Both mtDNA and Y chromosome lineages are distributed by broad regions of Africa due to different migration events, being only possible to analyze the patterns resulting from the different frequencies observed. However, we offer an estimation of the most likely sources of the lineages observed in Feixo and Restinga along with the existent historical records to contribute to the knowledge of the history of these communities.

Acknowledgments

We are grateful to the *quilombolas* from Feixo and Restinga for their participation in this study, specially to their leaderships Adão Gomes dos Anjos, Adriana da Silva Alexandre, and Claudia Ferreira Santos Rocha for their collaboration. We are also grateful to Professor C. M. Bravi for assistance with mtDNA genotyping protocols.

REFERENCES

- Aidoo L (2018) Slavery Unseen: Sex, Power, and Violence in Brazilian History. Duke University Press, Durham
- Álvarez-Iglesias V, Jaime JC, Carracedo Á, Salas A (2007) Coding region mitochondrial DNA SNPs: Targeting East Asian and Native American haplogroups. *Forensic Science International: Genetics* 1:44–55. <https://doi.org/10.1016/j.fsigen.2006.09.001>
- Anderson S, Bankier AT, Barrell BG, et al (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465. <https://doi.org/10.1038/290457a0>
- Andrews RM, Kubacka I, Chinnery PF, et al (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics* 23:147–147. <https://doi.org/10.1038/13779>
- Bailliet G, Ramallo V, Muzzio M, et al (2009) Brief communication: Restricted geographic distribution for Y-Q* paragroup in South America. *American Journal of Physical Anthropology* 140:578–582. <https://doi.org/10.1002/ajpa.21133>
- Bandelt HJ, Macaulay V, Richards M (2006) Human Mitochondrial DNA and the Evolution of Homo Sapiens, 1st edn. Springer, Berlin/Heidelberg
- Barbieri C, Vicente M, Oliveira S, et al (2014) Migration and interaction in a contact zone: mtDNA variation among Bantu-speakers in Southern Africa. *PLoS ONE* 9:e99117. <https://doi.org/10.1371/journal.pone.0099117>
- Batini C, Coia V, Battaggia C, et al (2007) Phylogeography of the human mitochondrial L1c haplogroup: Genetic signatures of the prehistory of Central Africa. *Molecular Phylogenetics and Evolution* 43:635–644. <https://doi.org/10.1016/j.ympev.2006.09.014>
- Bisso-Machado R, Cátila Bortolini M, Salzano FM (2012) Uniparental genetic markers in South Amerindians
- Bortolini MC, Da Silva WA, De Guerra DC, et al (1999) African-derived South American populations: A history of symmetrical and asymmetrical matings according to sex revealed by bi- and uniparental genetic markers. *American Journal of Human Biology* 11:551–563. [https://doi.org/10.1002/\(SICI\)1520-6300\(1999\)11:4<551::AID-AJHB15>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1520-6300(1999)11:4<551::AID-AJHB15>3.0.CO;2-Z)
- Bucciferro JR (2013) A Forced Hand: Natives, Africans, and the Population of Brazil, 1545–1850. *Revista de Historia Económica / Journal of Iberian and Latin American Economic History* 31:285–317. <https://doi.org/10.1017/S0212610913000104>
- Cardena MMSGSG, Ribeiro-dos-Santos Â, Santos S, et al (2013) Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. *PLoS ONE* 8:e62005. <https://doi.org/10.1371/journal.pone.0062005>
- Carvalho BM, Bortolini MC, dos Santos SEB, Ribeiro-dos-Santos ÂKC (2008) Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations. *Genetics and Molecular Biology* 31:12–22. <https://doi.org/10.1590/S1415-47572008000100002>
- Cigolini AA, Da Silva M (2018) Comunidades Remanescentes Quilombolas: Iconografias e circulações na comunidade da Restinga – Lapa-Pr, Brasil. *Revista Geografar* 13:98–118
- Coelho M, Sequeira F, Luiselli D, et al (2009) On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evolutionary Biology* 9:80. <https://doi.org/10.1186/1471-2148-9-80>
- Cruciani F, La Fratta R, Santolamazza P, et al (2004) Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. *American Journal of Human Genetics* 74:1014–1022. <https://doi.org/10.1086/386294>
- D'Atanasio E, Trionfetti F, Bonito M, et al (2020) Y Haplotype Diversity of the Dominican Republic: Reconstructing the Effect of the European Colonization and the Trans-Atlantic Slave Trades. *Genome biology and evolution* 12:1579–1590. <https://doi.org/10.1093/gbe/eva176>
- Davis AY (1944) Women, race, and class. Vintage Books, New York
- de Filippo C, Barbieri C, Whitten M, et al (2011) Y-chromosomal variation in sub-Saharan Africa: Insights into the history of Niger-Congo groups. *Molecular Biology and Evolution* 28:1255–1269. <https://doi.org/10.1093/molbev/msq312>

- de Souza AM, Resende SS, de Sousa TN, de Brito CFA (2019) A systematic scoping review of the genetic ancestry of the Brazilian population. *Genetics and Molecular Biology* 42:495–508. <https://doi.org/10.1590/1678-4685-gmb-2018-0076>
- Eltis D, Richardson D (2010) Voyages: The trans-Atlantic slave trade database.
- Fajardo S, Cunha LAG (2021) Paraná: Desenvolvimento e Diferenças Regionais, 1st edn. Atena, Ponta Grossa
- Fortes-Lima C, Verdu P (2021) Anthropological genetics perspectives on the transatlantic slave trade. *Human Molecular Genetics* 30:R79–R87. <https://doi.org/10.1093/hmg/ddaa271>
- Gomes FDS (2015) Mocambos e Quilombos: Uma história do campesinato negro no Brasil, 1st edn. Claro Enigma, São Paulo
- Grugni V, Raveane A, Ongaro L, et al (2019) Analysis of the human Y-chromosome haplogroup Q characterizes ancient population movements in Eurasia and the Americas. *BMC Biology* 17:1–14. <https://doi.org/10.1186/s12915-018-0622-4>
- Gutiérrez H (2006) Donos de terras e escravos no Paraná: padrões e hierarquias nas primeiras décadas do século XIX. *História* (São Paulo) 25:100–122. <https://doi.org/10.1590/s0101-90742006000100005>
- Hall GM (2005) Slavery and African ethnicities in the Americas: Restoring the links. The University of North Carolina Press, North Carolina
- Henn BM, Gignoux C, Lin AA, et al (2008) Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. *Proceedings of the National Academy of Sciences* 105:10693–10698. <https://doi.org/10.1073/pnas.0801184105>
- IBGE (2007) Brasil: 500 anos de Povoamento. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro
- Jobling MA, Tyler-Smith C (2003) The human Y chromosome: An evolutionary marker comes of age. *Nature Reviews Genetics* 4:598–612. <https://doi.org/10.1038/nrg1124>
- Joerin-Luque IA, Augusto DG, Calonga-Solís V, et al (2022) Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil. *Molecular Genetics and Genomics*. <https://doi.org/10.1007/s00438-022-01857-7>
- Joerin-Luque IA, Sukow NM, Bucco ID, et al (2023) Ancestry, diversity, and genetics of health-related traits in African-derived communities (quilombos) from Brazil. *Functional & Integrative Genomics* 23:74. <https://doi.org/10.1007/s10142-023-00999-0>
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281. <https://doi.org/10.7717/peerj.281>
- Kent M, Wade P (2015) Genetics against race: Science, politics and affirmative action in Brazil. *Social Studies of Science* 45:816–838. <https://doi.org/10.1177/0306312715610217>
- Kimura L, Nunes K, Macedo-Souza LI, et al (2017) Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes. *American Journal of Human Biology* 29:1–11. <https://doi.org/10.1002/ajhb.22930>
- Kimura L, Ribeiro-Rodrigues EM, De Mello Auricchio MTB, et al (2013) Genomic ancestry of rural African-derived populations from Southeastern Brazil. *American Journal of Human Biology* 25:35–41. <https://doi.org/10.1002/ajhb.22335>
- Lê S, Josse J, Husson F (2008) FactoMineR : An R package for multivariate analysis. *Journal of Statistical Software* 25:253–258. <https://doi.org/10.18637/jss.v025.i01>
- Lemes RB, Nunes, Kelly, Meyer, Diogo, et al (2014) Estimation of Inbreeding and Substructure Levels in African-Derived Brazilian Quilombo Populations. *Human Biology* 86:276. <https://doi.org/10.13110/humanbiology.86.4.0276>
- Lima-Costa MF, Rodrigues LC, Barreto ML, et al (2015) Genomic ancestry and ethnoracial self-classification based on 5,871 community-dwelling Brazilians (The Epigen Initiative). *Scientific Reports* 5:9812. <https://doi.org/10.1038/srep09812>
- Marrero AR, Silva-Junior WA, Bravi CM, et al (2007) Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *American Journal of Physical Anthropology* 132:301–310. <https://doi.org/10.1002/ajpa.20515>
- Martínez B, Simão F, Gomes V, et al (2020) Searching for the roots of the first free African American

- community. *Scientific Reports* 10:20634. <https://doi.org/10.1038/s41598-020-77608-8>
- Medina LSJ, Muzzio M, Schwab M, et al (2014) Human Y-chromosome SNP characterization by multiplex amplified product-length polymorphism analysis. *ELECTROPHORESIS* 35:2524–2527. <https://doi.org/10.1002/elps.201400020>
- Motti JMB, Schwab ME, Beltramo J, et al (2017) Diferenciación regional de poblaciones nativas de América a partir del análisis de los linajes maternos. *Intersecciones en Antropología* 18:271–282
- Moura C (1993) Quilombos: Resistência ao escravismo, 3rd edn. Ática, São Paulo
- Moura C (1992) História do Negro Brasileiro, 2nd edn. Ática, São Paulo
- Nascimento A (2016) O genocídio do Negro brasileiro: Processo de um racismo mascarado, 2nd edn. Perspectiva, São Paulo
- Navarro-López B, Granizo-Rodríguez E, Palencia-Madrid L, et al (2021) Phylogeographic review of Y chromosome haplogroups in Europe. *International Journal of Legal Medicine* 135:1675–1684
- Ongaro L, Sclar MO, Flores R, et al (2019) The genomic impact of European colonization of the Americas. *Current Biology* 29:3974–3986.e4. <https://doi.org/10.1016/j.cub.2019.09.076>
- Paradis E (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Pereira L, Macaulay V, Torroni A, et al (2001) Prehistoric and historic traces in the mtDNA of Mozambique: Insights into the Bantu expansions and the slave trade. *Annals of Human Genetics* 65:439–458. <https://doi.org/10.1046/j.1469-1809.2001.6550439.x>
- Pierron D, Heiske M, Razafindrazaka H, et al (2017) Genomic landscape of human diversity across Madagascar. *Proceedings of the National Academy of Sciences of the United States of America* 114:E6498–E6506. <https://doi.org/10.1073/pnas.1704906114>
- Poletto MM, Malaghini M, Silva JS, et al (2019) Mitochondrial DNA control region diversity in a population from Paraná state—increasing the Brazilian forensic database. *International Journal of Legal Medicine* 133:347–351. <https://doi.org/10.1007/s00414-018-1886-5>
- Quintana-Murci L, Quach H, Harmant C, et al (2008) Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *Proceedings of the National Academy of Sciences* 105:1596–1601. <https://doi.org/10.1073/pnas.0711467105>
- Raggio AZ, Bley RB, Trauczynski SC (2018) Abordagem histórica sobre a população negra no Estado de Paraná. SEJU, Curitiba
- Resque R, Gusmão L, Geppert M, et al (2016) Male lineages in Brazil: Intercontinental admixture and stratification of the European background. *PLoS ONE* 11:1–17. <https://doi.org/10.1371/journal.pone.0152573>
- Ribas KCS (2019) A constituição de identidade quilombola: um olhar acerca da comunidade quilombola do município da Lapa-PR. Graduate thesis, UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
- Ribeiro D (1995) O povo brasileiro: A formação e o sentido do Brasil. Companhia das Letras, São Paulo
- Rosa A, Brehm A (2011) African human mtDNA phylogeography at-a-glance. *Journal of Anthropological Sciences* 89:25–58. <https://doi.org/10.4436/jass.89006>
- Rowold D, Garcia-Bertrand R, Calderon S, et al (2014) At the southeast fringe of the Bantu expansion: Genetic diversity and phylogenetic relationships to other sub-Saharan tribes. *Meta Gene* 2:670–685. <https://doi.org/10.1016/j.mgene.2014.08.003>
- Ruiz-Linares A, Adhikari K, Acuña-Alonso V, et al (2014) Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. *PLoS Genetics* 10:e1004572. <https://doi.org/10.1371/journal.pgen.1004572>
- Salas A, Richards M, De la Fe T, et al (2002) The making of the African mtDNA landscape. *The American Journal of Human Genetics* 71:1082–1111. <https://doi.org/10.1086/344348>
- Salas A, Richards M, Lareu M-V, et al (2004) The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade. *The American Journal of Human Genetics* 74:454–465. <https://doi.org/10.1086/382194>
- Schaan AP, Costa L, Santos D, et al (2017) mtDNA structure: the women who formed the Brazilian Northeast. *BMC Evolutionary Biology* 17:185. <https://doi.org/10.1186/s12862-017-1027-7>
- Schaan AP, Gusmão L, Jannuzzi J, et al (2020) New insights on intercontinental origins of paternal

- lineages in Northeast Brazil. *BMC Evolutionary Biology* 20:1–9. <https://doi.org/10.1186/s12862-020-1579-9>
- Schönherr S, Weissensteiner H, Kronenberg F, Forer L (2023) Haplogrep 3 - an interactive haplogroup classification and analysis platform. *Nucleic Acids Research* 1–6. <https://doi.org/10.1093/nar/gkad284>
- Schwab ME (2019) Tesis doctoral Marisol Elisabet Schwab 2019 Director : Claudio M . Bravi Co-Director : Graciela Bailliet
- Silva M, Alshamali F, Silva P, et al (2015) 60,000 years of interactions between Central and Eastern Africa documented by major African mitochondrial haplogroup L2. *Scientific Reports* 5:12526. <https://doi.org/10.1038/srep12526>
- Silva WA, Bortolini MC, Schneider MPC, et al (2006) mtDNA haplogroup analysis of Black Brazilian and Sub-Saharan populations: Implications for the Atlantic Slave Trade. *Human Biology* 78:29–41. <https://doi.org/10.1353/hub.2006.0028>
- Soares P, Ermini L, Thomson N, et al (2009) Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *The American Journal of Human Genetics* 84:740–759. <https://doi.org/10.1016/j.ajhg.2009.05.001>
- Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* 38:3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Tavares GM, Reales G, Bortolini MC, Fagundes NJR (2019) Measuring the impact of European colonization on Native American populations in Southern Brazil and Uruguay: Evidence from mtDNA. *American Journal of Human Biology* 31:1–16. <https://doi.org/10.1002/ajhb.23243>
- Trombetta B, D'Atanasio E, Massaia A, et al (2015) Phylogeographic Refinement and Large Scale Genotyping of Human Y Chromosome Haplotype E Provide New Insights into the Dispersal of Early Pastoralists in the African Continent. *Genome Biology and Evolution* 7:1940–1950. <https://doi.org/10.1093/gbe/evv118>
- Underhill PA, Kivisild T (2007) Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annual Review of Genetics* 41:539–564. <https://doi.org/10.1146/annurev.genet.41.110306.130407>
- van Oven M (2015) PhyloTree Build 17: Growing the human mitochondrial DNA tree. *Forensic Science International: Genetics Supplement Series* 5:e392–e394. <https://doi.org/10.1016/j.fsigss.2015.09.155>
- Walsh R (1830) Notices of Brazil in 1828 and 1829 Volume I. Frederick Westley and A.H Davis, London
- Westphalen CM (1972) A introdução de escravos novos no Litoral Paranaense. *Revista de História* 44:139. <https://doi.org/10.11606/issn.2316-9141.rh.1972.131319>

6.4 CAPÍTULO IV: G6PD African variants prevalence correlates with ancestry and is higher in *Quilombos* from Southern Brazil (in preparation)

Iriel A. Joerin-Luque^{1,2*}, Maria Eduarda da Silveira Costa², Priscila Ianzen dos Santos^{2,3}, Ana Cecília Guimarães Alves^{1,2}, Natalie Mary Sukow^{1,2}, Ana Carolina Malanczyn de Oliveira², Aymee Fernanda Gros², Maria Letícia Saatkamp², Victor Dobis Barros², Joana Gehlen Tessaro², Luana Leonardo Garcia², Isabela Dall Ogglio Bucco², Denise Raquel de Moura Bones^{1,2}, Sarah Elisabeth Cupertino², Letícia Boslooper Gonçalves^{1,2}, Angela Rodrigues do Amaral Bispo², Danillo G. Augusto^{1,4}, Rodrigo Coutinho de Almeida^{1,5}, Maria Luiza Petzl-Erler¹, Aline Borsato Hauser⁶, Adriana Inês de Paula⁷, Claudemira Vieira Gusmão Lopes⁸, Marcia Holsbach Beltrame^{1,2}

¹Programa de Pós-graduação em Genética, Departamento de Genética, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ²Laboratório de Genética Molecular Humana, Departamento de Genética, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ³Programa de Pós-graduação em Medicina Interna, Hospital de Clínicas (HC), Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ⁴University of California, San Francisco, Estados Unidos. ⁵Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. ⁶Laboratório Escola de Análises Clínicas, Departamento de Farmácia, Universidade Federal do Paraná, Curitiba, Brasil.

⁷Departamento de Educação Física, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, 81531-990, Brasil. ⁸Setor Litoral, Universidade Federal do Paraná (UFPR), Matinhos, Paraná, Brasil.

*Correspondence E-mail: irieljoerin@gmail.com

Abstract

Glucose-6-phosphate dehydrogenase deficiency (G6PDd) is a common red blood cell enzymopathy linked to protection against malaria and can trigger hemolytic crises induced by certain food and drugs. Here, we studied the prevalence of the main African G6PDd variants *G6PD*A* and *G6PD*A-* in urban (Central-west $n = 294$, Southeast $n = 38$, South $n = 71$) and Amerindian ($n = 36$) populations from Brazil and two African-derived communities (*quilombos*) from the South of Brazil ($n = 159$), Restinga and Feixo. Besides, we studied 25 additional single nucleotide variants (SNVs) of the *G6PD* gene in the urban and Amerindian datasets. We found a cline of frequency between the Brazilian regions (Central-west *G6PD*A* = 6.24%, *G6PD*A-* = 3.53%; Southeast *G6PD*A* = 5.45%, *G6PD*A-* = 1.79%; South *G6PD*A-* = 1.77%), lower and probably allochthonous prevalence in Amerindians (*G6PD*A-* = 1.74%), and higher frequency in the *quilombos* (*G6PD*A* = 10.95%, *G6PD*A-* = 1.82%). In Brazil, the prevalence of these variants correlates positively with African ancestry and Black phenotype. It could have implications for other complex traits more prevalent in Blacks, such as type 2 diabetes. Other severe *G6PD* variants were not detected, such as *G6PD*B-* (Mediterranean), Kaiping, and Montalbano.

Keywords: glucose-6-phosphate dehydrogenase, African diaspora, erythrocyte, diabetes

Introduction

Glucose-6-phosphate dehydrogenase deficiency (G6PDd, MIM: 300908) is an X-linked enzymopathy caused by the presence of deleterious variants in the *G6PD* gene (HOWES et al., 2013). *G6PD* is located proximal to the telomeric region in the long arm of the X chromosome Xq28 (GÓMEZ-MANZO et al., 2016). Due to its X-linked recessive mode of inheritance, the G6PDd symptoms are usually more severe in males and homozygous females (DALLOL et al., 2012). However, compound heterozygosity or skewed X-inactivation of mutated chromosomes can also produce symptoms in females (TIBERIO, 1994). Glucose-6-phosphate dehydrogenase (G6PD) is a crucial enzyme in all cells as it catalyzes the first reaction of the pentose phosphate pathway (PPP), allowing the conversion of glucose-6-phosphate to 6-phosphogluconolactone. This reaction reduces nicotinamide adenine dinucleotide phosphate (NADP) to NADPH, a fundamental cofactor for protecting cells against physiologically high levels of oxidative damage (MASON; BAUTISTA; GILSANZ, 2007). In the mature erythrocyte, the PPP is the only source of NADPH, so in the absence of normal levels of G6PD, oxidative damage ultimately results in hemolysis. Thus, G6PD deficient individuals are asymptomatic until the consumption of specific substances whose process of assimilation generates oxidizing agents, as well as infection. These events can then precipitate three main clinical outcomes: neonatal jaundice, chronic non-spherocytic hemolytic anemia (CNSHA), and acute hemolytic anemia (AHA) (MASON; BAUTISTA; GILSANZ, 2007).

More than 1,300 genetic variants have been described in the *G6PD* gene in humans (GECK; POWELL; DUNHAM, 2023). Nevertheless, only 400 have a functional interpretation (GECK; POWELL; DUNHAM, 2023). So far, 221 variants have been associated with enzymatic deficiency (GECK; POWELL; DUNHAM, 2023). Most of them alter the amino acid sequence, but there are also a few regulatory variants (LUZZATTO; ALLY; NOTARO, 2021).

It is known that the frequency of deleterious *G6PD* alleles rises in malaria-endemic regions in a balanced way, as they confer relative protection against severe infection by the parasite (LUZZATTO; USANGA; REDDY, 1969). Furthermore, different *G6PD* variants have emerged in unrelated populations, representing an example of convergent evolution in response to malaria pressure in humans (VERLAG; SPRING, 1992). The malaria resistance mechanism is still unknown (LUZZATTO; ALLY; NOTARO, 2021). One hypothesis is that, as the red cells infected by malaria produce high amounts of oxidizing agents as hemoglobin is degraded in the parasites' vacuoles, deficient cells are rapidly damaged and phagocytosed along with the parasites (MASON; BAUTISTA; GILSANZ, 2007). Hence, the global distribution of G6PDd overlaps the geographic distribution of *Plasmodium* sp., supporting the selective pressure hypotheses for maintaining the deleterious alleles. The highest prevalence of this deficiency has been observed in Sub-Saharan Africa and the Arabian Peninsula (HOWES et al., 2012). Aminoquinoline-derived anti-malarial drugs are, unfortunately, some of the hemolytic agents (DOMBROWSKI et al., 2017). In Brazil, *Plasmodium* sp. occurs in the northern region. As a consequence, malaria is endemic in this region. Its transmission occurs residually in the rest of the country (PREFEITURA DE SÃO PAULO, 2018).

*G6PD*A* and *A-* are some of the most common variants associated with enzyme deficiency. *G6PD*A-*, along with *G6PD*Med*, are the main ones responsible for the occurrence of hemolytic events (MONTEIRO et al., 2014). At the molecular level, *G6PD*A-* contains two mutations: rs1050829T>C, which defines the *A* variant, in addition to a second mutation which is almost always rs1050828C>T. Alternatively to rs1050828C>T, the mutation rs76723693A>G (called the Betica) or rs137852328C>A also produces enzymes with the characteristics of *G6PD*A-* (BEUTLER et al., 1989).

In Brazil, the testing of G6PD deficiency is not massive nor mandatory before the treatment with anti-malarial drugs. Hence, monitoring G6PD variants in the population is essential, especially in people of African descent, where the *A* and *A-* variants are more prevalent. *G6PD*A* and *A-* variants are presumed to have arisen only once, with *A* being the most ancient and *A-* the most recent (TISHKOFF et al., 2001).

According to the World Health Organization, *G6PD* variants are classified into four categories based on variant activity and severity of anemia. However, a recent study has reviewed the outcomes of the variants described in papers and databases, updating the phenotypic interpretation of several variants (GECK; POWELL; DUNHAM, 2023). It is worth mentioning that, for the same genetic variant, the *G6PD* activity and clinical manifestation can vary widely between individuals and over time (GECK; POWELL; DUNHAM, 2023). This last situation is particularly relevant for heterozygous females, where X-inactivation as red blood cell turnover can vary over time (GECK; POWELL; DUNHAM, 2023).

However, *G6PD* variants are not only linked to the mentioned outcomes. Their minor alleles have also been associated with other phenotypes (GECK; POWELL; DUNHAM, 2023), such as the variation of glycated hemoglobin levels (HbA1c). HbA1c measurement is used to assess glycemic control and to diagnose type 2 diabetes (T2D) (WHEELER et al., 2017). In a recent transethnic genome-wide meta-analysis, the *G6PD* G202A variant (rs1050828*A) was shown to explain 14.4% of the variance in HbA1c, associating with a decrease of 0.81%-units in HbA1c in hemizygous men and of 0.68%-units in homozygous women (WHEELER et al., 2017). As suggested by Wheeler et al., the potential of the *G6PD**A- variant to shorten red blood cells' lifespan could reduce the HbA1c levels independently of the blood glucose level. Hence, the authors estimated that the presence of that variant could lead to approximately 2% of African-American adults with T2D remaining undiagnosed when screened with glycated hemoglobin (WHEELER et al., 2017). In Brazil, as in other African ancestry populations, Blacks suffered more diabetes and diabetes-linked conditions, such as obesity and hypertension, than Whites (BRITO; LOPES; ARAÚJO, 2001), and the measurement of HbA1c is one of the criteria to diagnose T2D, together with fasting glucose levels (FG) (SOCIEDADE BRASILEIRA DE DIABETES, 2020). Furthermore, diabetic ketoacidosis (DKS), a severe complication in diabetic patients, can also trigger a hemolytic crisis in *G6PDd* individuals (ANSARI et al., 2022).

For those mentioned above, the study of the *G6PD* variants in admixed populations such as Brazilians is required, especially in individuals with African ancestry, which, in Brazil, represent almost the whole population (JOERIN-LUQUE et al., 2022). In this work, we analyzed the frequency of 27

G6PD SNVs, including the *G6PD*A* and A- variants, in urban and Amerindian populations from the Central-west, Southeast, and South of Brazil, comparing their frequency to X-chromosome ancestry and phenotype. In addition, we assayed the prevalence of *G6PD*A* and A- variants in two *Quilombos* from Southern Brazil, Feixo and Restinga.

Subjects and Methods

Human subjects

This work has been approved by the national ethics committee for research with human subjects (CONEP- Comissão Nacional de Ética em Pesquisa, protocol number 02727412.4.0000.0096), and from the local institutional review board (CEP/SD- Comitê de Ética em Pesquisa em seres humanos, Setor de Ciências da Saúde, CAAE number 42761521.9.0000.0102), according to Brazilian Federal Laws. All participants gave written informed consent for sample and anthropometric measures collection and interviews. All data analyses were performed anonymously.

In Brazil, the Institute of Geography and Statistics implements phenotypic categories to population censuses, mainly based on skin color and other phenotypic traits (IBGE, <https://www.ibge.gov.br>). This classification has five categories: Amerindian, White, Black, Admixed, and Yellow (in Portuguese: *Indígena, Branca, Preta, Parda, and Amarela*). In this work, we used the IBGE term “Pardo,” which is the Portuguese term used by IBGE, meaning “Admixed” for self-declared individuals and, for individuals hetero-identified (phenotypic classification performed by the researcher at the moment of sample collection), we merge the categories Black and Admixed in a unique category, “Blacks,” as demanded by the Brazilian Black movement (DAFLON; CARVALHAES; FERES JÚNIOR, 2017).

Sample collection and DNA extraction

Blood samples from 446 individuals from the Central-Western ($n = 249$), Southeastern ($n = 38$), and Southern ($n = 71$) geographical regions of Brazil were previously collected and described (AUGUSTO et al., 2021; FARIAZ et al., 2018). For that set, samples were subdivided into Blacks, Whites, and Amerindians based on self-classification, phenotype, and a careful assessment of individual family history. Samples from the Amerindian Kaingang ($n = 10$), Guarani Mbya ($n = 7$), and Guarani Kaiowa ($n = 7$) ethnic groups were also previously collected (Fig. S1) (PETZL-ERLER; LUZ; SOTOMAIOR, 1993; TSUNETO et al., 2003). Samples from two *Quilombos* from the state of Paraná, in Southern Brazil, Feixo ($n = 111$) and Restinga ($n = 46$), were collected for this study. Briefly, blood samples were collected separately for DNA extraction from the buffy coat using a commercial kit (Wizard® Genomic DNA Purification Kit, Promega) to measure glycated hemoglobin and for serum separation to perform additional laboratory analyses (including blood glucose and whole hemogram). After DNA extraction, all samples were aliquoted to a final concentration of 20 ng/uL for the PCR experiments.

Along with the blood samples, we collected ethnographic information (name, sex, age, place of birth, and self-declared ancestry), including family (self-declared ancestry and place of birth of parents) and medical history (diagnosed diseases and undiagnosed symptoms, as well as medication usage). Questionnaires were based on the official ones applied by the Brazilian National Health Research (Pesquisa Nacional em Saúde -PNS-, <https://www.gov.br/>). Blood glucose levels were measured by the capillary method using a G-tech Lite® kit and from serum with the enzymatic method (Hexokinase) in an automatic Labmax 400 equipment. Hematological profiles were analyzed from smear blood in a microscope (Panthera - Red Dot Award 2017) and from whole blood by the impedance method with an ABX MICROS 60 equipment and analyzed according to the Brazilian laboratory standards (available at: <https://www.pncq.org.br/uploads/2019/VNH2019.pdf>).

Genotyping

Urban and Amerindian datasets were genotyped as described elsewhere (AUGUSTO et al., 2021; FARIAS et al., 2018). Briefly, the platform Illumina CoreExome-24 v1.1 (Illumina, San Diego, CA) was used for microarray hybridization. A standard quality control procedure was performed as described elsewhere (ANDERSON et al., 2010). Additionally, filtering of SNPs from the non-pseudoautosomal region of the X chromosome present in the dataset (positions 2699520-154931044 according to the human genome assembly GRCh37) was performed with PLINK v1.9.0 (CHANG et al., 2015). Additional filters were used to eliminate SNPs and samples with less than 90% call rate and to match the reference populations datasets (AUTON et al., 2015; CANN et al., 2002). Finally, 3625 SNPs were used for X chromosome ancestry and *G6PD* SNVs analysis.

The *Quilombola* dataset was genotyped by allele-specific PCR with primers designed for the African *G6PD* variants rs1050828 and rs1050829 for this study (Table S1, Figure S2) and visualized using agarose gel electrophoresis.

Ancestry and population genetics analysis

X chromosome ancestry was estimated using the maximum likelihood algorithm implemented no software ADMIXTURE v1.3, using an unsupervised approach testing for clustering values from 3 to 12 and the command --haploid="male:23" for the correct analysis of X chromosome SNPs both in females and males (ALEXANDER; NOVEMBRE; LANGE, 2009). Allele and genotype frequencies from all datasets and Hardy-Weinberg equilibrium deviation *p*-values were calculated with PLINK v1.9.0. Subsequent analyses were performed in the R programming environment (R CORE TEAM, [s.d.]). Extraction of unambiguous *G6PD* haplotypes from males, haplotype parsimony networks, and indices of haplotype (*H*) and nucleotide diversity (π) were calculated with the R package *pegas* (PARADIS, 2010). The linkage phase from female's diploid genotypes was inferred using the expectation-maximization algorithm implemented in the R function *haplo_em* from *haplo.stats* package (SINNWELL; SCHAIID, 2023). G6PD haplotype frequencies in reference populations from the 1000 Genomes dataset were extracted using the *LDhap* function from *LDlinkR* package (MYERS;

CHANOCK; MACHIELA, 2020). Linkage disequilibrium was calculated with *LDheatmap* function from the *LDheatmap* package (SHIN et al., 2006). Correlation tests were performed with the R package *ggpubr* (KASSAMBARA, 2023). Graphical visualization was made using the *ggplot2* package (WICKHAM, 2009).

Results and discussion

***G6PD* variants prevalence in urban admixed populations**

All variants evaluated were in Hardy-Weinberg equilibrium in the evaluated populations (Table S2). In the urban dataset, we assessed the frequency of 27 SNVs in the *G6PD* gene according to Brazilian geographical region (Central-west, Southeast, and South) and hetero-classified phenotype (or self-declared ancestry, in the case of Brazilian Amerindians) (Figure 1, Table S2 and S3).

Only 8 out of 27 variants were polymorphic in the Brazilian population (Figure 1, Table S2 and S3). Among them, the 3' UTR variant rs1050757*G, of African origin, and the synonymous allele rs2230037*A were present in all the groups tested. Despite that little is known about not exonic or intron-exon boundary *G6PD* SNPs, both variants, in combination with the rs2071429*C allele (not tested here), have been previously associated with acute hemolytic anemia due to favism in Palestinian individuals without other exonic or exonic-intronic boundary variants known to cause G6PDd (SIRDAH et al., 2017). These patients showed 26-78% G6PD enzymatic activity (SIRDAH et al., 2017). In our study, 34 hemizygous males (21.8% of Blacks, 16.9% of Whites, and 15.4% of Amerindians) and 11 homozygous females showed both rs1050757*G and rs2230037*A alleles. Both alleles were more frequent in Blacks than Whites (Chi-square *p*-value = 4.15E-06 and 0.03117, respectively).

The rs2230036*A allele, a synonymous variant without known phenotypic consequences, was present in all groups (Figure 1, Table S2 and S3). Another two benign variants, the rs5986990*A and rs77214077*A alleles of African origin, were also detected in the urban dataset. Whereas the

rs5986990*A was detected in all groups, the rs77214077*A allele was not detected in the Southeast nor in Amerindians, presenting a higher frequency in Blacks than in Whites (Figure 1, Table S2 and S3).

On the other hand, the Seattle variant, rs137852318*A, reported to cause G6PD deficiency and acute hemolytic anemia (GECK; POWELL; DUNHAM, 2023), was detected at low frequencies (0.22-0.88 %) in the Central-west and Southern regions and was only observed in a low proportion of White individuals (0.44%). The *G6PD**B- variant (also known as *G6PD**Med, rs5030868*A) of Mediterranean origin, known to cause a severe G6PD deficiency, as well as other pathogenic variants tested (the Montalbano - rs74575103*A -, and the Kaiping variant - rs72554664*A -) were not detected in this dataset (GECK; POWELL; DUNHAM, 2023). *G6PD**B- has been previously found in Salvador, Bahia, in the Brazilian Northeast, and Rio Grande do Sul, in Southern Brazil (DE MOURA NETO et al., 2008; WEIMER et al., 1993).

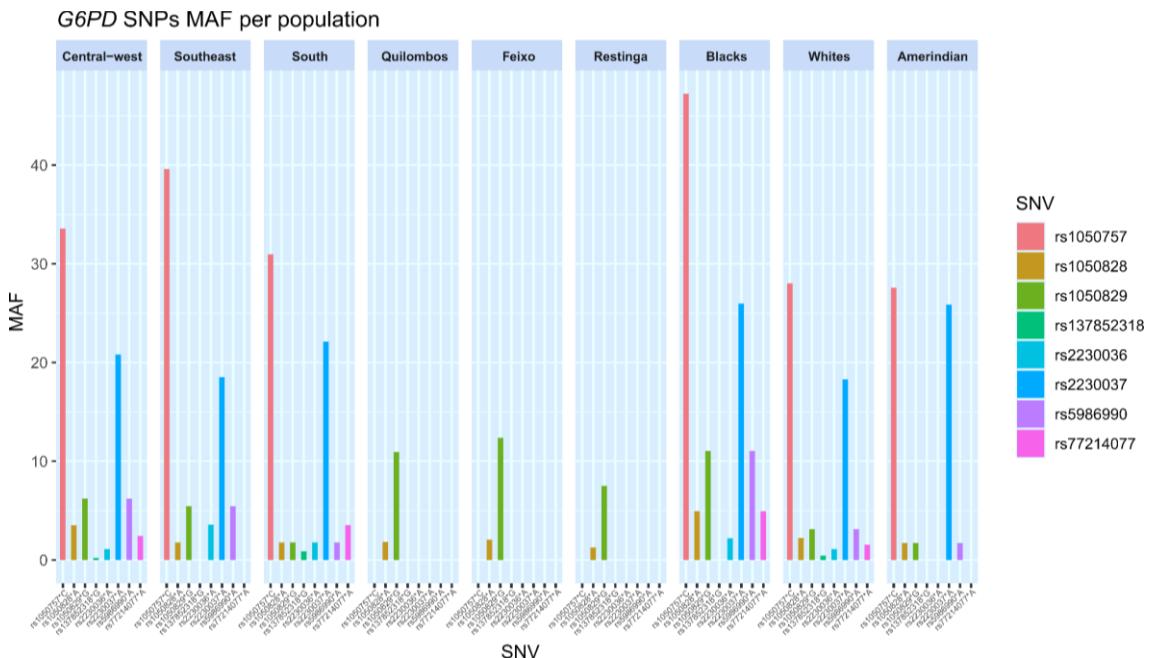


Figure 1. Minor allele frequency (MAF, in percentage) of *G6PD* single nucleotide variants (SNVs) in Brazilians according to hetero-classified phenotype (Blacks and Whites) or self-declared identification (Amerindian and *Quilombolas*). Monomorphic SNVs are not shown.

G6PD*A and A- prevalence correlates with ancestry and are more frequent in Black Brazilians

The *G6PD*A* and *A-* variants (rs1050829*G and rs1050829*G + rs1050828*G, respectively) were present in all the groups tested (Figure 1 and 3, Table S2 and S3). However, we observed a higher frequency of *G6PD*A* in Blacks than Whites or Amerindians (*p*-value = 0.00007383 and 0.02902, respectively). We also observed a positive correlation between African ancestry and the presence of these variants (Figure 2A and B), whereas European ancestry is negatively correlated (Figure 2C and D). In addition, and as the combination with other alleles not tested here can produce G6PD variants with the characteristics of *G6PD*A-* it is important to reinforce that those individuals carrying the *G6PD*A* variant could also have a more severe deficiency if they carry those additional alleles (BEUTLER et al., 1989). In this way, and considering the complex implications of these variants in phenotypic traits linked to health, such as T2D complications and drug and food responses, these results highlight the need to test G6PD enzymatic activity in the Brazilian population, especially Blacks.

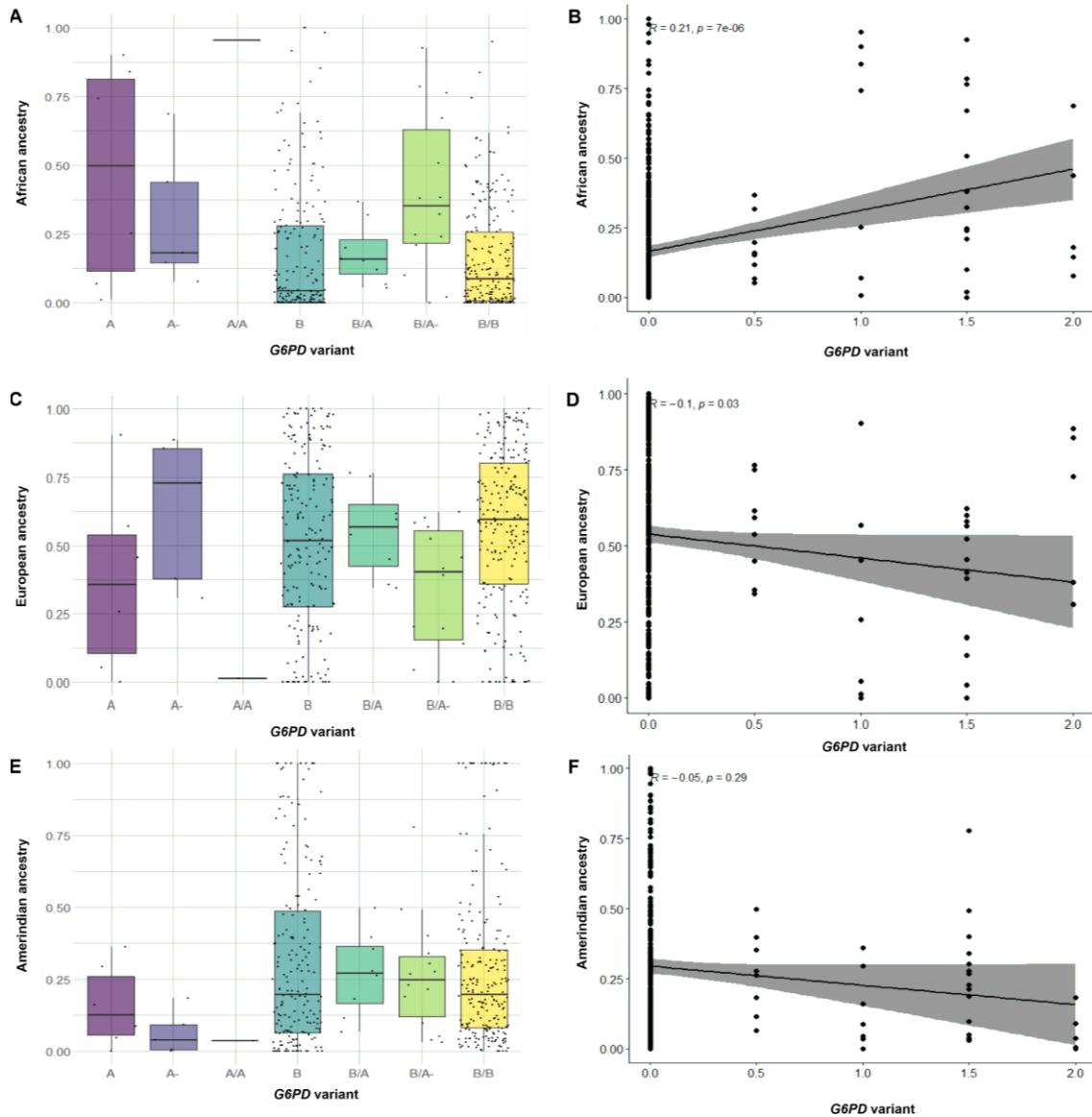


Figure 2. The prevalence of *G6PD* variants A and A- correlates positively with African ancestry and negatively with European ancestry in the Brazilian urban dataset. A) and B): African ancestry percentages according to the *G6PD* genotype. C) and D): European ancestry percentages according to the *G6PD* genotype. E) and F): Amerindian ancestry percentages according to the *G6PD* genotype. For the correlation analysis, *G6PD* genotypes were ranked as follows: 0 = wild type B or B/B, 0.5 = heterozygous B/A individuals, 1 = hemizygous A or homozygous A/A, 1.5 = heterozygous B/A-, 2 = homozygous A-/A-. R and p-values were calculated with Spearman's correlation method.

In Amerindians, the *G6PD**A- variant has possibly been introduced by admixture

Only one self-declared Amerindian female presented the *G6PD* A- variant in heterozygosis in combination with the benign rs5986990*A African allele. This participant also presented low European

and African ancestral contributions (20 and 1.9%, respectively) in her X chromosome, being that variant possibly introduced by admixture.

On the other hand, Amerindians from Guarani and Kaingang ethnic groups living in reservations did not present any *G6PD* variant other than the rs1050757*G and rs2230037*A alleles, which were in high frequency (Figure 1, Table S3). The detected alleles have not been described as causing G6PDd. Among the Amerindian individuals presenting these variants, 69.2% did not show signals of European or African ancestries. Thus, these variants are likely part of the common Amerindian genetic variation. The haplotype containing these alleles was the second most common in the Admixed Brazilian population, Amerindians, and Europeans (Figure 3 and Table S4, haplotype number 2 in the whole dataset). On the other hand, that haplotype was the most frequent in Africans (Table S4).

Even though these alleles alone have not been linked to G6PDd, it would be essential to test for the presence of the rs2071429*C allele (previously associated with AHA in combination with rs1050757*G and rs2230037*A), as well as for G6PD enzymatic activity in Brazilian populations (SIRDAH et al., 2017).

The studies of G6PDd variants prevalence in Amerindian populations are scarce (FERRELL et al., 1980; WEIMER et al., 1998). Only among the Chilean Aymara a private protein variant has been described (G6PD Aymara). However, it was not subjected to a detailed characterization (FERRELL et al., 1980). Other populations presented predominantly the wild-type gene with low frequencies of the G6PD*A variant, probably introduced by admixture (WEIMER et al., 1998).

Haplotype diversity indices are higher in Blacks

Nei's haplotype (H) and nucleotide diversity (π) indices were calculated based on male (obtained directly from the urban and Amerindian datasets) and female haplotypes (obtained through inference with the Expectation-Maximization algorithm).

As expected, considering that Blacks have, on average, higher African ancestry and that in the African continent exists the highest human genetic diversity, that group showed the highest diversity

indices from the whole dataset (Table 2, Figure 3, Table S4) (GOMEZ; HIRBO; TISHKOFF, 2014). It is important to highlight here that this higher genetic diversity and the long evolutionary history of African populations compared with those from out of Africa implies that there is a higher probability that Blacks also have additional variants still unknown, as well as other variants not tested in this work that could be associated to G6PDd (GOMEZ; HIRBO; TISHKOFF, 2014). It is known that additional variants are frequently detected in different populations. For example, in Salvador, Bahia (in the Brazilian Northeast), in a study developed with neonates (354 females and 301 males), an additional missense variant (153764222:A>T, c.197T>A) was found in combination with *G6PD**A- (3 females and 1 male) and also with rs1050828*A alone (1 female). The variant, called “Bahia”, as detected when the *G6PD* exon 4 (where rs1050828 is located) was sequenced, and it was not present among neonates without deficiency (313 females and 276 males) of the same sample, suggesting that it could be linked to deficiency (DE MOURA NETO et al., 2008).

In our dataset, the haplotype containing the *G6PD**A- variant (number VI in Figure 3, Table S4) in urban admixed and Amerindian Brazilians also includes the rs5986990*A allele. That haplotype occurs in African populations from the 1000 Genomes Project at 13.12%, being the only haplotype containing the *G6PD**A- variant (Table S4). In European populations from the same Project dataset, the haplotype containing *G6PD**A- was not detected (Table S4), confirming the African origin of this haplotype in Brazilian populations. These alleles are in complete linkage disequilibrium in the Brazilian populations studied here (Figure S2).

Group	N	Nhap	H	Variance	π	Variance
Blacks	179	7	0.6488	0.0008	0.0565	0.0014
Whites	440	9	0.4483	0.0007	0.0340	0.0007
Amerindians	57	3	0.4204	0.0033	0.0337	0.0007

Table 2. Diversity indices in the three Brazilian phenotypic groups studied in the urban male and female dataset. N represents the total number of chromosomes studied. All parameters were calculated excluding individuals with missing genotypes. Nhap: number of haplotypes, H: Nei's haplotype diversity index, π : nucleotide diversity.

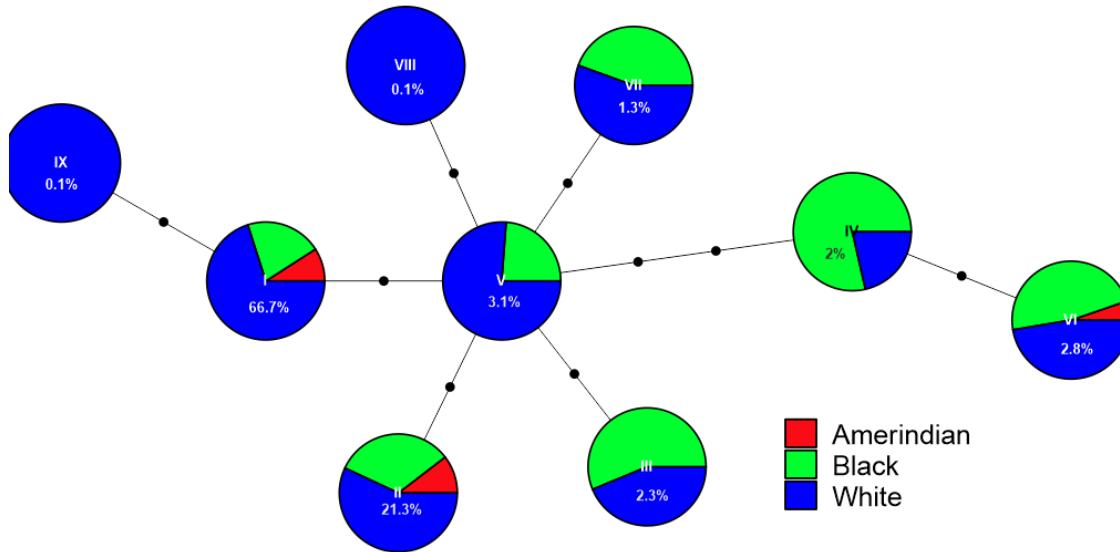


Figure 3. *G6PD* haplotypes in Blacks ($n = 117$) (A), Whites ($n = 285$) (B), and Amerindians ($n = 36$) (C) from the Brazilian urban dataset. Haplotypes are identified with Roman numerals (see Table S4). Mutational steps between each haplotype are shown as small black balls. Each haplotype's frequency (in percentage) is shown inside the node. Haplotypes are shown in Table S4.

***G6PD*^A* and *A-* are more frequent in *Quilombos* from Southern Brazil than in their urban counterpart**

In the *quilombos* Feixo and Restinga, we genotyped 159 individuals (Feixo, $N = 113$; Restinga, $N = 46$). Considering both communities, 29 individuals carry at least one *G6PDd* allele (18.24%). We detected 6 heterozygous individuals carrying the *G6PD*^A* variant and 2 carrying the *G6PD*^{A-}* variant. We also observed 5 hemizygous (11.36% of males) and one homozygous *G6PD*^A* individual (0.86% of females). The allele frequency was 10.95% *G6PD*^A* and 1.83% *G6PD*^{A-}* (Figure 1 and Table S5).

In both communities, 8.2% of the participants declared themselves Whites, 55.3% Pardos, and 36.5% Blacks. No association with self-declared ancestry was detected in this population. When compared, Feixo and Restinga presented homogeneous frequencies of both variants (Chi-square p -value = 0.3746 and 1, respectively, for the *G6PD*^A* and *G6PD*^{A-}* variants, Table S5). However, when we compared the frequency of these variants with the urban Southern Brazilian dataset (which is mainly formed by participants from the state of Paraná, where these *quilombos* are located), we observed that a

higher frequency of *G6PD**A exists in the *quilombos* (Chi-square *p*-value = 0.006485). *Quilombos* are populations formed by Africans and their descendants, thus having a higher African ancestry when compared with their urban counterparts (JOERIN-LUQUE et al., 2023). That can explain the higher frequency of the African *G6PD* variants. However, each *quilombo* has a particular evolutionary history with different contribution levels from neighbor populations of diverse ancestries (JOERIN-LUQUE et al., 2022). It is worth noting that the Southern Brazilian sample analyzed here was mainly formed by participants recruited in a public, academic context, which, in Brazil, mainly belong to the middle to upper-income population and is predominantly White and of European ancestry (average 71.1% European and 16.1% African ancestries) (JOERIN-LUQUE et al., 2022). Therefore, studying these variants in other, more representative samples of Southern Brazil remains fundamental.

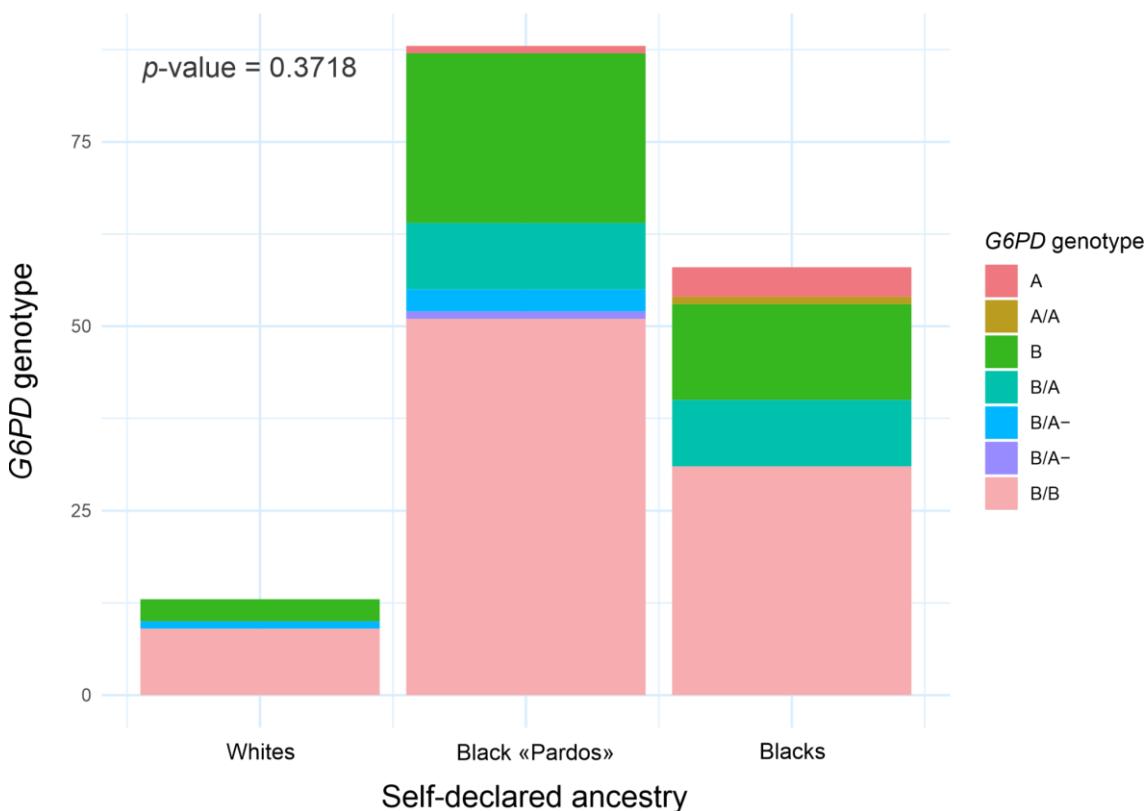


Figure 4. In *quilombos*, there is no association between *G6PD**A and A- variants and self-declared ancestry. The chi-square *p*-value is shown.

As it was highlighted when considering the urban Brazilian Black population, we cannot exclude the possibility that other variants of African origin (as well as from other ancestries, in a lesser extension) could be present in this population, increasing the severity of the G6PD deficiency caused by *G6PD*A* and *A-*.

In addition to the present work, only nine other *quilombos* from the Brazilian Amazon region (Northern Brazil) were studied for *G6PD* variants (OLIVEIRA et al., 2018). In those communities, the frequencies observed for the *G6PD*A* and *A-* variants were 10.4% and 6.1%, respectively (OLIVEIRA et al., 2018). While the frequency of the *G6PD*A* variant detected in Feixo and Restinga was similar to that observed in the Amazonian *quilombos*, the frequency of the *G6PD*A-* in the Southern Brazilian *quilombos* was lower. As previously mentioned, each *quilombo* has its unique evolutionary history, so they are not expected to be genetically similar. Thus, the monitoring of these variants' prevalence in every Brazilian *quilombo* and the wide circulation of their implications for health is essential.

The implications of *G6PD* variants for the health of the *quilombolas* from Feixo and Restinga

The deficiency of G6PD can be clinically silent until the consumption of certain food, drugs, and/or the occurrence of infection, among other events, that can trigger hemolytic crises of variable severity in G6PDd individuals (BEUTLER et al., 1989; GECK; POWELL; DUNHAM, 2023).

In Feixo and Restinga, no participant reported having hemolytic anemia. However, only the general term “anemia” was included in the interview. Among the participants that reported to have anemia, only 6 female participants (representing 15.8% of the participants who reported to have anemia in the genotyped dataset) carried one allele of the tested *G6PD* variants (four heterozygous B/A and one B/A-). However, *G6PD*A* and *A-* in heterozygosis are not expected to cause severe hemolytic crisis (BROEK; HEYLEN; AKKER, 2016). At the moment of sample collection or interview, none of the participants presented evident symptoms of severe hemolytic anemia, such as extreme tiredness, dizziness, jaundice, weakness, etc. (PHILLIPS; HENDERSON, 2018). So, as expected, none of the participants with one nor two copies of the G6PDd alleles presented a hemogram with the typical “Bite

cells” observed during the hemolytic crisis, nor a low level of cellular hemoglobin concentration mean (C.H.C.M.) at the moment of sample collection (BROEK; HEYLEN; AKKER, 2016). No association was detected between *G6PD**A or A- variants and low levels of Hemoglobin, C.H.C.M, or Red Blood Cell count (RBC) (Figure 5). Nevertheless, we cannot discard the hemolytic crises not reported by the participants before or after the interviews or sample collection. In addition, as the more prevalent variant detected in the *quilombos*, *G6PD**A is known to cause mild deficiency, and the severity of G6PDd can vary broadly between individuals, it is possible that the participants only had mild symptoms if some (BROEK; HEYLEN; AKKER, 2016).

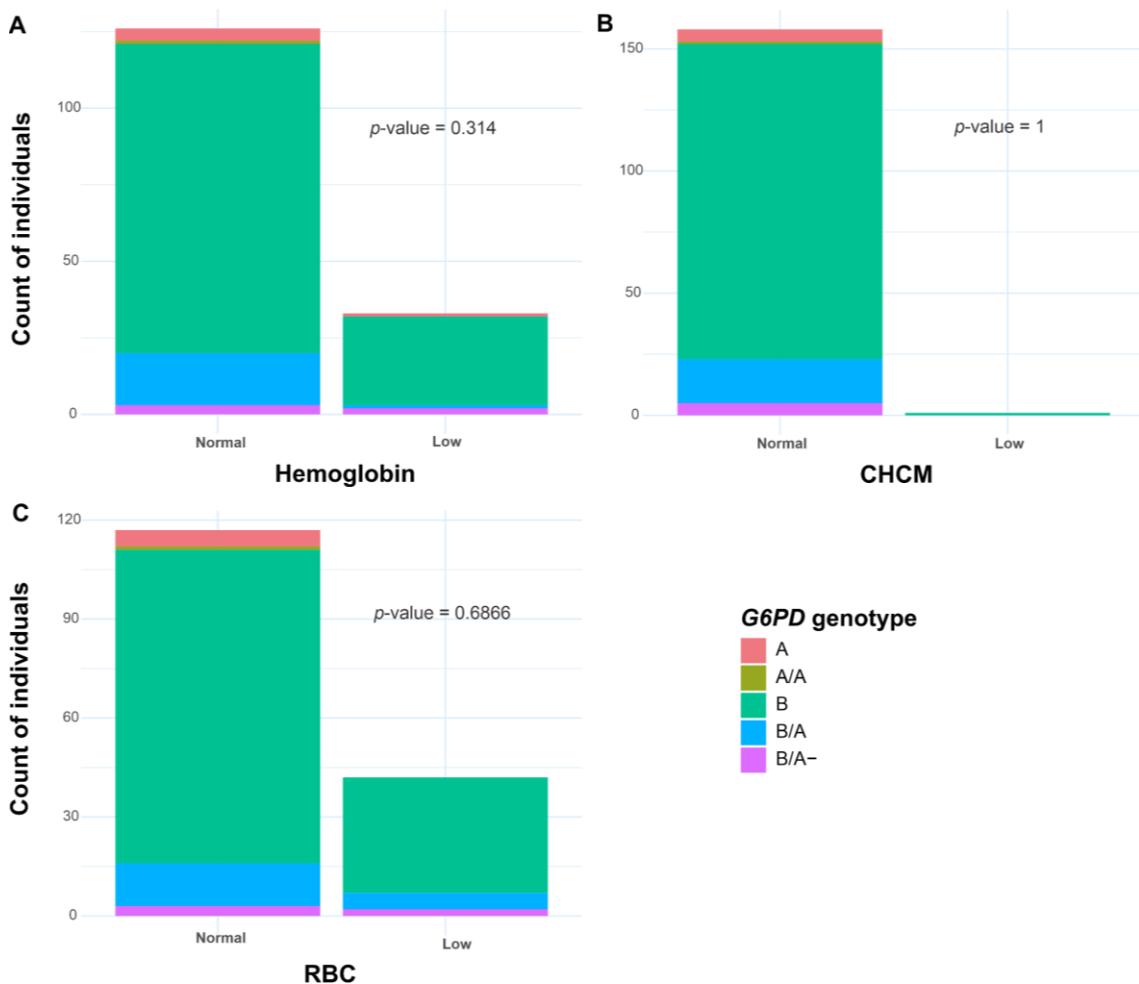


Figure 5. No association was detected between the presence of *G6PD**A and A- and low levels of hemoglobin (A), cellular hemoglobin concentration mean (CHCM), and red blood cell count (RBC). Fisher's exact test *p*-values are shown.

Epidemiological data suggested that G6PDd could be a risk factor for diabetes (CARETTE et al., 2011; ZHANG et al., 2010). In both *quilombos* ($n = 227$), forty individuals declared to have a medical diagnosis of diabetes (17.62%). Among the participants genotyped for *G6PD* variants in this study ($n = 159$), 32 declared to be diagnosed diabetic (20.12%). Among these patients, there was only 1 homozygous A/A (3.12%) and 6 heterozygous B/A (18.75%). No association between the presence of the *G6PD**A variant and diabetes was found (Chi-square test p -value = 0.4151).

On the other hand, as the presence of the rs1050828*A allele has been associated with a reduction of HbA1c levels independently of blood glucose levels in African Americans, we decided to test if some association could be detected in our sample (WHEELER et al., 2017). According to Wheeler et al., an imperfect correlation of HbA1c and glycemia may indicate the impact of reduced RBC lifespan on HbA1c in individuals with the rs1050828*A allele (WHEELER et al., 2017). Thus, we decided to evaluate the level of correlation between blood glucose level and HbA1c in our sample. We observed that the correlation between those two parameters was not perfect and that the value of that correlation was reduced when comparing individuals carrying one or two copies of the A and A- variants to individuals with the wild type B and B/B genotypes (Figure 6). Then, we evaluated if a correlation existed between the presence of the *G6PD* variants and the levels of HbA1c. To do that, the genotypes of the individuals were ranked according to the severity of the variant carried, as follows: 0 = wild type B or B/B genotypes, 0.5 = heterozygous B/A individuals, 1 = hemizygous A or homozygous A/A individuals, 1.5 = heterozygous B/A- individuals. We did not observe any significant correlation between HbA1c levels and the presence of the *G6PD**A and A- variants. However, considering our small sample size, we cannot discard that a correlation could be detected in a larger sample.

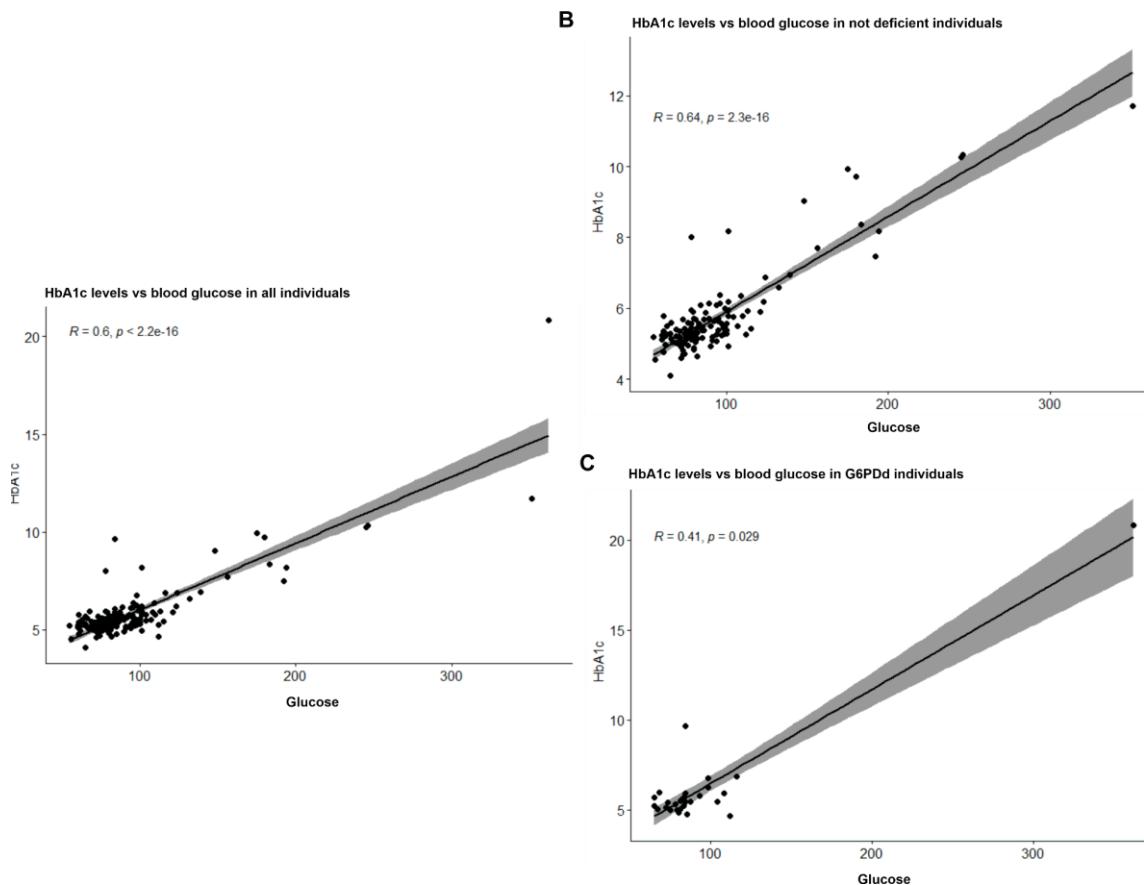


Figure 6. The correlation of blood glucose levels (Glucose) and glycated hemoglobin (HbA1c) is reduced in G6PD deficient individuals with the A and A- variants from *quilombos*. R and p-values were calculated with Spearman's correlation method.

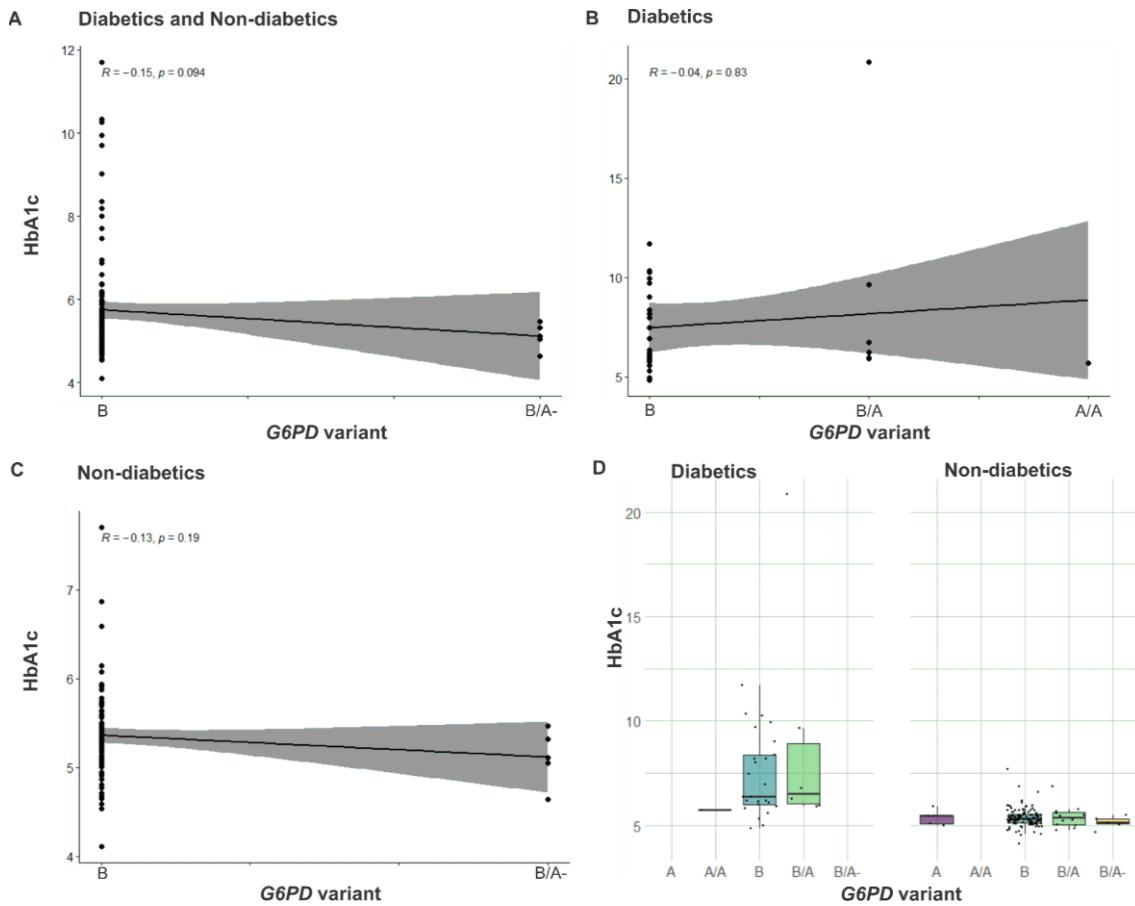


Figure 7. The negative correlation between glycated hemoglobin and G6PD variants was not significant in the *quilombola* sample set. R and p-values were calculated with Spearman's correlation method.

Last, we evaluated the usage of medical drugs among the *quilombolas* according to the recommendation of substances to be avoided by G6PD deficient individuals (RANGEL; RANGEL, 2022). We observed that one individual with the A/A genotype used Glibenclamide, a drug used to control blood glucose levels in type 2 diabetic individuals. This drug, used to control blood sugar levels, has been shown to induce AHA in diabetic patients carrying the A- variant (MCDONAGH et al., 2012).

Conclusion

Our results demonstrate a relatively low prevalence of G6PDd African variants in urban populations from the Central-West, Southeast, and South of Brazil, with a clear cline between these regions. Blacks showed significantly higher frequencies of *G6PD*A* than Whites or Amerindian

Brazilians. Compared with the predominantly European urban population from Paraná studied, the *quilombos* Feixo and Restinga showed higher frequencies of the *G6PD*A* variant. Despite the small size of the African-derived *quilombola* sample, we observed that HbA1c diminishes its correlation with blood glucose levels in individuals carrying *G6PD*A* and A- variants. Also, it is essential to evaluate if there is a significant negative correlation between these variants and HbA1c levels in a larger Brazilian dataset, as that could lead to undiagnosed diabetic patients.

Final considerations and perspectives

The presence of *G6PD* variants in higher frequency in *quilombos* and Black Brazilians highlights the need to increase the comprehension of these populations' particularities by the political and public health organisms. Furthermore, it stresses the importance of socially responsible political leadership. *Quilombos* are populations historically marginalized, with low access to health services, low schooling, high economic stress, and high food insecurity. Thus, they are more susceptible to wrongly designed health service policies. For example, the intense stimulus of the preventive use of chloroquine and hydroxy-chloroquine by the previous Brazilian government during the COVID-19 pandemic or the prescription of other drugs contraindicated, even with knowledge of the risk of hemolytic anemia for G6PD deficient individuals (DA ROCHA et al., 2021; HOFFMANN et al., 2020; MELO et al., 2021). G6PDd is a silent condition as it is not tested massively in Brazil and is only discovered after exposure to hemolytic agents (WHEELER et al., 2017). In addition, it must be kept in mind the implications of G6PDd in diabetic individuals, the higher prevalence of both conditions in Black Brazilians, and the higher susceptibility of G6PD individuals to infections (ALRAHMANY et al., 2023).

Thus, monitoring these variants in the Brazilian population and the wide circulation of their medical implications is critical for preventing adverse drug reactions and worsening other complex disease outcomes.

REFERENCES

- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19:1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Alrahmany D, Omar AF, Hafez W, et al (2023) Infections in Glucose-6-Phosphate Dehydrogenase G6PD-Deficient Patients; Predictors for Infection-Related Mortalities and Treatment Outcomes. *Antibiotics* 12:494. <https://doi.org/10.3390/antibiotics12030494>
- Anderson CA, Pettersson FH, Clarke GM, et al (2010) Data quality control in genetic case-control association studies. *Nature Protocols* 5:1564–1573. <https://doi.org/10.1038/nprot.2010.116>
- Ansari U, Bhardwaj P, Quadri H, et al (2022) Diabetic Ketoacidosis Unmasking a Diagnosis of Glucose-6-Phosphate Dehydrogenase Deficiency: A Case Report and Literature Review. *Cureus* 14:10–15. <https://doi.org/10.7759/cureus.23842>
- Augusto DG, de Almeida RC, Farias TDJ, et al (2021) Unsuspected Associations with Variants within the Genes NOTCH4 and STEAP2-ASI Uncovered by a Genome-Wide Association Study in Endemic Pemphigus Foliaceus. *Journal of Investigative Dermatology*. <https://doi.org/10.1016/j.jid.2021.04.017>
- Auton A, Abecasis GR, Altshuler DM, et al (2015) A global reference for human genetic variation. *Nature* 526:68–74. <https://doi.org/10.1038/nature15393>
- Beutler E, Kuhl W, Vives-Corrons J, Prchal J (1989) Molecular heterogeneity of glucose-6-phosphate dehydrogenase A-. *Blood* 74:2550–2555. <https://doi.org/10.1182/blood.v74.7.2550.2550>
- Brito IC, Lopes AA, Araújo LMB (2001) Associação da cor da pele com diabetes mellitus tipo 2 e intolerância à glicose em mulheres obesas de Salvador, Bahia. *Arquivos Brasileiros de Endocrinologia & Metabologia* 45:475–480. <https://doi.org/10.1590/S0004-27302001000500011>
- Broek L, Heylen E, Akker M (2016) Glucose-6-phosphate dehydrogenase deficiency: not exclusively in males. *Clinical Case Reports* 4:1135–1137. <https://doi.org/10.1002/ccr3.714>
- Cann HM, De Toma C, Cazes L, et al (2002) A Human Genome Diversity Cell Line Panel. *Science* 296:261b–2262. <https://doi.org/10.1126/science.296.5566.261b>
- Carette C, Dubois-Laforgue D, Gautier JF, Timsit J (2011) Diabète sucré et déficit en glucose-6-phosphate déshydrogénase: d'une crise à l'autre. *Diabetes and Metabolism* 37:79–82. <https://doi.org/10.1016/j.diabet.2010.09.004>
- Chang CC, Chow CC, Tellier LCAM, et al (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4:1–16. <https://doi.org/10.1186/s13742-015-0047-8>
- Cigolini AA, Da Silva M (2018) Comunidades Remanescentes Quilombolas: Iconografias e circulações na comunidade da Restinga – Lapa-Pr, Brasil. *Revista Geografar* 13:98–118
- da Rocha JEB, Othman H, Tiemessen CT, et al (2021) G6PD distribution in sub-Saharan Africa and potential risks of using chloroquine/hydroxychloroquine based treatments for COVID-19. *The Pharmacogenomics Journal*. <https://doi.org/10.1038/s41397-021-00242-8>
- Daflon VT, Carvalhaes F, Feres Júnior J (2017) Sentindo na Pele: Percepções de Discriminação Cotidiana de Pretos e Pardos no Brasil. *Dados* 60:293–330. <https://doi.org/10.1590/001152582017121>
- Dallol A, Banni H, Gari MA, et al (2012) Five novel glucose-6-phosphate dehydrogenase deficiency haplotypes correlating with disease severity. *Journal of Translational Medicine* 10:199. <https://doi.org/10.1186/1479-5876-10-199>
- de Moura Neto JP, Dourado MV, dos Reis MG, Gonçalves MS (2008) A novel c.197T → A variant among Brazilian neonates with glucose-6-phosphate dehydrogenase deficiency. *Genetics and Molecular Biology* 31:33–35. <https://doi.org/10.1590/s1415-47572008000100006>
- Dombrowski JG, Souza RM, Curry J, et al (2017) G6PD deficiency alleles in a malaria-endemic region in the Western Brazilian Amazon. *Malaria Journal* 16:1–9. <https://doi.org/10.1186/s12936-017-1889-6>
- Farias TDJ, Augusto DG, Almeida RCD, Malheiros D (2018) Screening the full leucocyte receptor complex genomic region revealed associations with pemphigus that might be explained by gene regulation. *Immunology* 156:86–93. <https://doi.org/10.1111/imm.13003>
- Ferrell RE, Bertin T, Barton SA, et al (1980) The multinational Andean genetic and health program. IX.

- Gene frequencies and rare variants of 20 serum proteins and erythrocyte enzymes in the Aymara of Chile. *American Journal of Human Genetics* 32:92–102
- Geck RC, Powell NR, Dunham MJ (2023) Functional interpretation, cataloging, and analysis of 1,341 glucose-6-phosphate dehydrogenase variants. *American Journal of Human Genetics* 110:228–239. <https://doi.org/10.1016/j.ajhg.2023.01.003>
- Gomez F, Hirbo J, Tishkoff SA (2014) Genetic variation and adaptation in Africa: Implications for human evolution and disease. *Cold Spring Harbor Perspectives in Biology* 6:6-a008524. <https://doi.org/10.1101/cshperspect.a008524>
- Gómez-Manzo S, Marcial-Quino J, Vanoye-Carlos A, et al (2016) Glucose-6-Phosphate Dehydrogenase: Update and Analysis of New Mutations around the World. *International Journal of Molecular Sciences* 17:2069. <https://doi.org/10.3390/ijms17122069>
- Hoffmann M, Mösbauer K, Hofmann-Winkler H, et al (2020) Chloroquine does not inhibit infection of human lung cells with SARS-CoV-2. *Nature* 585:588–590. <https://doi.org/10.1038/s41586-020-2575-3>
- Howes RE, Battle KE, Satyagraha AW, et al (2013) G6PD Deficiency. In: *Advances in Parasitology*. Elsevier, pp 133–201
- Howes RE, Piel FB, Patil AP, et al (2012) G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map. *PLoS Medicine* 9:e1001339. <https://doi.org/10.1371/journal.pmed.1001339>
- Joerin-Luque IA, Augusto DG, Calonga-Solís V, et al (2022) Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil. *Molecular Genetics and Genomics*. <https://doi.org/10.1007/s00438-022-01857-7>
- Joerin-Luque IA, Sukow NM, Bucco ID, et al (2023) Ancestry, diversity, and genetics of health-related traits in African-derived communities (quilombos) from Brazil. *Functional & Integrative Genomics* 23:74. <https://doi.org/10.1007/s10142-023-00999-0>
- Kassambara A (2023) ggpubr: “ggplot2” Based Publication Ready Plots. R package version 0.6.0
- Luzzatto L, Ally M, Notaro R (2021) Glucose-6-phosphate dehydrogenase deficiency. *Blood* 136:1225–1240. <https://doi.org/10.1182/blood.2019000944>
- Luzzatto L, Usanga EA, Reddy S (1969) Glucose-6-phosphate dehydrogenase deficient red cells: Resistance to infection by malarial parasites. *Science* 164:839–842. <https://doi.org/10.1126/science.164.3881.839>
- Mason PJ, Bautista JM, Gilsanz F (2007) G6PD deficiency: the genotype-phenotype association. *Blood Reviews* 21:267–283. <https://doi.org/10.1016/j.blre.2007.05.002>
- McDonagh EM, Thorn CF, Bautista JM, et al (2012) PharmGKB summary: very important pharmacogene information for G6PD. *Pharmacogenetics and Genomics* 22:219–228. <https://doi.org/10.1097/FPC.0b013e32834eb313>
- Melo JRR, Duarte EC, de Moraes MV, et al (2021) Automedicação e uso indiscriminado de medicamentos durante a pandemia da COVID-19. *Cadernos de Saúde Pública* 37:. <https://doi.org/10.1590/0102-311x00053221>
- Monteiro WM, Val FFA, Siqueira AM, et al (2014) G6PD deficiency in Latin America: systematic review on prevalence and variants. *Memórias do Instituto Oswaldo Cruz* 109:553–568. <https://doi.org/10.1590/0074-0276140123>
- Myers TA, Chanock SJ, Machiela MJ (2020) LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Frontiers in Genetics* 11:1–5. <https://doi.org/10.3389/fgene.2020.00157>
- Oliveira HSS de, da Silva ANLM, Andrade GB, et al (2018) Molecular genotyping of G6PD mutations and Duffy blood group in Afro-descendant communities from Brazilian Amazon. *Genetics and Molecular Biology* 41:758–765. <https://doi.org/10.1590/1678-4685-gmb-2017-0253>
- Paradis E (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Petzl-Erler LM, Luz R, Sotomaior VS (1993) The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. *Tissue Antigens* 41:227–237. <https://doi.org/10.1111/j.1399-0039.1993.tb02011.x>
- Phillips J, Henderson AC (2018) Hemolytic anemia: Evaluation and differential diagnosis. *American*

- Family Physician 98:354–361
- Prefeitura de São Paulo (2018) Informe Técnico sobre Malária. 22
- R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria URL <https://wwwR-project.org/>
- Rangel A, Rangel L (2022) Medicamentos e substâncias contraindicados e/ou que exigem uso em assistência hospitalar: Cartilha de substâncias, 1st edn. São Gonçalo
- Shin J-H, Blay S, Graham J, McNeney B (2006) LDheatmap : An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. Journal of Statistical Software 16:1–9. <https://doi.org/10.18637/jss.v016.c03>
- Sinnwell J, Schaid D (2023) haplo.stats: Routines for the analysis of indirectly measured haplotypes
- Sirdah MM, Shubair ME, Al-Kahlout MS, et al (2017) Possible association of 3' UTR +357 A>G, IVS11-nt 93 T>C, c.1311 C>T polymorphism with G6PD deficiency. Hematology 22:370–374. <https://doi.org/10.1080/10245332.2016.1276117>
- Sociedade Brasileira de Diabetes (2020) Sociedade Brasileira de Diabetes. Tipos de diabetes
- Tiberio G (1994) MZ female twins discordant for X-linked diseases: A review. Acta Geneticae Medicae et Gemellologiae 43:207–214. <https://doi.org/10.1017/s0001566000001963>
- Tishkoff SA, Varkonyi R, Cahinhinan N, et al (2001) Haplotype diversity and linkage disequilibrium at human G6PD: Recent origin of alleles that confer malarial resistance. Science 293:455–462. <https://doi.org/10.1126/science.1061573>
- Tsuneto LT, Probst CM, Hutz MH, et al (2003) HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. Tissue Antigens 62:512–526. <https://doi.org/10.1046/j.1399-0039.2003.00139.x>
- Verlag T, Spring C (1992) The molecular basis of gluc0se-6-phosphate dehydr0genase deficiency. World 1–3
- Weimer TA, Salzano FM, Westwood B, Beutler E (1993) Molecular characterization of glucose-6-phosphate dehydrogenase variants from Brazil. Human biology; an international record of research 65:41–47
- Weimer TA, Salzano FM, Westwood B, Beutler E (1998) G6PD Variants in Three South American Ethnic Groups: Population Distribution and Description of Two New Mutations. Human Heredity 48:92–96. <https://doi.org/10.1159/000022787>
- Wheeler E, Leong A, Liu C-T, et al (2017) Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transetnic genome-wide meta-analysis. PLOS Medicine 14:e1002383. <https://doi.org/10.1371/journal.pmed.1002383>
- Wickham H (2009) ggplot2 Book. Media 35:211
- Zhang Z, Liew CW, Handy DE, et al (2010) High glucose inhibits glucose-6-phosphate dehydrogenase, leading to increased oxidative stress and β-cell apoptosis. The FASEB Journal 24:1497–1505. <https://doi.org/10.1096/fj.09-136572>

7 CONCLUSÕES

1. Foram constatados e revisados 57 artigos científicos abordando os aspectos genéticos de populações quilombolas brasileiras, incluindo o estudo de ancestralidade e saúde nas mesmas. Apenas 93 dos 5972 quilombos do Brasil foram incluídos em estudos genéticos e, dentre eles, 80% são originários das regiões Norte e Nordeste, ressaltando a necessidade de se estudarem mais quilombos das Regiões Centro-Oeste, Sudeste e Sul do Brasil.
2. Os centros urbanos das Regiões Centro-Oeste e Sudeste apresentaram ancestralidade materna predominantemente ameríndia e africana, respectivamente. No Sul, a ancestralidade materna ameríndia e a europeia tiveram a mesma proporção. Nas mesmas populações, a ancestralidade paterna europeia foi a predominante. No quilombo Feixo, a ancestralidade materna ameríndia foi predominante, enquanto na Restinga predominou a africana. Em ambos os quilombos a ancestralidade paterna predominante foi europeia.
3. Todas as populações brasileiras estudadas apresentaram um padrão de miscigenação enviesado pelo sexo, com predomínio de ancestralidade paterna europeia e ancestralidade materna ameríndia ou africana. Na amostra urbana da Região Sul do Brasil, o somatório das proporções de ancestralidade materna africana e indígena sobrepassou a proporção observada para a ancestralidade europeia, demonstrando o mesmo padrão.
4. A análise de haplogrupos maternos africanos nas três regiões urbanas sugeriu uma contribuição predominantemente Niger-Congo não Bantu para o Sudeste e Sul do Brasil enquanto uma contribuição predominantemente centro-oeste africana Bantu foi descrita pela primeira vez para a Região Centro-Oeste do Brasil. Os haplogrupos paternos africanos das Regiões Centro-Oeste e Sul, analisados em conjunto, indicaram uma ancestralidade predominantemente centro-oeste africana. A análise filogeográfica de haplogrupos africanos no Feixo sugeriu a presença de linhagens maternas fundadoras do Centro-Oeste/Sudeste Africano Bantu e de linhagens do Oeste Africano na Restinga. Por outro lado, nas duas comunidades quilombolas, as linhagens paternas sugerem uma possível contribuição do Centro-Leste e Sul Africanos.

5. A análise de haplogrupos mitocondriais nativos-americanos nas populações urbanas das três regiões brasileiras sugeriu contribuições predominantes de grupos ameríndios andinos e do Brasil central nas Regiões Centro-Oeste e Sudeste brasileiras e de grupos ameríndios do Sul na mesma região. Já nos quilombos, as frequências de haplogrupos nativo-americanos sugerem uma contribuição de grupos Kaingang no Feixo e Guarani na Restinga.
6. O estudo de variantes no gene *G6PD* demonstrou uma diminuição da frequência das variantes *G6PD*A* e *G6PD*A-* do Centro-Oeste para o Sudeste e Sul do Brasil. Foi também identificada uma correlação positiva na prevalência das variantes *G6PD*A-* e *G6PD*A* com a ancestralidade africana e o fenótipo naquelas populações urbanas. Nos quilombos Feixo e Restinga, a variante *G6PD*A-* foi encontrada em maior frequência que nas populações urbanas do Sudeste e Sul, enquanto a variante *G6PD*A* foi encontrada em maior frequência que nas três regiões urbanas avaliadas.

8 CONSIDERAÇÕES FINAIS

Dada a importância do estudo da ancestralidade para preencher as lacunas de conhecimento na história da diáspora africana no Brasil e também como marco teórico para uma adequada implementação da medicina de precisão nessa população, este trabalho teve como objetivo contribuir com o esclarecimento da ancestralidade genética de populações afro-brasileiras, com especial ênfase nas origens subcontinentais dos seus fundadores. Ademais, é de especial relevância para a população paranaense tirar o véu da história hegemônica do estado e ajudar na visibilização dos africanos e afrodescendentes que participaram ativamente no desenvolvimento da região.

9 REFERÊNCIAS BIBLIOGRÁFICAS

- AIDOO, L. **Slavery Unseen: Sex, Power, and Violence in Brazilian History**. Durham: Duke University Press, 2018. v. 47
- ALMEIDA, S. L. DE. **Racismo estrutural**. 1. ed. São Paulo: Polen, 2019.
- ÁLVAREZ-IGLESIAS, V. et al. Coding region mitochondrial DNA SNPs: Targeting East Asian and Native American haplogroups. **Forensic Science International: Genetics**, v. 1, n. 1, p. 44–55, 2007.
- ANDERSON, S. et al. Sequence and organization of the human mitochondrial genome. **Nature**, v. 290, n. 5806, p. 457–465, abr. 1981.
- ANDREWS, R. M. et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. **Nature Genetics**, v. 23, n. 2, p. 147–147, out. 1999.
- ANSARI POUR, N.; PLASTER, C. A.; BRADMAN, N. Evidence from Y-chromosome analysis for a late exclusively eastern expansion of the Bantu-speaking people. **European Journal of Human Genetics**, v. 21, n. 4, p. 423–429, 2013.
- ANSARI, U. et al. Diabetic Ketoacidosis Unmasking a Diagnosis of Glucose-6-Phosphate Dehydrogenase Deficiency: A Case Report and Literature Review. **Cureus**, v. 14, n. 4, p. 10–15, abr. 2022.
- AUGUSTO, D. G. et al. Unsuspected Associations with Variants within the Genes NOTCH4 and STEAP2-ASI Uncovered by a Genome-Wide Association Study in Endemic Pemphigus Foliaceus. **Journal of Investigative Dermatology**, maio 2021.
- AUTON, A. et al. A global reference for human genetic variation. **Nature**, v. 526, n. 7571, p. 68–74, 2015.
- AVISE, J. C. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. **Annual review of ecology and systematics**. Vol. 18, v. 18, n. 1987, p. 489–522, 1987.
- AVISE, J. C. Phylogeography: retrospect and prospect. **Journal of Biogeography**, v. 36, n. 1, p. 3–15, jan. 2009.
- BAILLIET, G. et al. Brief communication: Restricted geographic distribution for Y-Q* paragroup in South America. **American Journal of Physical Anthropology**, v. 140, n. 3, p. 578–582, 2009.
- BANDELT, H. J.; FORSTER, P.; RÖHL, A. Median-joining networks for inferring intraspecific phylogenies. **Molecular Biology and Evolution**, v. 16, n. 1, p. 37–48, 1999.
- BANDELT, H. J.; MACAULAY, V.; RICHARDS, M. **Human Mitochondrial DNA and the Evolution of Homo Sapiens**. 1. ed. Berlin/Heidelberg: Springer, 2006.
- BARBIERI, C. et al. Ancient Substructure in Early mtDNA Lineages of Southern Africa. **The American Journal of Human Genetics**, v. 92, n. 2, p. 285–292, fev. 2013.
- BARBIERI, C. et al. Unraveling the complex maternal history of Southern African Khoisan populations. **American Journal of Physical Anthropology**, v. 153, n. 3, p. 435–448, 2014a.
- BARBIERI, C. et al. Migration and interaction in a contact zone: mtDNA variation among Bantu-speakers in Southern Africa. **PLoS ONE**, v. 9, n. 6, p. e99117, jun. 2014b.
- BATINI, C. et al. Phylogeography of the human mitochondrial L1c haplogroup: Genetic signatures of the prehistory of Central Africa. **Molecular Phylogenetics and Evolution**, v. 43, n. 2, p. 635–644, 2007.
- BATISTA, L. E.; KALCKMANN, S. **Seminário Saúde da População Negra Estado de São Paulo 2004**. First ed. São Paulo: Instituto de Saúde de São Paulo, 2005.
- BATTAGLIA, V. et al. The first peopling of South America: New evidence from Y-chromosome haplogroup Q. **PLoS ONE**, v. 8, n. 8, p. e71390, ago. 2013.
- BEHAR, D. M. et al. The dawn of human matrilineal diversity. **The American Journal of Human Genetics**, v. 82, n. 5, p. 1130–1140, maio 2008.
- BEHAR, D. M. et al. A “copernican” reassessment of the human mitochondrial DNA tree from its root. **American Journal of Human Genetics**, v. 90, n. 4, p. 675–684, 2012.
- BELBIN, G. M. et al. Toward a fine-scale population health monitoring system. **Cell**, v. 184, n. 8, p. 2068–2083.e11, 2021.
- BISSO-MACHADO, R. et al. Distribution of Y-chromosome q lineages in native americans. **American Journal of Human Biology**, v. 23, n. 4, p. 563–566, jul. 2011.
- BISSO-MACHADO, R.; CATIRA BORTOLINI, M.; SALZANO, F. M. **Uniparental genetic markers**

- in South Amerindians.** [s.l.: s.n.]. Disponível em: <www.sbg.org.br>.
- BODNER, M. et al. Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes. **Genome Research**, v. 22, n. 5, p. 811–820, maio 2012.
- BORTOLINI, M. C. et al. African-derived South American populations: A history of symmetrical and asymmetrical matings according to sex revealed by bi- and uni-parental genetic markers. **American Journal of Human Biology**, v. 11, n. 4, p. 551–563, 1999.
- BRITO, I. C.; LOPES, A. A.; ARAÚJO, L. M. B. Associação da cor da pele com diabetes mellitus tipo 2 e intolerância à glicose em mulheres obesas de Salvador, Bahia. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 45, n. 5, p. 475–480, out. 2001.
- BUCCIFERRO, J. R. A Forced Hand: Natives, Africans, and the Population of Brazil, 1545–1850. **Revista de Historia Económica / Journal of Iberian and Latin American Economic History**, v. 31, n. 2, p. 285–317, set. 2013.
- CALONGA-SOLÍS, V. et al. Variation in genes implicated in B-cell development and antibody production affects susceptibility to pemphigus. **Immunology**, v. 162, n. 1, p. 58–67, set. 2020.
- CAMPBELL, M. C.; TISHKOFF, S. A. African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. **Annual Review of Genomics and Human Genetics**, v. 9, n. 1, p. 403–433, set. 2008.
- CAMPBELL, M. C.; TISHKOFF, S. A. The Evolution of Human Genetic and Phenotypic Variation in Africa. **Current Biology**, v. 20, n. 4, p. 1–15, 2010.
- CANN, H. M. et al. A Human Genome Diversity Cell Line Panel. **Science**, v. 296, n. 5566, p. 261b–2262, abr. 2002.
- CANN, R. L.; STONEKING, M.; WILSON, A. C. **Mitochondrial DNA and human evolution**. **Nature**, 1987.
- CARDENA, M. M. S. G. S. G. et al. Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. **PLoS ONE**, v. 8, n. 4, p. e62005, abr. 2013.
- CARDOSO, G. C. et al. Analysis of eight mtDNA coding region polymorphisms for characterization of the female lineages ancestry in Alagoas, Brazil. **Forensic Science International: Genetics Supplement Series**, v. 4, n. 1, p. 131–132, 2013.
- CARETTE, C. et al. Diabète sucré et déficit en glucose-6-phosphate déshydrogénase: d'une crise à l'autre. **Diabetes and Metabolism**, v. 37, n. 1, p. 79–82, 2011a.
- CARETTE, C. et al. Diabetes mellitus and glucose-6-phosphate dehydrogenase deficiency: From one crisis to another. **Diabetes and Metabolism**, v. 37, n. 1, p. 79–82, 2011b.
- CARVALHO, B. M. et al. Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations. **Genetics and Molecular Biology**, v. 31, n. 1, p. 12–22, 2008.
- CEREZO, M. et al. Reconstructing ancient mitochondrial DNA links between Africa and Europe. **Genome Research**, v. 22, n. 5, p. 821–826, maio 2012.
- CHIARONI, J.; UNDERHILL, P. A.; CAVALLI-SFORZA, L. L. Y chromosome diversity, human expansion, drift, and cultural evolution. **Proceedings of the National Academy of Sciences**, v. 106, n. 48, p. 20174–20179, dez. 2009.
- CIGOLINI, A. A.; DA SILVA, M. Comunidades Remanescentes Quilombolas: Iconografias e circulações na comunidade da Restinga – Lapa-Pr, Brasil. **Revista Geografar**, v. 13, n. 1, p. 98–118, 2018.
- COELHO, M. et al. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. **BMC Evolutionary Biology**, v. 9, n. 1, p. 80, 2009.
- CRUCIANI, F. et al. Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. **American Journal of Human Genetics**, v. 74, n. 5, p. 1014–1022, 2004.
- CRUCIANI, F. et al. Human y chromosome haplogroup R-V88: A paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. **European Journal of Human Genetics**, v. 18, n. 7, p. 800–807, 2010.
- DA ROCHA, J. E. B. et al. G6PD distribution in sub-Saharan Africa and potential risks of using chloroquine/hydroxychloroquine based treatments for COVID-19. **The Pharmacogenomics Journal**,

jul. 2021.

- DALLOL, A. et al. Five novel glucose-6-phosphate dehydrogenase deficiency haplotypes correlating with disease severity. **Journal of Translational Medicine**, v. 10, n. 1, p. 199, 2012.
- D'ATANASIO, E. et al. Y Haplogroup Diversity of the Dominican Republic: Reconstructing the Effect of the European Colonization and the Trans-Atlantic Slave Trades. **Genome biology and evolution**, v. 12, n. 9, p. 1579–1590, set. 2020.
- DAVIS, A. Y. **Women, race, and class**. New York: Vintage Books, 1944.
- DE FILIPPO, C. et al. Y-chromosomal variation in sub-Saharan Africa: Insights into the history of Niger-Congo groups. **Molecular Biology and Evolution**, v. 28, n. 3, p. 1255–1269, 2011.
- DE KNIJFF, P. Messages through Bottlenecks: On the Combined Use of Slow and Fast Evolving Polymorphic Markers on the Human Y Chromosome. **The American Journal of Human Genetics**, v. 67, n. 5, p. 1055–1061, nov. 2000.
- DE SAINT PIERRE, M. et al. Arrival of Paleo-Indians to the Southern Cone of South America: New Clues from Mitogenomes. **PLoS ONE**, v. 7, n. 12, p. e51311, dez. 2012.
- DE SOUZA, A. M. et al. A systematic scoping review of the genetic ancestry of the Brazilian population. **Genetics and Molecular Biology**, v. 42, n. 3, p. 495–508, set. 2019.
- DOMBROWSKI, J. G. et al. G6PD deficiency alleles in a malaria-endemic region in the Western Brazilian Amazon. **Malaria Journal**, v. 16, n. 1, p. 1–9, 2017.
- DÜR, A.; HUBER, N.; PARSON, W. Fine-Tuning Phylogenetic Alignment and Haplogrouping of mtDNA Sequences. **International Journal of Molecular Sciences**, v. 22, n. 11, p. 5747, maio 2021.
- ELTIS, D.; RICHARDSON, D. **Voyages: The trans-Atlantic slave trade database**. , 2010. Disponível em: <www.slavevoyages.org/assessment/estimates>. Acesso em: 7 jul. 2020
- ESCHER, L. M. et al. Challenges in selecting admixture models and marker sets to infer genetic ancestry in a Brazilian admixed population. **Scientific Reports**, v. 12, n. 1, p. 21240, dez. 2022.
- FAJARDO, S.; CUNHA, L. A. G. **Paraná: Desenvolvimento e Diferenças Regionais**. 1. ed. Ponta Grossa: Atena, 2021.
- FORTES-LIMA, C.; VERDU, P. Anthropological genetics perspectives on the transatlantic slave trade. **Human Molecular Genetics**, v. 30, n. R1, p. R79–R87, abr. 2021.
- FUNDAÇÃO CULTURAL PALMARES. **Certificação quilombola**. Fundação Cultural Palmares, 2021. Disponível em: <http://www.palmares.gov.br/?page_id=37551>. Acesso em: 9 mar. 2021
- GAMPIO GUEYE, N. S. et al. An update on glucose-6-phosphate dehydrogenase deficiency in children from Brazzaville, Republic of Congo. **Malaria Journal**, v. 18, n. 1, p. 4–9, 2019.
- GASPARRE, G.; PORCELLI, A. M. **The Human Mitochondrial Genome: From basic biology to disease**. London: Academic Press, 2020.
- GECK, R. C.; POWELL, N. R.; DUNHAM, M. J. Functional interpretation, cataloging, and analysis of 1,341 glucose-6-phosphate dehydrogenase variants. **American Journal of Human Genetics**, v. 110, n. 2, p. 228–239, 2023.
- GOMES, F. D. S. **Mocambos e Quilombos: Uma história do campesinato negro no Brasil**. 1. ed. São Paulo: Claro Enigma, 2015.
- GOMEZ, F.; HIRBO, J.; TISHKOFF, S. A. Genetic variation and adaptation in Africa: Implications for human evolution and disease. **Cold Spring Harbor Perspectives in Biology**, v. 6, n. 7, p. 6-a008524, jul. 2014.
- GÓMEZ-MANZO, S. et al. Glucose-6-Phosphate Dehydrogenase: Update and Analysis of New Mutations around the World. **International Journal of Molecular Sciences**, v. 17, n. 12, p. 2069, dez. 2016.
- GONZÁLEZ, A. M. et al. Mitochondrial DNA affinities at the Atlantic fringe of Europe. **American Journal of Physical Anthropology**, v. 120, n. 4, p. 391–404, abr. 2003.
- GONZÁLEZ, A. M. et al. Mitochondrial lineage M1 traces an early human backflow to Africa. **BMC Genomics**, v. 8, p. 1–12, 2007.
- GONZÁLEZ, M. et al. The genetic landscape of Equatorial Guinea and the origin and migration routes of the y chromosome haplogroup R-V88. **European Journal of Human Genetics**, v. 21, n. 3, p. 324–331, 2013.
- GRUGNI, V. et al. Analysis of the human Y-chromosome haplogroup Q characterizes ancient population movements in Eurasia and the Americas. **BMC Biology**, v. 17, n. 1, p. 1–14, 2019.

- GRUPO DE TRABALHO CLÓVIS MOURA. **Grupo de Trabalho Clóvis Moura.** Disponível em: <<https://www.administracao.pr.gov.br/ArquivoPublico/Pagina/Grupo-de-Trabalho-Clovis-Moura-Apresentacao>>. Acesso em: 2 ago. 2023.
- GURDASANI, D. et al. The African Genome Variation Project. **Cell**, v. 179, n. 4, p. 984- 1002.e36, out. 2019.
- GUTIÉRREZ, H. Donos de terras e escravos no Paraná: padrões e hierarquias nas primeiras décadas do século XIX. **História (São Paulo)**, v. 25, n. 1, p. 100–122, 2006.
- HABER, M. et al. A rare deep-rooting D0 African Y-chromosomal haplogroup and its implications for the expansion of modern humans out of Africa. **Genetics**, v. 212, n. 4, p. 1421–1428, 2019.
- HALL, G. M. **Slavery and African ethnicities in the Americas: Restoring the links**. North Carolina: The University of North Carolina Press, 2005.
- HENN, B. M. et al. Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. **Proceedings of the National Academy of Sciences**, v. 105, n. 31, p. 10693–10698, ago. 2008.
- HERNÁNDEZ, C. L. Mitochondrial DNA in Human Diversity and Health: From the Golden Age to the Omics Era. **Genes**, v. 14, n. 8, p. 1534, jul. 2023.
- HOFFMANN, M. et al. Chloroquine does not inhibit infection of human lung cells with SARS-CoV-2. **Nature**, v. 585, n. 7826, p. 588–590, set. 2020.
- HOWES, R. E. et al. G6PD Deficiency. Em: **Advances in Parasitology**. [s.l.] Elsevier, 2013. v. 81p. 133–201.
- IBGE. **Brasil: 500 anos de Povoamento**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2007.
- JOBLING, M. et al. **Human Evolutionary Genetics**. 2. ed. New York and London: Garland Science, Taylor & Francis Group, LLC, 2014.
- JOBLING, M. A.; TYLER-SMITH, C. The human Y chromosome: An evolutionary marker comes of age. **Nature Reviews Genetics**, v. 4, n. 8, p. 598–612, 2003.
- JOBLING, M. A.; TYLER-SMITH, C. Human Y-chromosome variation in the genome-sequencing era. **Nature Reviews Genetics**, v. 18, n. 8, p. 485–497, 2017.
- JOERIN-LUQUE, I. A. et al. Uniparental markers reveal new insights on subcontinental ancestry and sex-biased admixture in Brazil. **Molecular Genetics and Genomics**, jan. 2022.
- JOERIN-LUQUE, I. A. et al. Ancestry, diversity, and genetics of health-related traits in African-derived communities (quilombos) from Brazil. **Functional & Integrative Genomics**, v. 23, n. 1, p. 74, mar. 2023.
- KAMVAR, Z. N.; TABIMA, J. F.; GRÜNWALD, N. J. Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. **PeerJ**, v. 2, n. 1, p. e281, mar. 2014.
- KARAFET, T. M. et al. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. **Genome Research**, v. 18, n. 5, p. 830–838, fev. 2008.
- KEHDY, F. S. G. et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. **Proceedings of the National Academy of Sciences**, v. 112, n. 28, p. 8696–8701, jul. 2015.
- KENT, M.; WADE, P. Genetics against race: Science, politics and affirmative action in Brazil. **Social Studies of Science**, v. 45, n. 6, p. 816–838, dez. 2015.
- KIMURA, L. et al. Genomic ancestry of rural African-derived populations from Southeastern Brazil. **American Journal of Human Biology**, v. 25, n. 1, p. 35–41, 2013.
- KIMURA, L. et al. Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes. **American Journal of Human Biology**, v. 29, n. 2, p. 1–11, mar. 2017.
- KIMURA, M. The neutral theory of molecular evolution and the world view of the neutralists. **Genome**, v. 31, n. 1, p. 24–31, jan. 1989.
- KIVISILD, T. et al. The Emerging Limbs and Twigs of the East Asian mtDNA Tree. **Molecular Biology and Evolution**, v. 19, n. 10, p. 1737–1751, out. 2002.
- KLEIN, H. S. As origens africanas dos escravos brasileiros. Em: PENA, S. D. J. (Ed.). **Homo Brasilis: Aspectos genéticos, linguísticos, históricos e socioantropológicos da formação do povo brasileiro**. 1. ed. Ribeirão Preto: FUNPEC, 2002. p. 93–112.
- KOILE, E. et al. Phylogeographic analysis of the Bantu language expansion supports a rainforest route.

- Proceedings of the National Academy of Sciences of the United States of America**, v. 119, n. 32, 2022.
- KORUNES, K. L.; GOLDBERG, A. Human genetic admixture. **PLOS Genetics**, v. 17, n. 3, p. e1009374, mar. 2021.
- KOSOY, R. et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. **Human Mutation**, v. 30, n. 1, p. 69–78, 2009.
- LACOMBE, A. J.; SILVA, E.; BARBOSA, F. DE A. **Rui Barbosa e a queima dos arquivos**. 1. ed. Rio de Janeiro: Fundação Casa de Rui Barbosa, 1988. v. 1
- LÊ, S.; JOSSE, J.; HUSSON, F. FactoMineR: An R package for multivariate analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 253–258, fev. 2008.
- LEITE, I. B. Quilombos e quilombolas: cidadania ou folclorização? **Horizontes Antropológicos**, v. 5, n. 10, p. 123–149, 1999.
- LEMES, R. B. et al. Estimation of Inbreeding and Substructure Levels in African-Derived Brazilian Quilombo Populations. **Human Biology**, v. 86, n. 4, p. 276, 2014.
- LIMA-COSTA, M. F. et al. Genomic ancestry and ethnорacial self-classification based on 5,871 community-dwelling Brazilians (The Epigen Initiative). **Scientific Reports**, v. 5, n. 1, p. 9812, set. 2015.
- LUZZATTO, L.; ALLY, M.; NOTARO, R. Glucose-6-phosphate dehydrogenase deficiency. **Blood**, v. 136, n. 11, p. 1225–1240, 2021.
- MACAULAY, V. et al. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. **Science**, v. 308, n. 5724, p. 1034–1036, 2005.
- MARRERO, A. R. et al. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. **American Journal of Physical Anthropology**, v. 132, n. 2, p. 301–310, fev. 2007.
- MARTÍNEZ, B. et al. Searching for the roots of the first free African American community. **Scientific Reports**, v. 10, n. 1, p. 20634, dez. 2020.
- MASON, P. J.; BAUTISTA, J. M.; GILSANZ, F. G6PD deficiency: the genotype-phenotype association. **Blood Reviews**, v. 21, n. 5, p. 267–283, 2007.
- MATHIESON, I.; SCALLY, A. What is ancestry? **PLOS Genetics**, v. 16, n. 3, p. e1008624, 2020.
- MEDINA, L. S. J. et al. Human Y-chromosome SNP characterization by multiplex amplified product-length polymorphism analysis. **ELECTROPHORESIS**, v. 35, n. 17, p. 2524–2527, set. 2014.
- MELO, J. R. R. et al. Automedicação e uso indiscriminado de medicamentos durante a pandemia da COVID-19. **Cadernos de Saúde Pública**, v. 37, n. 4, 2021.
- MOTTI, J. M. B. et al. Diferenciación regional de poblaciones nativas de América a partir del análisis de los linajes maternos. **Intersecciones en Antropología**, v. 18, n. 3, p. 271–282, 2017.
- MOURA, C. **História do Negro Brasileiro**. 2nd. ed. São Paulo: Ática, 1992.
- MOURA, C. **Quilombos: Resistência ao escravismo**. 3. ed. São Paulo: Ática, 1993.
- MYRES, N. M. et al. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. **European Journal of Human Genetics**, v. 19, n. 1, p. 95–101, jan. 2011.
- NASCIMENTO, A. **O genocídio do Negro brasileiro: Processo de um racismo mascarado**. 2. ed. São Paulo: Perspectiva, 2016.
- NASLAVSKY, M. S. et al. Whole-genome sequencing of 1,171 elderly admixed individuals from São Paulo, Brazil. **Nature Communications**, v. 13, n. 1, p. 1004, mar. 2022.
- NAVARRO-LÓPEZ, B. et al. **Phylogeographic review of Y chromosome haplogroups in Europe**. **International Journal of Legal Medicine** Springer Science and Business Media Deutschland GmbH, , set. 2021.
- NOVELLETTO, A. Y chromosome variation in Europe: Continental and local processes in the formation of the extant gene pool. **Annals of Human Biology**, v. 34, n. 2, p. 139–172, 2007.
- OLIVEIRA, H. S. S. DE et al. Molecular genotyping of G6PD mutations and Duffy blood group in Afro-descendant communities from Brazilian Amazon. **Genetics and Molecular Biology**, v. 41, n. 4, p. 758–765, dez. 2018.
- ONGARO, L. et al. The genomic impact of European colonization of the Americas. **Current Biology**, v. 29, n. 23, p. 3974- 3986.e4, dez. 2019.
- PARADIS, E. Pegas: An R package for population genetics with an integrated-modular approach. **Bioinformatics**, v. 26, n. 3, p. 419–420, 2010.
- PARSON, W. et al. DNA Commission of the International Society for Forensic Genetics: Revised and

- extended guidelines for mitochondrial DNA typing. **Forensic Science International: Genetics**, v. 13, p. 134–142, nov. 2014.
- PATIN, E. et al. Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. v. 546, n. May, p. 543–546, 2017.
- PATIN, E.; QUINTANA-MURCI, L. The demographic and adaptive history of central African hunter-gatherers and farmers. **Current Opinion in Genetics & Development**, v. 53, p. 90–97, dez. 2018.
- PENA, S. D. J. et al. DNA tests probe the genomic ancestry of Brazilians. **Brazilian Journal of Medical and Biological Research**, v. 42, n. 10, p. 870–876, set. 2009.
- PEPRAH, E. et al. Genome-Wide Association Studies in Africans and African Americans: Expanding the Framework of the Genomics of Human Traits and Disease. **Public Health Genomics**, v. 18, n. 1, p. 40–51, 2015.
- PEREIRA, L. et al. Prehistoric and historic traces in the mtDNA of Mozambique: Insights into the Bantu expansions and the slave trade. **Annals of Human Genetics**, v. 65, n. 5, p. 439–458, 2001.
- PFAFFELHUBER, P. et al. How to choose sets of ancestry informative markers: A supervised feature selection approach. **Forensic Science International: Genetics**, v. 46, n. February, p. 102259, 2020.
- PFENNIG, A. et al. Evolutionary Genetics and Admixture in African Populations. **Genome Biology and Evolution**, v. 15, n. 4, p. 1–20, abr. 2023.
- PIERRON, D. et al. Genomic landscape of human diversity across Madagascar. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, n. 32, p. E6498–E6506, 2017.
- PINHEIRO, A. A.; ASSUMPÇÃO, J. E.; FOLLMANN, J. I. **África e afrodescendentes no sul do Brasil: história, religião e educação**. São Leopoldo: Casa Leiria, 2015.
- PINOTTI, T. et al. Y chromosome sequences reveal a short beringian standstill, rapid expansion, and early population structure of native american founders. **Current Biology**, v. 29, n. 1, p. 149–157.e3, 2019.
- POLETO, M. M. et al. Mitochondrial DNA control region diversity in a population from Paraná state—increasing the Brazilian forensic database. **International Journal of Legal Medicine**, v. 133, n. 2, p. 347–351, 2019.
- POZNIK, G. D. et al. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. **Nature Genetics**, v. 48, n. 6, p. 593–599, jun. 2016.
- PREFEITURA DE SÃO PAULO. **Informe Técnico sobre Malária**. São Leopoldo, RSPrefeitura de São Paulo, , 2018. Disponível em: <www.prefeitura.sp.gov.br/covisa>
- PRIMATIVO, G. et al. Bight of Benin: a maternal perspective of four Beninese populations and their genetic implications on the American populations of African ancestry. **Annals of Human Genetics**, v. 81, n. 2, p. 78–90, 2017.
- PRIORI, A. et al. **História do Paraná (séculos XIX e XX)**. 1. ed. Maringá: Eduem, 2012.
- QUARTEY, P. et al. Type 2 diabetes mellitus in glucose-6-phosphate dehydrogenase deficient individuals in a Ghanaian population. **International Journal of Research in Medical Sciences**, v. 8, n. 12, p. 4348, 2020.
- QUINTANA-MURCI, L. et al. Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. **Proceedings of the National Academy of Sciences**, v. 105, n. 5, p. 1596–1601, fev. 2008.
- QUINTERO FERRER, J. M.; PARDO GOVEA, T. C.; BORJAS FUENTES, L. B. Frecuencia de heteroplasmia en las regiones hipervariables HVI y HVII del ADN Mitocondrial en una muestra de la población de Maracaibo, Venezuela. **Revista de Ciencias Forenses de Honduras**, v. 5, n. 2, p. 14–24, 2019.
- RAGGIO, A. Z.; BLEY, R. B.; TRAUCZYNSKI, S. C. **Abordagem histórica sobre a população negra no Estado de Paraná**. Curitiba: SEJU, 2018. v. 2
- RANGEL, A.; RANGEL, L. **Medicamentos e substâncias contraindicados e/ou que exigem uso em assistência hospitalar: Cartilha de substâncias**. 1. ed. São Gonçalo: [s.n.].
- REECE, K. **West African Kingdoms: Empires Of Gold and Trade**. primeira ed. [s.l.] Rourke Pub Group, 2005.
- REIS, J. J. Presença negra: conflitos e encontros. Em: IBGE (Ed.). **Brasil: 500 anos de povoamento**. 1. ed. Rio de Janeiro: IBGE, 2007. p. 80–99.
- RESQUE, R. et al. Male lineages in Brazil: Intercontinental admixture and stratification of the European

- background. **PLoS ONE**, v. 11, n. 4, p. 1–17, 2016.
- RIBAS, K. C. S. **A constituição de identidade quilombola: um olhar acerca da comunidade quilombola do município da Lapa-PR**. Graduate thesis—[s.l.] UNIVERSIDADE FEDERAL DA FRONTEIRA SUL, 2019.
- RIBEIRO, D. **O povo brasileiro: A formação e o sentido do Brasil**. São Paulo: Companhia das Letras, 1995.
- RITO, T. et al. The First Modern Human Dispersals across Africa. **PLoS ONE**, v. 8, n. 11, p. e80031, nov. 2013.
- RODNEY, W. **Como a Europa subdesenvolveu a África**. 1975. ed. Lisboa: [s.n.].
- ROEWER, L. et al. Continent-Wide Decoupling of Y-Chromosomal Genetic Variation from Language and Geography in Native South Americans. **PLoS Genetics**, v. 9, n. 4, p. e1003460, abr. 2013.
- ROSA, A.; BREHM, A. African human mtDNA phylogeography at-a-glance. **Journal of Anthropological Sciences**, v. 89, n. Mar, p. 25–58, 2011.
- ROWOLD, D. et al. At the southeast fringe of the Bantu expansion: Genetic diversity and phylogenetic relationships to other sub-Saharan tribes. **Meta Gene**, v. 2, p. 670–685, 2014.
- RUIZ-LINARES, A. et al. Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. **PLoS Genetics**, v. 10, n. 9, p. e1004572, set. 2014.
- SÁ JUNIOR, M. T. Não é feitiçaria! É tecnologia! Africanos e descendentes na mineração do Mato Grosso setecentista. **Revista Eletrônica História em reflexão**, v. 10, n. 19, p. 81–98, 2016.
- SALAS, A. et al. The making of the African mtDNA landscape. **The American Journal of Human Genetics**, v. 71, n. 5, p. 1082–1111, 2002.
- SALAS, A. et al. The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade. **The American Journal of Human Genetics**, v. 74, n. 3, p. 454–465, mar. 2004.
- SANTOS, F. R.; TYLER-SMITH, C. Reading the human Y chromosome: the emerging DNA markers and human genetic history. **Brazilian Journal of Genetics**, v. 19, n. 4, p. 665–670, 1996.
- SANTOS, H. C. et al. A minimum set of ancestry informative markers for determining admixture proportions in a mixed American population: The Brazilian set. **European Journal of Human Genetics**, v. 24, n. 5, p. 725–731, 2016a.
- SANTOS, S. F. et al. Associação da deficiência da glicose-6-fosfato desidrogenase (G6PD) em população brasileira afrodescendente. **Boletim Informativo Geum**, v. 7, n. 1, p. 106–111, 2016b.
- SCHAAN, A. P. et al. mtDNA structure: the women who formed the Brazilian Northeast. **BMC Evolutionary Biology**, v. 17, n. 1, p. 185, dez. 2017.
- SCHAAN, A. P. et al. New insights on intercontinental origins of paternal lineages in Northeast Brazil. **BMC Evolutionary Biology**, v. 20, n. 1, p. 1–9, 2020.
- SCHÖNHERR, S. et al. Haplogrep 3 - an interactive haplogroup classification and analysis platform. **Nucleic Acids Research**, p. 1–6, abr. 2023.
- SCHWAB, M. E. Tesis doctoral Marisol Elisabet Schwab 2019 Director : Claudio M . Bravi Co-Director : Graciela Bailliet. 2019.
- SECHER, B. et al. The history of the North African mitochondrial DNA haplogroup U6 gene flow into the African, Eurasian and American continents. **BMC Evolutionary Biology**, v. 14, n. 1, p. 109, dez. 2014.
- SENGUPTA, D. et al. Genetic substructure and complex demographic history of South African Bantu speakers. **Nature Communications**, v. 12, n. 1, p. 2080, abr. 2021.
- SILVA, M. et al. 60,000 years of interactions between Central and Eastern Africa documented by major African mitochondrial haplogroup L2. **Scientific Reports**, v. 5, n. 1, p. 12526, dez. 2015.
- SILVA, W. A. et al. mtDNA haplogroup analysis of Black Brazilian and Sub-Saharan populations: Implications for the Atlantic Slave Trade. **Human Biology**, v. 78, n. 1, p. 29–41, 2006.
- SIRUGO, G.; WILLIAMS, S. M.; TISHKOFF, S. A. The Missing Diversity in Human Genetic Studies. **Cell**, v. 177, n. 1, p. 26–31, 2019.
- SLENES, R. W. “Malungu, ngoma vem!”: África coberta e descoberta do Brasil. **Revista USP**, v. 12, n. 12, p. 48, fev. 1992.
- SOARES, P. et al. Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. **The American Journal of Human Genetics**, v. 84, n. 6, p. 740–759, jun. 2009.

- SOARES, P. et al. The expansion of mtDNA haplogroup L3 within and out of Africa. **Molecular Biology and Evolution**, v. 29, n. 3, p. 915–927, mar. 2012.
- STAVENHAGEN, R. **Los pueblos originarios: el debate necesario**. 1. ed. Buenos Aires: CLACSO, 2010. v. 16
- TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, v. 38, n. 7, p. 3022–3027, 2021.
- TARAZONA SANTOS, E. et al. Brasil e a idiossincrasia da miscigenação. **Revista da Universidade Federal de Minas Gerais**, v. 22, n. 1.2, p. 232–249, 2016.
- TATONETTI, N. P.; ELHADAD, N. Fine-scale genetic ancestry as a potential new tool for precision medicine. **Nature Medicine**, v. 27, n. 7, p. 1152–1153, jul. 2021.
- TAVARES, G. M. et al. Measuring the impact of European colonization on Native American populations in Southern Brazil and Uruguay: Evidence from mtDNA. **American Journal of Human Biology**, v. 31, n. 3, p. 1–16, 2019.
- TORRONI, A. et al. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. **Genetics**, v. 130, n. 1, p. 153–162, 1992.
- TORRONI, A. et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. **American Journal of Human Genetics**, v. 53, n. 3, p. 563–590, 1993.
- TORRONI, A. et al. Harvesting the fruit of the human mtDNA tree. **Trends in Genetics**, v. 22, n. 6, p. 339–345, jun. 2006.
- TROMBETTA, B. et al. Phylogeographic Refinement and Large Scale Genotyping of Human Y Chromosome Haplogroup E Provide New Insights into the Dispersal of Early Pastoralists in the African Continent. **Genome Biology and Evolution**, v. 7, n. 7, p. 1940–1950, jul. 2015.
- UNDERHILL, P. A.; KIVISILD, T. Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. **Annual Review of Genetics**, v. 41, n. 1, p. 539–564, 2007.
- VAN OVEN, M. PhyloTree Build 17: Growing the human mitochondrial DNA tree. **Forensic Science International: Genetics Supplement Series**, v. 5, p. e392–e394, 2015.
- VAN OVEN, M.; KAYSER, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. **Human mutation**, v. 30, n. 2, p. 386–394, 2009.
- VÁZQUEZ, E. et al. Avances metodológicos para el estudio conjunto de la información genética , genealógica y geográfica en análisis evolutivos y de distribución. **Revista Chilena de Historia Natural**, v. 2, n. 1, p. 277–297, 2009.
- WALSH, R. **Notices of Brazil in 1828 and 1829 Volume I**. London: Frederick Westley and A.H Davis, 1830.
- WESTPHALEN, C. M. A introdução de escravos novos no Litoral Paranaense. **Revista de História**, v. 44, n. 89, p. 139, 1972.
- WHEELER, E. et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transthectic genome-wide meta-analysis. **PLOS Medicine**, v. 14, n. 9, p. e1002383, set. 2017.
- WOOD, E. T. et al. Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes. **European Journal of Human Genetics**, v. 13, n. 7, p. 867–876, 2005.
- ZHANG, Z. et al. High glucose inhibits glucose-6-phosphate dehydrogenase, leading to increased oxidative stress and β-cell apoptosis. **The FASEB Journal**, v. 24, n. 5, p. 1497–1505, 2010.

ANEXO I - Projeto de pesquisa aprovado pelo Comitê de Ética em Pesquisa e Ciências da Saúde

Projeto de

Pesquisa

Descrição da Pesquisa

Título do Projeto	Genes, ancestralidade e saúde: traçando um perfil de populações afro-brasileiras do Paraná
Pesquisador Principal	Profa. Dra. Marcia Holsbach Beltrame
Colaboradores	Profa. Dra. Claudemira Vieira Gusmão Lopes Profa. Dra. Adriana Inês de Paula Profa. Dra. Karin Braun Prado Profa. Dra. Maria Luiza Petzl-Erler Prof. Dr. Danillo Gardenal Augusto Priscila Ianzen dos Santos Iriel Araceli Joerin Luque Eduardo Delabio Auer Fernando Luiz Ramos Brock Ana Cecília Guimarães Alves Marluci Aparecida da Silva Natalie Mary Sukow
Local de Realização	Laboratório de Genética Molecular Humana, departamento de Genética, Setor de Ciências Biológicas, UFPR
Período da Pesquisa	01/06/2021 a 31/05/2023

Resumo

Estudos de genética humana raramente envolvem populações africanas ou de ancestralidade africana, embora sejam essas as populações geneticamente mais diversas do mundo. Tal diversidade está associada principalmente à origem do homem moderno na África há cerca de 200 mil anos e também pela adaptação a uma grande diversidade de ambientes. O estudo da ancestralidade genômica de uma população, além de contribuir com o estabelecimento das raízes históricas da mesma, também apresenta importantes implicações na área de saúde. Variações de frequências alélicas (i.e., variantes genéticas) entre populações e a presença de variantes população-específicas influem tanto na susceptibilidade a doenças quanto no diagnóstico e na resposta a tratamentos. No Brasil, mais da metade da população é negra, porém a escassez de estudos impacta diretamente no nosso conhecimento e habilidade de promover um atendimento à saúde adequado para a população afro-brasileira. Neste estudo, temos como objetivo investigar a diversidade genética e a saúde da população afro-brasileira, para melhor compreender suas origens e sua diversidade, assim como a forma como sua diversidade genética foi moldada pela evolução e como ela impacta a saúde. Iremos realizar levantamento de dados socioeconômicos, pessoais, históricos familiares e médicos, aferir medidas antropométricas e realizar exames laboratoriais, coletar amostras biológicas para exames e para extração de DNA genômico para análise de marcadores genéticos informativos de saúde e também de ancestralidade. Com esses dados poderemos traçar um perfil genético das populações estudadas, bem como investigar associações entre variantes genéticas e doenças. Além disso, serão realizados exames laboratoriais e medidas antropométricas consideradas como indicadores de saúde, pois estas ainda não foram validadas para a população estudada, e os parâmetros aplicados ainda são os estabelecidos para populações de ancestralidade europeia. Isso é de extrema relevância para a atenção à saúde da população, pois a acurácia de previsões feitas com base no genoma europeu é de apenas cerca de 20%, o que mostra a importância de se estudar essas populações. Com esse estudo iremos contribuir para o desenvolvimento de indicadores de saúde específicos da população e assim aumentar a qualidade do atendimento à saúde dos afro-brasileiros no Brasil. Além disso, os dados genéticos poderão contribuir na elucidação das origens africanas subcontinentais de nossa população.

1 Objetivo da Pesquisa

1.1 Objetivo Geral

- 1.1.1 Investigar o estado de saúde da população Afro-brasileira quilombola no Paraná e possíveis variantes genéticas de risco para doenças.
- 1.1.2 Investigar as origens genômicas da ancestralidade subcontinental Africana nos brasileiros residentes em comunidades quilombolas no Paraná;

1.2 Objetivos Específicos

- 1.2.1 Investigar a epidemiologia de doenças diagnosticadas na população afro-brasileira, descrevendo suas prevalências;
- 1.2.2 Investigar a associação entre doenças comuns encontradas na população e exames laboratoriais realizados e medidas antropométricas, para identificar possíveis fatores de risco na população;
- 1.2.3 Investigar a associação entre doenças comuns encontradas na população e polimorfismos genéticos previamente associados em outras populações;
- 1.2.4 Investigar a associação de resultados de exames laboratoriais e medidas antropométricas com os polimorfismos genéticos analisados.
- 1.2.5 Identificar a origem mais provável, com a máxima precisão possível, das variantes genéticas encontradas em afro-brasileiros (não necessariamente o país de origem, mas principalmente o grupo étnico, como Bantu, Yoruba, etc.);
- 1.2.6 Quantificar a proporção dos alelos identificados que pertencem a cada grupo de origem, para cada indivíduo amostrado;
- 1.2.7 Quantificar a proporção média dos alelos identificados que pertencem a cada grupo de origem para a população afro-brasileira;

2 Relevância Social

A população brasileira é resultado da miscigenação de populações de três ancestralidades principais: os mais de 3 milhões de nativos americanos que viviam aqui antes da chegada dos Europeus, os 500 mil portugueses que chegaram aqui no ano de 1500 e os quase 5 milhões de africanos que foram trazidos para o Brasil entre 1930 e 1856, escravizados pelos europeus (PENA et al., 2009). A miscigenação foi violenta e com viés entre os sexos, com a linhagem paterna quase exclusivamente de origem europeia (98%) (CARVALHO-SILVA et al., 2001) e a linhagem materna predominantemente ameríndia (33%) e africana (28%) (ALVES-SILVA et al., 2000). Essa ancestralidade genômica revela a história de colonização europeia que não é documentada ou passada através das gerações nas famílias (ABEL et al., 2018). A escravidão no Brasil durou até 1888 e entre 1872 e 1975, cerca de 5,5 milhões de imigrantes europeus e asiáticos vieram para o Brasil, tornando nossa história de miscigenação ainda mais complexa (PENA et al., 2009). Hoje, de acordo com o IBGE (Instituto Brasileiro de Geografia e Estatística), 7,6% da população brasileira se autodeclara preta e outros 43,1% Parda ou miscigenada, enquanto 47,7% se declaram brancos, embora as proporções variem em diferentes regiões do país. A grande maioria dos africanos trazidos para as Américas veio das regiões oeste e centro-oeste da África, mas a localização precisa de origem é desconhecida (SCHROEDER et al., 2015). Considerando que mesmo dentro de um mesmo país africano existem diversos grupos étnicos distintos, que raramente são incluídos em estudos genéticos, a ancestralidade específica se torna ainda mais difícil de ser determinada (TISHKOFF et al., 2009). Recentemente, análises genômicas foram utilizadas para se determinar a origem genética de três indivíduos do século 17 na ilha caribenha São Martinho, originados de um grupo Bantu do norte de Camarões e um grupo não Bantu que vivia na região onde hoje se localiza a Nigéria e Gana (SCHROEDER et al., 2015). No Brasil, Congo e Angola são consideradas as principais origens (SCHROEDER et al., 2015). Devido à história da nossa população, a ancestralidade genética no Brasil vai muito além da biologia, com implicações sociais e políticas que afetam profundamente o status socioeconômico, a segurança e a saúde dos afro-brasileiros (ABEL, 2018; KENT and WADE, 2015), e pesquisas genéticas e médicas podem trazer informações sobre a história e a relevância do impacto de nossa ancestralidade em nossa saúde, incluindo susceptibilidade a doenças e resposta a medicamentos. Existem doenças que são mais comuns entre africanos e indivíduos de ancestralidade africana, como doenças cardiovasculares (FLACK et al., 2003), diabetes (BRANCATI et al., 2000), hipertensão (BELFORT et al., 2017), câncer de próstata (ROMERO et al., 2012), função pulmonar reduzida (MENEZES et al., 2015), anemia falciforme (BORGES et al., 2018). Por outro lado, outras doenças são menos prevalentes em africanos, como a doença de Alzheimer (SCHLESINGER et al., 2013). Entretanto, a ciência tradicionalmente ignora essas populações e precisamos de mais estudos para compreender melhor a prevalência e a epidemiologia dessas doenças em afro-brasileiros e qual a relação entre

essas doenças e os resultados de exames laboratoriais e medidas antropométricas, não assumindo que os estudos realizados em europeus ou populações de ancestralidade europeia sejam informativos também para africanos e populações de ancestralidade africana, uma vez que os próprios valores de referência podem ser diferentes entre populações. O objetivo deste estudo é investigar a população afro-brasileira em sua ancestralidade genômica e miscigenação incluindo as populações africanas de origem, sua diversidade genética e a relação dessa diversidade com fenótipos complexos, incluindo doenças e outros fenótipos relacionados direta ou indiretamente à saúde, como exames laboratoriais e medidas antropométricas.

3 Hipóteses a Serem Testadas

Baseando-se no fato de que as populações africanas apresentam alta diversidade genética e variantes genéticas exclusivas que interferem na susceptibilidade à doenças, nossa hipótese é de que a população afro-brasileira, que é sabidamente miscigenada (com ancestralidade africana, indígena e europeia), porém com importante componente de ancestralidade africana principalmente em comunidades remanescentes quilombolas, deve apresentar variantes genéticas em frequências significativamente diferentes da população euro-brasileira e também da própria população afrobrasileira urbana (não quilombola), devido ao histórico de fundação e isolamento das populações quilombolas. As variantes genéticas podem influenciar diretamente na saúde da população, como já relatado em estudos que mostraram que em populações africanas as mesmas doenças podem ser devidas a variantes genéticas distintas (levando inclusive ao subdiagnóstico das mesmas na população). Em doenças multifatoriais, os marcadores genéticos de susceptibilidade também podem ser distintos e as próprias características consideradas indicadores de saúde podem variar entre populações. Assim, nossa hipótese é que a população afro-brasileira e possivelmente as populações remanescentes quilombolas em específico apresentem diferentes marcadores para as mesmas doenças comuns (como hipertensão e diabetes) do que os marcadores considerados no atendimento de saúde à população brasileira. Além disso, nossa hipótese é de que cada comunidade quilombola seria constituída de uma história de ancestralidade única que pode ser revelada através da análise de marcadores genéticos, o que contribuirá também para a elucidação da história da população brasileira, muito pouco documentada principalmente quanto à sua origem africana.

4 Antecedentes Científicos

O Brasil tem uma longa história de pesquisas de ancestralidade genômica da sua população, muitas das quais se valeram do uso de marcadores informativos de ancestralidade autossônicos (MIAs) ou de marcadores

uniparentais (os presentes no cromossomo Y e no DNA mitocondrial) (PENA et al., 2009; SOUZA et al., 2019). O uso de marcadores uniparentais fornece uma ferramenta útil para estudos de ancestralidade e processos demográficos, tendo ajudado na determinação da filogeografia de linhagens materna e paterna em diversas regiões brasileiras e ao entendimento de processos demográficos (BORTOLINI et al., 1999; SALAS et al., 2004; SILVA et al., 2006; KIMURA et al., 2017; SCHAAN et al., 2017). Através do uso destes marcadores pôde-se verificar um forte viés relacionado ao sexo nos cruzamentos ocorridos no Brasil pós-colônia, resultado do acasalamento preferencial entre homens europeus e mulheres ameríndias e africanas. Isto é evidenciado pela maior proporção de ancestralidade europeia nos cromossomos Y enquanto ameríndia e africana na maioria das linhagens de mtDNA (BORTOLINI et al., 1999; ALVES-SILVA et al., 2000; GONÇALVES et al., 2007; SCHAAN et al., 2017). No caso particular da população negra do Paraná, porém, o número de estudos de ancestralidade genética é escasso e insuficiente para uma adequada compreensão das origens africanas da população (POLETO et al., 2019).

As comunidades remanescentes de quilombos foram formadas no período em torno à abolição da escravidão no Brasil por negros e negras como uma forma de organização e resistência contra o sistema escravocrata, e tem se mantido relativamente isoladas desde aquela época, conservando um alto grau de ancestralidade genômica africana (SOUZA et al., 2019). Contudo, a história de cada comunidade é particular e com graus diferentes de interação com europeus e nativos americanos (SOUZA et al., 2019).

Decorrente da origem do ser humano moderno na África, populações africanas exibem a maior diversidade genética e fenotípica do mundo, produto da sua adaptação a múltiplas pressões seletivas como: ambiente, clima, dieta e exposição a patógenos (BELTRAME et al., 2016). Estudos recentes têm demonstrado que a variação genética encontrada em populações africanas é amplamente correlacionada com geografia, classificação linguística e modo de subsistência. Por conta disso, estudos genômicos de populações africanas diversas são cruciais para a compreensão da história evolutiva da humanidade e como a mesma tem um impacto na saúde das populações de hoje (GOMEZ et al., 2014). Isto é especialmente importante para populações de ancestralidade africana recente, como afro-brasileiros.

Apesar do viés eurocêntrico existente nos estudos médicos, existem já trabalhos sobre variantes e doenças de maior prevalência em populações de ancestralidade africana (PEPRAH et al., 2015). Um dos casos mais conhecidos é a anemia falciforme, a mais frequente de todas as hemoglobinopatias no Brasil (ZAGO; COSTA, 1985). A alta prevalência desse traço no país é associada à maior frequência dos alelos HbS e HbC, causadores da doença, nas regiões da África onde a malária é endêmica, já que esses alelos conferem resistência à infecção pelo parasita (GOMEZ et al., 2014). Porém, a distribuição destes alelos é heterogênea entre as diversas populações do

Brasil (SILVA, 2009; OLIVEIRA et al., 2017). Embora as variantes causadoras primárias da anemia falciforme sejam conhecidas, estudos de associação têm identificado genes implicados na severidade desta doença em afro-estadunidenses (PEPRAH et al., 2015).

Doenças complexas como hipertensão, diabetes mellitus tipo 2, asma, câncer de próstata e doença crônica do rim, são conhecidas por apresentar uma prevalência maior em populações de ancestralidade africana nos Estados Unidos, em comparação pessoas de ancestralidade europeia, sugerindo a associação de fatores genéticos população-específicos (FLACK et al., 2003; GOMEZ et al., 2014; PEPRAH et al., 2015). O termo hipertensão designa uma condição patológica de pressão arterial elevada crônica, que se associa frequentemente a alterações funcionais e/ou estruturais de órgãos como o coração, encéfalo, rins e vasos sanguíneos. No Brasil, a hipertensão é considerada um problema de saúde pública. Diversos estudos reportam que a prevalência desta doença é de 1,2 até 2,3 vezes maior em negros do que em brancos em diversas populações da Bahia, Rio de Janeiro e São Paulo (MINISTÉRIO DA SAÚDE, 2001). Ademais, estudos realizados em diversas comunidades remanescentes de quilombos (CRQs) apontam a valores de prevalência que vão de 26 a 52,9 %, valores elevados com relação à média da população total do Brasil (21%, estimada em 2013) (BEZERRA et al., 2013; IBGE, 2013; SOUZA et al., 2014; SANTOS et al., 2019). O diabetes mellitus tipo 2 é um distúrbio metabólico de etiologia múltipla, que resulta em defeitos na secreção de insulina, quase sempre com uma importante contribuição da resistência insulínica. Diferença de outros países, a última Pesquisa Nacional em Saúde realizada no Brasil não mostrou uma prevalência maior desta doença em negros com relação a brancos. Porém, é reconhecido o fato de que os dados epidemiológicos nessa população são limitados e que existe um controle metabólico inadequado decorrente de um menor acesso aos serviços de saúde, principalmente em populações quilombolas (MINISTÉRIO DA SAÚDE, 2001; SOUZA et al., 2014). Entretanto, é observado no Brasil um excesso de complicações associadas a esta doença em negros, inclusive morbidade, em parte associada à maior prevalência de hipertensão nessa população (MINISTÉRIO DA SAÚDE, 2001). Tanto a diabetes quanto a hipertensão são considerados fatores de risco para o surgimento de doenças cardiovasculares, cuja taxa de mortalidade entre negros no Brasil é expressivamente maior do que entre os brancos (LOTUFO; BENENOR, 2013). Dentre os traços mencionados, a doença falciforme, hipertensão, diabetes, tuberculose foram registrados como doenças prevalentes em comunidades quilombolas de São Paulo, além de outras, como obesidade e leishmaniose cutânea (SILVA, 2009; FREITAS et al., 2011). Todas estas doenças possuem um componente genético, seja para o seu desenvolvimento ou como fator agravante (KEATON et al., 2014; PEPRAH et al., 2015). Algumas das doenças de alta prevalência na população afro-brasileira, como diabetes e hipertensão, têm sido associadas, entre outros, a alelos dos genes que codificam o

antígeno leucocitário humano (HLA, do inglês *Human Leucocyte Antigen*) em diversas populações (LUQUE OTERO et al., 1983; SAXENA et al., 2012; NG et al., 2014; LAUGESEN et al., 2015). Estes genes são os mais polimórficos do genoma humano (ROBINSON et al., 2015), tendo seus produtos gênicos um papel central nas respostas imunes adaptativa, através da apresentação de抗ígenos às células T (LEE et al., 2007; NEEFJES et al., 2011), e inata, através da sua ligação a receptores de tipo imunoglobulina em células assassinas naturais (KIR, do inglês *killer cell immunoglobulin-like receptors*) (SAUNDERS et al., 2015; HUDSON; ALLEN, 2016).

Fatores comuns a várias comunidades afro-brasileiras no estado de São Paulo no que se refere à saúde são os relatos de doenças com causas mal definidas, problemas de saúde não diagnosticados e erro médico (SANTOS; TATTO, 2008). Esses achados reforçam a necessidade de uma abordagem conjunta da ancestralidade africana e a identificação das variantes população-específicas na determinação de fenótipos complexos relacionados à saúde assim como também a necessidade de uma atualização dos dados epidemiológicos focados na população negra brasileira e, em particular, na população negra do Paraná, onde há uma limitação de estudos que abordem estas problemáticas.

Um outro traço que apresenta uma base genética diferente entre a população europeia e a africana é a não persistência da enzima lactase em adultos ou hipolactasia primária. A enzima lactase é a responsável por digerir o açúcar do leite e a sua não persistência no indivíduo adulto é um fenótipo comum em humanos de todo o mundo (MATTAR et al., 2009). A atividade desta enzima diminui naturalmente após o desmame na maioria das populações humanas, porém, em populações do norte da Europa (ou indivíduos com ancestralidade dessa região) e populações pastoralistas da África, península Arábica e da Ásia Central observa-se uma alta frequência do fenótipo de persistência da lactase (LP, do inglês *Lactase Persistence*) na vida adulta (RANCIARO et al., 2014). A presença do fenótipo LP nessas populações, cuja sobrevivência requer o consumo de leite, foi demonstrado ser um traço adaptativo recente, resultado de uma forte pressão seletiva (TISHKOFF et al., 2007). Foram descritas diferentes variantes genéticas fortemente associadas ao fenótipo LP nas populações pastoralistas para as quais o contexto genômico tem sugerido um surgimento independente, assim como também um rápido aumento de frequência na população decorrente de uma intensa pressão de seleção. Assim, populações africanas pastoralistas possuem variantes genéticas exclusivas que conferem LP provavelmente surgidas em forma independente às variantes responsáveis por este traço em populações europeias, demonstrando um exemplo de evolução convergente do traço da LP (TISHKOFF et al., 2007). Embora a variante mais comum em europeus associada a LP foi encontrada em várias populações ao longo do Brasil, variantes de provável origem africana foram também detectadas em indivíduos afro-brasileiros de Porto Alegre e Recife, associadas a uma ampla diversidade de

haplótipos, sugerindo que a prevalência do traço de LP em negros pode ser também explicado pelas diversas contribuições genéticas africanas (FRIEDRICH et al., 2012). Poucos estudos têm reportado dados sobre a prevalência desse traço na população brasileira. Contudo, trabalhos têm indicado uma prevalência menor do traço da LP em afro-brasileiros das regiões sul e sudeste com relação a indivíduos brancos, sugerindo a menor frequência de variantes que conferem LP nessa população (SPARVOLI, 1990; ESCOBOZA et al., 2004). Dada a variabilidade da prevalência da LP entre populações, estudos genéticos na população afro-brasileira são necessários para determinar a prevalência e as origens da intolerância à lactose nessa população. Estudos dessa índole podem também contribuir no conhecimento da origem evolutiva desse traço nas populações africanas e avaliar qual ou quais marcadores genéticos devem ser utilizados para o seu diagnóstico na população brasileira.

5 Casuística

Segundo dados do IBGE, a população afro-brasileira constituía, em 2010, 50,7% da população total do Brasil e 28,2% da população do estado do Paraná (IBGE, 2010).

Segundo o Conselho Estadual de Saúde do Paraná, a prevalência de diabetes na população do estado foi de 19,7% no período 2016-2019, porém na população negra a prevalência foi de mais de 30% (SECRETARIA DE ESTADO DA SAÚDE DO PARANÁ, 2016). Ainda, dados do Sistema Único de Saúde para o ano 2012 mostram que a porcentagem de óbitos por causa de doenças cerebrovasculares foi maior em negros do que em brancos (8,4 e 10,1% em pardos e pretos respectivamente e 8,3% em brancos). Similarmente, o número de óbitos causados por diabetes mellitus é maior em negros do que em brancos (4,8 e 5,8% pardos e pretos respectivamente para 4,6% em brancos). No caso da hipertensão, essa doença apresentou maior número de óbitos por 100.000 habitantes em negros do que em brancos (25 e 32,3 óbitos entre pardos e pretos respectivamente, enquanto 17 entre brancos). A doença falciforme, que atinge principalmente pessoas de ancestralidade africana, também apresentou um número maior de óbitos nessa população, sendo 0,28 entre pardos, 0,73 entre pretos e de 0,08 em brancos a cada 100.000 habitantes (MINISTÉRIO DA SAÚDE, 2016). No estado de Paraná foi encontrada uma prevalência na população negra de 2,7% dessa doença, dados coletados do Instituto Paranaense de Hemoterapia e Hematologia (LIDANI et al., 2015). Não existem até o momento dados da prevalência da hipolactasia na população negra do Paraná. Desta forma este trabalho constituiria o primeiro estudo na matéria.

6 Material e Metodologia

Serão incluídos no estudo aproximadamente 250 participantes das comunidades remanescentes de quilombo (CRQ) Feixo e Restinga, do município da Lapa, no estado de Paraná. Eles deverão ser maiores de 18 anos para poder dar consentimento à participação no projeto, e, caso forem do sexo feminino, não serem gestantes.

Após a aprovação do projeto pelo Comitê de Ética em Pesquisa, Ciências da Saúde, iremos divulgar o mesmo nas duas comunidades que são o foco do nosso estudo. Os participantes serão abordados por meio da explicação dos objetivos do estudo, dos procedimentos necessários para a participação e coleta de amostras, incluindo coleta de sangue e medidas antropométricas, e dos potenciais incômodos por esta provocados, assim como os resultados que lhes serão retornados. O contato será realizado de forma virtual, através de e-mail, instagram e WhatsApp, com auxílio dos líderes das duas comunidades. Caso eles aceitem participar, serão convidados a assinar o Termo de Consentimento Livre e Esclarecido (TCLE) e serão aplicados os questionários seguidos da coleta de amostras, como será descrito a continuação.

Toda a informação será coletada através de questionários, incluindo informação etnográfica (nome, sexo, idade, local de nascimento), história familiar (incluindo a ancestralidade auto-declarada dos pais e avós), historial médico (incluindo sintomas não diagnosticados), hábitos de exercício, dieta habitual e hábito de sono. Cada participante receberá um código anônimo para a identificação da sua amostra que será utilizado durante o estudo. Será feita, ademais, a coleta de medidas antropométricas. Para isso será solicitado aos participantes o uso de roupas leves (não casacos ou roupas pesadas). Será medido: altura em pé e sentado, usando um estadiômetro; altura do joelho, comprimento e circunferência do braço superior, circunferência da cintura, abdômen e quadril, usando uma fita antropométrica; força da mão direita ou esquerda (em função da mão dominante) usando um dinamômetro; pressão sanguínea e pulso serão medidas usando um tensiômetro e bioimpedânciâ (incluindo peso, porcentagem de gordura corporal, massa muscular e volume de líquido corporal) será medida usando uma escala digital. O índice de massa corporal (IMC) será calculado usando os dados de altura e peso. Todas as medidas serão feitas em um quarto privado e será solicitado o consentimento do participante para cada uma das medidas feitas, podendo a ou o participante consentir a tomada de certas medidas e não de outras. Os níveis de glicose serão medidos através do uso de um Accu-Chek Active Monitor kit. Este kit utiliza uma lanceta para extrair uma pequena gota de sangue do dedo do participante, a qual é logo absorvida por uma fita de papel para a realização da leitura no aparelho. Será feita também a coleta de sangue dos participantes em dois tubos de 4ml para a extração de DNA e separação de plasma, e em um tubo de 8ml para a coleta de soro. As amostras serão centrifugadas frescas, aliquotadas em microtubos estéreis de 2 ml e congeladas a -80°C até as próximas análises. O DNA será extraído da camada

leucocitária isolada do sangue total, usando um kit comercial. As amostras de DNA serão armazenadas na sua concentração original (que será determinada pela medida no aparelho nanodrop) a -80°C e diluídas a 20 ng/ul em água ultrapura para o seu uso futuro e armazenada a -20°C. A fase de coleta está prevista com 6 meses de duração, enquanto que os seguintes 6 meses serão para genotipagem, análise de perfil lipídico e análises estatísticas. Será utilizada a técnica de SSP-PCR (do inglês *sequence-specific-primer polymerase chain reaction, or allelic specific amplification*, que significa reação em cadeia da polimerase com cebadores sequência-específicos) para a genotipagem de 48 marcadores genéticos conhecidos em outras populações por estarem associados à saúde e ancestralidade assim como também 6 marcadores de ancestralidade materna (do genoma mitocondrial) e paterna (do cromossomo Y). Esta reação será realizada usando um método custo-efetivo que emprega marcadores fluorescentes que serão detectados em um aparelho de PCR em tempo real através da análise da curva de fusão (AMORIM et al., 2018). A análise de perfil lipídico incluirá colesterol total, colesterol HDL e quantificação de triglicerídeos e será realizado no laboratório escola, o qual realiza essas análises de forma rotineira e com emissão de laudo (Laboratório escola da UFPR). Todos os dados coletados serão analisados através da linguagem de programação R para a geração de estatísticas descritivas e gráficos, análises de associação genótipo-fenótipo e para a comparação de frequências alélicas com populações ancestrais de referência africanas, europeias e nativas americanas para a análise de miscigenação e inferência de origens ancestrais.

7 Resultados Esperados do Estudo

Neste projeto iremos gerar novos conhecimentos sobre a genética e a saúde das populações investigadas. Esperamos que estes resultados beneficiem aos participantes diretamente, com os resultados dos exames individuais, e a toda a população afro-brasileira por meio da aquisição de novos conhecimentos sobre a sua história e saúde, assim como também com possíveis aplicações futuras destes dados no cuidado da sua saúde.

Nós planejamos dar um retorno às comunidades participantes por meio da devolução dos resultados de exames laboratoriais e medidas coletadas a cada participante, como também referências e ajuda na sua interpretação. Retornaremos os resultados gerais da população à toda a comunidade participante de uma forma simples que permita a formulação e resposta às dúvidas dos participantes. Ao mesmo tempo, também planejamos divulgar os resultados deste estudo dentro e fora da universidade assim como também através da publicação de artigos em revistas científicas internacionais da área.

O progresso deste estudo será monitorizado de acordo ao nosso cronograma, número de participantes, sucesso na realização dos experimentos laboratoriais propostos e publicações de pesquisa finais, as quais serão submetidas à revisão por pares.

8 Análise Crítica de Riscos e Benefícios

8.1 Quais os benefícios, diretos ou indiretos, para a população e a sociedade?

Este projeto de pesquisa será em grande parte desenvolvido por estudantes de pós-graduação, tendo desta forma um impacto direto no seu desenvolvimento profissional. Além disso, queremos chamar a atenção à necessidade de estudos orientados à população afro-brasileira, incluindo-os na pesquisa em humanos e para a compreensão das suas necessidades específicas de saúde e doença. Esperamos também contribuir com o conhecimento da história da nossa comunidade e as suas consequências no presente, e que este conhecimento seja usado para a promoção da equidade e a eliminação das desigualdades presentes na nossa sociedade atual.

8.2 Quais os riscos inerentes ou decorrentes da pesquisa?

Os riscos são possível constrangimento, tanto no preenchimento do questionário quanto na aferição de medidas, principalmente medidas antropométricas; o risco associado à coleta de sangue, incluindo a dor da picada da agulha, possíveis hematomas no braço e em pessoas com medo de agulha ou sangue, a possibilidade de desmaio; e o risco de se descobrir informações de relevância sobre sua saúde e sua ancestralidade, que podem mudar a percepção de si mesmo e de sua origem, com impacto não apenas no participante da pesquisa mas também em seus familiares. Durante a pesquisa podemos inclusive descobrir accidentalmente casos de filhos não biológicos, porém esse tipo de informação sensível accidental não será revelada aos participantes.

8.3 Qual a possibilidade da ocorrência?

Uma vez que há um alto comprometimento da equipe com a coleta e processamento dos dados, a possibilidade de ocorrência de inconvenientes relacionados à coleta de sangue e medidas antropométricas é baixa, assim como a de quebra de confidencialidade. Em relação aos possíveis constrangimentos durante a aplicação do questionário, tal possibilidade de ocorrência torna-se imprevisível, uma vez que é pautada em características pessoais, inerentes ao indivíduo, porém todas as perguntas do questionário serão avaliadas previamente pelos líderes das comunidades e adequadas, se necessário, à cultura da comunidade.

8.4 Quais as medidas para sua minimização e proteção do participante da pesquisa?

Para minimizar o risco de qualquer tipo de constrangimento, todo e qualquer procedimento (pergunta, aferição, coleta) será feita apenas após a explicação do procedimento oralmente para cada participante e o seu consentimento e tomar-se-á todos os cuidados para a não exposição do paciente, protegendo sua integridade e bem-estar. O questionário on-line será aplicado individualmente para cada participante, evitando que outros membros da comunidade tenham acesso às informações do participante a não ser que sejam compartilhadas pelo próprio participante.

Para minimizar o risco de constrangimento na aferição de medidas elas poderão ser feitas em uma sala fechada, de modo que apenas os pesquisadores estejam presentes, a menos que o participante prefira a presença de outras pessoas.

Para minimizar os riscos associados à coleta de sangue, a coleta será realizada apenas por profissionais com formação e experiência na área de saúde, participantes da equipe, e será realizada de acordo com todos os protocolos padrões, com os participantes sentados com as costas apoiadas, o braço estendido em um suporte apropriado, com seringas estéreis novas e descartadas a cada coleta em recipiente adequado portado pelos pesquisadores da equipe.

Ainda, todos os pesquisadores da equipe usarão máscaras durante todo o período da visita, que serão trocadas a cada duas horas ou sempre que ficarem úmidas, usarão também luvas de procedimento trocadas a cada coleta. Todo o material utilizado que não for descartável (garrote, apoio de braço, cadeira, etc.) será limpo com álcool 70% após cada uso. Além disso, a equipe presente será restrita ao menor número de pessoas possível para realização do trabalho, que deverão todos serem testados para COVID-19 por exame de PCR em tempo real, apresentando resultado negativo nos dias anteriores à coleta.

9 Duração Total da Pesquisa – Cronograma

Identificação da Etapa	Início (01/06/2021)	Término (31/05/2024)
Divulgação da pesquisa nas comunidades	01/06/2021	06/06/2021
Aplicação de TCLE e questionários on-line	07/06/2021	30/06/2021
Análise de dados dos questionários	01/07/2021	31/07/2021
Retorno de informações aos participantes	01/08/2021	07/08/2021

Agendamento de coletas de amostras	08/08/2021	21/08/2021
Coleta de dados e amostras biológicas	22/08/2021	15/09/2021
Análises e exames laboratoriais	23/08/2021	30/09/2021
Elaboração de laudos e envio aos participantes	01/10/2021	08/10/2021
Extração de DNA e análises genéticas	09/10/2021	31/03/2022
Análise de dados e redação de resultados	01/04/2022	31/08/2022
Apresentação dos resultados para a população	01/09/2022	10/09/2022
Publicação e divulgação dos resultados em revistas científicas e de divulgação científica	11/09/2022	31/05/2023

10 Critérios para Suspender ou Encerrar a Pesquisa

- Intercorrelências ou algum risco/ dano à saúde do sujeito participante da pesquisa, não previsto anteriormente;
- Retirada do termo de consentimento informado livre e esclarecido;
- Solicitação pelo Comitê que aprovou ou pela CONEP.

11 Local de Realização da Pesquisa

A primeira etapa da pesquisa será realizada on-line, de forma remota, com a divulgação do projeto para a comunidade através do instagram e WhatsApp. Os questionários on-line serão enviados através de e-mail ou WhatsApp. O retorno das orientações sobre as condições de saúde detectadas também será feito de forma remota, on-line através de e-mail ou WhatsApp. Na segunda etapa da pesquisa, o recrutamento e as coletas serão realizados nas comunidades quilombolas Feixo e Restinga, Município de Lapa, Paraná. O material coletado será analisado no Laboratório Escola e no Departamento de Genética, ambos da UFPR. Os resultados individuais dos participantes serão enviados on-line (e-mail e WhatsApp) e entregues pessoalmente (impressos) nas mesmas comunidades. Os resultados gerais da pesquisa serão comunicados pessoalmente através de palestras e reuniões na própria comunidade e on-line (e-mail, WhatsApp e Instagram).

12 Demonstrativo da Existência de Infraestrutura

O projeto será realizado parte on-line, através dos computadores de trabalho e celulares dos pesquisadores da equipe, parte nas comunidades, que serão visitadas pelos pesquisadores utilizando seus próprios carros para deslocamento e carregando consigo todos os materiais necessários para coleta e medidas.

A parte laboratorial do projeto será realizada principalmente no Laboratório de Genética Molecular Humana, no departamento de Genética, setor de Ciências Biológicas da UFPR (onde trabalha a professora coordenadora do projeto, Marcia Holsbach Beltrame), onde estão localizados os equipamentos necessários para o processamento das amostras, extração de DNA e todas as análises genéticas a serem desenvolvidas (centrífugas, geladeiras, freezers, balanças, agitadores, termocicladores, pipetas e todos os materiais plásticos e descartáveis necessários, detalhados a seguir). Infraestrutura de informática, acesso à internet e bibliotecas também se encontram disponíveis.

Principais equipamentos e materiais permanentes disponíveis no Departamento de Genética, onde será executada a maior parte do trabalho:

1 Aparelho de PCR em tempo real ViiA™ 7, Applied Biosystem

1 Centrífuga refrigerada

1 Banho-maria

1 Destiladores de água

2 Fornos de micro-ondas

2 Estufas de secagem e esterilização

1 Capela de exaustão

2 Refrigeradores

1 Ultra-congelador / freezer -80°C

2 Congeladores / freezer -20°C

17 Micropipetas

2 Impressoras

2 Agitadores magnéticos

1 Purificador de água Mili-Q

2 Balanças analíticas

1 Sistema de fotodocumentação

2 Espectrofotômetros

2 pHmetros

4 Microcentrífugas

4 Termocicladores

3 Computadores

A realização dos exames de hemograma e perfil lipídico serão realizados no Laboratório Escola da UFPR (declaração em anexo), onde há equipamentos específicos que são utilizados na realização dos exames de rotina dos servidores da UFPR, com certificação para emissão de laudos.

13 Propriedade das Informações

As informações coletadas na pesquisa, sejam a partir de questionários on-line, questionários presenciais, medidas, exames rápidos ou exames laboratoriais, serão de responsabilidade dos pesquisadores da equipe do projeto, supervisionadas pelas professoras coordenadoras do projeto. Já no início da coleta de dados, os pesquisadores atribuirão um código a cada participante para a identificação de seus dados e amostras biológicas. Após a coleta de dados e organização em planilhas em formato excel, apenas as professoras responsáveis pela pesquisa terão acesso aos dados de identificação dos indivíduos (arquivos eletrônicos, termos e questionários impressos), sendo que os demais pesquisadores trabalharão apenas com os códigos atribuídos no início da pesquisa. Os arquivos eletrônicos serão armazenados no computador de trabalho das pesquisadoras com cópia em discos externos também armazenados nas salas de cada professora na UFPR. Os resultados de exames serão imediatamente vinculados aos participantes da pesquisa para envio de resultados que posteriormente serão armazenados também nos computadores de trabalho e discos externos das professoras responsáveis pelo projeto. Os resultados serão gerados pelos pesquisadores da equipe, e serão armazenados em seus computadores e discos externos com identificação de códigos para cada participante.

14 Informações Relativas ao Participante da Pesquisa e Características da População a Ser Estudada

Espera-se que participem do estudo aproximadamente 250 moradores das comunidades rurais quilombolas incluídas no projeto. Espera-se, pelas características da população, que a maioria dos participantes serão autodeclarados negros (pardos e/ou pretos). Essas pessoas deverão ter mais de 18 anos e estar em condições de consentir a participação no estudo. No caso das mulheres, não poderão estar grávidas. Espera-se um número semelhante de homens e mulheres participando do estudo.

15 Grupos Vulneráveis

A pesquisa não envolve os seguintes grupos vulneráveis: crianças, idosos, deficientes, gestantes, pessoas incapazes de fornecer consentimento ou pessoas debilitadas fisicamente não serão selecionadas para a pesquisa. Entretanto é provável que pessoas de baixa renda ou pouca escolaridade sejam incluídas no estudo devido a serem

moradores da região onde as coletas serão realizadas. É importante destacar que a população negra no Brasil se encontra historicamente em situação de desigualdade social com relação à população branca, com maiores índices de pobreza com relação a esta, o que também justifica o fato deste estudo ser realizado nas populações quilombolas. Ao mesmo tempo, a inclusão dessas pessoas na pesquisa é essencial pois é o grupo historicamente negligenciado nos âmbitos médico e científico e para o qual este projeto busca corrigir essa disparidade.

16 Fontes de Material da Pesquisa

O material utilizado nesta pesquisa será o seguinte:

- Respostas aos questionários: os questionários serão aplicados de forma on-line através do uso de Formulários do Google. As respostas serão coletadas a fim de obter informação etnográfica, hábitos de vida e histórico médico.
- Amostras biológicas: sangue, soro, plasma e DNA. Amostras de sangue serão coletadas diretamente do participante por pessoas capacitadas para essa tarefa que integram a equipe de pesquisa. Amostras de soro, plasma e DNA serão purificadas a partir das amostras de sangue coletadas. As amostras de soro e plasma serão utilizadas em exames laboratoriais e amostras de DNA serão utilizadas para a análise de variantes genéticas informativas de ancestralidade ou de importância em saúde.
- Medidas antropométricas: serão aferidas por membros da nossa equipe devidamente capacitados para essa tarefa. Essas medidas são importantes para a verificação da associação das mesmas como fatores de risco para doenças.

Este material será utilizado exclusivamente para os propósitos desta pesquisa.

17 Planos para o Recrutamento dos Participantes da Pesquisa

Inicialmente contactaremos os líderes comunitários das duas comunidades quilombolas para revisar todo o material a ser enviado para a comunidade e auxiliar no contato com os futuros participantes da pesquisa. Inicialmente será enviado um curto vídeo de divulgação do projeto (Anexo 22.1), com cerca de 5 minutos, apresentando as pesquisadoras e explicando de forma geral os objetivos do projeto, sua relevância, etapas e por fim o convite para participar do projeto e enviar dúvidas e sugestões para a pesquisa. A divulgação será feita principalmente através de WhatsApp, Instagram e e-mail, e com a ajuda dos líderes comunitários, o vídeo também será exibido em computadores locais na comunidade. As pesquisadoras ficarão então disponíveis para contato por WhatsApp, e-mail e telefone para responder dúvidas, receber sugestões e discutir o projeto.

Após a divulgação e discussão do projeto de pesquisa e as possíveis adequações que forem solicitadas pela comunidade, iniciaremos o envio também on-line dos TCLEs e dos questionários iniciais (Anexo 22.2) para participação na pesquisa. O questionário será no formato Formulários do Google, sendo apresentado inicialmente o TCLE on-line para consentimento e em seguida as perguntas iniciais para cadastro dos participantes, incluindo uma identificação geral de saúde reportada pelas próprias pessoas. As pessoas que optarem por participar do estudo deverão dar consentimento no TCLE on-line, responder o questionário inicial e enviar eletronicamente.

Após a primeira fase de cadastro e levantamento de dados da pesquisa, cada participante receberá eletronicamente de forma privada (e-mail ou WhatsApp) as informações e orientações sobre as doenças ou condições de saúde relatadas em seu questionário. Serão também divulgados eletronicamente na comunidade os resultados do levantamento preliminar de saúde acompanhados de curtos vídeos e cartazes explicando sobre a origem, o risco, o diagnóstico e o tratamento das doenças mais comuns presentes na comunidade (modelo apresentado no Anexo 22.3). A seguir, cada participante será contatado por WhatsApp, e-mail ou telefone para a continuação da pesquisa, com o agendamento da coleta de material biológico, realização de medidas e exames e entrevista mais detalhada sobre seu quadro de saúde, seus hábitos e suas origens. Os participantes que aceitarem participar da segunda etapa da pesquisa deverão assinar um segundo TCLE (impresso) onde será solicitado o consentimento para a realização de coletas de material biológico e medidas antropométricas, assim como exames laboratoriais e análises de DNA. As coletas serão realizadas pela equipe de pesquisa em visita à comunidade (local comunitário ou moradia de cada participante). Nessa mesma visita serão solicitados aos próprios participantes eventuais resultados de exames anteriores e/ou diagnósticos médicos recebidos anteriormente.

As medidas e exames rápidos realizados na visita serão registrados em uma ficha impressa que será imediatamente entregue a cada participante, e posteriormente enviadas eletronicamente, na forma de uma declaração, por WhatsApp ou e-mail, de acordo com a preferência de cada participante. Os resultados dos exames realizados posteriormente no laboratório também serão enviados a cada participante eletronicamente, sempre acompanhados de interpretações e recomendações em texto e em vídeo, acompanhadas do contato das pesquisadoras responsáveis para responder as dúvidas.

Ao final da pesquisa a equipe irá agendar com os líderes das comunidades uma nova visita à comunidade para a apresentação dos resultados gerais e um momento de discussão com a comunidade. Serão também enviados eletronicamente aos participantes da pesquisa e demais membros da comunidade o material de divulgação dos resultados da pesquisa em texto e vídeo, assim como todos os relatos, artigos científicos, resumos, pôsteres e apresentações de trabalho realizadas com os resultados da pesquisa.

18 Critérios de Inclusão e Exclusão

Serão incluídos na pesquisa todos os indivíduos pertencentes às comunidades quilombolas incluídas no estudo.

Critérios de inclusão

- Ser voluntária ou voluntário para participar do estudo.
- Pertencer à comunidade quilombola incluída no estudo.
- Maiores de 18 anos.

Critérios de exclusão

- Gestantes.
- Infecção grave ou qualquer outra doença ativa que dificulte a participação no estudo ou a coleta de material biológico.
- Doença psiquiátrica que possa interferir com o consentimento ou a participação no estudo.

19 Medidas de Proteção ou Minimização de Qualquer Risco Eventual

Listados no item 8.

20 Previsão de Ressarcimento de Gastos aos Participantes da Pesquisa

Como não haverá gastos para os participantes da pesquisa, não haverá também ressarcimento.

21 Referências Bibliográficas

ABEL, S. Of African Descent? Blackness and the Concept of Origins in Cultural Perspective. *Genealogy*, v. 2, n. 1, p. 11, 2018. Disponível em: <<https://www.mdpi.com/2313-5778/2/1/11>>..

ALVES-SILVA, J.; SILVA SANTOS, M. DA; GUIMARÃES, P. E. M.; et al. The Ancestry of Brazilian mtDNA Lineages. *The American Journal of Human Genetics*, v. 67, n. 2, p. 444–461, 2000. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0002929707626548>>..

AMORIM, L. M.; SANTOS, T. H. S.; HOLLENBACH, J. A.; et al. Cost-effective and fast KIR gene-content genotyping by multiplex melting curve analysis. *HLA*, v. 92, n. 6, p. 384–391, 2018. Disponível em:

- <<http://doi.wiley.com/10.1111/tan.13430>>..
- BELFORT, I. K. P.; AVELAR FERNANDES, M.; COSTA NUNES, J. D.; MOUTINHO MONTEIRO, S. C. Elevação de níveis pressórios em uma comunidade quilombola. **Revista Brasileira em Promoção da Saúde**, v. 30, n. 3, p. 1–8, 2017. Disponível em: <<http://periodicos.unifor.br/RBPS/article/view/6135>>..
- BELTRAME, M. H.; RUBEL, M. A.; TISHKOFF, S. A. Inferences of African evolutionary history from genomic data. **Current Opinion in Genetics & Development**, v. 41, p. 159–166, 2016. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.gde.2016.10.002>>..
- BEZERRA, V. M.; ANDRADE, A. C. DE S.; CÉSAR, C. C.; CAIAFFA, W. T. Comunidades quilombolas de Vitória da Conquista, Bahia, Brasil: Hipertensão arterial e fatores associados. **Cadernos de Saude Pública**, v. 29, n. 9, p. 1889–1902, 2013.
- BORGES, R. P. F.; CANUTTO, R. F. C.; GOMES, K. A.; et al. Sickle Cell Anemia in Quilombolas: a Literature Review. **Revista de Patologia do Tocantins**, v. 5, n. 1, p. 34, 2018.
- BORTOLINI, M. C.; SILVA, W. A. DA; GUERRA, D. C. DE; et al. African-derived South American populations: A history of symmetrical and asymmetrical matings according to sex revealed by bi- and uni-parental genetic markers. **American Journal of Human Biology**, v. 11, n. 4, p. 551–563, 1999. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1520-6300\(1999\)11:4%3C551::AID-AJHB15%3E3.0.CO;2-Z](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1520-6300(1999)11:4%3C551::AID-AJHB15%3E3.0.CO;2-Z)>..
- BRANCATI, F. L.; KAO, W. H. L.; FOLSOM, A. R.; WATSON, R. L.; SZKLO, M. Incident Type 2 Diabetes Mellitus in African American and White Adults. **JAMA**, v. 283, n. 17, p. 2253, 2000. Disponível em: <<http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.283.17.2253>>..
- CARVALHO-SILVA, D. R.; SANTOS, F. R.; ROCHA, J.; PENA, S. D. J. The Phylogeography of Brazilian Y-Chromosome Lineages. **The American Journal of Human Genetics**, v. 68, n. 1, p. 281–286, 2001.
- ESCOBOZA, P. M. L.; FERNANDES, M. I. M.; PERES, L. C.; EINERHAND, A. W. C.; GALVÃO, L. C. Adult-type hypolactasia: clinical, morphologic and functional characteristics in brazilian patients at a university hospital. **Journal of Pediatric Gastroenterology and Nutrition**, v. 39, n. 4, p. 361–365, 2004.
- FLACK, J. M.; FERDINAND, K. C.; NASSER, S. A. Epidemiology of Hypertension and Cardiovascular Disease in African Americans. **The Journal of Clinical Hypertension**, v. V, n. I, p. 5–11, 2003.
- FREITAS, D. A.; CABALLERO, A. D.; MARQUES, A. S.; HERNÁNDEZ, C. I. V.; ANTUNES, S. L. N. O. Saúde e comunidades quilombolas: uma revisão da literatura. **Revista CEFAC**, v. 13, n. 5, p. 937–943, 2011.
- FRIEDRICH, D. C.; SANTOS, S. E. B.; RIBEIRO-DOS-SANTOS, Â. K. C.; HUTZ, M. H. Several Different

- Lactase Persistence Associated Alleles and High Diversity of the Lactase Gene in the Admixed Brazilian Population. (D. C. Crawford, Ed.)**PLoS ONE**, v. 7, n. 9, p. e46520, 2012. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0046520>>..
- GOMEZ, F.; HIRBO, J.; TISHKOFF, S. A. Genetic Variation and Adaptation in Africa: Implications for Human Evolution and Disease. **Cold Spring Harbor Perspectives in Biology**, v. 6, n. 7, p. 6-a008524, 2014. Disponível em: <<http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a008524>>..
- GONÇALVES, V. F.; CARVALHO, C. M. B.; BORTOLINI, M. C.; BYDLOWSKI, S. P.; PENA, S. D. J. The phylogeography of African Brazilians. **Human Heredity**, v. 65, n. 1, p. 23–32, 2007.
- HUDSON, L. E.; ALLEN, R. L. Leukocyte Ig-like receptors - A Model for MHC class I disease associations. **Frontiers in Immunology**, v. 7, n. JUL, p. 1–8, 2016.
- IBGE. Censo Demográfico 2010. Disponível em: <<https://sidra.ibge.gov.br/Tabela/3175>>. Acesso em: 18/8/2019.
- IBGE, I. B. DE G. E. **Pesquisa Nacional em Saúde 2013 - Percepção do estado de saúde, estilos de vida e doenças crônicas**. Primeira ed. Rio de Janeiro: IBGE, 2013.
- KEATON, J. M.; COOKE BAILEY, J. N.; PALMER, N. D.; et al. A comparison of type 2 diabetes risk allele load between African Americans and European Americans. **Human Genetics**, v. 133, n. 12, p. 1487–1495, 2014. Disponível em: <<http://link.springer.com/10.1007/s00439-014-1486-5>>..
- KENT, M.; WADE, P. Genetics against race: Science, politics and affirmative action in Brazil. **Social Studies of Science**, v. 45, n. 6, p. 816–838, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/0306312715610217>>..
- KIMURA, L.; NUNES, K.; MACEDO-SOUZA, L. I.; et al. Inferring paternal history of rural African-derived Brazilian populations from Y chromosomes. **American Journal of Human Biology**, v. 29, n. 2, p. 1–11, 2017.
- LAUGESEN, E.; ØSTERGAARD, J. A.; LESLIE, R. D. G. Latent autoimmune diabetes of the adult: current knowledge and uncertainty. **Diabetic Medicine**, v. 32, n. 7, p. 843–852, 2015. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1111/dme.12700>>..
- LIDANI, K. C. F.; BARROS, R. F.; BOVO, F. Relationship between the prevalence of hemoglobin S and the ethnic background of blood donors in Paraná state. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 51, n. 4, p. 212–217, 2015. Disponível em: <<http://www.gnresearch.org/doi/10.5935/1676-2444.20150035>>..
- LOTUFO, P. A.; BENENOR, I. J. M. Race and stroke mortality in Brazil. **Revista de Saúde Pública**, v. 47, n. 6, p. 1201–1204, 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102013000601201&lng=pt&nrm=iso&tlang=en>.

LUQUE OTERO, M.; MARTELL CLAROS, N.; LLORENTE PÉREZ, L.; FERNÁNDEZ PINILLA, C.; FERNÁNDEZ-CRUZ, A. Severe hypertension in the Spanish population. Association with specific HLA antigens. **Hypertension**, v. 5, n. 6, 1983.

MATTAR, R.; MONTEIRO, M. S.; VILLARES, C. A.; et al. Frequency of LCT -13910C>T single nucleotide polymorphism associated with adult-type hypolactasia/lactase persistence among Brazilians of different ethnic groups. **Nutrition Journal**, v. 8, n. 1, p. 46, 2009. Disponível em: <<http://nutritionj.biomedcentral.com/articles/10.1186/1475-2891-8-46>>..

MENEZES, A. M. B.; WEHRMEISTER, F. C.; HARTWIG, F. P.; et al. African ancestry, lung function and the effect of genetics. **European Respiratory Journal**, v. 45, n. 6, p. 1582–1589, 2015. Disponível em: <<http://dx.doi.org/10.1183/09031936.00112114>>..

MINISTÉRIO DA SAÚDE. Manual de Doenças Mais Importantes, por Razões Étnicas, na População Brasileira Afro-Descendente. 1st ed. Brasília: Ministério da Saúde, 2001.

MINISTÉRIO DA SAÚDE. Temático Saúde da População Negra. 1st ed. Brasília: Ministério da Saúde, 2016.

NG, M. C. Y.; SHRINER, D.; CHEN, B. H.; et al. Meta-Analysis of Genome-Wide Association Studies in African Americans Provides Insights into the Genetic Architecture of Type 2 Diabetes. **PLOS Genetics**, v. 10, n. 8, p. e1004517, 2014.

OLIVEIRA, S. F.; PEDROSA, M. A. F.; SOUSA, S. M. B.; et al. Heterogeneous Distribution of HbS and HbC Alleles in Afro-derived Brazilian Populations. **International Journal of Human Genetics**, v. 2, n. 3, p. 153–159, 2017.

PENA, S. D. J.; BASTOS-RODRIGUES, L.; PIMENTA, J. R.; BYDLOWSKI, S. P. DNA tests probe the genomic ancestry of Brazilians. **Brazilian Journal of Medical and Biological Research**, v. 42, n. 10, p. 870–876, 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-879X2009001000001&lng=en&tlang=en>..

PEPRAH, E.; XU, H.; TEKOLA-AYELE, F.; ROYAL, C. D. Genome-Wide Association Studies in Africans and African Americans: Expanding the Framework of the Genomics of Human Traits and Disease. **Public Health Genomics**, v. 18, n. 1, p. 40–51, 2015. Disponível em: <<https://www.karger.com/Article/FullText/367962>>..

POLETTO, M. M.; MALAGHINI, M.; SILVA, J. S.; BICALHO, M. G.; BRAUN-PRADO, K. Mitochondrial DNA control region diversity in a population from Parana state—increasing the Brazilian forensic database. **International Journal of Legal Medicine**, v. 133, n. 2, p. 347–351, 2019. International Journal of Legal Medicine.

RANCIARO, A.; CAMPBELL, M. C.; HIRBO, J. B.; et al. Genetic origins of lactase persistence and the spread

of pastoralism in africa. **American Journal of Human Genetics**, v. 94, n. 4, p. 496–510, 2014. The American Society of Human Genetics. Disponível em: <<http://dx.doi.org/10.1016/j.ajhg.2014.02.009>>..

ROBINSON, J.; HALLIWELL, J. A.; HAYHURST, J. D.; et al. The IPD and IMGT/HLA database: allele variant databases. **Nucleic Acids Research**, v. 43, n. D1, p. D423–D431, 2015. Disponível em: <<http://academic.oup.com/nar/article/43/D1/D423/2438496/The-IPD-and-IMGT-HLA-database-allele-variant>>..

ROMERO, F. R.; ROMERO, A. W.; ALMEIDA, R. M. S. DE; TAMBARA FILHO, R. The prevalence of prostate cancer in Brazil is higher in Black men than in White men: systematic review and meta-analysis. **International braz j urol**, v. 38, n. 4, p. 440–447, 2012. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1677-55382012000400002&lng=en&tlang=en>..

SALAS, A.; RICHARDS, M.; LAREU, M. V.; et al. The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade. **American Journal of Human Genetics**, v. 74, n. 3, p. 454–465, 2004.

SANTOS, D. M. S.; PRADO, B. S.; OLIVEIRA, C. C. DA C.; ALMEIDA-SANTOS, M. A. Prevalence of Systemic Arterial Hypertension in Quilombola Communities, State of Sergipe, Brazil. **Arquivos Brasileiros de Cardiologia**, , n. April 2017, 2019. Disponível em: <<http://www.gnresearch.org/doi/10.5935/abc.20190143>>..

SANTOS, K. M. P.; TATTO, N. **Agenda Socioambiental de Comunidades Quilombolas do Vale do Ribeira**. First ed. São Paulo: Instituto Socioambiental, 2008.

SAUNDERS, P. M.; VIVIAN, J. P.; O'CONNOR, G. M.; et al. A bird's eye view of NK cell receptor interactions with their MHC class I ligands. **Immunological Reviews**, v. 267, n. 1, p. 148–166, 2015.

SAXENA, R.; ELBERS, C. C.; GUO, Y.; et al. Large-Scale Gene-Centric Meta-Analysis across 39 Studies Identifies Type 2 Diabetes Loci. **The American Journal of Human Genetics**, v. 90, p. 410–425, 2012.

SCHAAN, A. P.; COSTA, L.; SANTOS, D.; et al. mtDNA structure: the women who formed the Brazilian Northeast. **BMC Evolutionary Biology**, v. 17, n. 1, p. 185, 2017. BMC Evolutionary Biology. Disponível em: <<http://bmcevolbiol.biomedcentral.com/articles/10.1186/s12862-017-1027-7>>..

SCHLESINGER, D.; GRINBERG, L. T.; ALBA, J. G.; et al. African ancestry protects against Alzheimer's disease-related neuropathology. **Molecular Psychiatry**, v. 18, n. 1, p. 79–85, 2013. Nature Publishing Group. Disponível em: <<http://dx.doi.org/10.1038/mp.2011.136>>..

SCHROEDER, H.; ÁVILA-ARCOS, M. C.; MALASPINAS, A.-S.; et al. Genome-wide ancestry of 17th-century enslaved Africans from the Caribbean. **Proceedings of the National Academy of Sciences**, v. 112, n. 12, p. 201421784, 2015.

SECRETARIA DE ESTADO DA SAÚDE DO PARANÁ. **Plano Estadual de Saúde Paraná 2016-2019**. Curitiba, 2016.

- SILVA, A. M. DA. **Saúde nos Quilombos.** First ed. São Paulo: Instituto de Saúde – SESSP, 2009.
- SILVA, W. A.; BORTOLINI, M. C.; SCHNEIDER, M. P. C.; et al. mtDNA Haplogroup Analysis of Black Brazilian and Sub-Saharan Populations: Implications for the Atlantic Slave Trade. **Human Biology**, v. 78, n. 1, p. 29–41, 2006. Disponível em: <http://muse.jhu.edu/content/crossref/journals/human_biology/v078/78.1silva.html>..
- SOUZA, A. M. DE; RESENDE, S. S.; SOUSA, T. N. DE; BRITO, C. F. A. DE. A systematic scoping review of the genetic ancestry of the Brazilian population. **Genetics and Molecular Biology**, v. No prelo, 2019. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572019005020104&tlang=en>..
- SOUZA, C. L.; BARROSO, S. M.; GUIMARÃES, M. D. C. Oportunidade perdida para diagnóstico oportunista de diabetes mellitus em comunidades quilombolas do sudoeste da Bahia, Brasil. **Ciência & Saúde Coletiva**, v. 19, n. 6, p. 1653–1662, 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232014000601653&tlang=pt&tlang=pt>..
- SPARVOLI, A. C. **Malabsorção de lactose do adulto. Prevalencia na população sulina. Aspectos genéticos e evolutivos do polimorfismo da atividade da lactose**, 1990. Universidade Estadual de Campinas.
- TISHKOFF, S. A.; REED, F. A.; FRIEDELAENDER, F. R.; et al. The Genetic Structure and History of Africans and African Americans. (H. Bönig, Ed.)**Science**, v. 324, n. 5930, p. 1035–1044, 2009. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0178059>>..
- TISHKOFF, S. A.; REED, F. A.; RANCIARO, A.; et al. Convergent adaptation of human lactase persistence in Africa and Europe. **Nature Genetics**, v. 39, n. 1, p. 31–40, 2007. Disponível em: <<http://www.nature.com/articles/ng1946>>..
- ZAGO, M. A.; COSTA, F. F. Hereditary haemoglobin disorders in brazil. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 79, n. 3, p. 385–388, 1985.

22 Anexos

22.1 Vídeo de divulgação do projeto e convite para participação

22.2 Questionário inicial

Questionário 1: Informações pessoais

Aqui pedimos algumas informações sobre você e sua família. Seus dados serão usados apenas para a pesquisa e você será identificado através de um código da pesquisa, seu nome e dados não serão divulgados. Caso você não saiba alguma resposta, coloque apenas "não sei".

1. Nome e sobrenome:

2. Nome de solteiro(a) ou outro nome caso tenha alterado seu nome por qualquer motivo:

3. Sexo biológico:

- 1. Feminino
- 2. Masculino

4. Idade:

5. Data de nascimento (dia/mês/ano):

6. Local de nascimento (comunidade, cidade, estado):

7. Cor ou raça:

- 1. Amarela
- 2. Branca
- 3. Indígena
- 4. Parda
- 5. Preta

8. Nome e sobrenome da mãe biológica:

9. Nome de solteira da mãe (caso o nome tenha sido alterado):

10. Qual a cor / raça da sua mãe?

- 1. Amarela
- 2. Branca
- 3. Indígena
- 4. Parda
- 5. Preta
- 6. Não sei

11. Qual o local de nascimento da sua mãe (comunidade, cidade e estado se souber)?

12. Nome e sobrenome do pai biológico:

13. Qual a cor / raça do seu pai:

- 1. Amarela
- 2. Branca
- 3. Indígena
- 4. Parda
- 5. Preta
- 6. Não sei

14. Qual o local de nascimento do seu pai (comunidade, cidade e estado se souber)?

Questionário 2: Doenças crônicas

As perguntas deste módulo são sobre doenças crônicas. Vamos fazer perguntas sobre diagnóstico de doenças, uso dos serviços de saúde e tratamento dos problemas. Caso não se lembre exatamente, responda a sua melhor estimativa.

1. Toma algum medicamento? Se sim, qual ou quais?

2. Quando foi a última vez que o(a) sr(a) teve sua pressão arterial medida?

- 1. Há menos de 6 meses

2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. 3 anos ou mais
6. Nunca

3. Algum médico já lhe deu o diagnóstico de hipertensão arterial (pressão alta)?

1. Sim (Passa para o questionário específico)
2. Apenas durante a gravidez (só para mulheres)
3. Não

4. Quando foi a última vez que você fez exame de sangue para medir a glicemia, isto é, o açúcar no sangue?

1. Há menos de 6 meses
2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. Há 3 anos ou mais
6. Nunca fez

5. Algum médico já lhe deu o diagnóstico de diabetes?

1. Sim (Passa para o questionário específico)
2. Não

6. Nas duas últimas semanas, com que frequência você teve problemas no sono, como dificuldade para adormecer, acordar frequentemente à noite ou dormir mais do que de costume?

1. Nenhum dia
2. Menos da metade dos dias
3. Mais da metade dos dias
4. Quase todos os dias

7. Nas duas últimas semanas, com que frequência você teve problemas por não se sentir descansado(a) e disposto(a) durante o dia, sentindo-se cansado(a), sem ter energia?

1. Nenhum dia
2. Menos da metade dos dias
3. Mais da metade dos dias
4. Quase todos os dias

8. Nas duas últimas semanas, com que frequência você teve problemas na alimentação, como ter falta de apetite ou comer muito mais do que de costume?

1. Nenhum dia
2. Menos da metade dos dias
3. Mais da metade dos dias
4. Quase todos os dias

9. Quando foi a última vez que você fez exame de sangue para medir o colesterol e triglicerídeos?

1. Há menos de 6 meses
2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. Há 3 anos ou mais
6. Nunca fez

10. Algum médico já lhe deu o diagnóstico de colesterol alto?

1. Sim (**Passa para o questionário específico**)
2. Não

11. Algum médico já lhe deu o diagnóstico de uma doença do coração, tais como infarto, angina, insuficiência cardíaca ou outra?

1. Infarto
2. Angina
3. Insuficiência cardíaca
4. Outra (*Especifique:* _____)

Se afirmativo em alguma das opções acima passa para o questionário específico.

12. Algum médico já lhe deu o diagnóstico de AVC (Acidente Vascular Cerebral) ou derrame? 1.

Sim (**Passar para o questionário específico**)

2. Não

13. Algum médico já lhe deu o diagnóstico de asma (ou bronquite asmática)?

1. Sim (**Passar para o questionário específico**)
2. Não

14. Algum médico já lhe deu o diagnóstico de alguma doença no pulmão, tais como enfisema pulmonar, bronquite crônica ou DPOC (Doença Pulmonar Obstrutiva Crônica)?

1. Enfisema pulmonar
2. Bronquite crônica
3. Outro (*Especifique:* _____)

Se afirmativo em alguma das opções acima passa para o questionário específico.

15. Algum médico já lhe deu o diagnóstico de artrite ou reumatismo?

1. Sim (**Passar para o questionário específico**)
2. Não

16. Algum médico já lhe deu algum diagnóstico de câncer?

1. Sim (**Passar para o questionário específico**)
2. Não

17. Algum médico já lhe deu o diagnóstico de insuficiência renal crônica?

1. Sim (**Passar para o questionário específico**)
2. Não

Questionário específico: Hipertensão

1- Que idade você tinha no primeiro diagnóstico de hipertensão arterial (pressão alta)?

1. _____ anos
2. Menos de 1 ano

2. Você vai ao médico/serviço de saúde regularmente por causa da hipertensão arterial (pressão alta)?

1. Sim
2. Não, só quando tem algum problema
3. Nunca vai

3. Quando foi a última vez que você recebeu assistência médica por causa da hipertensão arterial?

1. Há menos de 6 meses
2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. Há 3 anos ou mais
6. Nunca recebeu

4. Em algum dos atendimentos para hipertensão, algum médico ou outro profissional de saúde lhe deu alguma dessas recomendações?

1. Manter uma alimentação saudável (com frutas e vegetais)
2. Manter o peso adequado
3. Ingerir menos sal
4. Praticar atividade física regular
5. Não fumar
6. Não beber em excesso
7. Fazer o acompanhamento regular
8. Outro (*Especifique:* _____)

5. Alguma vez você ficou internado por causa da hipertensão ou de alguma complicaçāo?

1. Sim
2. Não

Questionário específico: Diabetes

1. Que idade você tinha no primeiro diagnóstico de diabetes?

1. _____ anos

2. Menos de 1 ano

2. Você vai ao médico/serviço de saúde regularmente por causa do diabetes?

1. Sim

2. Não, só quando tem algum problema

3. Nunca vai

3. Nas duas últimas semanas, por causa do diabetes, você:

1. Tomou medicamentos orais para baixar o açúcar?

2. Usou insulina?

4. Quando foi a última vez que você recebeu assistência médica por causa do diabetes?

1. Há menos de 6 meses

2. Entre 6 meses e menos de 1 ano

3. Entre 1 ano e menos de 2 anos

4. Entre 2 anos e menos de 3 anos

5. Há 3 anos ou mais

6. Nunca recebeu

5. Em algum dos atendimentos para diabetes, algum médico ou outro profissional de saúde lhe deu alguma dessas recomendações?

1. Manter uma alimentação saudável (com frutas e vegetais)

2. Manter o peso adequado

3. Praticar atividade física regular

4. Não fumar

5. Não beber em excesso

6. Diminuir o consumo de carboidratos (massas, pães etc.)

7. Medir a glicemia em casa

8. Examinar os pés regularmente

9. Outro (*Especifique:* _____)

6. Quando foi a última vez que realizaram um exame de vista ou fundo de olho em que dilataram sua pupila?

1. Há menos de 6 meses

2. Entre 6 meses e menos de 1 ano

3. Entre 1 ano e menos de 2 anos

4. Entre 2 anos e menos de 3 anos

5. Há 3 anos ou mais

6. Nunca fez

7. Quando foi a última vez que um médico ou profissional de saúde examinou seus pés para verificar sensibilidade ou presença de feridas ou irritações?

1. Há menos de 6 meses
2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. Há 3 anos ou mais
6. Nunca teve os pés examinados

8. Você tem ou teve alguma destas complicações?

1. Infarto
2. AVC (Acidente Vascular cerebral) ou derrame
3. Outro problema circulatório
4. Problema nos rins
5. Úlcera/ferida nos pés
6. Problemas na vista
7. Amputação de membros (pés, pernas, mãos ou braços)
8. Coma diabético
9. Outro (*Especifique:* _____)

9. Alguma vez você se internou por causa do diabetes ou de alguma compilação?

1. Sim
2. Não

10. Há quanto tempo foi a última internação por causa do diabetes ou de alguma compilação?

1. Há menos de 6 meses
2. Entre 6 meses e menos de 1 ano
3. Entre 1 ano e menos de 2 anos
4. Entre 2 anos e menos de 3 anos
5. Há 3 anos ou mais

11. Em geral, em que grau o diabetes ou alguma compilação do diabetes limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: Dislipidemias

1. Que idade você tinha no primeiro diagnóstico de colesterol alto?

1. _____ anos
2. Menos de 1 ano

2. Algum médico ou outro profissional de saúde lhe deu algumas das seguintes recomendações por causa do colesterol alto?

1. Manter uma alimentação saudável (com frutas e vegetais)
2. Manter o peso adequado
3. Prática de atividade física
4. Tomar medicamentos
5. Não fumar
6. Fazer acompanhamento regular

Questionário específico: Cardiopatias**1. Que idade você tinha no primeiro diagnóstico da doença do coração?**

1. _____ anos
2. Menos de 1 ano

2. O que você faz atualmente por causa da doença do coração?

1. Dieta
2. Prática de atividade física
3. Toma medicamentos
4. Outro (Especifique _____)

3. Você já fez alguma cirurgia de ponte de safena ou colocação de stent ou angioplastia?

1. Sim
2. Não

4. Em geral, em que grau a doença do coração limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: Acidente Cérebro Vascular**1. Quantos derrames (ou AVC) você já teve?**

Quantos _____

2. Que idade você tinha no primeiro diagnóstico do derrame (ou AVC)?

1. _____ anos
2. Menos de 1 ano

3. Por causa do derrame (ou AVC), você realizou tomografia ou ressonância da cabeça?

1. Sim
2. Não

4. O que você faz atualmente por causa do derrame (ou AVC)?

1. Dieta
2. Fisioterapia
3. Outras terapias de reabilitação
4. Toma aspirina
5. Toma outros medicamentos
6. Outro (*Especifique:* _____)

5. Em geral, em que grau o derrame (ou AVC) limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: Asma

1. Que idade você tinha no primeiro diagnóstico de asma?

1. _____ anos
2. Menos de 1 ano

2. Nos últimos 12 meses, você teve alguma crise de asma?

1. Sim
2. Não

3. O que você faz atualmente por causa da asma?

1. Usa medicamentos (inaladores, aerossol ou comprimidos)
2. Outro (*Especifique:* _____)

4. Em geral, em que grau a asma limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita

2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: Artrite e reumatismo

1. Que idade você tinha no primeiro diagnóstico de artrite ou reumatismo?

1. _____ anos

2. O que você faz atualmente por causa da artrite ou reumatismo?

1. Exercício ou atividade física
2. Fisioterapia
3. Usa medicamentos ou injeções
4. Faz acupuntura
5. Outro (*Especifique:* _____)

3. Você já fez alguma cirurgia por causa da artrite ou reumatismo?

1. Sim
2. Não

4. Em geral, em que grau a artrite ou reumatismo limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: doenças pulmonares

1. Que idade você tinha no primeiro diagnóstico da doença no pulmão?

1. _____ anos
2. Menos de 1 ano

2. O que você faz atualmente por causa da doença no pulmão?

1. Usa medicamentos (inaladores, aerosol ou comprimidos)
2. Usa oxigênio
3. Fisioterapia respiratória
3. Outro (*Especifique:* _____)

3. Em geral, em que grau a doença do pulmão limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: Câncer

1. No primeiro diagnóstico de câncer, que tipo de câncer você tem ou teve?

1. Pulmão
2. Intestino
3. Estômago
4. Mama (*só para mulheres*)
5. Colo de útero (*só para mulheres*)
6. Próstata (*só para homens*)
7. Pele
8. Outro (*Especifique:* _____)

2. Que idade você tinha no primeiro diagnóstico de câncer?

1. _____ anos
2. Menos de 1 ano

3. Em geral, em que grau o câncer ou algum problema provocado pelo câncer limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

Questionário específico: doença renal crônica

1. Que idade você tinha no primeiro diagnóstico de insuficiência renal crônica?

1. _____ anos
2. Menos de 1 ano

2. O que você faz ou fez por causa da insuficiência renal crônica?

1. Toma medicamentos
2. Hemodiálise
3. Diálise peritoneal
4. Fez transplante de rim
5. Outro (*Especifique:* _____)

3. Em geral, em que grau a insuficiência renal crônica limita as suas atividades habituais (tais como trabalhar, realizar afazeres domésticos, etc.)?

1. Não limita
2. Um pouco
3. Moderadamente
4. Intensamente
5. Muito intensamente

22.3 Modelo de vídeo explicando uma doença

23 Qualificação das Pesquisadoras e dos Pesquisadores

Profa. Dra. Marcia Holsbach Beltrame : <http://lattes.cnpq.br/0412289083800375>

Profa. Dra. Claudemira Vieira Gusmão Lopes : <http://lattes.cnpq.br/1342667420502622>

Profa. Dra. Adriana Inês de Paula : <http://lattes.cnpq.br/0006669852273365>

Profa. Dra. Karin Braun Prado : <http://lattes.cnpq.br/0480268360549702>

Profa. Dra. Maria Luiza Petzl-Erler : <http://lattes.cnpq.br/8771103629794712>

Prof. Dr. Danillo Gardenal Augusto: <http://lattes.cnpq.br/2094051271434073>

Priscila Ianzen dos Santos: <http://lattes.cnpq.br/0651935154755446>

Fernando Luiz Ramos Brock: <http://lattes.cnpq.br/9508362137398369>

Iriel Araceli Joerin Luque : <http://lattes.cnpq.br/4820779308934525>

Ana Cecília Guimarães Alves : <http://lattes.cnpq.br/5854527240227168>

Natalie Mary Sukow : <http://lattes.cnpq.br/1204632838605419>

Marluci Aparecida da Silva : <http://lattes.cnpq.br/4006146316908899>

Eduardo Delabio Auer : <http://lattes.cnpq.br/7121996429118604>

24 Orçamento financeiro

Item de despesa: Aquisição Material de Consumo	Quantidade	Valor unitário (médio) (R\$)	Valor Total (R\$)
Accu-Chek Active stripes 50 unidades (6 reagentes)	21	80	1680
Accu Chek lancetas 25 unidades	42	25	1050
Lactol solução de lactose 50g	250	6	1500
Tubo de coleta de sangue a vácuo com EDTA 100 un (04ml-K3-EDTA)	6	50	300
Tubo de coleta de sangue a vácuo para soro - gel SST II 8,5 ml 100 un	3	160	480
Agulha para coleta de sangue a vácuo 25x7mm 100 un	3	40	120
Lâmina para microscopia 26 X 76 mm ponta fosca lapidada K5-7105-1 pacote com 50 unidades	10	20	200
Corante May grunwald 1000 mL	1	65	65
Corante Giemsa 1000 mL	1	85	85
Qualitrol 1H - Calibrador Labtest	1	110	110
Qualitrol 2H - Calibrador Labtest	1	110	110
HDL Cholesterol- Labtest 250 reações	2	350	700
Liquiform Cholesterol - Labtest 1212 reações	1	250	250
Liquiform Triglycerides - Labtest 1212 reações	1	370	370
DNA extraction kit Illustra blood genomicPrep Mini Spin Kit, 250 rxns -GE Healthcare	1	2700	2700
Algodão 95g	5	9	45
Álcool etílico 70% 1 L	3	10	30
Band-aid redondo 500 un	1	30	30
Luvas de procedimento 100 un	10	40	400
Máscara descartável de proteção individual - 100 un	5	130	650
Microtubos 2 ml - 1000 un	3	100	300
Pipette tips 1000 ul -T-1000 Y caixa com 96	6	100	600
Síntese de oligonucleotídeos iniciações (25 bases)	10	50	500
Enzima Taq DNA Polymerase 500U	2	1200	2400
dNTP Set Invitrogen/10297018 100mM	1	1500	1500
Placas de 96 poços para PCR Axygen / PCR-96-ABC 50un	2	980	1960
Selante platemax Axygen Caixa com 100 unidades	1	100	100
Ponteira 1-200 ul, sem filtro Axygen/T-200-Y-PCT	1	85	85
Ponteira em Polipropileno Axygen 0,5-20ul Longa PT/1000 Axygen/T-400	4	110	440
Enzima de restrição PvuII 2500U	1	180	180
Enzima de restrição HincII 2500U	1	820	820
Enzima de restrição MspI 3000U	1	240	240
Valor Total		R\$	20000

ANEXO II - Artigo de divulgação: Saúde de brasileiros negros em tempos de COVID-19

27/09/2021 10:31

Ciência Hoje | Saúde de brasileiros negros em tempos de covid-19

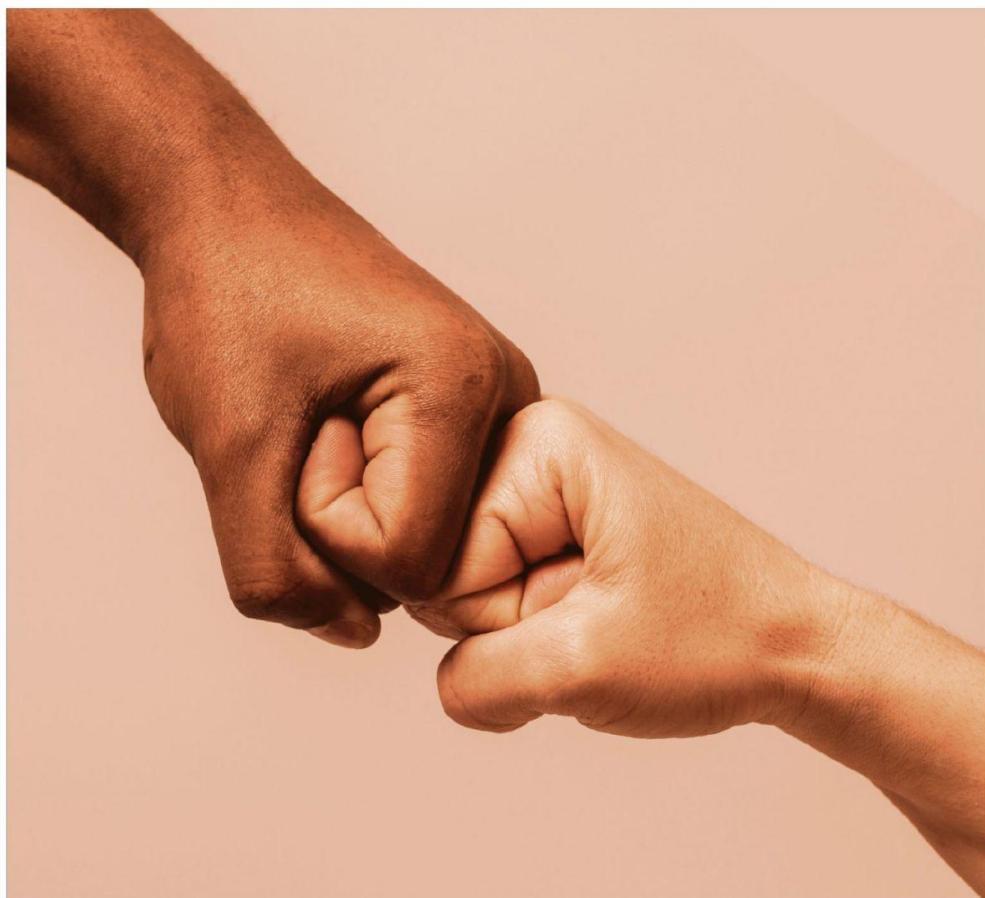


SAÚDE DE BRASILEIROS NEGROS EM TEMPOS DE COVID-19

[Página Inicial](#) > [Artigo](#)

A crise sanitária desencadeada pela pandemia de covid-19 ressaltou as profundas desigualdades étnico-raciais em matéria de saúde no Brasil e no mundo. Diante desse panorama, pesquisadores e instituições governamentais de fomento à pesquisa brasileiras encontram-se desafiados a eliminar o viés eurocêntrico dos estudos biomédicos e efetivar a equidade em saúde no dia a dia de sua população.

CRÉDITO: FOTO ADOBE STOCK



Em 2006, o Sistema Único de Saúde (SUS) implementou a Política Nacional de Saúde Integral da População Negra (PNSIPN), formulada por ativistas e intelectuais negros. Produto de décadas de luta, a PNSIPN tem como principal objetivo a desconstrução do racismo institucional que os negros enfrentam ao utilizar o SUS e a inclusão das práticas de cura de matriz afro-brasileira na saúde pública, efetivando, na prática, a equidade em saúde.



27/09/2021 10:31

Céncia Hoje | Saúde de brasileiros negros em tempos de covid-19

Outro ponto importante coberto pela PNSIPN é o incentivo à produção de conhecimento científico e tecnológico em saúde da população negra. A implementação da lei, entretanto, não está livre das barreiras impostas pelo racismo, como apontam Andrea Maila Voss Kominek e Ana Christina Vanali, organizadoras do livro Roteiros temáticos da diáspora: caminhos para o enfrentamento ao racismo no Brasil, de 2018.

O advogado, filósofo e professor Silvio Almeida define em seu livro Racismo estrutural, de 2019, três níveis de expressão do racismo na sociedade: o individual, que se reconhece por atos de preconceito contra o indivíduo; o estrutural, que, fundamentado em uma sociedade racista, tem um caráter político e é exercido por meio de práticas sociais e institucionais que atuam na imposição de regras e padrões sociais que atribuem privilégios a um grupo não racializado (aos brancos, no caso brasileiro); e, por último, essas práticas se traduzem no racismo institucional, originando desvantagens para o grupo discriminado.

Racismo estrutural

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), 54% da população brasileira é negra. Mesmo com uma população majoritariamente negra, o país continua extremamente desigual e racista. São várias as modalidades de racismo que imperam no Brasil desde tempos coloniais; porém, o racismo estrutural tem se sobressaído.

Essa modalidade de racismo está presente na estrutura da sociedade brasileira, nas instituições de ensino e pesquisa, nas escolas, no esporte, no acesso às políticas de geração de renda, de saúde e de educação. E tem relegado população negra a piores condições econômicas, laborais e de saneamento básico, quando comparada à população branca.

Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), realizada anualmente, o IBGE mostrou a persistência dos piores indicadores de renda e ocupação da população negra brasileira ao longo dos anos. Por exemplo, em 2019, os negros tiveram maior taxa de desocupação e ocupação informal, menor renda mensal, maior representatividade no grupo populacional abaixo da linha de pobreza e com domicílios inadequados, entre outros.

A DESIGUALDADE SOCIAL REPERCUTE DIRETAMENTE NA SAÚDE DESSA POPULAÇÃO QUE, MUITAS VEZES, AINDA DEVE ENFRENTAR O PRECONCEITO DE PROFISSIONAIS BRANCOS NOS ATENDIMENTOS MÉDICOS

A desigualdade social repercute diretamente na saúde dessa população que, muitas vezes, ainda deve enfrentar o preconceito de profissionais brancos nos atendimentos médicos. A gravidade dessa situação é visível nas estatísticas epidemiológicas brasileiras: maior taxa de mortalidade materna e infantil, maiores complicações e morbidade associada a doenças complexas, como diabetes, hipertensão arterial, doenças cardiovasculares, maior prevalência de infecções, entre outras, segundo dados do Ministério da Saúde, de 2016 (figura 1).



Figura 1. Indicadores sociais de trabalho, renda e moradia mostram as desigualdades entre negros e brancos. Dados obtidos de <https://www.ibge.gov.br/> (2021). Fonte: Iriel A. Joerin Luque (2021)

Vieses preconceituosos

Ao longo da história da ciência ocidental, pesquisadores têm-se utilizado de diferenças fenotípicas das populações (variações na aparência das pessoas) para justificar o racismo, fundamentando suas próprias crenças enviesadas com dados biológicos de pouco ou nenhum rigor científico. Hoje sabemos que as diferenças que observamos entre populações de distintos continentes refletem apenas uma pequena parte da diversidade genética humana que existe no mundo. A variação genética que existe entre indivíduos dentro de um continente é sempre maior que a variação entre continentes. Dentro de cada continente, por sua vez, existe uma maior semelhança entre populações nativas de locais próximos.

A variação entre populações humanas decorre, em parte, da adaptação aos diferentes ambientes, desde a origem dos humanos modernos no continente africano (cerca de 200 mil anos atrás) até a sua migração mais recente para a América (cerca de 15 mil anos atrás!). Assim, as populações que hoje habitam os demais continentes possuem apenas frações da diversidade genética e fenotípica encontrada na África. Dessa forma, as variações intercontinentais (entre os continentes) são apenas o acúmulo de mutações mais recentes dos humanos modernos que surgiram em cada novo continente.

É um fato conhecido que a medicina tem visto enormes avanços nas últimas décadas. Porém, um grande problema das pesquisas biomédicas é que elas são feitas, em sua maior parte, com base em populações europeias ou eurodescendentes, com baixa representação de populações de outras regiões do mundo.

A FALTA DE ESTUDOS EM POPULAÇÕES DISTINTAS DA EUROPEIA NÃO SÓ PROVOCA O DESCONHECIMENTO DE GRANDE PARTE DA DIVERSIDADE HUMANA, MAS LEVA TAMBÉM A UMA ATENÇÃO INADEQUADA DA SAÚDE DE POPULAÇÕES NÃO EUROPEIAS OU MISCIGENADAS, COMO A BRASILEIRA



27/09/2021 10:31

Ciência Hoje | Saúde de brasileiros negros em tempos de covid-19

Assim, a falta de estudos em populações distintas da europeia não só provoca o desconhecimento de grande parte da diversidade humana, mas leva também a uma atenção inadequada da saúde de populações não europeias ou miscigenadas, como a brasileira. Isso porque os resultados obtidos em uma população não são sempre extrapoláveis a outras, o que pode derivar na administração inadequada de medicamentos, mau entendimento das causas de certa doença ou análise incorreta de sintomas e de fatores de risco (características ou condições que aumentam a probabilidade de se ter uma doença).

De acordo com estudo de Célia Landmann Szwarcwald e colaboradores, publicado na Revista Brasileira de Epidemiologia em 2019, no Brasil, os valores de referência empregados para análises laboratoriais são obtidos com base em estudos realizados em populações de outros países ou fornecidos pelos fabricantes dos testes, não sendo estes verificados pela sua aplicabilidade na população brasileira (figura 2).

Figura 2. Porcentagem de estudos de associação de variantes genéticas com doenças segundo a ancestralidade da população estudada. Dados obtidos de Sirugo e colaboradores (2019). Fonte: Iriel A. Joerin Luque (2021)

O exemplo do oxímetro

Um dos exemplos mais recentes do viés eurocêntrico na área de saúde é o funcionamento do oxímetro de pulso. O aparelho é utilizado de forma rotineira para a avaliação do nível de oxigênio no sangue (também chamado de saturação de oxigênio ou SaO_2) durante o diagnóstico da covid-19, uma vez que a infecção e propagação do vírus SARS produz uma gradativa diminuição do mesmo.

O funcionamento do oxímetro de pulso baseia-se na diferença de absorção de luz vermelha e infravermelha pela hemoglobina (a proteína que transporta o oxigênio no sangue) em suas formas oxigenada e desoxigenada. A hemoglobina oxigenada absorve mais luz infravermelha, permitindo a luz vermelha atravessar o tecido onde é feita a medição. A hemoglobina desoxigenada, entretanto, absorve mais luz vermelha, deixando passar a infravermelha.

O aparelho utiliza dois LEDs emissores de luz, um para cada tipo, os quais são emitidos contra um local do corpo que deve ter uma certa translucidez e um bom fluxo sanguíneo. Os locais mais utilizados para a medição são, portanto, os dedos ou o lóbulo da orelha. Enquanto a luz atravessa o tecido, os sinais são detectados por um fotorreceptor e uma relação luz vermelha/infravermelha é calculada e comparada com uma tabela de valores para obter a medida de saturação de oxigênio. A detecção da hipóxia (situação em que esses níveis se situam abaixo de 92%) pode indicar a necessidade de um tratamento, mesmo quando o paciente ainda não sente desconforto respiratório (também chamado de dispneia ou falta de ar) e, por isso, deve ser confiável.

Porém, estudos como o de [A. C. Ralston](#) demonstram que alguns fatores – o uso de esmalte de unhas, a luz do ambiente, a presença de sinais eletromagnéticos e a pigmentação da pele, entre outros – podem afetar a precisão da leitura do aparelho. Isso ocorre porque a medida de SaO_2 que é calculada



27/09/2021 10:31

Cência Hoje | Saúde de brasileiros negros em tempos de covid-19

pelo dispositivo deve considerar fatores de correção que são determinados experimentalmente através da comparação entre a leitura do oxímetro e leituras diretas do oxigênio gasoso no sangue de pacientes.

Esses fatores de correção servem, assim, para a calibração do aparelho. No entanto, como apontado por [P. E. Bickler e colaboradores](#), a maioria dos oxímetros do mercado têm sido calibrados com base nas medidas realizadas em pacientes de pele clara, sem incluir as variações que podem decorrer de uma pigmentação mais escura.

Com o crescente uso dos oxímetros na atenção de pacientes de covid-19, um grupo de pesquisadores liderados por [M. W. Sioding](#) avaliou a precisão da medida de SaO₂ em pacientes negros e brancos dos Estados Unidos por meio da comparação entre a medida feita pelo aparelho e a leitura *in vivo* (a medição do nível do oxigênio diretamente no sangue da pessoa). A pesquisa mostrou que a taxa de erro é três vezes maior em pessoas negras, levando a uma superestimação do nível de oxigênio (figura 3). Isso é preocupante, pois pode levar à demora na busca de atendimento médico (em pacientes que utilizam o oxímetro em casa) ou em diagnósticos errados em centros de saúde.

Figura 3. A medição com o oxímetro de pulso apresenta uma taxa de erro maior em negros.

Vulnerabilidade social

NO BRASIL, A COVID-19 É CINCO VEZES MAIS LETAL EM NEGROS DO QUE EM BRANCOS

No Brasil, a pandemia da covid-19 escancarou o racismo estrutural porque evidenciou a vulnerabilidade social da população pobre e negra. Um exemplo é a situação precária de habitação dessa faixa da população que tem facilitado a disseminação do coronavírus. Outro fator importante são as doenças que atingem as pessoas negras – diabetes, hipertensão arterial, anemia falciforme e outras comorbidades – que, aliadas à demora na vacinação e à dificuldade de essa população acessar o sistema público de saúde, muito tem contribuído para elevar o número de mortos por covid-19 nessa população. Refletindo esse panorama, estudos, como o de [Bianca Santana](#), de 2021, já apontaram que, no Brasil, a covid-19 é cinco vezes mais letal em negros do que em brancos.



27/09/2021 10:31

Ciência Hoje | Saúde de brasileiros negros em tempos de covid-19

Para que a equidade em saúde seja atingida no Brasil, a comunidade científica deve enfrentar esse desafio. Quando procuramos por estudos científicos realizados no país sobre comorbidades comuns da população negra, como diabetes e hipertensão, encontramos que a maioria não realiza um recorte por raça/cor dos participantes do estudo e, quando o mesmo existe, observa-se que essa população é a menos representada.

Portanto, um fator importante é produzir conhecimento científico em saúde da população negra e incluir o recorte por raça/cor no planejamento dos estudos científicos. Porém, isso depende em grande parte, também, da existência de editais específicos que forneçam os recursos necessários para que pesquisas voltadas à saúde dessa população – e de outros grupos não eurodescendentes, como povos indígenas – sejam desenvolvidas.

Voltando a nossa pergunta inicial: qual a responsabilidade da ciência na falta de equidade em saúde no Brasil? Poderão os cientistas brasileiros assumir o seu papel na eliminação do viés eurocêntrico que existe nas pesquisas biomédicas e na desconstrução do racismo estrutural que afeta a saúde da sua população?

Buscando fornecer dados para melhorar os cuidados com a saúde da população negra, que, de acordo com o IBGE, corresponde a 33% da população paranaense, a Universidade Federal do Paraná lançou o projeto de pesquisa 'Genética e saúde da população negra'. O objetivo é estudar as bases genéticas das doenças que mais afetam a população negra, como diabetes e hipertensão arterial, assim como verificar se os fatores de risco comumente empregados na atenção à saúde (e que foram obtidos a partir de estudos em populações europeias ou eurodescendentes) se aplicam à população afro-brasileira. O projeto já teve início nas populações quilombolas da Restinga e do Feixo, no município de Lapa, Paraná e será desenvolvido até o final de 2022.

Iriel A. Joerin Luque

Marcia Holsbach Beltrame

Programa de Pós-graduação em Genética

Laboratório de Genética Molecular Humana

Universidade Federal do Paraná

Claudemira Vieira Gusmão Lopes

Câmara de Educação do Campo

Universidade Federal do Paraná

 [Leia mais](#)

Matéria publicada em 22.09.2021

COMENTÁRIOS



