medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Genomic Exploration of Essential Hypertension in African-Brazilian Quilombo Populations: A Comprehensive Approach with Pedigree Analysis and Family-Based Association Studies

Vinícius Magalhães Borges, Ph.D.[1,2], Andrea R.V.R. Horimoto, Ph.D.[3], Ellen Marie Wijsman, Ph.D.[3]; Lilian Kimura, Ph.D.[1]; Kelly Nunes, Ph.D.[1], Alejandro Q. Nato, Jr., Ph.D.[2], and Regina Célia Mingroni-Netto, Ph.D.[1]

[1]Centro de Estudos sobre o Genoma Humano e Células Tronco, Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo 05508-090, Brazil

[2]Department of Biomedical Sciences, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV 25755, USA

[3]Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA, 98105 USA

Corresponding authors:

Vinícius Magalhães Borges
Address: 1700 Third Ave (BBSC 239), Department of Biomedical Sciences, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV 25755, USA.

Regina Célia Mingroni-Netto
Address: Centro de Estudos sobre o Genoma Humano e Células Tronco, Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo 05508-090, Brazil.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Genomic Exploration of Essential Hypertension in Brazilian Quilombo Populations: A Comprehensive Approach with Pedigree Analysis and Family-Based Association Studies

ABSTRACT

Background: Essential Hypertension (EH) is a global health issue. Despite extensive research, much of EH heritability remains unexplained. We investigated the genetic basis of EH in African-derived individuals from partially isolated quilombo populations in Vale do Ribeira (SP-Brazil). **Methods:** Samples from 431 individuals (167 affected, 261 unaffected, 3 unknown) were genotyped using a 650k SNP array. Estimated global ancestry proportions were 47% African, 36% European, and 16% Native American. We constructed six pedigrees using additional data from 673 individuals and created three non-overlapping SNP subpanels. We phased haplotypes and performed local ancestry analysis to account for admixture. Genome-wide linkage analysis (GWLA) and fine-mapping via family-based association studies (FBAS) were conducted, prioritizing EH-associated genes through systematic approach involving databases like PubMed, ClinVar, and GWAS Catalog. **Results:** Linkage analysis identified 22 regions of interest (ROIs) with LOD scores ranging 1.45-3.03, encompassing 2,363 genes. Fine-mapping (FBAS) identified 60 EH-related candidate genes and 117 suggestive/significant variants. Among these, 14 genes, including *PHGDH*, *S100A10*, *MFN2*, and *RYR2*, were strongly related to hypertension harboring 29 suggestive/significant SNPs. **Conclusions:** Through a complementary approach — combining admixture-adjusted GWLA based on Markov chain Monte Carlo methods, FBAS on known and imputed data, and gene prioritizing — new loci, variants, and candidate genes were identified. These findings provide targets for future research, replication in other populations, facilitate personalized treatments, and improve public health towards African-derived underrepresented populations. Limitations include restricted SNP coverage, self-reported pedigree data, and lack of available EH genomic studies on admixed populations for independent validation, despite the performed genetic correlation analyses using summary statistics.

Keywords: High Blood Pressure; Genome Scan; Gene Mapping; Admixed Populations; Complex Trait

NOVELTY AND RELEVANCE

What Is New?
This study applies a multi-level computational approach integrating admixture-adjusted genome-wide linkage analysis (GWLA), family-based association studies (FBAS), and fine-mapping strategies to investigate the genetic basis of essential hypertension (EH) in Brazilian Quilombo populations, a historically underrepresented group in genomic research.

What Is Relevant?
By focusing on admixed populations with high African ancestry, our findings address gaps in hypertension genetics by identifying 22 regions of interest (ROIs), 60 candidate genes, and 117 suggestive/significant variants, highlighting population-specific genetic factors that may contribute to EH risk. The study also emphasizes the need for ancestry-aware genomic analyses to improve the precision of genetic risk assessment in underrepresented populations.

What Question Should Be Addressed Next?
Future studies should focus on replicating these findings in independent admixed cohorts, conducting functional validation of prioritized genes, and integrating polygenic risk scores (PRS) adjusted for ancestry to enhance clinical applications for hypertension prevention and treatment in diverse populations.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

NON-STANDARD ABBREVIATIONS AND ACRONYMS

ABDR	Abobral (quilombo population)
AGT	angiotensinogen
ALPK2	alpha kinase 2
AN	André Lopes (quilombo population)
ANG II	angiotensin II
ANP	atrial natriuretic peptide
ARMC	armadillo repeat containing 5
BMI	body mass index
BP	arterial blood pressure
CACNA1D	calcium voltage-gated channel subunit alpha1 d
CADD	combined annotation dependent depletion
CEU	Northern Europeans from Utah
CLM	Colombian in Medellin, Colombia
CVDs	cardiovascular diseases
CYP11B2	cytochrome p450 family 11 subfamily b member 2
DBP	diastolic blood pressure
EDARADD	edar associated via death domain
EH	essential hypertension
eNOS	endothelial nitric oxide synthase
GA	Galvão (quilombo population)
GLMM	generalized linear mixed model
GRK4	g protein-coupled receptor kinase 4
GRM	genetic relationship matrix
HGDP	human genome diversity project
IV	Ivaporunduva (quilombo population)
IVs	inheritance vectors
KCNK3	potassium two pore domain channel subfamily k member 3
KCNT1	potassium sodium-activated channel subfamily t member 1
KKS	kallikrein-kinin system
LD	linkage disequilibrium
LINC01256	long intergenic non-protein coding rna 1256
LOD	logarithm of the odds
LPP	lim domain containing preferred translocation partner in lipoma
MCMC	markov chain monte carlo
MFN2	mitofusin 2
MLX	Mexican Ancestry in Los Angeles, California
MOBP	myelin associated oligodendrocyte basic protein
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase
NH	Nhunguara (quilombo population)
NO	nitric oxide
NOS3	nitric oxide synthase 3
NPR3	natriuretic peptide receptor 3
P2RX1	purinergic receptor p2x 1
P2RY1	purinergic receptor p2y1
PC	Pedro Cubas (quilombo population)
PCA	principal components analysis
PCs	principal components
PEL	Peruvian in Lima, Peru
PHGDH	phosphoglycerate dehydrogenase
PHOX2	paired like homeobox 2
PUR	Puerto Rican in Puerto Rico
RAAS	renin-angiotensin-aldosterone system
ROIs	regions of interest

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RPA1	replication protein a1
RYR2	ryanodine receptor 2
S100A10	s100 calcium binding protein a10
SBP	systolic blood pressure
SCG2	secretogranin ii
SCNN1B	sodium channel epithelial 1 subunit beta
SCNN1G	sodium channel epithelial 1 subunit gamma
SERTAD2	serta domain containing 2
SNS	sympathetic nervous system
SP	São Pedro (quilombo population)
TENM4	teneurin transmembrane protein 4
TF	transcription factor
TU	Sapatu (quilombo population)
YRI	Yoruba in Ibadan, Nigeria
ZZEF1	zinc finger zz-type and ef-hand domain containing 1

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Essential Hypertension (EH), OMIM: 145500, is a pervasive and sustained raise in arterial blood pressure (BP) and a major cause of premature death worldwide, responsible for approximately 9.4 million deaths annually.[1] Classified as the primary preventable risk factor for cardiovascular diseases (CVDs),[2] EH is defined as systolic BP (SBP) ≥ 140 mmHg and/or diastolic BP (DBP) ≥ 90 mmHg.[3–5]

EH, which affects 1.3 billion people worldwide annually,[6] demands comprehensive investigation. Global prevalence is 33%,[6] 23.9% in Brazil.[7] EH is a multifactorial chronic condition, intricately weaving together environmental factors, social determinants, and greatly genetic/epigenetic influences, with 30-60% of estimated heritability.[8–10] Its prevalence varies across different regions, affecting 35% of the population in the Americas, 28% in the Western Pacific, 37% in Europe, 32% in South-East Asia, 38% in the Eastern Mediterranean, and 36% in Africa.[6] The risk for BP traits varies among ethnic groups[11] and genetic ancestry significantly influences hypertension risk,[12] particularly in African-derived populations.[1,13–16] In the U.S., data shows that the prevalence of EH among non-Hispanic Black individuals is approximately 48.8%, notably higher than the prevalence among non-Hispanic White individuals at 37.6% and Hispanic individuals at 27.9%.[17] Mortality rates due to EH and related diseases are also disproportionately higher, with African American individuals experiencing 4 to 5 times greater mortality than White adults.[18]

BP regulation involves complex interactions between the cardiovascular, renal, neural, and endocrine systems.[19] The renin-angiotensin-aldosterone system (RAAS) plays a central role, with angiotensin II and aldosterone driving vasoconstriction and sodium retention.[10,20] Baroreceptors and the sympathetic nervous system (SNS) provide rapid responses to BP changes,[21,22] while endothelial cells release nitric oxide (NO) for vasodilation,[23] The atrial natriuretic peptide (ANP) and the kallikrein-kinin system (KKS) counteract vasoconstriction and promote natriuresis.[24,25] Sodium concentration also affects vascular tone through calcium exchange.[26]

The complex regulation is particularly relevant in populations as the African-derived.[27] Genetic ancestry plays a crucial role in EH risk, with studies indicating that African-derived individuals generally develop higher BP starting in childhood. They excrete sodium more slowly and less completely than those of European descent, leading to volume-loading, which suppresses the RAAS and contributes to early-onset hypertension.[28,29] Consequently, African-derived individuals often present a biochemical profile characterized by low or high plasma aldosterone, suppressed plasma renin activity or direct renin concentration, and reduced levels of angiotensin I and II.[30–32] This results in a higher lifetime incidence of hypertension in African-derived populations compared with other populations.

Several genes have been found to be associated with an increased risk of EH specifically in African-derived populations, such as *ARMC5* (involved in RAAS[15]), *NOS3-GRK4* (involved in NO production and regulation[33]), *SCNN1B* and *SCNN1G* (involved in sodium channel[34]), *GRK4* (involved in sodium and water retention[13]), *SCG2* (confers regulation by PHOX2 transcription factors[35]), *AGT* (involved in angiotensin[36]) and *CYP11B2* (involved in alterations in aldosterone synthase production[13]).

Among social-environmental factors, EH risk is also influenced by adverse determinants of health, including overweight, smoking, physical inactivity,[9,37] lower socioeconomic and educational status, concentrated poverty, and limited access to affordable, high-quality fresh food. Dietary patterns, alcohol consumption,[38] particularly high-sodium[39–41] and low-potassium intake, further elevate the risk.[42] African-derived populations in lower social strata, such as those in the United States and South Africa, tend to experience higher rates of hypertension than expected based solely on anthropometric and socioeconomic factors.[43]

Yet, the genetic etiology of hypertension — encompassing genes, variants, susceptibility loci, and population disparities — remains elusive.[44,45] Despite advancements from the common disease-common variant and common disease-rare variant hypotheses, methodologies such as Genome-Wide Linkage Analysis (GWLA) and Genome-Wide Association Studies (GWAS) face limitations.[46] This gap leaves a segment of EH heritability unexplained by known genetic factors. Moreover, the existing underrepresentation (data as of December 2024) of African (0.19%), African American or Afro-Caribbean (0.45%), Hispanic or Latin American (0.34%), Other/Mixed (0.56%) populations[47] in worldwide genomic

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

investigations imposes constraints on the generalizability of results to admixed populations,[48,49] such as the Brazilian one (68.1% European, 19.6% African, and 11.6% Native American).[50]

This study focuses on the tri-hybrid admixed populations known as "quilombo remnants" in Vale do Ribeira region, São Paulo, Brazil. Quilombo remnants are communities established by runaway or abandoned African enslaved individuals, often exhibiting intricate mixtures with European and Native American ancestry. These populations represent a unique model for the study of diseases. They are marked by a high prevalence of EH,[51] well-defined clinical characterization, semi-isolation, background relatedness, high gene flow between populations, and founder effects.[52] They also exhibit relatively homogeneous environmental influences, including lifestyle, dietary habits, and natural habitat, thereby minimizing confounding factors found in larger urban populations. Studying EH in quilombo remnants helps to reduce biases associated with population heterogeneity. This approach improves the signal-to-noise ratio, thereby enhancing statistical power, while providing better representation of admixed populations in genomic studies. It also allows us to uncover chromosomal regions, genes, and variants that may contribute to EH.

In this study, we employed a multi-level computational approach that combined both pedigree-based and population-based methodologies on family analysis to account for the unique admixture of African, European, and Native American ancestries in the quilombo population, along with a two-step fine-mapping strategy based on family-based association studies (FBAS) and investigation of EH-related genes. The family analysis using GWLA has been successfully applied in previous studies[53] to address challenges such as population stratification and the complex genetic architecture of traits, enhancing the power to identify loci associated with rare variants that may go undetected in population-based studies. The pedigree- and population-based imputation methods we employed have also been rigorously developed and tested.[54,55] FBAS enables detailed investigation of EH-related genes while controlling for population structure and relatedness, reducing the problem of multiple testing — an essential factor in admixed populations like the quilombo.[56,57] Overall, by applying this strategy, we were able to fine-map candidate regions and variants associated with hypertension in a way that leverages the strengths of both family- and population-based genetic studies, providing insights into EH heritability in underrepresented populations.

METHODS

Anonymized summary statistics data have been made publicly available at the GWAS Catalog (study GCST90454187) and can be accessed at https://www.ebi.ac.uk/gwas/studies/GCST90454187.

SAMPLES AND SNP GENOTYPING

We conducted 51 trips to Vale do Ribeira (São Paulo, Brazil) from 2000-2020 to obtain samples (peripheral blood for DNA extraction), clinical data (average blood pressure, height, weight, waist circumference and hip circumference), and collect information (sex, age, family relationships, medical history, demographic information, and daily physical activity levels) from 431 consenting individuals aged 17 or older (Figure 1, box A). This study was approved by institutional review committees (USP/Institute of Biomedical Sciences 111/2001 and USP/Institute of Biosciences 012/2004 and 034/2005) and blood samples were drawn after subjects provided informed consent. To reduce bias, we excluded individuals who reported diabetes, severe kidney disease, and pregnancy. Additionally, no samples needed to be removed due to very low blood pressure (hypotensive readings), as no such cases were observed in the dataset.

Genomic DNA was extracted and quantified from each of the 431 blood samples and prepared for SNP genotyping through Axiom Genome-Wide Human Origins 1 Array SNPs (Figure 1, box B) according to Affymetrix requirements (details in Data S1). Raw data was processed, annotated, and subjected to quality control according to Affymetrix Human v.5a threshold[58] using the commercial software Axiom Analysis Suite v.3.1.

From the combined collected data, we used the commercial software GenoPro v.3.0.1.4 – tool for creating and managing family trees and genealogical data – to construct six extended pedigrees from 8 different populations (Abobral [ABDR], André Lopes [AN], Galvão [GA], Ivaporunduva [IV], Nhunguara

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[NH], Pedro Cubas [PC], São Pedro [SP], and Sapatu [TU]), with detailed locations as previously described[59] (Figure 2). The six pedigrees were named: **ABDR** (Abobral population); **ANNH** (André Lopes and Nhunguara populations); **GASP** (Galvão and São Pedro populations); **IV** (Ivaporunduva population); **PC** (Pedro Cubas population); and **TU** (Sapatu population). These pedigrees encompass 1,104 individuals (Table 1): 431 genotyped (167 affected, 261 non-affected and 3 unknown phenotype) and 673 non-genotyped. The characteristics of the genotyped cohort are detailed in Table 2. Pedigree structures underwent validation through calculation of multi-step pairwise kinship coefficients (Φ) using several algorithms: KING-Robust v.2.2.8[60], which efficiently identifies relatedness in large datasets while accounting for population structure; MORGAN (Monte Carlo Genetic Analysis) v.3.4[61,62], a robust suite designed for performing genetic linkage analysis in pedigrees; and PBAP (Pedigree-Based Analysis Pipeline) v.1/v.2[63], a pipeline for file processing and quality control of pedigree data with dense genetic markers.

We filtered and trimmed the dataset (details in Data S1) using KING-Robust v.2.2.8[60] and PLINK v.1.9/v.2.0[64], a widely used tool for GWAS and population data analysis that efficiently handles large-scale data and performs quality control and kinship analysis. We excluded samples with a genotyping rate ≤ 95%, as well as SNPs with a genotyping rate ≤ 95%. Monomorphic SNPs and SNPs resulting in heterozygous haploid calls across all remaining individuals were removed. SNPs not adhering to Hardy-Weinberg equilibrium (p-value < 1×10^{-3}) were also removed. SNP identification followed the dbSNP standard format (rsID), and their genetic locations (cM) were obtained through the Rutgers Combined Linkage-Physical Map v.3[65]. EH was considered a binary outcome, categorizing individuals as hypertensive (SBP ≥ 140 and/or DBP ≥ 90 mmHg) or normotensive (SBP < 140 and DBP < 90 mmHg). Individuals diagnosed and/or under medication for EH were classified as hypertensive.

We analyzed age differences between affected and unaffected individuals across pedigrees using Welch two sample t-tests, one-way and two-way ANOVA, and multiple regression analysis. These analyses were conducted using in-house R scripts, leveraging the R stats package v.3.6.2 (R Core Team, 2024) for statistical computations.

LOCAL ANCESTRY ESTIMATION

We estimated local ancestry fractions (Figure 1, box F), which allowed us to determine individual- and pedigree-specific ancestries. The reference dataset comprised 189 samples from the 1000 Genomes Project[66] and Stanford HGDP SNP Genotyping[67] data: 63 European (CEU - Northern Europeans from Utah), 63 African (YRI - Yoruba in Ibadan, Nigeria), and 63 Native American (Colombia, Maya, and Pima populations) samples. Overlapping markers (145,467 SNPs) present in both reference (189 samples) and inference (431 samples) datasets were extracted and both datasets were merged and pruned for missingness (≤ 95% genotyping rate) using PLINK. Haplotypes were inferred using SHAPEIT2 v2.17,[68] a tool for phasing genotype data and efficiently reconstructs haplotypes from large-scale datasets. RFMix v.1.5.4[69], a tool for local ancestry inference, specializes in assigning ancestry to specific genomic segments in admixed populations was used. We used in-house scripts to estimate global ancestry fractions for each sample and pedigree ancestries by averaging the local ancestry calls across the entire genome (details in Data S2).

PEDIGREE ANALYSIS

In our multipoint pedigree linkage analyses, we employed MORGAN suite, leveraging its versatility and robust capabilities rooted in the Markov Chain Monte Carlo (MCMC) approach, allowing for simultaneously handling numerous markers and individuals within pedigrees through a sampling methodology. [61]

We implemented a redundant and semi-independent design to yield results that are both independent and comparable. From the complete inference marker set, we selected three distinct non-overlapping subsets (or subpanels) of markers (Figure 1, box E) employing specific selection criteria and parameters through PBAP (details in Data S3). The first subpanel, labeled the gold standard, included top-quality markers meeting specific criteria. The other two subpanels, while informative, might have slightly lower

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

quality and were applied for validation of the finding. All three subpanels were crucial for discovery analyses. Parametric linkage analysis was performed using an autosomal-dominant model with a risk allele frequency of 0.01, an incomplete penetrance of 0.70 (for genotypes with 1 or 2 copies of the risk allele), and a phenocopy rate of 0.05. The penetrance was estimated by assessing affected and unaffected individuals within pedigrees. All subsequent steps were performed concurrently for each pedigree (details in Data S3.1).

We estimated allelic frequencies for each SNP marker (Figure 1, box G) using the specialized script ADMIXFRQ v.1.[53] This involved the generation of unique pedigree-specific files organized by ancestry based on local ancestry calls and subpanel information.

To compute LOD scores, we employed the approach in MORGAN outlined as:

$$P_{\gamma}(Y_T|Y_M) = \sum_{S_M} P_{\gamma}(Y_T|_M)\ P(S_M|Y_M)$$

where Y_M denotes the genetic marker loci, Y_T the trait characteristics and SM denotes the meiosis indicators for all markers.[62] The LOD score calculation was a two-step process. In the initial step (Figure 1, box H), we sampled inheritance vectors (IVs) for the gold standard subpanel using alternate SNP markers (2,500-3,000 SNPs) in a genome-wide “scan" analysis. Preliminary candidate regions were identified as those with a peak LOD score ≥ 1.

In the second step, we conducted a "dense" analysis (Figure 1, box J) restricted to the entire chromosomes containing each preliminary candidate region using all three subpanels and sampling of IVs using all available SNP markers. LOD scores were calculated again for this second step, defining Regions of Interest (ROIs) as those with a maximum LOD score ≥ 1.50 in at least one subpanel, with mandatory positive results for subpanel 1. ROI boundaries were determined based on LOD score ≥ 1 marker position (details in Data S3.3). Both steps (“scan” and “dense” analysis) were conducted separately for each pedigree.

To ensure the convergence of the sampling process, we performed diagnostic analysis of the MCMC runs (Figure 1, box I), evaluating run length, autocorrelation of LOD scores, and run stability using three different graphical tools (Figure S1). The default setup for MCMC runs was 100,000 Monte Carlo (MC) iterations, 40,000 burn-in iterations, 25 saved realizations, 1,000 identity-by-descent graphs and output scores saved at every 25 scored MC iterations (details in Data S3.4).

Results of this analysis were used to determine the appropriate running conditions. Once the correct setup for each pedigree was established, we repeated the genome-wide scan and dense analysis, which included sampling of IVs and calculation of LOD scores, to determine the final ROIs (Figure 1, box K).

FINE-MAPPING STRATEGIES

Following pedigree analysis adjusted for admixture, we identified and fine-mapped 22 ROIs (Figure 1, box L-S). We further explored these results by analyzing the imputed SNP dataset using family-based association studies, individually evaluating each of the 22 ROIs while combining all pedigrees to increase statistical power. Additionally, we identified EH-related genes through *in-silico* investigation.

Family-Based Association Studies

Initially, we addressed population structure through principal components analysis (PCA) (Figure 1, box L), first employing MORGAN Checkped[61,62,70] and PBAP Relationship Check[63] algorithms for sample relatedness verification. Pairwise kinship coefficients (Φ) were calculated using KING-robust.[60] The subsequent steps involved the iterative use of PC-AiR and PC-Relate functions conducted using the GENESIS (Genetic Estimation and Inference In Structured Samples) v3.19 R package.[57] GENESIS provides tools for GWAS and offers advanced methods for handling relatedness and population structure in complex datasets. In the initial iteration, KING-robust estimates informed both kinship and ancestry divergence calculations, with resulting principal components (PCs) used to derive ancestry-adjusted kinship estimates, the 1st genetic relationship matrix (GRM), via PC-Relate. To further refine PCs for ancestry, a secondary PC-AiR run utilized the 1st GRM for kinship and KING-robust estimates for ancestry divergence, yielding new PCs. These new PCs informed a secondary PC-Relate run,

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

culminating in a 2nd GRM. Subsequently, we evaluated variation using the top 10 PCs, selected in accordance with the Kaiser criterion on the Kaiser-Guttman rule, which suggests retaining PCs with eigenvalues greater than 1, as these components explain more variance than a single original variable.[71]

We employed a comprehensive two-way independent imputation strategy (Figure 1, box N) for variants within each ROI identified through the family analysis (details in Data S4.2). We utilized linkage disequilibrium (LD) information to perform a population-based approach. This involved chromosome-wise dataset separation, phasing through SHAPEIT2+duoHMM method[72], imputation using the Minimac4 v.4.0[73] software and filtering ($r^2 \geq 0.3$) through BCFtools v1.15[74]. Minimac4 efficiently imputes genotypes from large datasets based on reference panels, while BCFtools provides utilities for processing and filtering variant call format (VCF) and binary call format (BCF) files. Concurrently a pedigree-based strategy incorporating IVs information was applied. This involved chromosome-wise dataset separation and imputation conducted using GIGI2 v.1[75] a software tool designed for genetic imputation based on pedigree data.

To account for the complex correlation structure of the data, we included the 2nd GRM as a random effect to fit the mixed models through fitNullModel function, conducted using the GENESIS R package[57]. Comprehensive testing included 10 ancestry principal components and all available covariates as fixed effects (details in Data S4.1). The final statistical model incorporated the top 8 PCs, along with sex, BMI (Body Mass Index), and age as significant variables. The next step was to fit the generalized linear mixed model (GLMM) (Figure 1, box M). Specifically for single-variant tests we employed a logistic mixed model expressed as:

$$\text{logit}(\pi) = X\alpha + G_j\beta_j + g$$

where $\pi = P(y = 1 \mid X, G_j, g)$ represents the $Nx1$ column vector of probabilities of being affected for the N individuals conditional to covariates, allelic dosages; and random effects; X is the vector of covariates; and α is the vector of fixed covariate effects. We assume that $g \sim N(0, \sigma_\alpha^2 \Phi)$ is a vector $g = (g_{1,\ldots,} g_N)$ of random effects for the N subjects, where σ_α^2 is the additive genetic variance and Φ is the GRM; G_j is a vector with the allelic dosages (0, 1, or 2 copies of the reference allele) or expected dose (in the case of imputed genotypes) at the locus j; and β_j is its corresponding effect size. The null hypothesis of $\beta_j = 0$ was assessed using a multivariate score test[76].

Finally, for each ROI, two independent FBAS tests were conducted (Figure 1, box O) using imputed variant datasets that are either (1) pedigree-based or (2) population-based. The dataset comprised all 431 samples from the combined 6 pedigrees. The single-variant association tests were conducted using the GENESIS R package[57], implementing the adjusted GLMM to perform Score tests.

We performed multiple testing correction (Figure 1, box P) using the effective number of independent markers (M_e) estimated using the Genetic Type I Error Calculator software v.0.1[77], a software tool designed to assess the likelihood of false positives in GWAS. We also evaluated adequacy of the analysis modeling through evaluation of the genomic inflation factor (λ) by dividing the median of the chi-square statistics by the median of the chi-square distribution with 1 degree of freedom[78]. Analysis was carried out for each imputed variant dataset.

Suggestive and significant association variants underwent curation, assessment, and compilation (Figure 1, box Q) into a comprehensive database using a custom R script. Data were extracted in March 2023 from NIH dbSNP[79], CADD[80], NCBI PubMed[81], Ensembl Variant Effect Predictor[82], ClinVar[83], Mutation Taster[84], SIFT[85], PolyPhen[86], and VarSome[87]. Annotation included chromosome and physical positions (GRCh37/hg19), dbSNP rsID, associated genes, genomic alterations, variant consequences, exonic functions, and pathogenicity classifications. This procedure was executed independently for each ROI and for both imputed variant datasets.

Furthermore, we analyzed linkage disequilibrium (LD) (Figure 1, box R) patterns for variants within a 500,000 base pair window surrounding each statistically suggestive or significant associated variant. Utilizing the Ensembl REST API in conjunction with the ensemblQueryR tool v2.0[88], an R package designed for efficient querying of the Ensembl database, we accessed and retrieved genomic data extracted during May 2024. We established thresholds of $r^2 \geq 0.7$ and $D' \geq 0.9$. This analysis incorporated

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

data from the 1000 Genomes Project[66] for multiple populations, including European (CEU: Utah Residents with Northern and Western European Ancestry), African (YRI: Yoruba in Ibadan, Nigeria), and American populations: CLM (Colombian in Medellin, Colombia), MXL (Mexican Ancestry in Los Angeles, California), PEL (Peruvian in Lima, Peru), and PUR (Puerto Rican in Puerto Rico). Additionally, we utilized the Ensembl REST API VEP to annotate and select LD patterns based on variant consequences.

Investigation of EH-related Genes

To elucidate the hypertension-related implications of genes identified within each ROI (Figure 1, box S), we first identified all genes mapped within ROIs using R package biomaRt v.2.6.0.1[89]. Then we obtained and annotated genes associated with "essential hypertension" or "high blood pressure" according to NCBI PubMed[81], MedGen[90], MalaCards[91], ClinVar[83], Ensembl BioMart[92], and GWAS Catalog[93]. The data were extracted in March 2023. Finally, we matched both lists, systematically annotating based on physical position (base pair), cytogenetic band, summary, molecular function, related phenotype, gene ontology, genetic location (cM; GRCh37/hg19), expression patterns, and publication data. Additionally, to prioritize these genes (details in Data S5), we utilized VarElect v.5.21[94], a tool that assesses the potential pathogenicity of genetic variants for prioritization.

Genetic Correlation Analysis

To validate our FBAS findings, we performed a genetic correlation analysis using summary statistics from our study and the publicly available GWAS Catalog dataset GCST90436066[95], selected due to its extensive SNP coverage. The analysis was conducted using the GenomicSEM R package[96], with LD scores calculated using LDSC scripts[97] from the HapMap3 b36 YRI (Yoruba in Ibadan, Nigeria) population as the reference[98]. Given the high African ancestry in our dataset, the YRI population was selected as the reference for LD scores. Attempts to use LD scores from other HapMap3 populations resulted in negative heritability estimates, preventing accurate genetic correlation computation.

RESULTS

Significant differences in mean ages were observed between affected and unaffected individuals (Figure 3) across all pedigrees (ABDR: 35.9 vs. 49.7 years, $p = 2.112\times10^{-3}$; ANNH: 36.4 vs. 57 years, $p = 1.291\times10^{-7}$; GASP: 37 vs. 54.3 years, $p = 4.771\times10^{-7}$; PC: 38.4 vs. 60.6 years, $p = 1.026\times10^{-4}$; IV: 33.2 vs. 64.3 years, $p = 4.963\times10^{-7}$; TU: 38.2 vs. 53.8 years, $p = 8.866\times10^{-4}$), as seen in Table S1. Overall, the combined data showed a mean age of 36.73 ± 15.4 years for unaffected individuals versus 55.11 ± 17.3 years for affected individuals ($p = 6.855\times10^{-25}$). One-way ANOVA was performed with 5 degrees of freedom for pedigrees and 161 for residuals, the sum of squares for pedigrees was 2783 and for residuals was 46617, resulting in mean squares of 556.6 and 289.5, respectively. Comparing the average age of affected individuals across pedigrees for both sexes, this analysis yielded an F value of 1.922 with a *p*-value of 0.093, indicating non-significant differences in average age across pedigrees. Two-way ANOVA highlighted a significant effect of affected status on average age ($p < 0.001$). However, neither sex nor the interaction between sex and affected status showed significant effects on average age, with *p*-values of 0.601 and 0.671 respectively. This suggests that while affected status significantly influences average age and sex, the interaction between sex and affected status does not have a significant impact. The multiple regression model showed that affected status significantly influenced average age ($p < 0.001$). The coefficients for sex ($p = 0.197$), pedigree ($p = 0.551$), BMI ($p = 0.328$), as well as for SBP ($p = 0.278$) and DBP ($p = 0.440$) are not statistically significant. The overall model's performance is high (strong explanatory power) as indicated by the R-squared value of 0.761, suggesting that about 76.11% of the variability in average age can be explained by the predictors in the model. Additionally, the F-statistic is significant ($p = 9.109\times10^{-5}$), indicating that the model is significant in predicting average age.

We calculated ancestry proportions for individual admixture segments (local ancestry) using data from 145,467 SNPs considering all the 431 samples (Figure 4A) and individually per pedigree (Figure 4B). The three top principal components (PCs), PC1, PC2 and PC3, explain 12.42%, 6.17%, and 1.68% of the variance, respectively (Figure 5A), allowing to visualize the genetic distance between the inference and

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the reference datasets (Figure 5B). All six quilombo remnants pedigrees studied have a high degree of admixture. Among all 431 individuals the estimates are 47.4%, 36.3%, and 16.1%, respectively, for African, European, and Native American ancestries (Table 3). The pedigrees from ABDR and PC had the highest (51.5%) estimates of African ancestral contribution. In comparison, TU had the highest (50.8%) estimate of European ancestral contribution, and ABDR had the highest estimate (16.9%) of Native American ancestral contribution. For comparison, the ancestry proportions for the general Brazilian population were estimated as 19.6%, 68.1%, and 11.6% for African, European, and Native American ancestral contribution[50], respectively, with vivid contrasts among each Brazilian region (Table S2-3).

The MCMC runs diagnosis were conducted, within each pedigree, on the smallest chromosome exhibiting a positive LOD score (Figure S2-7). Through pedigree analysis, we identified 30 ROIs (Table S4) of which 8 were discarded due to a lack of corroboration between the three subpanels of SNPs. The remaining 22 ROIs contain 2363 genes (Table S5) (example in Figure 6A).

We utilized family-based association studies (Figure 6B-C) to analyze a total of 1,612,754 SNPs, derived from both pedigree- and population-based imputation strategies, across the 22 ROIs. The summary statistics for the extended SNP set are publicly accessible in the GWAS Catalog (GCST90454187). From this analysis, we identified 117 variants (Table S6) with suggestive or significant associations with EH.

Furthermore, our investigation of EH-related genes, utilizing publicly available databases and supported by the literature, narrowed down the initial set of 2363 genes to 60 promising candidate genes (Table 4) located within the 22 ROIs.

Among the 22 remaining ROIs, 20 of them were single-supported by the investigation of EH-related genes as a fine-mapping strategy, and 17 of them were additionally supported by family-based association studies as a fine mapping strategy (double-supported). Highlighting the common ground between the double supported results (by both the fine-mapping strategies), 14 genes (highlighted in Table 4) were identified within the mapped regions with compelling evidence of association with the phenotype. These genes include *PHGDH* and *S100A10* (ROI1); *MFN2* (ROI2); *RYR2*, *EDARADD*, and *MTR* (ROI3); *SERTAD2* (ROI4); *LPP* (ROI5); *KCNT1* (ROI11); *TENM4* (ROI13); *P2RX1*, *ZZEF1*, and *RPA1* (ROI18); and *ALPK2* (ROI20). All these genes were broadly described in the supplementary section (SS7: Double-supported EH genes). These genes harbor 29 SNPs (highlighted in Table S6) implicated by our family-based association studies.

From our analysis, the estimated genetic correlation between our study and GCST90436066 was 0.427, with a genetic covariance of 0.0105 (p = 0.743). Heritability estimates for GCST90436066 showed a total observed scale heritability (h^2) of 0.0023 (Z = 3.15, $p = 7\times10^{-4}$), while for our study, h^2 was 0.2623 (Z = 0.271), reflecting a significant difference between the datasets. These results are likely influenced by the fact that GCST90436066 represents a European cohort (there is no available African or admixed EH summary statistics in public repositories such as the GWAS Catalog), in contrast to the admixed nature of our dataset, which contains a high proportion of African ancestry. Despite these limitations, the moderate genetic correlation observed provides some level of validation, indicating some level of shared genetic architecture between the two cohorts for this trait.

DISCUSSION

Despite significant financial investment and the substantial number of samples analyzed in GWLA and GWAS, much of the heritability of complex phenotypes such as EH remains elusive. Moreover, due to the underrepresentation of African and Native American populations in global genetic studies, much of what is currently known about complex phenotypes may not be fully applicable to admixed populations, such as the descendants of quilombo populations and, more broadly, the Afro-descendant Brazilians. By studying EH in semi-isolated quilombo remnants, we reduce biases from population and clinical-biological heterogeneity, improve the signal-to-noise ratio (and consequently the statistical power), and increase the representation of admixed populations (like the Brazilian population) in genomic studies.

In our complementary approach, we combined two main strategies, i.e., linkage analysis and family-based association studies. Together, these strategies minimize the need for multiple testing, reduce population stratification, and leverage chromosomal location information provided by meiotic events. This

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

combined approach resulted in 22 filtered ROIs with 60 single-supported genes (harboring 117 suggestive/significant associated variants) by the EH-genes investigation, and 14 double-supported genes by both the fine-mapping (investigation of EH-genes and FBAS) harboring 29 suggestive/significant associated variants.

The 14 double-supported genes imply EH through their roles in vascular, cardiac, and metabolic biological pathways. For instance, *LPP* harbors BP associated SNPs[99] and plays a role in maintaining vascular smooth muscle cell contractility[100], which helps prevent hypertension-induced arterial remodeling[101]. Similarly, *P2RX1*, identified in both African ancestry populations and our Brazilian Afro-descendant cohort, plays a key role in vasoconstriction in small arteries[102] and renal autoregulation, both of which are critical for maintaining blood pressure homeostasis[103]. Apparently, this gene controls glomerular hemodynamics in ANG II hypertension[104]. *RYR2* harbors several BP and EH-associated SNPs[105–107] and is key to calcium handling in cardiac muscle[102], directly affecting heart function and blood pressure regulation, while *MFN2* is involved in mitochondrial function[102], with mutations linked to vascular smooth muscle proliferation, which may contribute to hypertension[108–110]. Collectively, these genes are involved in mechanisms affecting vascular tone, muscle contraction, and energy metabolism, all central to blood pressure regulation.

Several genes related to oxidative stress, endothelial dysfunction, and nitric oxide (NO) regulation, were associated to EH in this study and supported by the literature. *RPA1* influences the repression of endothelial nitric oxide synthase (eNOS)[111], thus reducing NO levels and leading to endothelial dysfunction[112], a key player in hypertension. *EDARADD* has been linked to pulse pressure regulation[102,113], while *S100A10* is involved in cellular processes such as inflammation[114–116], which can contribute to vascular dysfunction and elevated blood pressure. *MTR*, responsible for methionine biosynthesis[102], has been associated with diastolic blood pressure[117], indicating its possible role in homocysteine metabolism and cardiovascular risk in EH patients[118], and has been suggested as a genetic marker to assess the action of antihypertensive drugs[119]. We have identified a blood pressure - related SNP in *MTR* (rs1805087) in linkage disequilibrium ($r^2 \geq 0.97$) with the suggestive SNP rs10925261 (p-value = 2.5×10^{-3}) identified in this study. This set of genes collectively suggests that endothelial health, inflammation, and metabolic regulation converge on pathways influencing blood pressure.

Lastly, genes such *as ZZEF1*[120,121], *PHGDH*[122,123], and *TENM4*[124,125] were associated with blood pressure regulation via several genetic polymorphisms. *PHGDH* has been linked to blood pressure through epigenetic regulation, with methylation changes influencing systolic and diastolic pressure[123]. *KCNT1* may influence vascular tone indirectly although it is mainly involved in potassium channel regulation[126,127]. Although *SERTAD2*[110] and *ALPK2*[128] present BP regulation-associated polymorphisms, they have less direct evidence connecting them to EH but their involvement in cellular signaling and protein interactions suggest possible contributions to blood pressure regulation.

Note that the SNPs linked to the EH phenotype identified in this study were genotyped using genomic arrays and are therefore common variants. Many of these variants are situated in non-coding or intergenic regions and may or may not have recognized functional or regulatory effects. Although these SNPs are not expected to affect the phenotype directly, several of these variants are in LD with variants of more significant impact. Notably, among these, some may be rare and remain ungenotyped. Specifically, we identified unique tag SNPs: 77 non-coding SNPs (3' UTR, 5' UTR, or TF binding), 196 regulatory, and 15 missense SNPs (Table S7).

Recent GWAS studies have highlighted significant findings in African-derived populations, revealing genetic variants associated with blood pressure traits that differ from those identified in European populations. For example, a large meta-analysis involving 80,950 individuals of African ancestry identified 10 variants for SBP and 9 for DBP, including a novel variant, rs562545 in the *MOBP* gene, associated with DBP[129]. The AWI-Gen study, which focused on sub-Saharan African populations, identified two genome-wide significant signals for blood pressure traits: one near *P2RY1* for SBP and another near *LINC01256* for pulse pressure[130]. These findings, absent in European ancestry cohorts, demonstrate the critical need for African-derived specific studies. Similarly, large-scale GWAS efforts, such as those utilizing the UK Biobank, discovered that only a fraction of loci identified in European populations replicate in African-derived populations, demonstrating that ancestry-specific genetic variants play a crucial role in

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

hypertension susceptibility[129,130]. African-derived populations often exhibit unique genetic associations due to differences in genetic architecture. For example, variants in genes such as *CACNA1D* and *KCNK3* that have been found to significantly contribute to blood pressure regulation in African populations are less prominent in other populations.

Admixture mapping has provided valuable insights into loci with potential effects on BP and EH, identifying genomic regions particularly relevant in populations with African ancestry. For instance, *NPR3*, associated with blood pressure regulation in both African and European populations, was uncovered through admixture mapping in African American[131]. Additional regions such as 1q21.2–21.3, 4p15.1, 19q12, and 20p13 have demonstrated significant associations with DBP (p-values $\leq 2.07\times10^{-4}$), while regions 1q21.2–21.3 and 19q12 were also associated (p-values $\leq 5.32\times10^{-4}$) with mean arterial pressure[132].

Several regions identified in the literature have been corroborated by findings from our study. The 3q27.3 region (ROI 5), containing the *AGT* gene, is strongly linked to EH, especially in African ancestry populations[133]. Likewise, the 17q23.2 region (ROI 19) has been implicated in blood pressure traits, particularly through its role in regulating the RAAS[134]. Furthermore, 6p24.1 (ROI 9) has shown previous associations with blood pressure regulation[133]. Linkage studies have also implicated 1q43 (ROI 3) in essential hypertension[135], while 8p23.1 (ROI 10) has been recognized for its relevance in blood pressure regulation, particularly in populations of African descent[134].

These findings underscore the value of including diverse populations in genetic studies to uncover novel loci and better understand the genetic basis of complex traits like hypertension. In this study, our approach identified 22 ROIs, with 14 relevant genes, from which we highlight *PHGDH*, *S100A10*, and *RYR2* implicated in hypertension susceptibility. The overlap between findings in African ancestry populations and those in Brazilian Afro-descendants, such as *P2RX1* highlighted in our study, further supports the role of GWAS and admixture mapping in identifying population-specific loci relevant to hypertension risk.

To prioritize our results, we developed a comprehensive score based on the weights assigned to each strategy (Table S8), including linkage analysis, EH genes investigation, and association studies. The scores allowed to classify the ROIs into three tiers based on their priority level (Table 5): high (top 20% of the ROIs), intermediate (30% of the ROIs), and low (50% of the ROIs).

As presented, our study has some limitations. One such limitation is that the investigation of EH-related genes, as part of the fine-mapping strategy, was focused on genes previously supported by literature as being related to blood pressure regulation. Consequently, it is possible that genes that have never been implicated in EH may contain rare or novel variants relevant to the origin of hypertension in quilombo remnant populations.

On the other hand, the strength of this study comes from the improvement and application of a unique multi-level computational approach that combined mapping strategies to deal with large family data, which provided reliable results. By using genome-wide linkage analyses based on MCMC methods and adjusted for admixture, association studies as the primary fine-mapping strategy, and limiting analyses to candidate genomic regions, this study took advantage of meiotic information provided by pedigrees while simultaneously reducing the need for multiple tests and avoiding population stratification. Therefore, the ROIs identified in this study are credible and provide valuable insights into the genetic basis of essential hypertension in the quilombo remnant populations.

Conducting analyses by merging all six pedigrees into only one would be a formidable challenge, and probably not feasible. However, the prospect of replicating these analyses using alternative computational packages is exciting and not an impossible task. Our study has demonstrated that blood pressure and hypertension in the quilombo remnant populations are likely influenced by multiple genes, in a polygenic or oligogenic mechanism of inheritance. We have identified several loci across different chromosomes that contain genes and variants involved in the development of hypertension. Additionally, we have identified genomic regions of interest not previously associated with EH and will therefore be important targets for future research.

To overcome the limitations, future steps of this investigation will involve the use of Whole Genome Sequencing (WGS) and Whole Exome Sequencing (WES) from this dataset. WGS and WES will enable

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the investigation of coding and non-coding variants within all ROIs, with a focus on rare variants that may have a higher impact on gene functioning. To optimize this process, the prioritization of ROIs as performed in Table 5 is essential, since high and intermediate ROIs will be addressed first when filtering variants detected after WES and WGS, and the low-priority ROIs afterwards.

Furthermore, our study has the potential to shed light on important genomic regions, genes, and variants that are specific to African-derived populations. We have provided insights into the genetic factors that contribute to hypertension in a group that has been often underrepresented in genetic studies and databases.

ACKNOWLEDGMENTS

Authors would like to thank the support of Dr. Diogo Meyer (IB/USP - Brazil), Dr. Julia M. Pavan Soler (IME/USP - Brazil), Dr. Suely Ruiz Giolo (UFPR - Brazil), Dr. Paulo Otto (IB/USP - Brazil) and Dr. Alexandre da Costa Pereira (INCOR - Brazil) for valuable suggestions. The authors also thank Dr. Christian Kubisch (University Medical Center Hamburg / Eppendorf - Germany) for many ideas which stimulated the development of this project.

FUNDING SOURCES

We acknowledge funding from CEPID-FAPESP (Research Center on the Human Genome and Stem Cells - São Paulo Research Foundation, grants 1998/14254-2 and 2013/08028-1, and FAPESP/INCT-CNPq (São Paulo Research Foundation/National Institutes of Science and Technology-National Council for Scientific and Technological Development) 2014/50931-3 led by Dr. Mayana Zatz). This research also received support from grant FAPESP (São Paulo Research Foundation) 2012/18010-0, led by Dr. Diogo Meyer (IB/USP-Brazil). Additionally, we are grateful to CAPES (Brazilian Federal Agency for Support and Evaluation) for the sandwich Ph.D. fellowship #88887.371219/2019-00 and CNPq (National Council for Scientific and Technological Development) for the Ph.D. fellowship #142193/2017-8.

We also acknowledge funding from the Marshall University Joan C. Edwards School of Medicine, Marshall University Data Science Core, WV-INBRE (West Virginia- IDeA Networks of Biomedical Research Excellence) grant (NIH P20GM103434), Bench-to-Bedside Pilot grant (Dr. Nato) under the West Virginia Clinical and Translational Science Institute (WV-CTSI) grant (NIH 5U54GM104942), and Dr. Nato startup fund.

DISCLOSURES

None.

SUPPLEMENTAL MATERIAL

Data S1-S6

Tables S1-S8

Figures S1-S7

References 135-154

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol*. 2020;16(4):223-237. doi:10.1038/s41581-019-0244-2
2. Fuchs FD, Whelton PK. High Blood Pressure and Cardiovascular Disease. *Hypertension*. 2020;75(2):285-292. doi:10.1161/hypertensionaha.119.14240
3. Barroso WKS, Rodrigues CIS, Bortolotto LA, et al. Diretrizes Brasileiras de Hipertensão Arterial – 2020. *Arq Bras Cardiol*. 2021;116(3):516-658. Accessed January 29, 2025. https://doi.org/10.36660/abc.20201238
4. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Pr. *J Am Coll Cardiol*. 2018;71(19):e127-e248. doi:10.1016/j.jacc.2017.11.006
5. Unger T, Borghi C, Charchar F, et al. Hypertension: New Guidelines from the International Society of Hypertension. *Am Fam Physician*. 2021;103(12):763-765.
6. World Health Organization. *Global Report on Hypertension The Race against a Silent Killer*. (World Health Organization, ed.).; 2023. Accessed November 29, 2023. https://www.who.int/publications/i/item/9789240081062
7. Instituto Brasileiro de Geografia e Estatística (IBGE). Percepção do estado de saúde, estilos de vida, doenças crônicas e saúde bucal. Ibge. 2019. Accessed December 19, 2023. http://www.pns.icict.fiocruz.br/arquivos/Portaria.pdf
8. Niiranen TJ, McCabe EL, Larson MG, et al. Risk for hypertension crosses generations in the community: Amulti-generational cohort study. *Eur Heart J*. 2017;38(29):2300-2308. doi:10.1093/eurheartj/ehx134
9. Arnett DK, Claas SA. Omics of blood pressure and hypertension. *Circ Res*. 2018;122(10):1409-1419. doi:10.1161/circresaha.118.311342
10. Patel RS, Masi S, Taddei S. Understanding the role of genetics in hypertension. *Eur Heart J*. 2017;38(29):2309-2312. doi:10.1093/eurheartj/ehx273
11. Keaton JM, Hellwege JN, Giri A, et al. Associations of biogeographic ancestry with hypertension traits. *J Hypertens*. 2021;39(4):633-642. doi:10.1097/hjh.02701
12. Aggarwal R, Chiu N, Wadhera RK, et al. Racial/Ethnic Disparities in Hypertension Prevalence, Awareness, Treatment, and Control in the United States, 2013 to 2018. *Hypertension*. 2021;78(6):1719-1726. doi:10.1161/hypertensionaha.121.17570
13. Jones ES, Spence JD, McIntyre AD, et al. High frequency of variants of candidate genes in black Africans with low renin-resistant hypertension. *Am J Hypertens*. 2017;30(5):478-483. doi:10.1093/ajh/hpw167
14. Marden JR, Walter S, Kaufman JS, Glymour MM. African Ancestry, Social Factors, and Hypertension among Non-Hispanic Blacks in the Health and Retirement Study. *Biodemography Soc Biol*. 2016;62(1):19-35. doi:10.1080/19485565.2015.1108836
15. Zilbermint M, Gaye A, Berthon A, et al. ARMC 5 Variants and Risk of Hypertension in Blacks: MH- GRID Study. *J Am Heart Assoc*. 2019;8(14). doi:10.1161/jaha.119.012508
16. Zhou B, Carrillo-Larco RM, Danaei G, et al. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *The Lancet*. 2021;398(10304):957-980. doi:10.1016/S0140-6736(21)01330-1

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

17. Sells ML, Blum E, Perry GS, Eke P, Presley-Cantrell L. Excess Burden of Poverty and Hypertension, by Race and Ethnicity, on the Prevalence of Cardiovascular Disease. *Prev Chronic Dis*. 2023;20. doi:10.5888/pcd20.230065
18. Ferdinand KC, Armani AM. The management of hypertension in African Americans. *Crit Pathw Cardiol*. 2007;6(2):67-71. doi:10.1097/hpc.0b013e318053da59
19. Beevers G, Lip GYH, O'Brien E. The pathophysiology of hypertension. *Br Med J*. 2001;322(7291):912–916. doi:10.1136/bmj.322.7291.912
20. Seidel E, Scholl UI. Genetic mechanisms of human hypertension and their implications for blood pressure physiology. *Physiol Genomics*. 2017;49(11):630-652. doi:10.1152/physiolgenomics.00032.2017
21. Ghitani N, Chesler AT. The anatomy of the baroreceptor reflex. *Cell Rep*. 2019;29(8):2121-2122. doi:10.1016/j.celrep.2019.11.031
22. Nardone M, Incognito A V., Millar PJ. Evidence for pressure-independent sympathetic modulation of central pulse wave velocity. *J Am Heart Assoc*. 2018;7(3):1-8. doi:10.1161/jaha.117.007971
23. Hermann M, Flammer A, Lüscher TF. Nitric oxide in hypertension. *J Clin Hypertens (Greenwich)*. 2006;8(12 Suppl 4):17-29. doi:10.1111/j.1524-6175.2006.06032.x
24. Špiranec Spes K, Chen W, Krebes L, et al. Heart-microcirculation connection: Effects of ANP (atrial natriuretic peptide) on pericytes participate in the acute and chronic regulation of arterial blood pressure. *Hypertension*. Published online 2020:1637-1648. doi:10.1161/hypertensionaha.120.15772
25. Hamid S, Rhaleb IA, Kassem KM, Rhaleb NE. Role of kinins in hypertension and heart failure. *Pharmaceuticals*. 2020;13(11):1-25. doi:10.3390/ph13110347
26. Grillo A, Salvi L, Coruzzi P, Salvi P, Parati G. Sodium intake and hypertension. *Nutrients*. 2019;11(9):1-16. doi:10.3390/nu11091970
27. Zilbermint M, Hannah-Shmouni F, Stratakis CA. Genetics of hypertension in African Americans and others of African descent. *Int J Mol Sci*. 2019;20(5):15-17. doi:10.3390/ijms20051081
28. Li R, Richey PA, DiSessa TG, Alpert BS, Jones DP. Blood aldosterone-to-renin ratio, ambulatory blood pressure, and left ventricular mass in children. *J Pediatr*. 2009;155(2):170-175. doi:10.1016/j.jpeds.2009.02.029
29. Brier ME, Luft FC. Sodium kinetics in white and black normotensive subjects: possible relevance to salt-sensitive hypertension. *Am J Med Sci*. 1994;307 Suppl 1:S38-42.
30. Gafane LF, Schutte R, Van Rooyen JM, Schutte AE. Plasma renin and cardiovascular responses to the cold pressor test differ in black and white populations: The SABPA study. *J Hum Hypertens*. 2016;30(5):346-351. doi:10.1038/jhh.2015.88
31. Van Rooyen JM, Poglitsch M, Huisman HW, et al. Quantification of systemic renin-angiotensin system peptides of hypertensive black and white African men established from the RAS-Fingerprint®. *J Renin Angiotensin Aldosterone Syst*. 2016;17(4):1-7. doi:10.1177/1470320316669880
32. Rayner BL, Owen EP, King JA, et al. A new mutation, R563Q, of the beta subunit of the epithelial sodium channel associated with low-renin, low-aldosterone hypertension. *J Hypertens*. 2003;21(5):921-926. doi:10.1097/00004872-200305000-00016
33. Kimura L, Angeli CB, Auricchio MTBM, et al. Multilocus family-based association analysis of seven candidate polymorphisms with essential hypertension in an African-derived semi-isolated brazilian population. *Int J Hypertens*. 2012;2012. doi:10.1155/2012/859219

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

34. Liu F, Yang X, Mo X, et al. Associations of epithelial sodium channel genes with blood pressure: the GenSalt study. *J Hum Hypertens*. 2015;29(4):224. doi:10.1038/jhh.2014.78
35. Wen G, Wessel J, Zhou W, et al. An ancestral variant of Secretogranin II confers regulation by PHOX2 transcription factors and association with hypertension. *Hum Mol Genet*. 2007;16(14):1752-1764. doi:10.1093/hmg/ddm123
36. Kumar A, Li Y, Patil S, Jain S. A haplotype of the angiotensinogen gene is associated with hypertension in african americans. *Clin Exp Pharmacol Physiol*. 2005;32(5-6):495-502. doi:10.1111/j.1440-1681.2005.04217.x
37. Carey RM, Muntner P, Bosworth HB, Whelton PK. Prevention and Control of Hypertension: JACC Health Promotion Series. *J Am Coll Cardiol*. 2018;72(11):1278-1293. doi:10.1016/j.jacc.2018.07.008
38. Roerecke M, Kaczorowski J, Tobe SW, Gmel G, Hasan OSM, Rehm J. The effect of a reduction in alcohol consumption on blood pressure: a systematic review and meta-analysis. *Lancet Public Health*. 2017;2(2):e108-e120. doi:10.1016/S2468-2667(17)30003-8
39. Samadian F, Dalili N, Jamalian A. Lifestyle modifications to prevent and control hypertension. *Iran J Kidney Dis*. 2016;10(5):237-263.
40. Mente A, O'Donnell M, Rangarajan S, et al. Urinary sodium excretion, blood pressure, cardiovascular disease, and mortality: a community-level prospective epidemiological cohort study. *The Lancet*. 2018;392(10146):496-506. doi:10.1016/S0140-6736(18)31376-X
41. Mill JG, Malta DC, Machado ÍE, et al. Estimation of salt intake in the Brazilian population: Results from the 2013 national health survey. *Revista Brasileira de Epidemiologia*. 2019;22(Suppl 2). doi:10.1590/1980-549720190009.supl.2
42. Howard G, Cushman M, Moy CS, et al. Association of Clinical and Social Factors With Excess Hypertension Risk in Black Compared With White US Adults. *JAMA*. 2018;320(13):1338-1348. doi:10.1001/jama.2018.13467
43. Cooper RS, Forrester TE, Plange-Rhule J, et al. Elevated hypertension risk for African-origin populations in biracial societies: modeling the Epidemiologic Transition Study. *J Hypertens*. 2015;33(3):473-481. doi:10.1097/hjh.0429
44. Ehret GB, Caulfield MJ. Genes for blood pressure: An opportunity to understand hypertension. *Eur Heart J*. 2013;34(13):951-961. doi:10.1093/eurheartj/ehs455
45. Seidel E, Scholl UI. Genetic mechanisms of human hypertension and their implications for blood pressure physiology. *Physiol Genomics*. 2017;49(11):630-652. doi:10.1152/physiolgenomics.00032.2017
46. Olczak KJ, Taylor-Bateman V, Nicholls HL, Traylor M, Cabrera CP, Munroe PB. Hypertension genetics past, present and future applications. *J Intern Med*. 2021;290(6):1130-1152. doi:10.1111/JOIM.13352
47. Mills MC, Rahal C. The GWAS Diversity Monitor tracks diversity by disease in real time. *Nat Genet*. 2020;52(3):242-243. doi:10.1038/S41588-020-0580-Y
48. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624):161-164. doi:10.1038/538161a
49. Buniello A, Macarthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005-D1012. doi:10.1093/nar/gky1120
50. de Souza AM, Resende SS, de Sousa TN, de Brito CFA. A systematic scoping review of the genetic ancestry of the brazilian population. *Genet Mol Biol*. 2019;42(3):495-508. doi:10.1590/1678-4685-gmb-2018-0076

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

51. Borges VM, Kimura L. Panorama of arterial hypertension in quilombos in Brazil: a narrative review. *Physis: Revista de Saúde Coletiva*. 2023;33(e33050). doi:https://doi.org/10.1590/S0103-7331202333050.en
52. Lemes RB, Nunes K, Meyer D, Mingroni-Netto RC, Otto PA. Estimation of inbreeding and substructure levels in African-derived Brazilian quilombo populations. *Hum Biol*. 2014;86(4):276-288. doi:10.13110/humanbiology.86.4.0276
53. Nafikov RA, Nato, A. Q. J, Sohi H, et al. Analysis of Pedigree Data in Populations with Multiple Ancestries: Strategies for Dealing with Admixture in Caribbean Hispanic Families from the ADSP. *Genet Epidemiol*. 2018;42(6):500-515. doi:10.1002/gepi.22133
54. Saad M, Wijsman EM. Combining family- and population-based imputation data for association analysis of rare and common variants in large pedigrees. *Genet Epidemiol*. 2014;38(7):579-590. doi:10.1002/gepi.21844
55. Ullah E, Mall R, Abbas MM, et al. Comparison and assessment of family- and population-based genotype imputation methods in large pedigrees. *Genome Res*. 2019;29(1):125-134. doi:10.1101/gr.236315.118
56. Chen H, Wang C, Conomos MP, et al. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. *Am J Hum Genet*. 2016;98(4):653-666. doi:10.1016/j.ajhg.2016.02.012
57. Gogarten SM, Sofer T, Chen H, et al. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics*. 2019;35(24):5346-5348. doi:10.1093/bioinformatics/btz567
58. Applied Biosystems. Axiom Analysis Suite 5.1 User Guide. 2020;(703307):33. https://www.thermofisher.com/us/en/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/axiom-analysis-suite.html
59. Kimura L, Ribeiro-Rodrigues EM, De Mello Auricchio MTB, Vicente JP, Batista Santos SE, Mingroni-Netto RC. Genomic ancestry of rural African-derived populations from Southeastern Brazil. *American Journal of Human Biology*. 2013;25(1):35-41. doi:10.1002/ajhb.22335
60. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867-2873. doi:10.1093/bioinformatics/btq559
61. Wijsman EM, Rothstein JH, Thompson EA. Multipoint linkage analysis with many multiallelic or dense diallelic markers: Markov chain-Monte Carlo provides practical approaches for genome scans on general pedigrees. *Am J Hum Genet*. 2006;79(5):846-858. doi:10.1086/508472
62. Tong L, Thompson E. Multilocus lod scores in large pedigrees: combination of exact and approximate calculations. *Hum Hered*. 2008;65(3):142-153. doi:10.1159/000109731
63. Nato AQ, Chapman NH, Sohi HK, Nguyen HD, Brkanac Z, Wijsman EM. PBAP: A pipeline for file processing and quality control of pedigree data with dense genetic markers. *Bioinformatics*. 2015;31(23):3790-3798. doi:10.1093/bioinformatics/btv444
64. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1):1-16. doi:10.1186/s13742-015-0047-8
65. Matise TC, Chen F, Chen W, et al. A second-generation combined linkage-physical map of the human genome. *Genome Res*. 2007;17(12):1783-1786. doi:10.1101/gr.7156307

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

66. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. *Nucleic Acids Res*. 2020;48(D1):D941-D947. doi:10.1093/nar/gkz836
67. Huang L, Jakobsson M, Pemberton TJ, et al. Haplotype variation and genotype imputation in African populations. *Genet Epidemiol*. 2011;35(8):766. doi:10.1002/gepi.20626
68. O'Connell J, Gurdasani D, Delaneau O, et al. A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet*. 2014;10(4). doi:10.1371/journal.pgen.1004234
69. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet*. 2013;93(2):278-288. doi:10.1016/j.ajhg.2013.06.020
70. Thompson E. The structure of genetic linkage data: From LIPED to 1M SNPs. *Hum Hered*. 2011;71(2):86-96. doi:10.1159/000313555
71. Kaiser HF. The Application of Electronic Computers to Factor Analysis. *Educ Psychol Meas*. 1960;20(1):141-151. doi:10.1177/001316446002000116
72. Delaneau O, Marchini J, McVeanh GA, et al. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nature Communications 2014 5:1*. 2014;5(1):1-9. doi:10.1038/ncomms4934
73. Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics*. 2015;31(5):782-784. doi:10.1093/bioinformatics/btu704
74. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2):1-4. doi:10.1093/gigascience/giab008
75. Cheung CYK, Thompson EA, Wijsman EM. GIGI: An approach to effective imputation of dense genotypes on large pedigrees. *Am J Hum Genet*. 2013;92(4):504-516. doi:10.1016/j.ajhg.2013.02.011
76. Horimoto ARVR, Boyken LA, Blue EE, et al. Admixture mapping implicates 13q33.3 as ancestry-of-origin locus for Alzheimer disease in Hispanic and Latino populations. *Human Genetics and Genomics Advances*. 2023;4(3):19. doi:10.1016/j.xhgg.2023.100207
77. Li MX, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131(5):747-756. doi:10.1007/s00439-011-1118-2
78. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004. doi:10.1111/j.0006-341X.1999.00997.x
79. Sherry ST, Ward M, Sirotkin K. dbSNP - database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res*. 1999;9(8):677-679. doi:10.1101/gr.9.8.677
80. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med*. 2021;13(1):1-12. doi:10.1186/s13073-021-00835-9
81. National Center for Biotechnology Information (NCBI). National Center for Biotechnology Information. Bethesda (MD): National Library of Medicine (US). 1988. Accessed January 9, 2024. https://www.ncbi.nlm.nih.gov/
82. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):1-14. doi:10.1186/s13059-016-0974-4

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

83. Landrum MJ, Lee JM, Benson M, et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062-D1067. doi:10.1093/nar/gkx1153
84. Schwarz JM, Cooper DN, Schuelke M, Seelow D. Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11(4):361-362. doi:10.1038/nmeth.2890
85. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc*. 2016;11(1):1-9. doi:10.1038/nprot.2015.123
86. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249. doi:10.1038/nmeth0410-248
87. Kopanos C, Tsiolkas V, Kouris A, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2019;35(11):1978-1980. doi:10.1093/bioinformatics/bty897
88. Fairbrother-Browne A, García-Ruiz S, Reynolds RH, Ryten M, Hodgkinson A. ensemblQueryR: fast, flexible and high-throughput querying of Ensembl LD API endpoints in R. *GigaByte*. 2023;2023:1. doi:10.46471/gigabyte.91
89. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;4(8):1184-1191. doi:10.1038/nprot.2009.97
90. National Library of Medicine. MedGen. Bethesda, MD: U.S. Accessed January 6, 2024. https://www.ncbi.nlm.nih.gov/medgen/
91. Rappaport N, Nativ N, Stelzer G, et al. MalaCards: An integrated compendium for diseases and their annotation. *Database*. 2013;2013:1-14. doi:10.1093/database/bat018
92. Kinsella RJ, Kähäri A, Haider S, et al. Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database*. 2011;2011:1-9. doi:10.1093/database/bar030
93. Sollis E, Mosaku A, Abid A, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res*. 2023;51(D1):D977-D985. doi:10.1093/nar/gkac1010
94. Stelzer G, Plaschkes I, Oz-Levi D, et al. VarElect: The phenotype-based variation prioritizer of the GeneCards Suite. *BMC Genomics*. 2016;17(Suppl 2). doi:10.1186/s12864-016-2722-2
95. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50(9):1335-1341. doi:10.1038/s41588-018-0184-y
96. Grotzinger AD, Rhemtulla M, de Vlaming R, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav*. 2019;3(5):513-525. doi:10.1038/s41562-019-0566-x
97. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi:10.1038/ng.3406
98. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010;467(7311):52-58. doi:10.1038/nature09298
99. Surendran P, Drenos F, Young R, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat Genet*. 2016;48(10):1151-1161. doi:10.1038/ng.3654
100. Hooper CL, Dash PR, Boateng SY. Lipoma preferred partner is a mechanosensitive protein regulated by nitric oxide in the heart. *FEBS Open Bio*. 2012;2:135-144. doi:10.1016/j.fob.2012.05.005
101. Nahar T. *Role of the Lim-Domain Proteins Lpp and Zyxin in Hypertension-Induced Cardiovascular Remodeling*. University of Heidelberger; 2017. doi:10.11588/heidok.00023649

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

102. O'Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;44(D1):D733-D745. doi:10.1093/nar/gkv1189
103. Gordienko D, Povstyan O, Sukhanova K, et al. Impaired P2X signalling pathways in renal microvascular myocytes in genetic hypertension. *Cardiovasc Res*. 2015;105(2):131-142. doi:10.1093/cvr/cvu249
104. Franco M, Bautista-Pérez R, Cano-Martínez A, et al. Physiopathological implications of P2x1 and P2x7 receptors in regulation of glomerular hemodynamics in angiotensin II-induced hypertension. *Am J Physiol Renal Physiol*. 2017;313(1):F9-F19. doi:10.1152/ajprenal.00663.2016
105. Burton PR, Clayton DG, Cardon LR, et al. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-678. doi:10.1038/nature05911.Genome-wide
106. Ehret GB, Morrison AC, Connor A a O, et al. Replication of the Wellcome Trust genome-wide association study of essential hypertension. *European journal of human genetics*. 2008;16(12):1507-1511. doi:10.1038/ejhg.2008.102.Replication
107. Kostis WJ, Cabrera J, Hooper WC, et al. Relationships between selected gene polymorphisms and blood pressure sensitivity to weight loss in elderly persons with hypertension. *Hypertension*. 2013;61(4):857-863. doi:10.1161/hypertensionaha.111.00712
108. Jin F, Li X, Wang Z, et al. Association of mitofusin 2 methylation and essential hypertension: a case-control study in a Chinese population. *Hypertension Research*. 2018;41(8):605-613. doi:10.1038/s41440-018-0057-x
109. Wang Z, Liu Y, Liu J, et al. HSG/Mfn2 gene polymorphism and essential hypertension: A case-control association study in Chinese. *J Atheroscler Thromb*. 2011;18(1):24-31. doi:10.5551/jat.5611
110. Sung YJ, Winkler TW, de las Fuentes L, et al. A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. *Am J Hum Genet*. 2018;102(3):375-400. doi:10.1016/j.ajhg.2018.01.015
111. Miyamoto Y, Saito Y, Nakayama M, et al. Replication protein A1 reduces transcription of the endothelial nitric oxide synthase gene containing a -786T→C mutation associated with coronary spastic angina. *Hum Mol Genet*. 2000;9(18):2629-2637. doi:10.1093/hmg/9.18.2629
112. Gamil S, Erdmann J, Abdalrahman IB, Mohamed AO. Association of NOS3 gene polymorphisms with essential hypertension in Sudanese patients: A case control study. *BMC Med Genet*. 2017;18(1):1-9. doi:10.1186/s12881-017-0491-7
113. Chen J, Wang W, Li Z, Xu C, Tian X, Zhang D. Heritability and genome-wide association study of blood pressure in Chinese adult twins. *Mol Genet Genomic Med*. 2021;9(11):1-17. doi:10.1002/mgg3.1828
114. Sattar Z, Lora A, Jundi B, Railwah C, Geraghty P. The S100 Protein Family as Players and Therapeutic Targets in Pulmonary Diseases. *Pulm Med*. 2021;2021. doi:10.1155/2021/5488591
115. Huan T, Esko T, Peters MJ, et al. A Meta-analysis of Gene Expression Signatures of Blood Pressure and Hypertension. *PLoS Genet*. 2015;11(3):1-29. doi:10.1371/journal.pgen.1005035
116. Zeller T, Schurmann C, Schramm K, et al. Transcriptome-wide analysis identifies novel associations with blood pressure. *Hypertension*. 2017;70(4):743-750. doi:10.1161/hypertensionaha.117.09458

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

117. Moskau S, Farmand S, Semmler A, et al. The methionine synthase polymorphism c.2756A>G (D919G) influences diastolic blood pressure. *J Hum Hypertens*. 2007;21(5):418-420. doi:10.1038/sj.jhh.1002165
118. Wang XF, Liao YH, Cai Y, Li JY. Gender-specific association of MSA2756G with hypertension in patients attending a health facility in Ningxia Province, China. *Tropical Journal of Pharmaceutical Research*. 2016;15(9):1995-1999. doi:10.4314/tjpr.v15i9.25
119. Zhang Y, Zhang M, Niu T, et al. D919G polymorphism of methionine synthase gene is associated with blood pressure response to benzepril in Chinese hypertensive patients. *J Hum Genet*. 2004;49(6):296-301. doi:10.1007/s10038-004-0149-0
120. Tong X, Wei C, Lu Q. Genome-wide joint analysis of singlenucleotide variant sets and gene expression for hypertension and related phenotypes. *BMC Proc*. 2016;10(Supp 7):1-5. doi:10.1186/s12919-016-0017-x
121. Hoffmann TJ, Ehret GB, Nandakumar P, et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat Genet*. 2017;49(1):54-64. doi:10.1038/ng.3715
122. Richard MA, Huan T, Ligthart S, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *Am J Hum Genet*. 2017;101(6):888-902. doi:10.1016/j.ajhg.2017.09.028
123. Syme C, Shin J, Richer L, et al. Epigenetic Loci of Blood Pressure. *Circ Genom Precis Med*. 2019;12(1):e002341. doi:10.1161/circgen.118.002341
124. Park YM, Kwock CK, Kim K, Kim J, Yang YJ. Interaction between single nucleotide polymorphism and urinary sodium, potassium, and sodium-potassium ratio on the risk of hypertension in Korean adults. *Nutrients*. 2017;9(3). doi:10.3390/nu9030235
125. Li Z, Wang W, Tian X, Duan H, Xu C, Zhang D. Bivariate genome-wide association study (GWAS) of body mass index and blood pressure phenotypes in northern Chinese twins. *PLoS One*. 2021;16(2 February):1-15. doi:10.1371/journal.pone.0246436
126. Kim GE, Kronengold J, Barcia G, et al. Human Slack Potassium Channel Mutations Increase Positive Cooperativity between Individual Channels. *Cell Rep*. 2014;9(5):1661-1672. doi:10.1016/j.celrep.2014.11.015
127. Li P, Halabi CM, Stewart R, et al. Sodium-activated potassium channels moderate excitability in vascular smooth muscle. *Journal of Physiology*. 2019;597(20):5093-5108. doi:10.1113/JP278279
128. Chauvet C, Crespo K, Ménard A, et al. α-Kinase 2 is a novel candidate gene for inherited hypertension in Dahl rats. *J Hypertens*. 2011;29(7):1320-1326. doi:10.1097/HJH.0b013e32834705e4
129. Udosen B, Soremekun O, Kamiza A, et al. Meta-Analysis and Multivariate GWAS Analyses in 80,950 Individuals of African Ancestry Identify Novel Variants Associated with Blood Pressure Traits. *Int J Mol Sci*. 2023;24(3). doi:10.3390/ijms24032164
130. Singh S, Choudhury A, Hazelhurst S, et al. Genome-wide association study meta-analysis of blood pressure traits and hypertension in sub-Saharan African populations: an AWI-Gen study. *Nat Commun*. 2023;14(1). doi:10.1038/s41467-023-44079-0
131. Franceschini N, Fox E, Zhang Z, et al. Genome-wide association analysis of blood-pressure traits in african-ancestry individuals reveals common associated genes in African and Non-African populations. *Am J Hum Genet*. 2013;93(3):545-554. doi:10.1016/j.ajhg.2013.07.010
132. Liu Z, Shriner D, Hansen NF, et al. Admixture mapping identifies genetic regions associated with blood pressure phenotypes in African Americans. *PLoS One*. 2020;15(4). doi:10.1371/journal.pone.0232048

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

133. Mani A. Update in genetic and epigenetic causes of hypertension. *Cellular and Molecular Life Sciences*. 2024;81(1). doi:10.1007/s00018-024-05220-4
134. Kraja AT, Hunt SC, Rao DC, Dávila-Román VG, Arnett DK, Province MA. Genetics of hypertension and cardiovascular disease and their interconnected pathways: Lessons from large studies. *Curr Hypertens Rep*. 2011;13(1):46-54. doi:10.1007/s11906-010-0174-7
135. Ehret GB, O'Connor AA, Weder A, Cooper RS, Chakravarti A. Follow-up of a major linkage peak on chromosome 1 reveals suggestive QTLs associated with essential hypertension: GenNet study. *European Journal of Human Genetics*. 2009;17(12):1650-1657. doi:10.1038/ejhg.2009.94
136. Frisancho A. *Anthropometric Standards for the Assessment of Growth and Nutritional Status*. University of Michigan Press; 1990.
137. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393
138. Li JZ, Absher DM, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008;319(5866):1100-1104. doi:10.1126/SCIENCE.1153717
139. Endre Bakken Stovner. Snpflip. December 2017. Accessed February 24, 2023. github.com/endrebak/snpflip
140. Martin AR, Gignoux CR, Walters RK, et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet*. 2017;100(4):635-649. doi:10.1016/j.ajhg.2017.03.004
141. Donnelly KP. The probability that related individuals share some section of genome identical by descent. *Theor Popul Biol*. 1983;23(1):34-63. doi:10.1016/0040-5809(83)90004-7
142. Horimoto A, Wang B, Nato AQ, et al. *Tips and Tricks for Genome-Wide Linkage Analysis in Large Pedigrees in the next Generation Sequencing Era*.; 2025.
143. Gogarten SM, Bhangale T, Conomos MP, et al. GWASTools: An R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics*. 2012;28(24):3329-3331. doi:10.1093/bioinformatics/bts610
144. Cheung CYK, Thompson EA, Wijsman EM. GIGI: An approach to effective imputation of dense genotypes on large pedigrees. *Am J Hum Genet*. 2013;92(4):504-516. doi:10.1016/j.ajhg.2013.02.011
145. Li MX, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131(5):747-756. doi:10.1007/s00439-011-1118-2
146. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004. doi:10.1111/j.0006-341X.1999.00997.x
147. Kinsella RJ, Kähäri A, Haider S, et al. Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database*. 2011;2011:1-9. doi:10.1093/database/bar030
148. Saloum de Neves Manta F, Pereira R, Vianna R, et al. Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLoS One*. 2013;8(9):1-11. doi:10.1371/journal.pone.0075145
149. Pena SDJ, di Pietro G, Fuchshuber-Moraes M, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One*. 2011;6(2). doi:10.1371/journal.pone.0017063

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

150. Kehdy FSG, Gouveia MH, Machado M, et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proc Natl Acad Sci U S A.* 2015;112(28):8696-8701. doi:10.1073/pnas.1504447112
151. Gontijo CC, Mendes FM, Santos CA, et al. Ancestry analysis in rural Brazilian populations of African descent. *Forensic Sci Int Genet*. 2018;36:160-166. doi:10.1016/j.fsigen.2018.06.018
152. Bortolini MC, Weimer TA, Salzano FM, et al. Evolutionary relationships between black South American and African populations. *Human biology; an international record of research*. 1995;67(4):547-559.
153. Gontijo CC, Guerra Amorim CE, Godinho NMO, et al. Brazilian quilombos: A repository of Amerindian alleles. *American Journal of Human Biology*. 2014;26(2):142-150. doi:10.1002/ajhb.22501
154. Maciel LGL, Rodrigues EM, dos Santos NPC, dos Santos AR, Guerreiro JF, Santos S. Afro-derived Amazonian populations: Inferring continental ancestry and population substructure. *Hum Biol*. 2011;83(5):627-636. doi:10.3378/027.083.0504

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

TABLES AND FIGURES

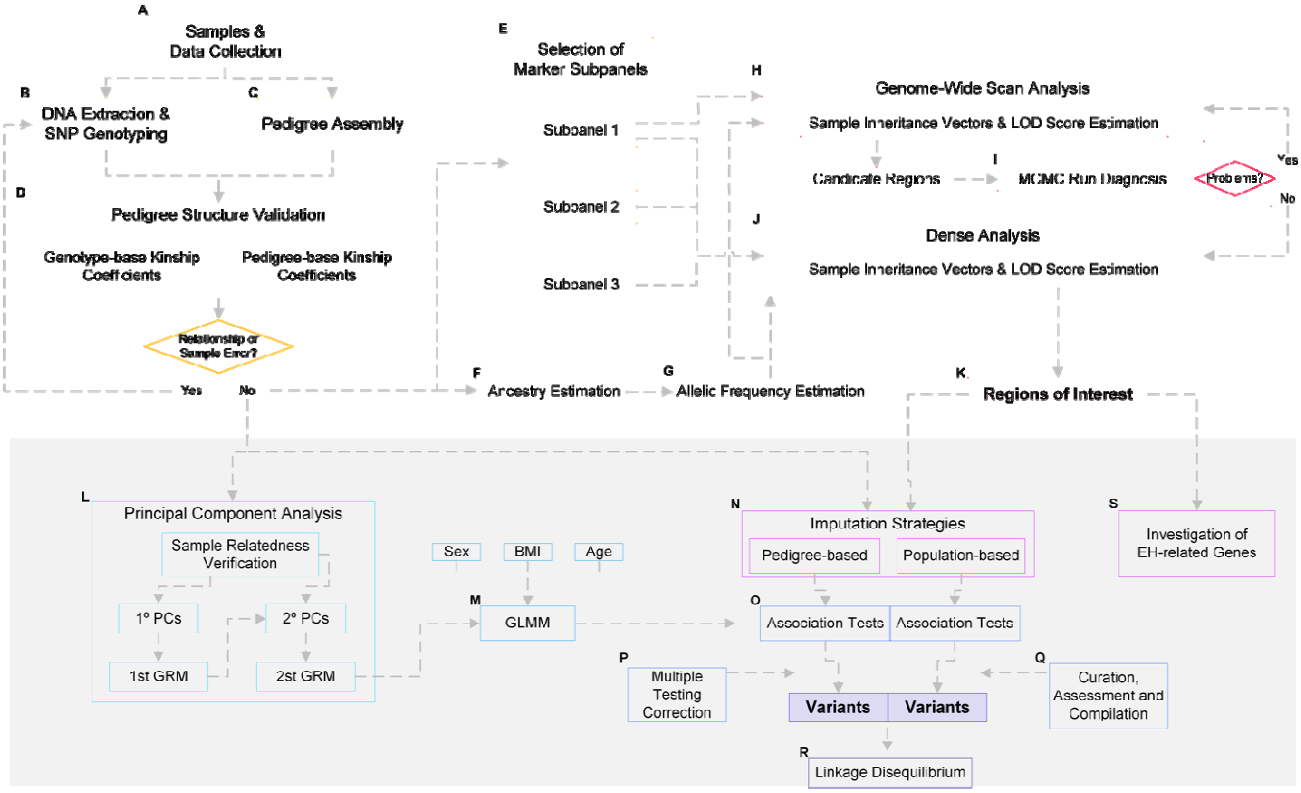

Figure 1. Schematic diagram of the main (boxes A-K) and fine-mapping (boxes L-S) strategy workflow. The color gradient progressively darkens to illustrate the advancement of the process. A) Sample collection and data preparation; B) DNA extraction, quantification, SNP genotyping, and quality control; C) Pedigree assembly; D) Pedigree structure validation; E) Marker selection for subpanels; F) Ancestry estimation; G) Allelic frequency estimation; H) Genome-wide scan analysis; I) Markov chain Monte Carlo (MCMC) convergence diagnostics; J) Dense mapping analysis; K) Identification of regions of interest (ROIs); L) Principal component analysis (PCA); M) Generalized linear mixed model fitting; N) Pedigree and population-based imputation; O) Family-based association tests; P) Multiple testing correction; Q) Variant curation and assessment; S) Investigation of EH-related genes.

Table 1. Distribution of samples across pedigrees. Total samples included by pedigree. Samples are separated into genotyped (affected, unaffected and unknown phenotype) and non-genotyped. Abobral (ABDR), André Lopes and Nhunguara (ANNH), Galvão and São Pedro (GASP), Ivaporunduva (IV), Pedro Cubas (PC) and Sapatu (TU).

Pedigree	**Genotyped**				**Non-genotyped**	**Total**
	Affected	Unaffected	Unknown	Subtotal		
ABDR	38	29	1	68	105	173
ANNH	37	54	-	91	117	208
GASP	45	49	1	95	88	183
IV	12	34	1	47	110	157
PC	16	51	-	67	130	197
TU	19	44	-	63	123	186
Total	**167**	**261**	**3**	**431**	**673**	**1104**

Table 2. Study cohort characteristics. Comparison of cohort characteristics between affected and unaffected individuals across six pedigrees Abobral (ABDR), André Lopes and Nhunguara (ANNH), Galvão and São Pedro (GASP), Ivaporunduva (IV), Pedro Cubas (PC) and Sapatu (TU). The table presents the absolute number and percentage of males, mean age, systolic and diastolic blood pressure (BP) in mmHg, height in meters (m), weight in kilograms (kg), waist girth (cm), and hip girth (cm) for both affected and unaffected groups. The total number of individuals in each group is also provided for each pedigree. Values are shown as mean ± standard deviation. Pedigrees ABDR, GASP, and IV each include additionally one individual with an unknown phenotype.

	ABDR		**ANNH**		**GASP**		**IV**		**PC**		**TU**	
	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected
Males	19 (50%)	9 (31%)	8 (21.6%)	26 (48.1%)	24 (53.3%)	24 (49%)	4 (33.3%)	12 (35.3%)	9 (56.2%)	20 (39.2%)	8 (42.1%)	19 (43.2%)
Age	49.7 ±19.5	35.9 ± 15.9	57 ± 16.6	36.4 ± 16.7	54.3 ± 16.8	37 ± 13.5	64.3 ± 12.7	33.2 ± 14.1	60.6 ± 17	38.4 ± 14.6	53.8 ± 15.1	38.2 ± 17.2
Systolic BP	147.2 ± 25.9	121.6 ± 9.4	145.7 ± 23.2	112.1 ± 14.5	145.2 ± 21	120 ± 15.3	143.8 ± 18.6	115.9 ± 9.9	143.2 ± 15.4	128.7 ± 24.9	141.3 ± 21.8	112.5 ± 15.3
Diastolic BP	93.3 ± 13.4	75.1 ± 9.7	87.2 ± 11.5	75.9 ± 8.2	87.6 ± 14.4	75.1 ± 9.4	85.4 ± 11.6	73.5 ± 7.3	82.1 ± 11.4	80.9 ± 15.5	86.3 ± 6.1	73.3 ± 7.8
Height	1.6 ± 0.1	1.61 ± 0.1	1.54 ± 0.09	1.6 ± 0.12	1.58 ± 0.1	1.62 ± 0.09	1.57 ± 0.1	1.61 ± 0.08	1.6 ± 0.08	1.6 ± 0.09	1.6 ± 0.1	1.61 ± 0.08
Weight	63.3 ± 10.3	63.3 ± 10.7	58.5 ± 12.4	62.6 ± 12.8	64 ± 11.5	65.6 ± 12.3	65.7 ± 12.4	64.8 ± 11	64.3 ± 9.8	61.8 ± 12.9	69.8 ± 15.4	57.6 ± 9.5
Waist Girth	84.1 ± 9.5	83.5 ± 9.9	82.9 ± 10	80.2 ± 10	86.9 ± 9.9	83 ± 10.2	89.8 ± 10.5	82.8 ± 9.9	85.7 ± 5.5	82 ± 9.7	89.5 ± 8.1	76.3 ± 8.3
Hip Girth	94.3 ± 10.9	93.9 ± 11.8	93.4 ± 10	89.3 ± 10.9	94.7 ± 10.2	92.4 ± 12.7	97.7 ± 11.7	93.7 ± 14.3	90.6 ± 7.9	92.2 ± 10.5	97.5 ± 9.8	87.1 ± 8.6
Total	**38 (55.9%)**	**29 (42.6%)**	**37 (40.7%)**	**54 (59.3%)**	**45 (47.4%)**	**49 (51.6%)**	**12 (25.5%)**	**34 (72.3%)**	**16 (23.9%)**	**51 (76.1%)**	**19 (30.2%)**	**44 (69.8%)**

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

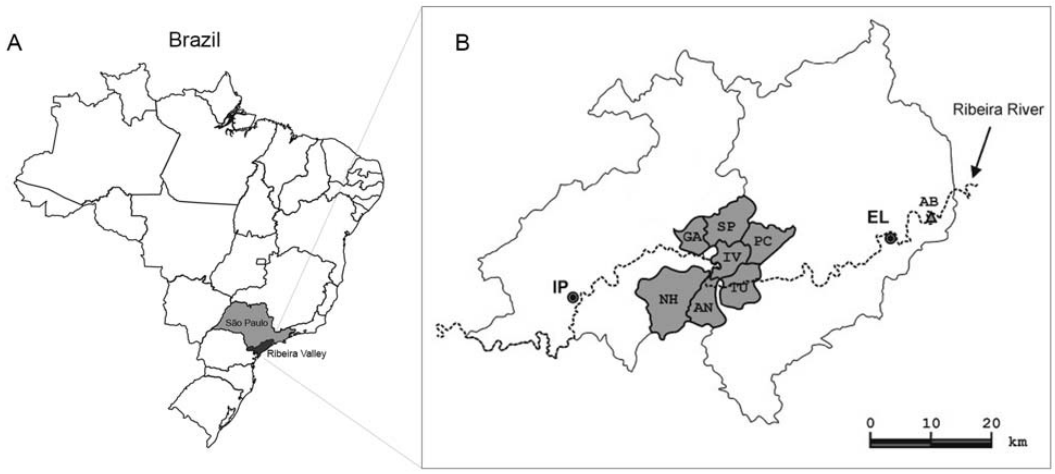

Figure 2. Geographical location of quilombo populations. A) Brazilian territory in South America, with the State of São Paulo highlighted in gray and the Ribeira Valley (*Vale do Ribeira*) in a darker shade of gray; B) Location of quilombo populations: AB (Abobral), AN (Andre Lopes), GA (Galvão), IV (Ivaporanduva), NH (Nhunguara), PC (Pedro Cubas), SP (São Pedro), and TU (Sapatu). The black dots denote the urban centers of Eldorado (EL) and Iporanga (IP). Adapted from Kimura L, *et al.* (2013) [59].

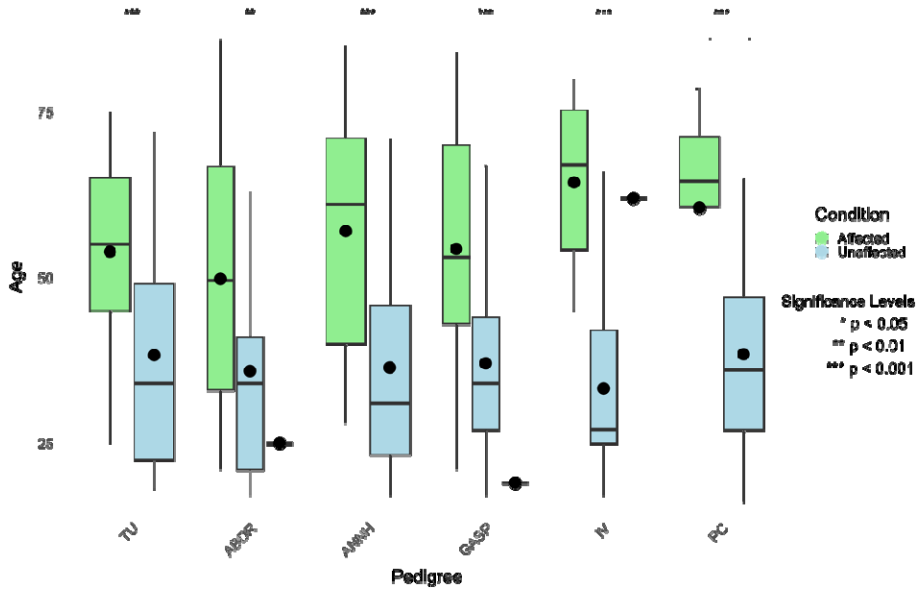

Figure 3. Distribution of ages for both affected and unaffected individuals across different pedigrees. The box plots represent the interquartile range (IQR) with the median age indicated by the line within each box. The small black dots indicate the mean ages for each group, and the error bars represent the standard deviation around these means. Affected individuals are shown in green, and unaffected individuals are shown in blue. Pedigrees Abobral (ABDR), Galvão and São Pedro (GASP), and Ivaporunduva (IV) each include additionally one non-represented individual with an unknown phenotype.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Ancestry Proportions. The ancestry proportions (%) are presented for each pedigree, detailing the contributions of African, European, and Native American ancestries. The corresponding sample sizes are included for reference. The total dataset is summarized under "Total." Pedigrees are identified as Abobral (ABDR), André Lopes and Nhunguara (ANNH), Galvão and São Pedro (GASP), Ivaporunduva (IV), Pedro Cubas (PC), and Sapatu (TU).

Pedigrees	Sample size	Ancestry Proportions		
		African	European	Native American
ABDR	68	51.5%	31.4%	16.9%
GASP	95	46.8%	36.9%	16.2%
ANNH	91	50.8%	33.2%	15.8%
IV	47	47.3%	35.9%	16.7%
TU	63	33.4%	50.8%	15.7%
PC	67	51.5%	32.4%	16.0%
Total	**431**	**47.4%**	**36.3%**	**16.1%**

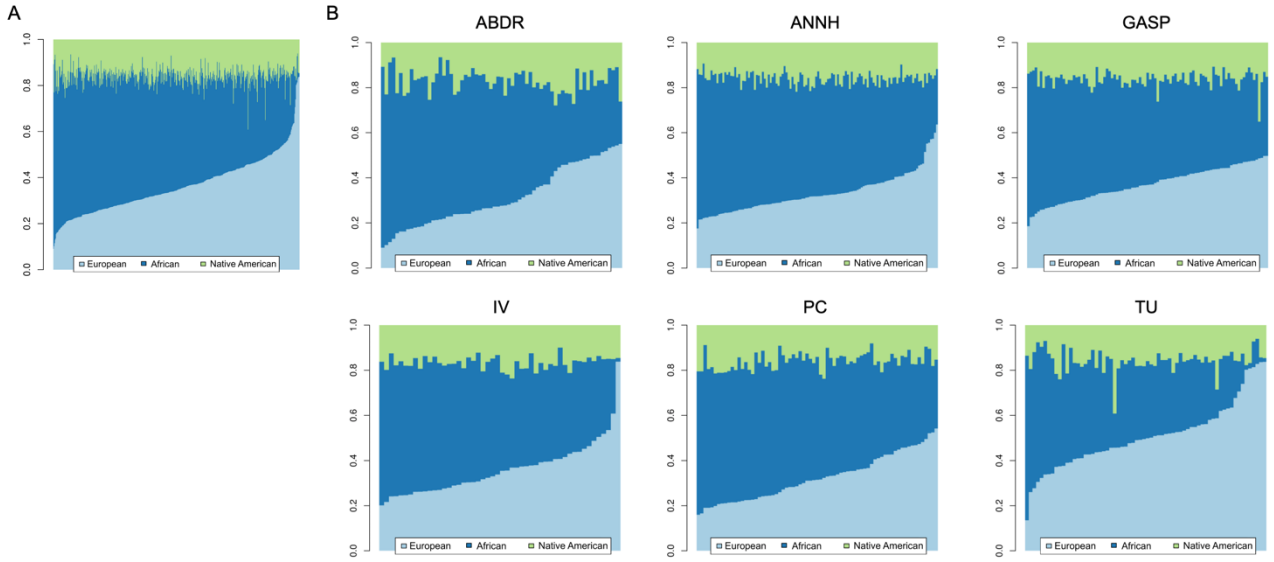

Figure 4. Graphic representation of global ancestry estimates. A) Ancestry estimates for the entire dataset (n = 431); B) Ancestry estimates stratified by pedigree. Colors represent ancestral contributions: green (Native American), dark blue (African), and light blue (European). Pedigrees include Abobral (ABDR), André Lopes and Nhunguara (ANNH), Galvão and São Pedro (GASP), Ivaporunduva (IV), Pedro Cubas (PC), and Sapatu (TU).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

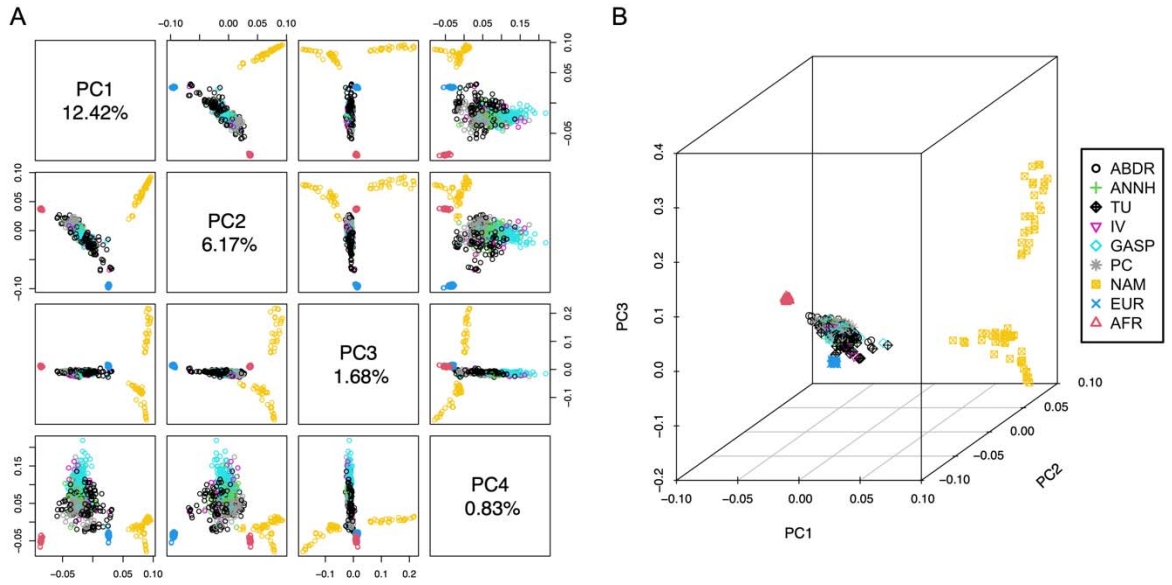

Figure 5. Principal Component Analysis (PCA) of Genetic Distance. A) PCA results for the first four principal components (PCs); B) Three-dimensional PCA representation. Reference samples from European (EUR), African (AFR), and Native American (NAM) populations are depicted in blue, red, and yellow, respectively. Pedigrees include Abobral (ABDR), André Lopes and Nhunguara (ANNH), Galvão and São Pedro (GASP), Ivaporunduva (IV), Pedro Cubas (PC), and Sapatu (TU).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 4. Key Genes Identified Across Analysis Strategies. Main genes identified from the three analysis strategies—linkage analysis, EH-related gene investigation (EH), and association studies (AS)—for each region of interest (ROI).

Gene	Cytoband	ROI	EH	AS
ALPK2	18q21.31-q21.32	20	●	●
EDARADD	1q42.3	3	●	●
KCNT1	9q34.3	11	●	●
LPP	3q27.3-q28	5	●	●
MFN2	1p36.22	2	●	●
MTR	1q43	3	●	●
P2RX1	17p13.2	18	●	●
PHGDH	1p12	1	●	●
RPA1	17p13.3	18	●	●
RYR2	1q43	3	●	●
S100A10	1q21.3	1	●	●
SERTAD2	2p14	4	●	●
TENM4	11q14.1	13	●	●
ZZEF1	17p13.2	18	●	●
ABCC9	12p12.1	14	●	
ACE	17q23.3	19	●	
ADIPOQ	3q27.3	5	●	
APOE	19q13.32	21	●	
ARHGEF17	11q13.4	13	●	
BORCS7	10q24.32	12	●	
CASP3	4q35.1	7	●	
CDH13	16q23.3	17	●	
CNNM2	10q24.32	12	●	
CORIN	4p12	6	●	
CYGB	17q25.1	19	●	
CYP17A1	10q24.32	12	●	
CYP2C19	10q23.33	12	●	
CYP2C9	10q23.33	12	●	
EDN1	6p24.1	9	●	
HTR2A	13q14.2	16	●	
KDR	4p12	6	●	
KIT	4p12	6	●	
KLKB1	4q35.2	7	●	
KNG1	3q27.3	5	●	
LPL	8p21.3	10	●	
MC4R	18q21.32	20	●	
MEIS1	2p14	4	●	
NT5C2	10q24.32-q24.33	12	●	
PDE3A	12p12.2	14	●	
PDE4D	5q11.2-q12.1	8	●	
PHACTR1	6p24.1	9	●	
PLCE1	10q23.33	12	●	
PRKCA	17q24.2	19	●	
RBM47	4p14	6	●	
RBP4	10q23.33	12	●	
SGCZ	8p22	10	●	
SLC1A4	2p14	4	●	
SLC39A14	8p21.3	10	●	
SOCS3	17q25.3	19	●	
SST	3q27.3	5	●	
TBX2	17q23.2	19	●	
TGFB1	19q13.2	21	●	
TIMP2	17q25.3	19	●	
TLR3	4q35.1	7	●	
TMOD4	1q21.3	1	●	
TNFRSF1B	1p36.22	2	●	
UCP2	11q13.4	13	●	
VEGFC	4q34.3	7	●	
WBP1L	10q24.32	12	●	
ZDHHC2	8p22	10	●	

medRxiv preprint doi: https://doi.org/10.1101/2024.06.26.24309531; this version posted February 4, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

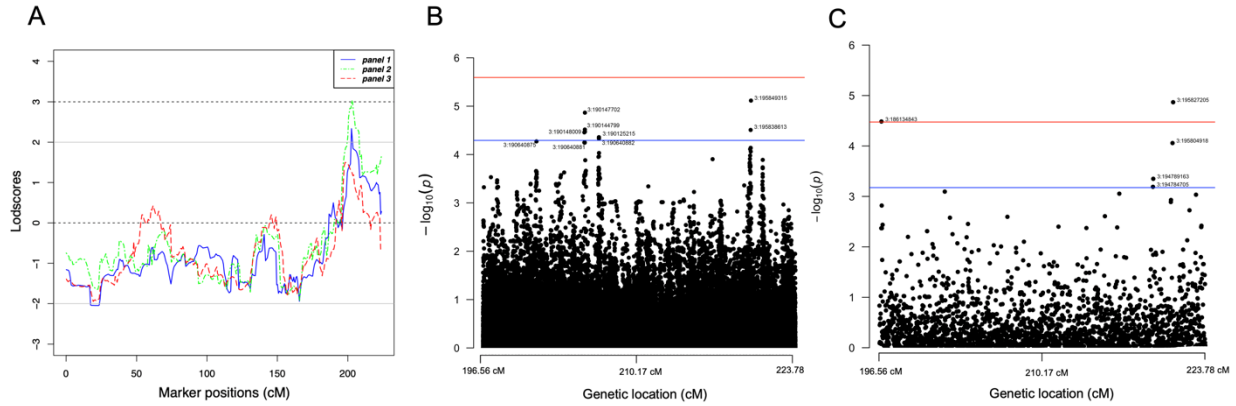

Figure 6. Illustration of the results for ROI5. A) Dense linkage analysis across the three subpanels; B) Association study using population-based imputed data; C) Association study using pedigree-based imputed data. Black dots represent SNPs (labeled by chromosome and physical position), blue lines indicate the suggestive *p*-value threshold (-log10(p)) and the red lines to the significant *p*-value threshold (-log10(p)).

Table 5. Ranked Score-Weighted Regions of Interest (ROIs). The ROIs are classified into three tiers according to their priority level: the top 20% are labeled as high priority (dark gray), the medium 30% as intermediate priority (medium gray), and the bottom 50% as low priority (light gray).

ROI	Cytoband	ROI size (Mb)	EH Genes	Linkage Analysis		Assoc. Studies Sugg./Sig.	Gene Relevance	TOTAL
				Peak LOD Score	Subpanels consensus			
5	3q27.3-q29	11.4 Mb	Yes | 4 genes (+1p)	3.036 (+6p)	3 (+3p)	Sig. (+3p)	Yes (+5p)	18
12	10q23.33-q25.1	11.2 Mb	Yes | 9 genes (+1p)	2.795 (+5p)	3 (+3p)	Sig. (+3p)	Yes (+5p)	17
13	11q13.4-q14.1	7.6 Mb	Yes | 3 genes (+1p)	2.414 (+4p)	3 (+3p)	Sugg. (+1p)	Yes (+5p)	14
19	17q23.2-25.3	19 Mb	Yes | 9 genes (+1p)	2.138 (+3p)	3 (+3p)	Sugg. (+2p)	Yes (+5p)	14
3	1q43	1.3 Mb	Yes | 3 genes (+1p)	2.161 (+3p)	2 (+2p)	Sugg. (+2p)	Yes (+5p)	13
9	6p24.1-p22.3	8 Mb	Yes | 2 genes (+1p)	2.621 (+5p)	2 (+2p)	Sugg. (+2p)	Yes (+3p)	13
10	8p23.1-p21.3	9.8 Mb	Yes | 4 genes (+1p)	2.658 (+5p)	3 (+3p)	Sugg. (+1p)	Yes (+3p)	13
7	4q32.3-q35.2	20.6 Mb	Yes | 4 genes (+1p)	2.322 (+4p)	2 (+2p)	Sugg. (+2p)	Yes (+3p)	12
21	19q13.12-13.32	10 Mb	Yes | 2 genes (+1p)	2.503 (+4p)	3 (+3p)	Sugg. (+1p)	Yes (+3p)	12
14	12p12.3-p11.23	11.1 Mb	Yes | 2 genes (+1p)	2.662 (+5p)	2 (+2p)	Sugg. (+1p)	Yes (+2p)	11
8	5q12.1-q13.2	11.1 Mb	Yes | 1 gene (+1p)	2.117 (+3p)	3 (+3p)	Sugg. (+2p)	Yes (+2p)	11
18	17p13.3-p13.2	4.3 Mb	Yes | 3 genes (+1p)	1.771 (+2p)	2 (+2p)	Sugg. (+1p)	Yes (+4p)	10
2	1p36.21-p36.22	3.2 Mb	Yes | 2 genes (+1p)	1.824 (+2p)	3 (+3p)	Sugg. (+1p)	Yes (+3p)	10
16	13q14.13-q21.33	25.6 Mb	Yes | 1 gene (+1p)	2.554 (+4p)	3 (+3p)	No (0p)	Yes (+2p)	10
1	1p12-q21.3	32.5 Mb	Yes | 3 genes (+1p)	1.503 (+1p)	2 (+2p)	Sugg. (+1p)	Yes (+4p)	9
4	2p14	3.4 Mb	Yes | 3 genes (+1p)	1.906 (+2p)	2 (+2p)	Sugg. (+1p)	Yes (+3p)	9
20	18q21.32	1.7 Mb	Yes | 2 genes (+1p)	1.608 (+1p)	3 (+3p)	Sugg. (+1p)	Yes (+3p)	9
6	4p15.1-q12	26.5 Mb	Yes | 4 genes (+1p)	1.938 (+2p)	3 (+3p)	No (0p)	Yes (+3p)	9
11	9q34.3	0.8 Mb	Yes | 1 gene (+1p)	1.418 (+1p)	3 (+3p)	Sugg. (+1p)	Yes (+2p)	8
17	16q23.3-q24.1	4.5 Mb	Yes | 1 gene (+1p)	2.175 (+3p)	2 (+2p)	No (0p)	Yes (+2p)	8
15	12q24.32-q24.33	4.6 Mb	No (0p)	2.087 (+3p)	2 (+2p)	Sugg. (+1p)	Yes (+1p)	7
22	19q13.41-13.42	3.2 Mb	No (0p)	1.860 (+2p)	3 (+3p)	Sugg. (+1p)	Yes (+1p)	7