**FACULDADE DE MEDICINA DA
UNIVERSIDADE DE SÃO PAULO**

MARI MAKI SÍRIA GODOY CARDENA

**MAPEAMENTO POR MISCIGENAÇÃO EM AMOSTRA DE
PACIENTES BRASILEIROS COM INSUFICIÊNCIA
CARDÍACA**

**São Paulo
2018**

MARI MAKI SÍRIA GODOY CARDENA

**Mapeamento por miscigenação em amostra de pacientes
brasileiros com insuficiência cardíaca**

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para obtenção do título
de Doutor em Ciências

Programa de Fisiopatologia Experimental
Orientadora: Prof^a. Dr^a. Cintia Fridman Rave

(Versão corrigida. Resolução CoPGr 6018, de 03 de outubro de 2011.
A versão original está disponível na Biblioteca FMUSP)

**São Paulo
2018**

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Cardena, Mari Maki Síria Godoy
Mapeamento por miscigenação em amostra de
pacientes brasileiros com insuficiência cardíaca /
Mari Maki Síria Godoy Cardena. -- São Paulo, 2018.
Tese(doutorado)--Faculdade de Medicina da
Universidade de São Paulo.
Programa de Fisiopatologia Experimental.
Orientadora: Cintia Fridman Rave.

Descritores: 1.Insuficiência cardíaca
2.Mortalidade 3.Marcadores genéticos 4.Hipertensão
5.Brasil 6.Mapeamento cromossômico

USP/FM/DBD-220/18

Responsável: Kátia Maria Bruno Ferreira - CRB-8/6008

Nome: CARDENA, Mari Maki Síria Godoy

Título: Mapeamento por miscigenação em amostra de pacientes brasileiros com insuficiência cardíaca.

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para obtenção do
título de Doutor em Ciências.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Dedico esta tese...

**Ao meu marido Leandro, a minha mãe Elzi e meu
Mundo Marribe.**

**Pelo amor, incentivo, compreensão e incansável apoio
ao longo do período de elaboração dessa tese.**

Essa conquista também é de vocês!

AGRADECIMENTOS

Primeiro de tudo, gostaria de agradecer a Deus por me guiar, iluminar e me dar tranquilidade para seguir em frente com os meus objetivos e não desanimar com as dificuldades;

À Prof^a. Dr^a. Cintia, pela atenção e apoio durante o processo de definição, pela paciência, compreensão e orientação;

Ao Dr Alexandre da Costa Pereira pela confiança e por acreditar nesse trabalho;

À Hadassa, Andréa, Marília, Mateus, Tiago e ao prof. Dr. Eduardo Tarazona por todo apoio e ajuda na programação das análises dessa tese. Sem vocês seria muito mais difícil;

Ao meu amado marido, Leandro Ferreira por ser tão importante na minha vida. Sempre a meu lado, me pondo para cima e me fazendo acreditar que posso mais que imagino. Graças ao seu companheirismo, amizade, paciência, compreensão, apoio, alegria e amor, este trabalho pôde ser concretizado;

À minha mais que amada mãe Elzi e irmã Marribe (Mundo), meu infinito agradecimento. Sempre acreditaram em minha capacidade e me acharam A MELHOR de todas, mesmo não sendo. Isso só me fortaleceu e me fez tentar não ser A MELHOR, mas a fazer o melhor de mim. Obrigada pelo amor incondicional!

Ao meu pai Silvio, minhas cunhadas Rosana e Claudia, e minha sogra Dona Cida pelo incentivo, carinho e atenção;

Aos meus irmãos Riber, Macrelli, seus companheiros, meu cunhado Anésio e meu padrasto Aauto que mesmo de longe sempre se mostraram interessados e presentes em meus estudos, sempre me dando forças para continuar;

Aos meus sobrinhos Marcello, Sara e Tarlei pelo carinho que nunca faltou;

As minhas amigas Felícia, Marina, Natalie, Liliane e Fernanda, pela amizade sincera, pelos ombros amigos e incentivo ao longo destes anos de estudos;

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de doutorado e pelo apoio financeiro para a realização desta pesquisa. Processo nº 2014/07348-5, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP);

À Faculdade de Medicina da Universidade de São Paulo, ao Programa de Fisiopatologia Experimental e ao LIM 40 do Departamento de Medicina Legal, Ética Médica, Medicina Social e do Trabalho pela disponibilização do espaço e pelo apoio concedido para a realização desse estudo;

À todas as pessoas que contribuíram de alguma forma para a realização deste trabalho.

MUITO OBRIGADA!

RESUMO

Cardena MMSG; *Mapeamento por miscigenação em amostra de pacientes brasileiros com insuficiência cardíaca* [tese]; São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2018.

As doenças cardiovasculares lideram as causas de morte em vários países, inclusive no Brasil, sendo a insuficiência cardíaca (IC) uma das enfermidades mais frequentes. Associações entre ancestralidade genética e susceptibilidade ao desenvolvimento da IC já foram relatadas em diferentes populações. O presente estudo teve como objetivo estimar as ancestralidades global e local de 492 pacientes com IC, identificar regiões e ancestralidades genômicas associadas à IC, seus fatores de risco (diabetes e hipertensão) e à mortalidade causada pela IC, por meio do mapeamento por miscigenação (AM, *admixture mapping*). Utilizando 182.090 Polimorfismos de Nucleotídeo Único (SNPs, *Single Nucleotide Polymorphisms*) comuns à nossa população e a três populações ancestrais (europeia, africana e ameríndia) foram realizadas as análises das ancestralidades global e local. Todos os pacientes apresentaram maior proporção de ancestralidade global europeia, média de $0,618 \pm 0,218$. As análises de AM da IC e da diabetes não mostraram nenhuma região associada, nas três ancestralidades avaliadas. No estudo de AM da hipertensão houve associação estatisticamente significativa com a ancestralidade local africana no cromossomo 2 (chr2p25.3) ($p=4,65 \times 10^{-5}$). Nessa região estão mapeados 7 RNAs não codificantes, 2 RNAs longos de região intergênica não codificadores de proteína, 44,93% do gene *Syntrophin Gamma 2* (*SNTG2*) e o gene que codifica uma proteína de membrana (*TMEM18*). A análise de AM da mortalidade por IC foi realizada com os mesmos 492 pacientes com IC, usando o modelo caso-controle, sendo o grupo de casos ($n=248$) composto por pacientes em óbito no final de 4 anos de avaliação, e o grupo de controles ($n=244$), de indivíduos que ainda estavam vivos no final desse mesmo período. Foi observada associação estatisticamente significativa com a ancestralidade local europeia no cromossomo 6 (chr6p22.3) ($p=6,805 \times 10^{-5}$). Nessa região estão mapeados 30,74% do gene *Ataxina 1* (*ATXN1*) e o gene *Guanosina Monofosfato Redutase* (*GMPR*). O mapeamento fino nessa região, com 7.916 SNPs, apresentou dois marcadores genéticos, rs1042391 e rs2142672, com resultados significativos ($p=1,140 \times 10^{-4}$, $p=3,921 \times 10^{-4}$, respectivamente). O alelo em homozigose TT do rs1042391 foi associado a um aumento de 21% ao risco de morte por IC (HR=1.21, $p=0,0013$), enquanto que o alelo em homozigose CC do rs2142672 foi associado a um aumento de 51% ao risco de morte por IC (HR=1.51, $p=0,0004$). Estes são achados iniciais e, como tal, devem ser considerados como hipóteses geradoras. Apesar disso, ficou demonstrado a utilidade do estudo de AM para identificar regiões com diferentes ancestralidades genômicas que podem contribuir para o risco de doenças complexas em populações geneticamente miscigenadas. Esses dados podem auxiliar na compreensão da etiologia genética da hipertensão e da mortalidade em pacientes com IC, relacionados à ancestralidade, podendo servir como ponto de partida para estudos funcionais na tentativa de aprofundar os conhecimentos dos processos biológicos que levam à hipertensão e à IC. Estudos futuros são necessários para replicar as associações encontradas, detectar variáveis causais que conduzem essas associações e explorar aplicações clínicas para as regiões de genes consistentemente associadas a mortalidade em pacientes com IC e à hipertensão.

Descritores: insuficiência cardíaca; mortalidade; marcadores genéticos; hipertensão; Brasil; mapeamento cromossômico.

ABSTRACT

Cardena MMSG; *Admixture mapping in sample of Brazilian patients with heart failure* [thesis]; São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2018.

Cardiovascular diseases lead the causes of death in several countries, including Brazil, with the heart failure (HF) being one of the most frequent diseases. Associations between genetic ancestry and susceptibility to the development HF have already been reported in different populations. The present study aimed to estimate the global and local ancestry of 492 HF patients, and identify genomic regions and ancestry associated with HF, HF risk factors (diabetes and hypertension) and HF mortality, through admixture mapping (AM). Using 182,090 Single Nucleotide Polymorphisms (SNPs) common to our population and to three ancestral populations (European, African and Amerindian) the analyses of global and local ancestry were performed. All patients showed a higher proportion of European global ancestry, mean of 0.618 ± 0.218 . The AM analysis of HF and diabetes did not show any associated regions, in the three ancestries evaluated. In AM of hypertension there was a statistically significant association with African local ancestry on chromosome 2 (chr2p25.3) ($p=4.65 \times 10^{-5}$). In this region are mapped 7 non-coding RNAs, 2 long intergenic non-protein coding RNAs, 44.93% of the *Syntrophin Gamma 2* gene (*SNTG2*) and the gene encoding a transmembrane protein (*TMEM18*). The AM analysis of HF mortality was performed with the same 492 HF patients, using case-control model, case group ($n=248$) was composed of patients who died at the end of 4 years of evaluation, and control group ($n=244$) included individuals who were still alive at the end of that period. A statistically significant association with European local ancestry on chromosome 6 (chr6p22.3) ($p=6.805 \times 10^{-5}$) was observed. In this region are mapped 30.74% of the gene *Ataxin 1* (*ATXN1*) and the *Guanosine Monophosphate Redutase* gene (*GMPR*). The fine mapping in this region, with 7,916 SNPs, presented two genetic markers, rs1042391 and rs2142672, with significant results ($p=1.140 \times 10^{-4}$, $p=3.921 \times 10^{-4}$, respectively). The TT homozygous allele of rs1042391 was associated with 21% increase risk of HF death ($HR=1.21$, $p=0.0013$), while the CC homozygous allele of rs2142672 was associated with 51% increase risk of HF death ($HR=1.51$, $p=0.0004$). These are initial findings and, as such, should be considered as generating hypotheses. Despite this, it was demonstrated the utility of the AM study to identify regions with different genomic ancestry that may contribute to the risk of complex diseases in genetically mixed populations. These data may contribute to understand the genetic etiology of hypertension and mortality in HF patients, related to ancestry, and may serve as a starting point for functional studies in an attempt to deepen the knowledge of the biological processes that lead to hypertension and HF. Future studies are needed to replicate the associations found, to detect causal variables that lead these associations, and to explore clinical applications for gene regions consistently associated with death in patients with HF and hypertension.

Descriptors: Heart Failure; Mortality; Genetic Markers; Hypertension; Brazil; Chromosome Mapping.

ABREVIATURAS

AIMs	Marcadores Informativos de Ancestralidade (<i>Ancestry Informative Markers</i>)
AMP	Adenosina Monofosfato
ATXN1	Gene <i>Ataxina 1</i>
CAPPesq	Comissão de Ética para Análise de Projetos de Pesquisa do Hospital das Clínicas da FMUSP
CEU	Residentes de UTAH com ancestralidade da Europa do Norte e Ocidental
Chr	Cromossomo
DAC	Doenças do Aparelho Circulatório
DM2	Diabetes Mellitus tipo 2
DP	Desvio Padrão
GMP	Guanosina Monofosfato
GMPR	Gene <i>guanosina monofosfato redutase</i> (<i>Guanosine Monophosphate Reductase</i>)
GWAS	Estudos de Associação Genômica Ampla (<i>Genome-Wide Association Study</i>)
HDL	Lipoproteína de Alta Densidade (<i>High Density Lipoprotein</i>)
HGDP-CEPH	<i>Human Genome Diversity Cell Line Panel</i>
HMM	Modelo Oculto de Markov (<i>Hidden Markov Model</i>)
HR	<i>hazard ratio</i>
IC	Insuficiência Cardíaca
IMC	Índice de Massa Corporal
INDEL	Inserção-Deleção
LD	Desequilíbrio de Ligação (<i>Linkage Disequilibrium</i>)
LDGH-UFMG	Laboratório de Diversidade Genética Humana da Universidade Federal de Minas Gerais
lincRNA	RNAs longos de região intergênica não codificadores de proteína (<i>long intergenic non-protein coding RNA</i>)
LDL	Lipoproteínas de Baixa Densidade (<i>Low Density Lipoprotein</i>)
LOD	Logaritmo da Probabilidade (<i>Logarithm of Odds score</i>)

LWK	População do Quênia
MAF	Frequência do Alelo Menor (<i>Minor Allele Frequency</i>)
mg/dL	Miligramas por Decilitro
MCMC	Método de Monte Carlo via cadeias de Markov (<i>Markov chain Monte Carlo</i>)
mtDNA	DNA mitocondrial
ncRNA	RNAs não codificantes (<i>non-coding RNA</i>)
OR	<i>Odds Ratio</i>
pb	Pares de base
PCA	Análise de Componentes Principais
SNP	Polimorfismo de Nucleotídeo Único (<i>Single Nucleotide Polymorphism</i>)
<i>SNTG2</i>	Gene <i>Syntrophin Gamma 2</i>
STR	Sequências Curtas Repetidas em Tandem (<i>Short Tandem Repeats</i>)
SUS	Sistema Único de Saúde
<i>TMEM18</i>	Gene <i>Transmembrane Protein 18</i>
TSI	População Toscana da Itália
VLDL	Lipoproteína de Muito Baixa Densidade (<i>Very Low Density Lipoprotein</i>)
VNTR	Repetições em Tandem de Número Variável (<i>Variable Number of Tandem Repeats</i>)
YRI	População da Nigéria
95% IC	Intervalo de confiança de 95%

LISTA TABELA

Tabela 1: Características gerais dos 492 pacientes com IC avaliados nesse estudo.	31
Tabela 2: Análise da associação entre a variável sexo e as variáveis hipertensão e diabetes.....	32
Tabela 3: Médias das ancestralidades genômicas em dois dos nossos estudos.	33
Tabela 4: Quantidade de SNPs utilizados na análise de ancestralidade local.	35
Tabela 5: Resultado do <i>admixture mapping</i> da hipertensão em pacientes com IC na ancestralidade africana.	50
Tabela 6: Lista dos 25 SNPs presentes na janela que apresentou diferença estatística.	51
Tabela 7: Análise das etiologias nos dados de mortalidade por IC.....	55
Tabela 8: Análise das ancestralidades globais do grupo de casos e controles.	56
Tabela 9: Resultado do <i>admixture mapping</i> caso-controle da ancestralidade europeia dos dados de mortalidade dos pacientes com IC.....	60
Tabela 10: Lista dos 25 SNPs presentes na janela que apresentou diferença estatística.	61
Tabela 11: SNPs que apresentaram diferenças estatisticamente significante no mapeamento fino da região chr6p22.3 na ancestralidade europeia em pacientes com IC do dado de mortalidade.	63
Tabela 12: Análise associação entre SNPs e o risco de morte por IC.....	64

LISTA FIGURA

Figura 1: Diagrama do processo de mistura.....	9
Figura 2: Ilustração de como um <i>locus</i> de susceptibilidade à determinada doença poderá ser visualizado em uma análise de <i>admixture mapping</i> em um estudo caso-controle para população miscigenada di-híbrida	10
Figura 3: Representação do gráfico Manhattan com os resultados do <i>admixture mapping</i> em estudo somente de casos com câncer de mama	13
Figura 4: Representação dos resultados do <i>admixture mapping</i> em estudo de casos e caso-controle com hipertensão em relação à ancestralidade africana e o calibre vascular da retina	14
Figura 5: Localização no cromossomo 2 da janela que apresentou diferença estatística no <i>admixture mapping</i> da ancestralidade africana para hipertensão em pacientes com insuficiência cardíaca.	51
Figura 6: Localização no cromossomo 6 da janela que apresentou diferença estatisticamente significativa no <i>admixture mapping</i> da ancestralidade europeia para mortalidade em pacientes com insuficiência cardíaca	61

LISTA GRÁFICO

Gráfico 1: Resultado do <i>admixture mapping</i> da insuficiência cardíaca na ancestralidade europeia.	37
Gráfico 2: Resultado do <i>admixture mapping</i> da insuficiência cardíaca na ancestralidade ameríndia.	38
Gráfico 3: Resultado do <i>admixture mapping</i> da insuficiência cardíaca na ancestralidade africana.	39
Gráfico 4: Resultado do <i>admixture mapping</i> da diabetes em pacientes com insuficiência cardíaca na ancestralidade europeia.	42
Gráfico 5: Resultado do <i>admixture mapping</i> da diabetes em pacientes com insuficiência cardíaca na ancestralidade ameríndia.	43
Gráfico 6: Resultado do <i>admixture mapping</i> da diabetes em pacientes com insuficiência cardíaca na ancestralidade africana.	44
Gráfico 7: Resultado do <i>admixture mapping</i> da hipertensão em pacientes com insuficiência cardíaca na ancestralidade europeia.	47
Gráfico 8: Resultado do <i>admixture mapping</i> da hipertensão em pacientes com insuficiência cardíaca na ancestralidade ameríndia.	48
Gráfico 9: Resultado do <i>admixture mapping</i> da hipertensão em pacientes com insuficiência cardíaca na ancestralidade africana.	49
Gráfico 10: Mapeamento fino com 7.682 SNPs da região chr2p25.3 na ancestralidade africana.	54
Gráfico 11: Resultado do <i>admixture mapping</i> da mortalidade em pacientes com insuficiência cardíaca na ancestralidade europeia.	57
Gráfico 12: Resultado do <i>admixture mapping</i> da mortalidade em pacientes com insuficiência cardíaca na ancestralidade ameríndia.	58
Gráfico 13: Resultado do <i>admixture mapping</i> da mortalidade em pacientes com insuficiência cardíaca na ancestralidade africana.	59
Gráfico 14: <i>Admixture mapping</i> das análises caso-controle e somente de casos da ancestralidade europeia para a mortalidade em pacientes com IC.	61
Gráfico 15: Mapeamento fino com 7.916 SNPs da região chr6p22.3 na ancestralidade europeia.	63

SUMÁRIO

RESUMO

ABSTRACT

1	INTRODUÇÃO	13
1.1	População Brasileira	2
1.2	Insuficiência Cardíaca.....	4
1.3	Insuficiência Cardíaca e Ancestralidade Genética.....	6
1.4	<i>Admixture Mapping</i>	7
2	OBJETIVOS	18
2.1	Objetivo Geral.....	18
2.2	Objetivos Específicos.....	18
3	CASUÍSTICA E MÉTODOS.....	20
3.1	Casuística	20
3.2	Genotipagem	21
3.3	Limpeza de Dados	21
3.4	População Ancestral	22
3.5	Ancestralidade Global.....	23
3.6	Ancestralidade Local e <i>Admixture Mapping</i>	24
3.7	Mapeamento Fino.....	27
3.8	Fluxograma dos Métodos	29
4	RESULTADOS E DISCUSSÃO	30
4.1	Caracterização Clínica dos Pacientes	31
4.2	Ancestralidade Global.....	32
4.3	Ancestralidade Local e <i>Admixture Mapping</i>	34
4.3.1	Insuficiência Cardíaca	36
4.3.2	Diabetes	41
4.3.3	Hipertensão.....	46
4.3.4	Mortalidade.....	55
5	CONCLUSÃO.....	70
6	REFERÊNCIAS	73
	ANEXO 1.....	87
	ANEXO 2.....	97

1 INTRODUÇÃO

1.1 População Brasileira

O processo de miscigenação ocorre quando indivíduos de duas ou mais populações geograficamente isoladas, com diferentes frequências alélicas, misturam-se e formam uma nova população (Mersha; Abebe, 2015; Reed, 1969). Esse fluxo gênico ao longo do tempo entre populações humanas geneticamente distintas, resulta em populações miscigenadas um mosaico de segmentos cromossômicos que individualmente representam cada uma das populações ancestrais (Hellenthal et al., 2014; Mersha; Abebe, 2015).

As populações nativas da América Latina apresentam diferentes graus de mistura genética devido a diferentes processos históricos que ocorreram desde o final do século 15, deixando vestígios genéticos, principalmente europeu e africano, nos genomas desses indivíduos (Salzano; Sans, 2014). A população brasileira não é uma exceção a esse padrão.

A ancestralidade genética original do Brasil é composta por nativos americanos (ameríndios) (Perego et al., 2010). No início do século 16, teve início o processo de miscigenação da população a partir da colonização do Brasil, quando os primeiros colonos portugueses (cerca de meio milhão) começaram a misturar-se com as populações indígenas (cerca de 2,5 milhões). O aumento dessa miscigenação se deu com a chegada dos escravos africanos (cerca de 4 milhões) em meados do século 16, e continuou após a criação da República do Brasil, no século 19, com um grande fluxo de imigrantes para o Brasil, principalmente de italianos, alemães e japoneses (IBGE, 2007). Essa mistura genética é diretamente influenciada pelo processo de colonização, resultando no Brasil com uma população geneticamente de ancestralidade tri-híbrida: europeia, africana e ameríndia (Cardena et al., 2013; Leite et al., 2011; Moura et al., 2015).

Essa herança genômica pode ser avaliada mediante análise de cromossomos autossômicos, cromossomos sexuais e análise do DNA mitocondrial (mtDNA) (Alves-Silva et al., 2000; Manta et al., 2013; Palha et al., 2012).

Atualmente, o Brasil possui cinco regiões sociogeográficas (Centro-Oeste, Norte, Nordeste, Sul e Sudeste) que são reconhecidas pelos censos oficiais. Estudos em quatro dessas regiões (com exceção do Centro-Oeste), por meio de análises do

cromossomo Y evidenciaram que a grande maioria da ancestralidade paterna, independente da região do país, é de origem europeia (Cardena et al., 2016; Carvalho-Silva et al., 2001; Carvalho-Silva et al., 2006; Resque et al., 2016; Simão et al., 2018).

Por outro lado, análises das regiões hipervariáveis do mtDNA revelaram uma realidade diferente: considerando o Brasil como um todo, 39% da ancestralidade materna é de origem europeia, 33% de origem ameríndia e 28% africana (Alves-Silva et al., 2000; Fridman et al., 2014; Schaan et al., 2017). Significativamente, a frequência de ancestralidade do mtDNA variou muito nas diferentes regiões: a maior porcentagem de ancestralidade materna ameríndia (54%) esteve presente na região Norte, enquanto que a ancestralidade africana é maior no Nordeste (44%) e a ancestralidade europeia é predominante no Sul (66%) (Alves-Silva et al., 2000).

Juntos, esses dados revelam que a contribuição genética europeia foi basicamente por meio dos homens, enquanto que a contribuição genética africana e ameríndia aconteceu, principalmente, pelas mulheres (Alves-Silva et al., 2000; Cardena et al., 2016; Carvalho-Silva et al., 2001; Carvalho-Silva et al., 2006; Fridman et al., 2014; Pena et al., 2011; Resque et al., 2016; Schaan et al., 2017; Simão et al., 2018).

Muitos estudos foram realizados na população brasileira para estimar a ancestralidade autossômica em diversas regiões do país. Em 2015, Moura e colaboradores, realizaram um estudo de revisão sistemática e meta-análise utilizando dados de 25 trabalhos de 38 populações brasileiras de diferentes regiões, utilizando somente amostras populacionais de indivíduos saudáveis e mostraram que a contribuição genética autossômica europeia predominou nas regiões Sudeste e Sul (média de 77%), enquanto a contribuição africana foi mais observada no Nordeste (média de 27%) e a ameríndia, na região Norte (média de 32%) (Moura et al., 2015).

Se avaliarmos os 25 estudos considerados por Moura et al. (2015) é possível observar discrepâncias nos valores das contribuições de cada ancestralidade autossômica apresentados nas diferentes regiões do país, provavelmente devido aos diferentes processos de colonização e migração ocorridos em cada região.

Contudo, se considerarmos os dados das análises do cromossomo Y, mtDNA e ancestralidade autossômica podemos evidenciar a alta miscigenação da população brasileira. Ao utilizar os dados desses três métodos de análise de ancestralidade

torna-se possível descrever a real estrutura étnica e fazer uma qualificação mais apropriada de populações muito miscigenadas, como a população brasileira.

Com populações melhor caracterizadas etnicamente, utilizando essas análises, é possível evidenciar o efeito de estratificação populacional em estudos de associação genética, principalmente em populações miscigenadas, e também avaliar a influência das variações étnicas no desenvolvimento de doenças, identificando quais ancestralidades estariam relacionadas ao risco aumentado e/ou diminuído para diversas enfermidades, possibilitando melhor conhecimento sobre a origem de diversas doenças complexas, incluindo a insuficiência cardíaca (IC) (Enoch et al., 2006; Leite et al., 2011; McKeigue et al., 2000, 2005; Patterson et al., 2004; Winkler; Nelson; Smith, 2010).

1.2 Insuficiência Cardíaca

A insuficiência cardíaca (IC) é uma síndrome complexa, de caráter progressivo e irreversível (Hunt et al., 2005). Diversas condições podem causar a IC, como: doença arterial coronariana, hipertensão arterial, diabetes, cardiomiopatias, valvulopatias, doenças congênitas do coração, arritmias, doenças do pericárdio, miocardites, hipertensão pulmonar e substâncias cardiotoxícas, incluindo álcool (Mosterd; Hoes, 2007), sendo que a cardiomiopatia isquêmica é a etiologia predominante no Brasil, seguida pelas cardiomiopatias idiopática dilatada, hipertensiva, valvular e chagásica (Bocchi et al., 2009; Bocchi, 2013).

Embora a IC não poupe crianças e jovens, essa enfermidade acomete principalmente indivíduos acima de 65 anos de idade, que respondem por mais de 80% do total de internações por essa doença no Brasil (Bocchi et al., 2009; Bocchi, 2013). Os avanços científicos e tecnológicos e as melhores condições socioeconômicas têm possibilitado o aumento da longevidade da população em geral, porém têm acarretado um aumento da incidência e prevalência da IC no mundo (Davenport et al., 2006).

As Doenças do Aparelho Circulatório (DAC) são as principais causas de internações no Sistema Único de Saúde (SUS) no Brasil. Em 2017, ocorreram 1.126.717 internações por DAC no SUS, 18,4% (n=207.587) foram decorrentes de IC,

e nesse mesmo ano as DAC foram responsáveis por 92.568 mortes ocorridas no país, sendo que a IC foi responsável por 24,8% (n=23.001) dessas mortes (DATASUS, 2017).

Estudos clínicos apontam que a incidência e a prevalência da IC podem ser díspares entre indivíduos com diferentes cores de pele (Bahrami et al., 2008; Bibbins-Domingo et al., 2009; Pereira et al., 2013; Taylor et al., 2014). Indivíduos negros, além de serem mais susceptíveis ao desenvolvimento da IC, apresentam maior mortalidade por essa enfermidade e também desenvolvem a IC em idades mais jovens. Isso se deve ao risco aumentado que esse grupo apresenta para desenvolver hipertensão, obesidade, doenças renais e diabetes, que são fatores de risco para a IC (Bahrami et al., 2008; Bibbins-Domingo et al., 2009; Gordon et al., 2010; Hunt et al., 2005; Okin et al., 2011).

Além desses dados clínicos, estudos genéticos que compararam pacientes descendentes de europeus com os pacientes afro-americanos apontaram a existência de diferenças nas frequências de alelos que codificam componentes importantes de células de sinalização neuro-hormonal na IC, apontando para o dado da ancestralidade genética como um fator presente no desenvolvimento da IC em indivíduos de cor de pele negra (Pereira et al., 2013; Taylor et al., 2014; Yancy et al., 2001).

No entanto, informações isoladas das diferenças de cor de pele associadas a fatores importantes para o desenvolvimento da IC não podem explicar totalmente a maior incidência e prevalência da IC entre os indivíduos negros. Tem sido reforçada a hipótese de que parte das diferenças étnicas encontradas na prevalência e desenvolvimento de doenças cardiovasculares deve ser mediada por fatores genéticos que podem levar a um aumento da susceptibilidade ao desenvolvimento à IC e determinar a gravidade da doença, sendo um desses fatores genéticos a ancestralidade (Bernardez-Pereira et al., 2016), lembrando que ancestralidade genética não está, necessariamente, associada diretamente à cor de pele (Cardena et al., 2013).

A ancestralidade de um indivíduo pode afetar o sistema cardiovascular devido a influência de fatores genéticos. Essa influência genética é explicada pelo fato de que pessoas com ancestralidade genética similar possuem perfis genéticos semelhantes, pois compartilham regiões no genoma e, conseqüentemente, possuem variantes

comuns entre elas (Bernardez-Pereira et al., 2016; Franceschini et al., 2011). O fato de muitas dessas variantes serem relacionadas com o risco de doenças cardiovasculares, resulta que a similaridade de ancestralidade pode conferir semelhança ao risco de diversas doenças, inclusive a IC.

1.3 Insuficiência Cardíaca e Ancestralidade Genética

As doenças complexas, como a IC, são causadas por diversos fatores genéticos e ambientais e suas interações. Um dos fatores genéticos que tem sido analisado em alguns estudos é a ancestralidade (Bernardez-Pereira et al., 2016; Cardena et al., 2013, 2014, 2016; Gallardo et al., 2012). Com a estimativa de ancestralidade genética é possível avaliar a influência das variações relacionadas à ancestralidade no desenvolvimento da IC.

Em estudo anterior, realizado pelo nosso grupo, foram avaliados 503 pacientes com IC e 188 controles saudáveis em relação à ancestralidade genética, com o uso de 48 Marcadores Informativos de Ancestralidade (AIMs, do inglês *Ancestry Informative Markers*) do tipo Inserção-Deleção (INDEL), além dos haplogrupos mitocondriais, obtidos pela análise das 3 regiões hipervariáveis do DNA mitocondrial (HV1, HV2 e HV3) (Cardena, 2013; Cardena et al., 2014, 2016).

A comparação da ancestralidade dos haplogrupos mitocondriais entre os dois grupos revelou uma associação entre haplogrupos africanos e o risco aumentado para o desenvolvimento da IC, enquanto que os haplogrupos ameríndios foram associados a um menor risco (Cardena, 2013).

Já a análise da ancestralidade autossômica demonstrou que, nos casos com etiologia hipertensiva, a maior contribuição da ancestralidade genômica africana conferiu risco aumentado, enquanto que a maior contribuição de ancestralidade genômica europeia conferiu risco diminuído para o desenvolvimento da IC (Cardena et al., 2016).

Além disso, nosso estudo evidenciou que os pacientes com maior contribuição de ancestralidade genômica ameríndia apresentaram maior sobrevida por IC por um período de 4 anos (Cardena et al., 2014).

Em relação à ancestralidade materna, Gallardo e colaboradores (2012) avaliaram influência da ancestralidade mitocondrial no desenvolvimento de IC em 450 indivíduos que receberam transplante cardíaco, 248 doadores cardíacos e 206 controles saudáveis, e observaram associação do haplogrupo europeu H com o risco aumentado a IC.

Analizando a ancestralidade autossômica, Bernardez-Pereira e colaboradores (2016) encontraram a ancestralidade genômica africana associada a piores parâmetros de função diastólica, enquanto que a ancestralidade ameríndia se correlacionou com um padrão pior de contratilidade ventricular em 362 pacientes com IC.

Embora não possamos excluir a interação dos fatores ambientais, a dieta e fatores socioeconômicos, os dados apresentados na literatura fornecem apoio adicional para a hipótese de que as diferenças de ancestralidades genéticas (uniparental e biparental) são importantes a serem avaliadas em estudos epidemiológicos para melhor entendimento de doenças complexas, como a IC. As análises de ancestralidade uniparental compreendem, mtDNA e cromossomo Y, e a análise biparental é realizada utilizando marcadores genéticos de cromossomos autossômicos.

1.4 *Admixture Mapping*

Alguns marcadores genéticos autossômicos são denominados de Marcadores Informativos de Ancestralidade (AIMs) e são conceituados como marcadores moleculares que apresentam diferenças significativas de frequências alélicas entre duas ou mais populações distintas (Chakraborty et al., 1992; Parra et al., 2003; Shriver et al., 1997).

Com o uso desses marcadores genéticos é possível realizar uma melhor classificação da ancestralidade de populações e de indivíduos, pois os AIMs são capazes de estimar as origens geográficas dos ancestrais de um indivíduo, geralmente expressa como a proporção de sua ancestralidade que vem de diferentes regiões continentais (Ding et al., 2011). Por meio dos AIMs compara-se os polimorfismos presentes em um indivíduo com as frequências alélicas desses

polimorfismos em populações ancestrais pouco miscigenadas, previamente analisadas em populações parentais cuja história ancestral é bastante conhecida e que servirão de referência. Alguns dados dos genótipos de diversas populações ancestrais ficam disponíveis gratuitamente em bancos de dados genômicos (Ding et al., 2011). Os bancos de dados genômicos mais utilizados são os do projeto 1000 Genomas (<http://www.1000genomes.org>), do projeto *International HapMap* (<http://hapmap.ncbi.nlm.nih.gov/>) e do projeto *Human Genome Diversity Cell Line Panel* (HGDP-CEPH) (http://www.cephb.fr/en/hgdp_panel.php).

Além disso, os AIMS são usados para determinar a ancestralidade de genes relacionados com doenças ou fenótipos em populações miscigenadas, pois eles permitem definir a origem ancestral do haplótipo que está sendo avaliado (Mersha, 2015). Deste modo, é possível avaliar a influência das variações de ancestralidade no desenvolvimento de doenças (Leite et al., 2011; Mckeigue et al., 2000, 2005).

Atualmente, estudos epidemiológicos e populacionais têm utilizado uma nova estratégia, denominada mapeamento por miscigenação (*admixture mapping*) ou mapeamento por miscigenação em desequilíbrio de ligação (*mapping by admixture linkage disequilibrium*), que é baseada na análise de um conjunto destes marcadores autossômicos (AIMs) espaçados mais ou menos uniformemente ao longo de todo o genoma, para avaliar a associação da ancestralidade genômica e encontrar regiões específicas e/ou genes relacionados com doenças (Zhu; Ashley-Koch; Dunson, 2013).

O argumento que fundamenta o *admixture mapping* é que, quando duas ou mais populações com arquiteturas genéticas diferentes se misturam, criam-se longos segmentos de DNA (haplótipos) que possuem origens ancestrais distinguíveis, portanto, formando vários blocos de ancestralidade (Figura 1) (Mersha, 2015; Patterson et al., 2004).

Em populações recentemente miscigenadas, como afro-americanas ou latino-americanas, cada cromossomo se assemelha a um mosaico com blocos das diferentes ancestralidades que formaram essa população (Figura 1) (Zhu; Ashley-Koch; Dunson, 2013).

Figura 1: Diagrama do processo de mistura, evidenciando o mosaico com blocos de diferentes ancestralidades no cromossomo de um indivíduo com ancestralidade miscigenada. Figura modificada de Hellenthal et al., 2014.

Deste modo, o *admixture mapping* é um método de mapeamento de genes aplicado em populações miscigenadas para identificar segmentos genômicos (blocos) em que a proporção de uma ancestralidade é notavelmente mais elevada ou mais baixa do que a observada em outros lugares no genoma avaliado ou em amostras de um grupo controle (Chen et al., 2014; McKeigue, 2005; Patterson et al., 2004; Winkler; Nelson; Smith, 2010). Portanto, o estudo da ancestralidade nesse segmento genômico pode revelar risco aumentado ou diminuído para os fenótipos ou doenças (Patterson et al., 2004; Winkler; Nelson; Smith, 2010).

Assim sendo, quando se está avaliando fenótipos ou doenças com a técnica de *admixture mapping* é possível observar que nas regiões cromossômicas que contém as variantes que contribuem para o risco à doença ou ao fenótipo haverá uma super-representação da ancestralidade que apresenta maior proporção de alelos de risco (Figura 2) (Patterson et al., 2004; Winkler; Nelson; Smith, 2010).

A premissa do *admixture mapping* é de que a doença ou o fenótipo em questão apresenta incidência diferente nas populações ancestrais, e que este fenômeno é devido às diferenças genéticas entre estas populações.

Figura 2: Ilustração de como um *locus* de susceptibilidade à determinada doença poderá ser visualizado em uma análise de *admixture mapping* em um estudo caso-controle para população miscigenada di-híbrida. Suponhamos que uma determinada doença é herdada da população de ancestralidade europeia (verde escuro). A barra amarela indica a localização do gene que confere susceptibilidade à doença. (A) Na região do gene de susceptibilidade à doença, há um excesso de blocos da ancestralidade europeia entre os casos, aparecendo como um pico no gráfico da média da ancestralidade europeia para esse grupo. (B) Entre o grupo de controles, na região do gene de susceptibilidade à doença, aparece tanto blocos de ancestralidade europeia quanto blocos da outra ancestralidade avaliada (verde claro). Figura modificada de Winkler; Nelson; Smith, 2010.

É possível estimar as proporções de ancestralidade global e local de um indivíduo ou população. Para se obter a estimativa da ancestralidade global calcula-se a média das contribuições de cada população ancestral em todo o genoma do indivíduo. Para obter a estimativa da ancestralidade local calcula-se a média das contribuições de um segmento específico de um cromossomo; assim, um indivíduo pode ter 0, 1 ou 2 cópias de um alelo derivado de cada população ancestral (Hohman et al., 2016; Padhukasahasram, 2014; Thornton; Bermejo, 2014).

Para estudos de doenças ou fenótipos a técnica de *admixture mapping* primeiramente analisa-se as variantes polimórficas presentes no indivíduo que são

comuns às populações ancestrais escolhidas. Em seguida, são realizadas as estimativas de ancestralidade dos segmentos cromossômicos, obtendo-se os resultados de ancestralidade global e local do indivíduo. Com os dados de ancestralidade realiza-se os cálculos para identificação de segmentos genômicos em que uma ancestralidade é significativamente mais elevada ou mais baixa do que a observada na ancestralidade global do genoma avaliado ou em uma amostra de um grupo controle. Uma vez identificada a região genômica com a ancestralidade significativamente associada com a variável de interesse biomédico, faz-se uma busca em bases de dados informatizados que gera uma lista dos genes presentes nessa região ou de regiões não gênicas envolvidas na regulação da expressão gênica (ENCODE et al., 2007; McCarthy et al., 2008).

Uma das vantagens da técnica de *admixture mapping* é que ela pode ser usada sem a necessidade de haver um grupo controle em relação à doença ou ao fenótipo, visto que a escolha dos controles para muitas doenças complexas é extremamente complicada (Hoggart et al., 2004). Em estudos que utilizam caso-controle, o *admixture mapping* é baseado na comparação de ancestralidade local entre casos e controles, enquanto que estudos que utilizam somente casos são baseados na comparação da ancestralidade estimada em cada região (ancestralidade local) com a média da ancestralidade global desse mesmo grupo (Bensen et al., 2014; Shriner, 2013).

Em estudos somente de casos a associação é verdadeira entre a ancestralidade e a doença, se por exemplo, a população ancestral europeia carrega o alelo de risco genético para a doença em uma frequência maior, então os genomas dos pacientes irão apresentar maior proporção da contribuição de ancestralidade europeia no *locus* de susceptibilidade à doença em comparação com a média de ancestralidade europeia em todo o genoma (Darvasi; Shifman, 2005; Mersha; Abebe, 2015; Winkler; Nelson; Smith, 2010).

Já, em estudos de caso-controle, a associação verdadeira da ancestralidade com a doença ocorre quando, por exemplo, a população ancestral africana carrega o alelo de risco genético para a doença em uma frequência maior, e é observado no *admixture mapping* uma maior proporção da ancestralidade africana no *locus* da doença, comparando as proporções de ancestralidade nesse local entre casos e controles. Mas se uma maior proporção de ancestralidade africana é observada em todo o genoma dos pacientes em relação aos controles, mas sem aumento

significativo de ancestralidade local em um lugar particular, isso pode apontar para um papel mais forte de fatores socioambientais (por exemplo, renda, educação, tabagismo e hábitos alimentares) independente da ancestralidade (Darvasi; Shifman, 2005; Mersha; Abebe, 2015; Winkler; Nelson; Smith, 2010).

Na análise de *admixture mapping*, tanto em estudos de casos como em estudos caso-controle, primeiro é realizada uma avaliação geral de todas as ancestralidades em todos os cromossomos, para definir se existe algum segmento genômico com desvio significativo de ancestralidade local do grupo de casos em relação ao grupo de controle, em estudos caso-controle, ou em relação à média da ancestralidade global, em estudos somente de casos. Essa análise normalmente é expressa em gráficos, já que se fosse apresentada em tabelas as mesmas seriam de grandes dimensões. As Figuras 3 e 4 exemplificam tipos de gráficos que podem resultar desta análise. Os resultados podem ser expressos em valores de Logaritmo da Probabilidade score (LOD, do inglês *Logarithm of Odds score*) ou Z score.

A análise de LOD score, em estudos com *admixture mapping* e em estudos de associação genômica ampla (GWAS, do inglês *Genome-Wide Association Study*), é empregada para estimar evidências de ligação entre uma região cromossômica ou um marcador genético com a característica de interesse (Ott; Wang; Leal, 2015; Morton, 1955). A presença de LOD score positivo é indicativo da existência da associação entre o marcador e o fenótipo estudado, enquanto que o LOD score negativo indica que a associação é menos provável (Morton, 1955). Tem sido utilizado para *loci* identificados no LOD score o nível de significância com valores de $-\log_{10} \geq 4$ (Cheng et al., 2009; Cheng et al., 2010).

O Z score pode ser usado para avaliar a evidência de associação entre uma região cromossômica ou uma variante com uma doença ou fenótipo em cada local do genoma (Zhu; Cooper; Elston; 2004; Zhu et al., 2006). Se não houver uma região genômica associada com a doença, o Z score está previsto para ser distribuído aproximadamente de acordo com uma distribuição normal. Tem sido utilizado para *loci* identificados no Z score o nível de significância com valores de Z score ≤ -4 ou ≥ 4 (Cheng et al., 2009; Cheng et al., 2010).

Figura 3: Representação do gráfico Manhattan com os resultados do *admixture mapping* em estudo somente de casos com câncer de mama. O eixo X representa as posições genômicas por cromossomo. No eixo Y estão os valores de p-valor de $-\log_{10}$ para a associação entre os segmentos genômicos da ancestralidade testada e o risco para a doença. Cada ponto no gráfico corresponde a um segmento específico no cromossomo. (A) Análise de *admixture mapping* em relação à ancestralidade ameríndia. (B) Análise de *admixture mapping* em relação à ancestralidade europeia. (C) Análise de *admixture mapping* em relação à ancestralidade africana. O círculo vermelho representa a localização dos segmentos cromossômicos com p-valor de $-\log_{10}$ mais significativos na ancestralidade ameríndia, associando essa ancestralidade, nesse local, com a doença. Figura modificada de Fejerman et al., 2012.

Figura 4: Representação dos resultados do *admixture mapping* em estudo de casos e caso-controle com hipertensão em relação à ancestralidade africana e o calibre vascular da retina. O eixo X representa as posições genômicas por cromossomo. O eixo Y do lado esquerdo são os valores de p-valor de $\log_{10}$ para a associação entre os segmentos genômicos e o risco para a doença. O eixo Y do lado direito são os valores de Z score para a associação entre os segmentos genômicos e o risco para a doença. Para a análise somente de casos (linha vermelha), o nível mais alto de um *locus*-específico de LOD score de 5,47 foi detectado na região chr6p21.1. O sinal gerado foi confirmado por análise de caso-controle (Z score, linha azul). No pico da região chr6p21.1, os casos tiveram maior contribuição de ancestralidade africana do que os controles (caso-controle com Z score = -5.26, $p=1,44 \times 10^{-7}$). Figura modificada de Cheng et al., 2010.

As regiões genômicas identificadas pelo *admixture mapping* são geralmente de tamanho grande, o que torna difícil identificar variantes casuais. Portanto, é realizada em seguida uma análise mais refinada (mapeamento fino, *fine-mapping*) para identificar as potenciais variações causais da doença ou do fenótipo avaliado. Deste modo, o mapeamento fino permite avaliar a possibilidade de cada variante presente na região onde houve associação ser a responsável pelo pico apresentado no *admixture mapping*.

A estratégia do mapeamento fino é ter os dados de genotipagem de um grande número de variantes na região onde houve a associação e, possivelmente identificar uma ou várias variantes que estejam com alto desequilíbrio de ligação (LD, do inglês *Linkage Disequilibrium*) na ancestralidade avaliada (Farh et al., 2015; Hormozdiari et al., 2014; Kichaev et al., 2014; Spain; Barrett, 2015).

Para a realização da análise de mapeamento fino é necessário fazer a genotipagem das variantes genéticas que não foram avaliadas na região onde houve o pico no *admixture mapping* ou faz-se a imputação do genótipo dessas variantes.

A imputação de dados genotípicos é uma estratégia estatística que usa padrões de desequilíbrio de ligação observados em painéis de referência, para computacionalmente inferir variantes genéticas não diretamente genotipadas (*in silico genotypes*) em um painel alvo (Howie; Marchini; Stephens, 2011; Morris, 2014).

Assim, o método de imputação preenche as lacunas das variantes que não foram genotipadas no painel de estudo. Logo, a imputação atribui as variáveis específicas de um painel de referência em outro painel e, posteriormente, cria um novo painel de referência com a união das variantes de todos os painéis, ou seja, reúne em um novo painel as variantes genotipadas com as variantes não genotipadas do banco de dados de genotipagem escolhido (Howie; Marchini; Stephens, 2011; Kichaev et al., 2014). A imputação pode, portanto, fornecer uma cobertura mais completa da variação genética de uma região, sem a necessidade de re-sequenciamento ou genotipagem adicional, mas é limitado às variantes que estão presentes no banco de dados genômicos de referência utilizado, ou banco de dados de genotipagem de outros grupos de estudo que for escolhido para a imputação (Howie; Marchini; Stephens, 2011; Morris, 2014). Os bancos públicos de dados genômicos mais utilizados são os do projeto 1000 Genomas (<http://www.1000genomes.org>) e do projeto *International HapMap* (<http://hapmap.ncbi.nlm.nih.gov/>).

Após a imputação é realizada uma análise estatística nesse novo painel com as variantes. Uma série de métodos diferentes foram desenvolvidos para a priorização de variantes causais para explicar os sinais de associação, podendo ser análise estatística baseada no p-valor ou LD e métodos bayesianos, bem como diversos programas podem ser utilizados para a realização das análises estatísticas (Spain; Barrett, 2015).

Enfim, a análise de estimativa de ancestralidade genética é útil para compreender a história da população humana, interpretação de variação genômica e também para estudos com diversas doenças e fenótipos (Bansal; Libiger, 2015).

Cabe ainda destacar que os estudos de associação entre estrutura genômica e doenças complexas têm sido desenvolvidos predominantemente em populações pouco miscigenadas (Galanter et al., 2014; McCarthy et al., 2008). Estudos em outras populações são importantes para revelar novos *loci* de susceptibilidade, examinar a consistência de associações já estabelecidas e fornecer novos conhecimentos sobre as disparidades de ancestralidade na prevalência e gravidade de doenças complexas (Bustamante et al., 2011; Galanter et al., 2014; McCarthy et al., 2008).

Nesse cenário, a população brasileira é de grande interesse pela sua natureza miscigenada. Essa característica imprime desafios específicos e abre novas

perspectivas para se conhecer a variabilidade do genoma, mapear a ancestralidade e explorar a associação entre a ancestralidade e doenças complexas, como a insuficiência cardíaca.

2 OBJETIVOS

2 OBJETIVOS

2.1 Objetivo Geral

Realizar a estimativa de ancestralidade genômica global e local de uma amostra de pacientes com insuficiência cardíaca.

2.2 Objetivos Específicos

- Avaliar a associação de ancestralidade local e identificar regiões genômicas associadas à insuficiência cardíaca, por meio do mapeamento por miscigenação (*admixture mapping*);
- Avaliar a associação de ancestralidade local e identificar regiões genômicas associadas com fatores de risco de IC, como diabetes e hipertensão;
- Avaliar a associação de ancestralidade local e identificar regiões genômicas associadas com a mortalidade por IC.

3 CASUÍSTICA E MÉTODOS

3 CASUÍSTICA E MÉTODOS

3.1 Casuística

Foram incluídos nesse estudo 492 pacientes, não aparentados, diagnosticados com IC de classe funcional III ou IV, de acordo com a *New York Heart Association* (The Criteria Committee, 1964). Esses pacientes são acompanhados pelo grupo do Laboratório de Genética e Cardiologia Molecular (LIM13) do Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo.

O diagnóstico de IC foi realizado de acordo com os critérios de *Framingham* (McKee et al., 1971), que se baseiam na presença simultânea de pelo menos dois critérios maiores ou um critério maior em conjunto com dois critérios menores.

As características consideradas como critérios maiores são: dispneia paroxística noturna; turgência jugular; crepitações pulmonares; cardiomegalia (à radiografia de tórax); edema agudo de pulmão; terceira bulha (galope); aumento da pressão venosa central; refluxo hepatojugular; perda de peso > 4,5 kg em 5 dias em resposta ao tratamento. As características consideradas como critérios menores são: edema de tornozelos bilateral; tosse noturna; dispneia a esforços ordinários; hepatomegalia; derrame pleural; diminuição da capacidade funcional em um terço da máxima registrada previamente; taquicardia (frequência cardíaca >120 bpm (batimentos por minuto)) (McKee et al., 1971).

A classificação da etiologia da IC foi baseada nos sinais e sintomas apresentados pelos pacientes e em exames realizados (ecocardiografia transtorácica com Doppler, eletrocardiograma, sorologia para chagas, Raio-X de tórax) de acordo com recomendações de Hunt et al., 2005 e Richardson et al., 1996.

Diabetes mellitus foi definida pela glicemia de jejum ≥ 126 miligramas por decilitro (mg/dL) (Sociedade Brasileira de Diabetes, 2014).

Hipertensão foi definida por valores de pressão arterial sistólica (PAS) e pressão arterial diastólica (PAD) de ≥ 140 e ≥ 90 milímetros de mercúrio (mmHg), respectivamente (Chobanian et al., 2003).

Os valores do Índice de Massa Corporal (IMC) foram calculados dividindo o peso (em quilogramas) pela altura (em metros) ao quadrado.

Para a realização do nosso estudo, houve uma seleção dentro do Banco de Pacientes com IC do LIM13 de indivíduos com as seguintes características: diagnosticados entre 1995 e 2015 com avaliação ecocardiográfica transtorácica sequencial, com a finalidade de estabelecer o diagnóstico etiológico da IC, acompanhamento clínico e informação de morte por um período de 4 anos.

Nos casos em que não foi possível ter a informação de morte, por meio de contato telefônico, email ou consultas, foi realizada uma busca no banco de dados de mortalidade de São Paulo (ProAim- Programa de Aprimoramento de Informações de Mortalidade do Município de São Paulo).

Todos os dados clínicos e informação de óbito foram obtidos a partir de prontuários.

Essa pesquisa obteve aprovação pela Comissão de Ética para Análise de Projetos de Pesquisa do Hospital das Clínicas da FMUSP (CAPPesq) sob o número CAAE: 29477414.9.0000.0068 e todos os participantes assinaram termo de consentimento livre e esclarecido.

3.2 Genotipagem

A genotipagem das amostras foi realizada pela empresa Affymetrix usando um chip personalizado, Affymetrix® Axiom™ Genome-Wide InCor_BB Array, que contém aproximadamente 800.000 variantes genéticas (Affymetrix, Santa Clara, CA, EUA).

A maior parte dessas variantes genéticas foi selecionada a partir dos 3,1 milhões de SNPs genotipados no Projeto *Internacional HapMap* (<http://hapmap.ncbi.nlm.nih.gov/>).

Nesse estudo avaliamos somente dados de SNPs.

3.3 Limpeza de Dados

Após a genotipagem das ~800.000 variantes genéticas, foi necessário realizar uma etapa de limpeza de dados. O programa PLINK (<http://pngu.mgh.harvard.edu/~purcell/plink/>) foi utilizado para os procedimentos de

limpeza de dados e de controle de qualidade sobre os marcadores genéticos. Foram realizadas as verificações e, se necessário, a exclusão de indivíduos com grande quantidade de dados de genótipos faltando com taxa superior 10%; exclusão de SNPs com falha na genotipagem com taxa superior a 10%; exclusão de marcadores que apresentaram desvio significativo no teste de equilíbrio de Hardy-Weinberg, com o limite de $1.00e^{-04}$ e exclusões de SNPs com frequência do alelo menor (MAF) $\leq 1\%$ (MAF refere-se à frequência na qual o alelo menos comum ocorre em uma determinada população).

Foi realizada a verificação para identificar indivíduos aparentados ou em duplicidade, com o objetivo de eliminar dados duplicados. Foi realizado um controle analítico de sexagem, onde foi verificada qualquer incompatibilidade entre o sexo informado e a genotipagem.

Após essa limpeza ficaram 759.896 SNPs para análise em 492 pacientes.

3.4 População Ancestral

Para determinar as ancestralidades global e local dos pacientes com IC foram usadas 3 populações ancestrais (europeia, africana e ameríndia) como referência, descritas a seguir.

Foram usados dados de genotipagem de 4 populações do projeto 1000 Genomas, sendo 2 populações de ancestralidade europeia (**CEU** - Residentes de UTAH com ancestralidade da Europa do Norte e Ocidental; e **TSI** - população Toscana da Itália) e 2 populações de ancestralidade africana (**LWK** – Quênia; e **YRI** - Nigéria) (<http://www.1000genomes.org/>).

Além dessas, foram usadas 2 populações de ancestralidade ameríndia (**Ashanincas** e **Shimaa**) genotipadas pelo grupo do Dr. Eduardo Martin Tarazona Santos do Laboratório de Diversidade Genética Humana da Universidade Federal de Minas Gerais (UFMG) (LDGH-UFMG) (Kehdy et al., 2015). Utilizamos essas populações ameríndia do Peru, pois eram as que apresentaram maior proximidade genética com a ancestralidade ameríndia da população brasileira, ao invés de usar as populações ameríndias dos projetos *HapMap*, 1000 Genomas ou HGDP-CEPH.

Assim, as populações ancestrais foram compostas por: 185 indivíduos com ancestralidade africana, 99 indivíduos com ancestralidade europeia e 89 indivíduos com ancestralidade ameríndia.

3.5 Ancestralidade Global

Para a análise de ancestralidade global foram utilizados os dados no formato *phased* a partir do projeto 1000 Genomas <ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/>, compreendendo todas as populações de interesse e todos os cromossomos autossômicos. Em seguida foi utilizado o programa SHAPEIT (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html) para transformar os arquivos do 1000 Genomas para os formatos ped/map.

Para fazer a preparação do nosso arquivo a partir dos arquivos das populações de ancestralidade europeia e africana selecionadas do 1000 Genomas, e dos arquivos das populações de ancestralidade ameríndia do grupo LDGH-UFMG, foram utilizados os programas PLINK, BEAGLE (<http://faculty.washington.edu/browning/beagle/beagle.html>) e o programa GTOOL (<http://www.well.ox.ac.uk/~cfreeman/software/gwas/gtool.html>), deixando todos os arquivos com a mesma formatação.

Foi utilizado o programa *Rstudio* (<http://www.rstudio.com/>) para obtenção dos SNPs genotipados em comum entre nossa população e as populações ancestrais do 1000 Genomas e do LDGH-UFMG. Portanto, dos 759.896 SNPs restantes da limpeza da genotipagem, foram utilizados 182.090 SNPs comuns.

Para realizar o cálculo de ancestralidade global para cada indivíduo foi utilizado o programa *ADMIXTURE* (<https://www.genetics.ucla.edu/software/admixture/>). O programa *ADMIXTURE* faz a estimativa de ancestralidade global baseado em uma estrutura de máxima verossimilhança (*maximum-likelihood estimation*). Foi utilizado o número de populações ancestrais fixado em 3 ($k=3$) (populações ancestrais africana, europeia e ameríndia). Assim, cada paciente apresentou uma média de ancestralidade global dividida para cada uma das 3 ancestralidades.

Os dados de ancestralidade global foram utilizados para caracterização da amostra de 492 pacientes com IC, para averiguação de estratificação amostral na análise caso-controle da mortalidade e para as análises de *admixture mapping* da IC, diabetes e hipertensão, que foram realizadas no modelo de somente casos. No modelo somente de casos, a análise é baseada na comparação da ancestralidade local estimada de cada região com a média da ancestralidade global desse mesmo grupo.

3.6 Ancestralidade Local e *Admixture Mapping*

Para a análise da ancestralidade local, primeiramente, os programas PLINK e SHAPEIT foram utilizados para transformar nossos arquivos dos formatos ped/map em arquivos para os formatos haps/samples e para separar os 22 cromossomos em diferentes arquivos.

Foram utilizados os dados no formato *phased* a partir do projeto 1000 Genomas ALL.integrated_phase1_v3.20101123.snps_indels_svsvs.genotypes.nomono.tgz, compreendendo todas as populações de interesse e todos os cromossomos autossômicos. Estes conjuntos de dados, separados por cromossomos, foram obtidos no formato haps/samples ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase1/analysis_results/shapeit2_phased_haplotypes/.

Para realizar o cálculo de ancestralidade local para cada indivíduo foi utilizado o programa PCAdmix (<https://sites.google.com/site/pcadmix/home>), ressaltando que a estimativa da ancestralidade local não é dada para cada SNPs mas para segmentos genômicos (janelas).

As inferências das ancestralidades locais foram realizadas com marcadores com $r^2 > 0,99$, ou seja, foram removidos marcadores que estavam em desequilíbrio de ligação com $r^2 < 0,99$ para evitar erros de inferência de ancestralidades causadas por um sobre-ajuste (*overfitting*), não houve a exclusão de nenhum SNPs. Em relação às janelas, considerando a nossa quantidade total de SNPs comuns a todas as populações (182.090 SNPs), foi definido um comprimento da janela com 25 SNPs, de acordo com o estudo de Moreno-Estrada et al. (2014), portanto, cada janela possuía

25 SNPs. A criação de janelas é por conta do método que o programa utiliza, que capta a não independência dos SNPs utilizando janelas ou blocos de SNPs (Kehdy et al., 2015).

Inicialmente, o programa PCAdmix agrupa os SNPs em janelas e depois faz dois cálculos para inferir a ancestralidade local dessas janelas.

Primeiro o programa aplica a análise de componentes principais (PCA), analisando cada janela e avaliando a probabilidade de que uma dada janela daquele indivíduo miscigenado venha de cada população de referência. Desta forma, as janelas com maior probabilidade de terem a mesma ancestralidade se agrupam.

Em seguida, o programa utiliza o modelo oculto de Markov (HMM, do inglês *Hidden Markov Model*) para melhor estimar as ancestralidades dessas janelas que se agruparam. Foram consideradas apenas as janelas inferidas pelo algoritmo com uma probabilidade $>0,90$ de ancestralidade.

O resultado final do PCAdmix é a probabilidade de que uma determinada janela de um cromossomo tenha cada uma das ancestralidades das populações ancestrais. Uma vez que o nosso estudo utilizou o número de populações ancestrais fixado em 3 ($k=3$) (populações ancestrais africana, europeia e ameríndia), assim sendo cada janela dos 22 cromossomos apresentou uma média de ancestralidade local dividida para as 3 ancestralidades.

Os dados da ancestralidade local de cada janela são usados somente para fazer a análise de *admixture mapping*.

Foi realizado em seguida, o teste estatístico de χ^2 (qui quadrado) para examinar a significância estatística de associação entre a variável sexo, e as variáveis hipertensão e diabetes, sendo que o nível de significância foi estabelecido para um valor de $p \leq 0,05$. Como não havia evidência de interação entre as variáveis, todas as análises de *admixture mapping* foram realizadas utilizando o sexo como uma co-variável. Essa análise foi realizada utilizando o software GraphPad Prism versão 7.03 (GraphPad Software, La Jolla California USA, www.graphpad.com).

Portanto, para a confecção do *admixture mapping* da IC e da mortalidade, as análises foram ajustadas para as co-variáveis sexo, diabetes, hipertensão, idade (contínua), IMC (contínua) e colesterol total (contínua). Para o *admixture mapping* da diabetes e hipertensão foram utilizadas quase todas as co-variáveis utilizadas no

admixture mapping da IC e mortalidade, com exceção das co-variáveis diabetes e hipertensão que foram retiradas das análises, respectivamente.

As análises de *admixture mapping* da IC, diabetes e hipertensão foram realizadas utilizando regressão logística e no modelo somente de casos. As análises de *admixture mapping* da mortalidade por IC foram realizadas usando regressão logística, utilizando os modelos de caso-controle e somente de casos. Nessa análise, o grupo de casos foi composto por pacientes em óbito no final do período de 4 anos de avaliação, contados a partir do momento da IC diagnosticada, e o grupo de controles, incluía os demais indivíduos que estavam vivos no final do tempo de acompanhamento.

Para realizar o *admixture mapping* dos dados de mortalidade, inicialmente foi realizado o teste de Kruskal-Wallis para examinar a significância estatística de associação entre as etiologias da IC apresentadas pelos pacientes desse estudo e a mortalidade por IC, assim avaliamos se alguma etiologia estaria associada ao óbito, sendo que o nível de significância foi estabelecido para um valor de $p \leq 0,05$.

Em seguida, foi realizado o teste estatístico *t student* das médias de ancestralidade global entre os grupos de casos e controles para identificar a existência de uma possível estratificação amostral, sendo que o nível de significância foi estabelecido para um valor de $p \leq 0,05$.

As duas análises estatísticas foram realizadas utilizando o software GraphPad Prism versão 7.03 (GraphPad Software, La Jolla California USA, www.graphpad.com).

Posteriormente, foi realizada o *admixture mapping* dos dados de mortalidade, utilizando o programa PLINK.

Para as análises de regressão entre as variáveis e a ancestralidade local de cada indivíduo, foi utilizado os programas R (<https://www.r-project.org/>) e o PLINK usando um modelo aditivo. Para testar a associação para cada ancestralidade, foi feito um *script* em R para codificar a ancestralidade local inferida de cada janela como 1 (ancestralidade a ser testada) e 2 (outras ancestralidades). Por exemplo, para testar a associação entre ancestralidade local africana e a IC, a ancestralidade local africana foi codificada como 1 e as ancestralidades europeia ou ameríndia foram codificadas como 2. Nesse caso, uma janela homozigota africana será 11, uma homozigota europeia ou ameríndia será 22 e uma janela heterozigota africana/europeia ou africana/ameríndia será 12, o que equivale, nesse caso, a ter 0, 1 ou 2 cópias de

ancestralidade local africana inferida para cada janela. Essa análise foi realizada para as três ancestralidades (africana, ameríndia e europeia) avaliadas nesse estudo, em relação a cada característica testada.

Para a definição do p-valor considerado estatisticamente significativo nas análises de *admixture mapping* foi utilizada a seguinte estratégia:

Uma vez que os segmentos ancestrais cromossômicos identificados pelo programa PCAdmix são bastante amplos, e existir uma correlação entre estimativas de ancestralidade genômica e os marcadores genéticos, para estipular o valor de p considerado estatisticamente significativo, o uso de uma correção de Bonferroni com base no número total de marcadores não foi apropriado para esse estudo. Portanto, para determinar um valor de p apropriado para verificar a significância de associação, primeiro estimamos o número total de testes independentes de acordo com Shriner et al. 2011 e, então, estabelecemos um limiar de significância de acordo com uma correção de Bonferroni para uma taxa de erro do tipo I de 5%. O nível de significância considerado limiar foi de $p \leq 9,78 \times 10^{-4}$ nas análises de *admixture mapping*, nos gráficos Manhattan o valor considerado estatisticamente significativo foi $-\log_{10} \geq 4$ e no dado de Z score o nível de significância foi definido como valores ≤ -4 .

Para a plotagem dos resultados das análises de regressão logística nos gráficos foram utilizados os programas PLINK e o programa R (<https://cran.r-project.org/web/packages/car/index.html>) conforme descrito por Clarke et al., 2011.

3.7 Mapeamento Fino

Com relação às análises de *admixture mapping* que apresentaram associação consistente com a doença em determinada ancestralidade, foi realizada análise mais detalhada da relação entre genótipo e o desfecho, aqui chamada de mapeamento fino.

Para o mapeamento fino, primeiramente os nossos dados genômicos, exclusivamente da região de interesse, foram imputados com o dados do painel de referência do projeto 1000 Genomas https://mathgen.stats.ox.ac.uk/impute/1000GP_Phase3.html, usando somente os dados das populações de interesse, ou seja, da ancestralidade que houve associação,

utilizando o programa IMPUTE2 v2.2.2
https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#home.

O programa IMPUTE2 é usado para imputação de genótipos faltantes. O IMPUTE2 usa o método de Monte Carlo via cadeias de Markov (MCMC, do inglês *Markov chain Monte Carlo*), que reúne, em um novo painel, os SNPs genotipados aos SNPs não genotipados do banco de dados de genotipagem escolhido. O IMPUTE2 foi executado com todas as opções padrões do programa de acordo com Howie; Marchini; Stephens, 2011.

Em seguida foi realizada uma limpeza desses dados, excluindo SNPs com frequência de MAF $\leq 1\%$, utilizando o programa PLINK.

Para avaliar se algum dos SNPs presentes na região de interesse corresponde ao pico detectado por meio do *admixture mapping*, foi realizada uma regressão logística com ajuste de idade e sexo, utilizando o programa SNPTEST v2.4.1 https://mathgen.stats.ox.ac.uk/genetics_software/snpTest/snpTest.html. Para todas as análises de associação foram utilizados o modelo aditivo de herança e o método "score" do programa SNPTEST.

O SNPTEST é um programa que realiza testes de associação de SNP com fenótipos ou doenças que levam em consideração a incerteza genotípica.

O p-valor considerado estatisticamente significativo foi de $p \leq 1,78 \times 10^{-2}$.

O gráfico do mapeamento fino foi produzido usando o programa LocusZoom <http://locuszoom.sph.umich.edu/locuszoom/>.

Para os SNPs detectados por meio do mapeamento fino em relação aos dados de mortalidade foi verificada a existência de associação entre o SNP e o risco de morte por IC, realizada pelo modelo de riscos proporcional de Cox, utilizando o programa R, com ajuste de idade e sexo. O intervalo de tempo de seguimento foi definido como o tempo entre a data da IC diagnosticada e a data da morte, ou o último contato do paciente que perdeu o seguimento ou o final do seguimento de 4 anos; foi considerado o que aconteceu primeiro. O p-valor considerado estatisticamente significativo foi de $p \leq 0,05$.

Os resultados serão apresentados pela medida de *hazard ratio* (HR).

Os métodos estão representados de forma resumida no fluxograma a seguir.

3.8 Fluxograma dos Métodos

4 RESULTADOS E DISCUSSÃO

4.1 Caracterização Clínica dos Pacientes

Nossa amostra foi composta por 295 (60%) homens e 197 (40%) mulheres. A idade média dos indivíduos foi de $58 \pm 14,4$ anos. De um total de 492 pacientes avaliados nesse estudo, pudemos observar que 73,2% (n=360) dos indivíduos se autodeclararam de cor de pele branca. A diabetes estava presente em 29,9% (n=147) dos pacientes, e a hipertensão foi observada em 63% (n=310) dos pacientes, 28,9% (n=142) dos pacientes apresentaram etiologia hipertensiva, seguido da etiologia isquêmica (n=141, 28,7%). Ao final de 4 anos de avaliação 50,4% (n=248) dos pacientes evoluíram para o óbito (Tabela 1).

Tabela 1: Características gerais dos 492 pacientes com IC avaliados nesse estudo.

Variáveis	Resultados
Idade (anos), média $\pm$ Desvio Padrão (DP)	58 $\pm$ 14,4
Sexo, n (%)	
Masculino	295 (60,0)
Feminino	197 (40,0)
Autodeclaração de Cor de Pele, n (%)	
Branca	360 (73,2)
Parda	68 (13,8)
Negra	51 (10,4)
Amarela/indígena	5 (1,0)
Não declarado	8 (1,6)
Diabetes, n (%)	
Sim	147 (29,9)
Não	345 (70,1)
Hipertensão, n (%)	
Sim	310 (63,0)
Não	182 (37,0)
Etiologia, n (%)	
Hipertensiva	142 (28,9)
Isquêmica	141 (28,7)
Valvar	72 (14,6)
Chagásica	59 (12,0)
Outras	78 (15,9)
Óbito, n (%)	248 (50,4)
IMC, média $\pm$ DP	25,6 $\pm$ 5,4
Colesterol Total, média $\pm$ DP mg/dL	190,4 $\pm$ 51,5
Triglicerídeo, média $\pm$ DP mg/dL	123,7 $\pm$ 67,0
HDL Colesterol, média $\pm$ DP mg/dL	45,3 $\pm$ 15,4
LDL Colesterol, média $\pm$ DP mg/dL	120,2 $\pm$ 42,1
VLDL Colesterol, média $\pm$ DP mg/dL	24,9 $\pm$ 13,7
Glicemia, média $\pm$ DP mg/dl	102,2 $\pm$ 50,2

Inicialmente foi verificado a significância estatística de associação entre a variável sexo, e as variáveis hipertensão e diabetes, que não apresentou valores estatisticamente significativos (Tabela 2). Assim, foram utilizadas as co-variáveis idade, sexo, IMC, colesterol total e diabetes ou hipertensão para ajustar as análises estatística de associação do *admixture mapping* por regressão logística e linear.

Tabela 2: Análise da associação entre a variável sexo e as variáveis hipertensão e diabetes.

Variável, n (%)	Sexo		p-valor*
	Masculino, n (%)	Feminino, n (%)	
Hipertensão	310 (63,0)	178 (61,0)	0,255
Não hipertensão	182 (37,0)	114 (39,0)	
Total	492	292	
Diabetes	147 (29,9)	77 (26,6)	0,061
Não diabetes	345 (70,1)	212 (73,4)	
Total	492	289	

* Nível de significância valor de $p \leq 0,05$.

4.2 Ancestralidade Global

A análise de ancestralidade global na amostra de 492 pacientes mostrou que a proporção da ancestralidade europeia variou de 0,025% a 0,994% (média $0,618 \pm 0,218$), a ancestralidade africana variou de 0,002% a 0,966% (média $0,249 \pm 0,223$), e a ancestralidade ameríndia variou de 0,002% a 0,563% (média $0,133 \pm 0,106$) (Tabela 3). A estimativa de ancestralidade global de cada paciente está apresentada no Anexo 1.

Comparando nossos dados atuais com os dados do nosso estudo anterior, onde utilizamos 48 marcadores do tipo INDEL em 503 pacientes com IC, observamos que não houve diferença estatística entre as médias das ancestralidades encontradas (Cardena, 2013; Cardena et al., 2014; 2016). Entretanto, é possível observar que houve maior contribuição de ancestralidade europeia e diminuição das contribuições ameríndias e africanas no grupo atual de pacientes com IC em relação ao estudo anterior (Tabela 3).

Tabela 3: Médias das ancestralidades genômicas em dois dos nossos estudos.

Ancestralidade	503 Pacientes (anterior) INDEL	492 Pacientes (atual) SNPs	p-valor*
Europeia	0,594±0,203	0,618±0,218	0,079
Africana	0,261±0,188	0,249±0,223	0,357
Ameríndia	0,145±0,100	0,133±0,106	0,069

* Nível de significância valor de $p \leq 0,05$.

Além de serem grupos distintos de pacientes, essa diferença de ancestralidade entre os nossos dois estudos pode ser consequência de três fatos:

1. Foram utilizados diferentes painéis de AIMs: no primeiro estudo foram usados 48 AIMs do tipo INDEL e no projeto atual estamos utilizando 182.090 AIMs do tipo SNP;
2. Os grupos ancestrais que foram utilizados para calcular as médias das ancestralidades foram distintos: no primeiro estudo foram construídos painéis de populações ancestrais, sendo que o grupo ancestral ameríndio foi composto por tribos indígenas provenientes da região amazônica do Brasil, o grupo ancestral europeu foi composto por populações naturais de Portugal e o grupo ancestral africano foi composto por populações naturais da África subsaariana (Angola, Moçambique, Zaire, Camarões e Costa do Marfim). No projeto atual estão sendo usadas 4 populações do projeto 1000 Genomas, sendo 2 populações de ancestralidade europeia (CEU e TSI) e 2 populações de ancestralidade africana (LWK e YRI), e 2 populações de ancestralidade ameríndia (Ashanincas e Shimaa) genotipadas pelo grupo do Dr. Eduardo Martin Tarazona (LDGH-UFMG);
3. Os programas utilizados para calcular a ancestralidade são díspares: no primeiro estudo foi utilizado o programa *STRUCTURE*, já no estudo atual está sendo utilizado o programa *ADMIXTURE*. O programa *ADMIXTURE* faz a estimativa de ancestralidade baseado em uma estrutura de máxima verossimilhança, enquanto o *STRUCTURE* utiliza uma abordagem Bayesiana com o método de Monte Carlo via cadeias de Markov (MCMC).

Um estudo recente de revisão sistemática e meta-análise que utilizou dados de 25 trabalhos de 38 populações brasileiras de diferentes regiões, utilizando somente amostras populacionais de indivíduos saudáveis (totalizando dados de 8.733 indivíduos), mostrou que as proporções de ancestralidades globais no Brasil são: europeia de 62%, africana 21%, e ameríndia 17% (Moura et al., 2015). Já os valores

em relação a região Sudeste são: ancestralidade europeia 67%, ancestralidade africana 23% e ancestralidade ameríndia 10% (Moura et al., 2015). Esses 25 estudos utilizaram os AIMs do tipo INDEL, SNP, sequências curtas repetidas em tandem (STR, do inglês *Short Tandem Repeats*) e repetições em tandem de número variável (VNTR, do inglês *Variable Number of Tandem Repeats*).

Tanto nosso estudo como o de Moura e colaboradores (2015) ressaltam a informação de uma ancestralidade genômica autossômica tri-híbrida da população brasileira (europeia, africana e ameríndia), com uma predominância da ancestralidade genômica europeia. A partir dos nossos resultados, demonstramos a existência de uma grande heterogeneidade de ancestralidade, o que reflete o alto grau de miscigenação da população brasileira.

4.3 Ancestralidade Local e *Admixture Mapping*

Para a análise da ancestralidade local, foi realizada a separação dos SNPs por cromossomo, a partir da homogeneização dos dados da nossa amostra em relação às 3 populações ancestrais, conforme descrito na Casuística e Métodos.

Dos 182.090 SNPs comuns, ficaram em cada cromossomo diferentes quantidades de SNPs utilizados e de janelas (Tabela 4).

Tabela 4: Quantidade de SNPs utilizados na análise de ancestralidade local.

Cromossomos	SNPs Total	SNPs utilizados	Janelas
chr 1	61.334	14.665	587
chr 2	59.029	15.057	602
chr 3	49.484	12.805	512
chr 4	44.689	11.725	469
chr 5	43.686	10.444	418
chr 6	51.833	10.471	419
chr 7	41.004	2.436	98
chr 8	36.527	9.522	381
chr 9	34.003	8.993	360
chr 10	36.231	9.635	385
chr 11	39.118	9.482	379
chr 12	36.190	9.457	378
chr 13	25.129	6.650	266
chr 14	25.099	6.431	257
chr 15	25.060	6.739	270
chr 16	28.819	7.054	282
chr 17	29.193	7.109	284
chr 18	21.466	6.061	242
chr 19	27.385	6.023	241
chr 20	19.631	5.398	216
chr 21	11.733	2.733	122
chr 22	13.253	3.200	128
TOTAL	759.896	182.090	7.296

É possível observar que o cromossomo 7 foi o que apresentou menor número de janelas (98), o que se explica pelo fato do menor número de SNPs genotipados em nossa população que eram comuns com os das populações de referência.

Feita a separação dos SNPs por cromossomos e realizada a preparação necessária dos arquivos para deixá-los compatíveis com o programa, para o cálculo da ancestralidade local, foi utilizado o programa PCAdmix para fazer a comparação entre os genótipos de cada indivíduo da nossa população com os genótipos presentes nas 3 populações ancestrais e calcular as proporções de ancestralidade local nas janelas dos 22 cromossomos dos 492 pacientes (Anexo 2).

Com os dados da ancestralidade local de todos os indivíduos foi possível a realização das análises de *admixture mapping* por meio de regressão linear e logística utilizando o programa PLINK, conforme descrito em Casuística e Métodos.

Para as análises de *admixture mapping*, o valor de p considerado estatisticamente significativo para análise das regressões foi $p \leq 9,78 \times 10^{-4}$ e nos

gráficos Manhattan o nível de significância foi definido como valores de $-\log_{10} \geq 4$ e no Z score, o nível de significância foi de ≤ -4 .

Nos gráficos Manhattan cada ponto representa uma janela de 25 SNPs. As cores azul e amarelo, ou azul claro e azul escuro, foram aplicadas somente para isolar os cromossomos, para uma melhor visualização dos resultados. O eixo X representa as posições genômicas por cromossomo. No eixo Y estão os valores de $-\log_{10}$ para a associação entre as janelas da ancestralidade testada e o risco para o dado que está sendo avaliado. A linha preta representa o nível do valor de $-\log_{10} \geq 4$ considerado estatisticamente significativo. O círculo vermelho representa a região que apresentou diferença estatisticamente significativa na análise de *admixture mapping*.

4.3.1 Insuficiência Cardíaca

Nessa análise de *admixture mapping* foi realizada a comparação da ancestralidade local dos 492 pacientes com IC com a média da ancestralidade global dos mesmos pacientes.

Os resultados da análise de *admixture mapping* da IC estão representados nos gráficos Manhattan (Gráficos 1, 2 e 3), para cada uma das ancestralidades, respectivamente.

Gráfico 1: Resultado do *admixture mapping* da insuficiência cardíaca na ancestralidade europeia.

Gráfico 2: Resultado do *admixture mapping* da insuficiência cardíaca na ancestralidade ameríndia.

Gráfico 3: Resultado do *admixture mapping* da insuficiência cardíaca na ancestralidade africana.

Na análise de *admixture mapping* da insuficiência cardíaca, para as três ancestralidades avaliadas, não foi observada nenhuma região com desvio significativo em relação à média da ancestralidade global (Gráficos 1, 2 e 3), ou seja, não houve associação de nenhuma região genômica em qualquer ancestralidade com a IC. Isso pode ser resultante do pequeno tamanho amostral, n=492 pacientes, ou também porque a IC apresenta diversos tipos de etiologias, o que pode resultar em uma amostra heterogênea quanto à etiologia da doença. Além de ser uma enfermidade de origem poligênica e multifatorial, o fator socioambiental apresenta uma contribuição significativa para o desenvolvimento e gravidade da IC.

A análise da estrutura da população com base na ancestralidade genômica é um componente cada vez mais importante em estudos genéticos. Têm sido relatados vários estudos que tiveram êxito com o uso da técnica de *admixture mapping* na identificação de *loci* de doenças complexas em populações miscigenadas, incluindo doenças que são fatores de risco para a IC, como a obesidade (Basu et al., 2009; Cheng et al., 2009), diabetes (Cheng et al., 2009; Elbein et al., 2010), hipertensão (Deo et al., 2007; Shriner et al., 2015; Zhu et al., 2005; Zhu; Cooper, 2007) e a doença arterial periférica (Gomez et al., 2015; Scherer et al., 2010).

Até o momento, nosso estudo foi o único a utilizar a técnica de *admixture mapping* para identificar ancestralidade genômica, regiões genômicas e/ou genes associados à IC. Outros estudos têm sido realizados utilizando a técnica de GWAS para encontrar variantes genéticas associadas com a IC, com alguns resultados significativos.

Já foi relatada associação de dois SNPs intrônicos (rs1739843 e rs6787362) no cromossomo 1 e 3, respectivamente, um SNP intergênico (rs10519210) na região chr15q22 e o SNP (rs2207418) na região chr20p12 associados com a IC em pacientes da ancestralidade europeia, em populações dos Estados Unidos (Cappola et al., 2010; Smith et al., 2010; Parsa et al., 2011).

Em pacientes com IC de ancestralidade africana foi relatado associação do *locus* (rs11172782) intergênico na região chr12q14, em populações dos Estados Unidos (Smith et al., 2010).

A insuficiência cardíaca tem sido apontada como um importante problema de saúde pública e considerada como epidemia com elevada mortalidade e morbidade (Go et al., 2014). As projeções mostram que a prevalência da IC no mundo aumentará

em 46% de 2012-2030, resultando em mais de 8 milhões de pessoas, acima dos 18 anos de idade, com IC. Esse fato se deve, provavelmente, ao aumento da expectativa de vida, uma vez que a IC acomete de forma preponderante faixas etárias mais elevadas (Go et al., 2014).

No último censo brasileiro (2010), observou-se um crescimento da população idosa e, portanto, com potencial crescimento de pacientes em risco ou portadores de IC (Gauí et al., 2014).

A IC consiste na via final comum de várias doenças, como hipertensão arterial, diabetes e coronariopatias (van Riet et al., 2016). Assim, muitos fatores, podem influenciar a gravidade da IC, como a ancestralidade genética, exposição ambiental e outros determinantes sociais da saúde que são fundamentais para melhor compreender a etiologia da IC.

4.3.2 Diabetes

Na análise de *admixture mapping* em relação à diabetes foi realizada a comparação da ancestralidade local dos 492 pacientes com IC com a média da ancestralidade global dos mesmos pacientes.

Os resultados da análise de *admixture mapping* da diabetes estão representados nos gráficos Manhattan (Gráficos 4, 5 e 6).

Gráfico 4: Resultado do *admixture mapping* da diabetes em pacientes com insuficiência cardíaca na ancestralidade europeia.

Gráfico 5: Resultado do *admixture mapping* da diabetes em pacientes com insuficiência cardíaca na ancestralidade ameríndia.

Gráfico 6: Resultado do *admixture mapping* da diabetes em pacientes com insuficiência cardíaca na ancestralidade africana.

A incidência de diabetes mellitus tipo 2 (DM2) vem aumentando globalmente, inclusive no Brasil, e recentemente o Estudo Longitudinal de Saúde do Adulto – Brasil (ELSA-Brasil) encontrou uma frequência de diabetes mellitus de 19,8% em adultos de 35-74 anos (Schmidt et al., 2014).

A DM2, juntamente com o câncer, as doenças cardiovasculares e as doenças respiratórias crônicas, representam cerca de 80% da mortalidade por doenças crônicas não-transmissíveis no mundo (Schmidt et al., 2014).

A susceptibilidade ao DM2 é atribuída a fatores comportamentais e socioeconômicos, mas também é influenciada por fatores genéticos e ancestralidade genética (Cheng et al., 2012; Ng, 2015).

Apesar do nosso estudo não ter identificado nenhum *locus* nas ancestralidades testadas associado ao risco para a DM2, mais de 30 *loci* já foram identificados em estudos de GWAS para o DM2 ou fenótipos relacionados (Cheng et al., 2012; Ng, 2015). Esses *loci* são comuns (frequência de alelos >25%) e conferem pequenos efeitos genéticos (OR 1,06-1,7) (Cheng et al., 2012). O maior efeito de um SNP relatado por GWAS é o rs7903146 (OR 1,45-2,41), localizado no gene *TCF7L2* mapeado no cromossomo 10 (Frayling, 2007). Outros genes também foram associados com a DM2 tais como *KCNJ11*, *IRS1*, *KCNQ1*, *CDKN2A/DKN2B*, *IGF2BP2* e *SLC30A8* (Grant et al., 2006; Scott et al., 2007; Sladek et al., 2007; Yasuda et al., 2008).

Embora todos estes estudos tenham identificado numerosos *loci* de risco para o DM2, eles utilizaram na maioria dos estudos, populações de ancestralidade europeia, e não conseguiram explicar o fato do acometimento da DM2 ser maior em indivíduos negros. Deste modo, o papel que a ancestralidade genética apresenta para o risco à DM2 ainda é mal compreendida.

Métodos alternativos têm sido utilizados, como o *admixture mapping*, na tentativa de identificar *loci* genéticos responsáveis pelo risco à DM2 e também para elucidar o papel que a ancestralidade genética apresenta na susceptibilidade dessa doença.

Por *admixture mapping*, no estudo de Elbein e colaboradores (2009), foram identificados dois *loci* nos cromossomos 2 e 13, associados a DM2 na ancestralidade africana (Elbein et al., 2009). A ancestralidade africana também foi associada com 33% de risco aumentado de desenvolver DM2, e dois *loci*, nas regiões 12p13.31 e

13q14.3, foram associados ao risco aumentado à DM2 em outro estudo (Cheng et al., 2012). Em 2014, Jeff e colaboradores também identificaram na ancestralidade africana uma região no cromossomo 11 associado ao risco à DM2 (Jeff et al., 2014).

Apesar desses resultados interessantes, ainda não foi determinado se existe um enriquecimento de alelos de risco em haplótipos de ancestralidade africana que possam influenciar o risco genético para DM2.

4.3.3 Hipertensão

Na análise de *admixture mapping* em relação à hipertensão foi realizada a comparação da ancestralidade local dos 492 pacientes com IC com a média da ancestralidade global dos mesmos pacientes.

Os resultados da análise de *admixture mapping* da hipertensão estão representados nos gráficos Manhattan (Gráficos 7, 8 e 9).

Gráfico 7: Resultado do *admixture mapping* da hipertensão em pacientes com insuficiência cardíaca na ancestralidade europeia.

Gráfico 8: Resultado do *admixture mapping* da hipertensão em pacientes com insuficiência cardíaca na ancestralidade ameríndia.

Gráfico 9: Resultado do *admixture mapping* da hipertensão em pacientes com insuficiência cardíaca na ancestralidade africana.

O círculo vermelho representa a região que apresentou diferença estatisticamente significativa na análise de *admixture mapping*.

A hipertensão arterial é um importante fator de risco para o desenvolvimento, morbidade e mortalidade de doenças cardiovasculares e tem causado preocupação crescente na área de saúde pública por causa de suas consequências sociais e econômicas. No Brasil, um estudo de análise transversal de 1980 a 2010 revelou que a prevalência da hipertensão arterial é de aproximadamente 31% na população adulta (Picon et al., 2012).

É conhecido o fato da hipertensão ser um traço biológico complexo que resulta da influência conjunta de fatores genéticos e ambientais. Importantes categorias de exposições ambientais já foram identificadas, mas a arquitetura genética dessa doença ainda não está bem definida (Shetty et al., 2012; Zhu et al., 2005). Sabe-se que os indivíduos negros têm taxas mais elevadas de hipertensão do que indivíduos brancos, e tem sido sugerido que essa diferença pode ser, em parte, devido a diferenças genéticas, como a ancestralidade (Tang et al., 2006).

Análises de *admixture mapping* em populações dos Estados Unidos encontraram associação da ancestralidade africana com a hipertensão, nas regiões chr6q24 e chr21q21 (Shetty et al., 2012; Zhu et al., 2005).

Em nosso estudo, na análise de *admixture mapping* da hipertensão arterial, também foi encontrado associação com a ancestralidade africana, entretanto, a região que apresentou super-representação está presente no cromossomo 2, na região p25.3 ($p=4,65 \times 10^{-5}$) (Gráfico 9, Tabela 5).

Tabela 5: Resultado do *admixture mapping* da hipertensão em pacientes com IC na ancestralidade africana.

Chr	Primeiro SNP	Último SNP	OR	95% IC	p-valor*
2	rs10190313	rs79572936	1.82	1.21 - 1.99	$4,650 \times 10^{-5}$
2	rs17041942	rs6738711	1.01	1.22 - 1.48	$3,903 \times 10^{-3}$
2	rs11892515	rs73184882	1.24	1.08 - 1.58	$2,709 \times 10^{-3}$
20	rs73123747	rs80147747	1.17	1.00 - 1.50	$2,504 \times 10^{-3}$

Chr = cromossomo; OR= *odds ratio*; 95% IC = Intervalo de confiança de 95%.

* Nível de significância valor de $p \leq 9,78 \times 10^{-4}$.

Como mencionado anteriormente, cada janela possui 25 SNPs, e a lista dos SNPs da janela que apresentou diferença estatística para hipertensão está apresentado na Tabela 6.

Tabela 6: Lista dos 25 SNPs presentes na janela que apresentou diferença estatística.

SNPs				
rs10190313	rs80026772	rs12714424	rs4854431	rs73908656
rs11689974	rs12470089	rs4484060	rs4428039	rs4488692
rs6707477	rs6738711	rs61151411	rs6548282	rs4586665
rs12987840	rs6714049	rs6760982	rs6752572	rs115922413
rs10192856	rs55993378	rs7421139	rs7605413	rs79572936

Essa janela apresenta 489.359 pares de base (pb) de tamanho e compreende o segmento do nucleotídeo 647.577 até 1.138.119 do cromossomo 2, com localização na região p25.3 (Figura 5).

Figura 5: Localização no cromossomo 2 da janela que apresentou diferença estatística no *admixture mapping* da ancestralidade africana para hipertensão em pacientes com insuficiência cardíaca. Modificado de <https://ghr.nlm.nih.gov/chromosome/2#ideogram>.

Nessa região estão presentes 7 RNAs não codificantes (ncRNA, do inglês *non-coding RNA*) (ncRNA LOC101060391, ncRNA LOC105373354, ncRNA LOC105373358, ncRNA LOC105373392, ncRNA LOC105373479, ncRNA LOC105373480, ncRNA LOC105373481), 2 RNAs longos de região intergênica não codificadores de proteína (lincRNA, do inglês *long intergenic non-protein coding RNA*) (lincRNA 1115, lincRNA 1939), 44,93% do gene *Syntrophin Gamma 2* (SNTG2) e o gene que codifica uma proteína de transmembrana (*Transmembrane Protein 18* (TMEM18)).

Os ncRNAs são uma classe de moléculas de RNA transcritos, mas sem potencial de codificação de proteínas. Muitos ncRNAs têm desempenhado papéis reguladores importantes em diversos processos biológicos, incluindo processos patológicos como câncer (Qu; Adelson, 2012).

Os lincRNAs são definidos como transcritos de RNA que têm mais de 200 nucleotídeos, localizados em regiões intergênicas e já foram identificados mais de 15.000 lincRNAs codificados pelo genoma humano. Os lincRNAs participam de uma

vasta variedade de processos biológicos, tais como: proliferação celular, morfogênese, pluripotência, processos neuronais, gametogênese, controle de expressão gênica e controle epigenético (Deniz; Erman, 2017).

No entanto, ainda não foram caracterizadas as funções dos 7 ncRNAs e dos 2 lincRNAs presentes na região que nosso estudo encontrou diferença estatística.

Dos 416.777 pb do gene *SNTG2*, somente 187.250 pb (44,93%) estão presentes na região que apresentou pico significativo no *admixture mapping*. O gene *SNTG2* está localizado entre os nucleotídeos 950.868 e 1.367.645, possui 29 éxons e codifica uma proteína pertencente à família da sintrofina (Rio et al., 2013).

As sintrofinas são proteínas de membrana citoplasmática que desempenham um papel fundamental como proteínas do esqueleto (*scaffolding proteins*), nas junções neuromusculares e no sistema nervoso central, onde interagem com distrofina e distrobrevina (Piluso et al., 2000; Rio et al., 2013; Yamakawa et al., 2007). A ausência ou redução das sintrofinas têm sido associadas à distrofia muscular de Duchenne e ao autismo (Piluso et al., 2000; Rio et al., 2013; Yamakawa et al., 2007).

Já o gene *TMEM18* está localizado entre os nucleotídeos 663.877 e 677.468, e possui sete éxons que codificam a proteína transmembranar 18, uma proteína com 140 aminoácidos e que se localiza na membrana do núcleo (Rask-Andersen et al., 2015).

A proteína TMEM18 apresenta três hélices transmembranares, propriedade relativamente incomum entre as proteínas da membrana humana, uma vez que menos de 5% delas partilham esta característica (Almén et al., 2010).

A proteína TMEM18 é expressa de forma ubíqua, porém com certas diferenças entre os tecidos. Os níveis mais elevados da proteína TMEM18 foram detectados no cérebro, músculo esquelético, fígado e tecido adiposo (Bernhard et al., 2013; Graff et al., 2017).

Estudos indicam variadas propriedades funcionais da proteína TMEM18 (Graff et al., 2017; Jurvansuu; Goldman, 2011; Rohde et al., 2014).

Em 2011, Jurvansuu e Goldman (2011) identificaram que a TMEM18 funciona como uma proteína de ligação ao DNA por meio de motivos específicos na sequência, permitindo potencializar a repressão ou ativação da transcrição de genes (Jurvansuu; Goldman, 2011), podendo afetar genes conhecidos por influenciar os níveis de glicose

em jejum, como por exemplo os genes: *MTNR1B*, *GCK* e *G6PC2* (Manning et al., 2012; Rohde et al., 2014).

Rask-Andersen e colaboradores (2012) apontaram a existência de correlação entre a expressão genética da proteína TMEM18 no córtex pré-frontal com o peso corporal em ratos, sugerindo assim um papel importante em funções cognitivas superiores relacionadas com o comportamento alimentar (Rask-Andersen et al., 2012).

Além disso, foi evidenciado que a proteína TMEM18 é abundantemente expressa no cérebro, principalmente no hipotálamo, região responsável pelo controle da homeostase energética e controle da saciedade (Bernhard et al., 2013; Srivastava et al., 2017; Rask-Andersen et al., 2015, 2012; Zandoná et al., 2017), concluindo-se que a função da proteína TMEM18 pode afetar vários genes ou vias metabólicas com impacto na homeostase da glicose ou na distribuição da gordura (Graff et al., 2017; Jurvansuu; Goldman, 2011; Kalnina et al., 2013; Rask-Andersen et al., 2015).

A partir dessas informações, diversos estudos tem associado tanto o gene *TMEM18* quanto alguns SNPs desse gene com a susceptibilidade à obesidade e à diabetes tipo 2, tanto em adultos quanto em crianças, em diversas populações, incluindo associação com ancestralidade africana e o aumento de IMC (Abadi et al., 2016; Almén et al., 2010; Bernhard et al., 2013; Gong et al., 2013; Graff et al., 2017; Jurvansuu; Goldman, 2011; Kalnina et al., 2013; Liu et al., 2016; Srivastava et al., 2017; Rask-Andersen et al., 2015; Zandoná et al., 2017).

Nosso resultado permite levantar a hipótese de que a associação do gene *TMEM18* com susceptibilidade à hipertensão na ancestralidade africana em pacientes com IC pode ser consequência da associação desse gene ao risco para obesidade, já que a obesidade aumenta o risco de hipertensão devido as alterações na hemodinâmica, prejuízo na homeostasia do sódio, disfunção renal, desequilíbrio do sistema nervoso autônomo, alterações endócrinas, estresse oxidativo e lesão vascular (Susic; Varagic, 2017). Entretanto, o gene *TMEM18* também pode contribuir para o desenvolvimento de hipertensão por outras vias que conduzem ao aumento do IMC.

Após encontrar em nosso estudo a região chr2p25.3 com super-representação na ancestralidade africana na hipertensão foi realizado o mapeamento fino com total 7.682 SNPs, contendo os 25 SNPs genotipados pelo nosso estudo juntamente com 7.657 SNPs imputados do projeto 1000 Genomas das populações africanas (LWK e

YRI), com MAF >1%. A associação dos SNPs na ancestralidade africana foi avaliada em uma regressão logística com ajuste de sexo e idade, entretanto nenhum SNPs apresentou diferença estatisticamente significativa (Gráfico 10).

Gráfico 10: Mapeamento fino com 7.682 SNPs da região chr2p25.3 na ancestralidade africana.

Cada SNP é representado como um ponto no contexto de sua localização genômica e é colorido com base no fato de ser diretamente genotipado (vermelho) ou imputado (azul). A posição genômica é dada de acordo com as coordenadas do projeto Genoma humano *Genome Reference Consortium Human Build 37* (GRCh37). Os picos em preto representam a taxa de recombinação na região de acordo com a posição no cromossomo. A associação foi definida com $-\log_{10} \geq 4$ indicada por uma linha vermelha tracejada, $p \leq 1,78 \times 10^{-2}$.

Apesar de não encontrar nenhum SNPs com diferença significativa nessa região como sendo o responsável pelo pico na análise de *admixture mapping*, muitas variantes genéticas já foram identificados e associadas à ancestralidade africana e à hipertensão, mas as associações resultantes são de efeito modesto, e poucos desses achados foram replicados com sucesso em outros estudos (Lima-Costa et al., 2016; Shetty et al., 2012; Tang et al., 2006; Zhu et al., 2005). A dificuldade em replicar achados significativos ilustra os desafios na identificação de variantes genéticas que afetam a pressão arterial na ancestralidade africana, principalmente devido ao fato da hipertensão ser uma doença multifatorial, e muito influenciada por fatores ambientais, socioeconômicos (Lima-Costa et al., 2016).

Deste modo, nossos resultados sugerem que o *admixture mapping* pode ser uma ferramenta de geração de hipóteses útil para identificar regiões genômicas que contribuem para doenças complexas em populações geneticamente miscigenadas. Esses dados podem contribuir para a compreensão da etiologia genética da hipertensão, na ancestralidade africana em pacientes com insuficiência cardíaca, e servem como ponto de partida para futuros estudos de sequenciamento e estudos moleculares para entender melhor os processos biológicos que levam à hipertensão e outras doenças complexas, como a IC.

4.3.4 Mortalidade

A análise de *admixture mapping*, em relação a mortalidade, foi realizada com o modelo caso-controle utilizando a ancestralidade local dos grupos, sendo que foram considerados no grupo de casos (n=248), pacientes em óbito no final do período de 4 anos de avaliação, e o grupo de controles (n=244), incluiu os demais indivíduos que ainda estavam livres desse evento no final do tempo de acompanhamento.

Inicialmente foi avaliado se alguma etiologia específica estaria associada com a mortalidade por IC (Tabela 7). Os resultados não mostraram nenhuma associação.

Tabela 7: Análise das etiologias nos dados de mortalidade por IC.

Etiologias	Total, n (%)	Óbitos totais, n (%)	p-valor*
Hipertensiva	142 (28,9)	59 (23,8)	0,063
Isquêmica	141 (28,7)	76 (30,6)	
Valvar	72 (14,6)	39 (15,7)	
Chagásica	59 (12,0)	39 (15,7)	
Outras	78 (15,9)	35 (14,1)	
Total	492	248	

* Nível de significância valor de $p \leq 0,05$.

Em seguida, foi analisada a estratificação populacional da amostra, utilizando os dados de ancestralidade global dos grupos de casos e controles (Tabela 8).

Tabela 8: Análise das ancestralidades globais do grupo de casos e controles.

Ancestralidades	Casos (média±DP)	Controles (média±DP)	p-valor*
Europeia	0,631±0,222	0,601±0,213	0,1269
Africana	0,244±0,231	0,255±0,214	0,5841
Ameríndia	0,125±0,107	0,144±0,120	0,0643

DP = desvio padrão;

* Nível de significância valor de $p \leq 0,05$.

A estratificação populacional refere-se à presença de uma diferença de frequências alélicas entre os casos e controles de um estudo, devido às diferenças de ancestralidade entre esses indivíduos. A estratificação é uma das principais fontes de resultados espúrios e de baixa reprodutibilidade em estudos caso-controle, podendo causar resultados falso-positivos ou falso-negativos. Esse tipo de resultado ocorre com frequência em estudos de associação em populações muito miscigenadas (Enoch et al., 2006; Hoggart et al., 2003; Lins et al., 2011).

Como os resultados não revelaram associação entre as etiologias e a mortalidade por IC ($p=0,063$, Tabela 7) e não evidenciou diferenças estatisticamente significativas entre os dois grupos, demonstrando a não estratificação populacional da amostra desse estudo (Tabela 8), foi possível realizar as análises de *admixture mapping* dos dados de mortalidade por IC.

Os resultados das análises de *admixture mapping* caso-controle estão representados nos gráficos Manhattan (Gráficos 11, 12 e 13).

Gráfico 11: Resultado do *admixture mapping* da mortalidade em pacientes com insuficiência cardíaca na ancestralidade europeia.

O círculo vermelho representa a região que apresentou diferença estatisticamente significativa na análise de *admixture mapping*

Gráfico 12: Resultado do *admixture mapping* da mortalidade em pacientes com insuficiência cardíaca na ancestralidade ameríndia.

Gráfico 13: Resultado do *admixture mapping* da mortalidade em pacientes com insuficiência cardíaca na ancestralidade africana.

Nessa análise de *admixture mapping* da mortalidade por IC observamos uma janela no cromossomo 6 com super-representação da ancestralidade local europeia ($p=6,805 \times 10^{-5}$) (Gráfico 11, Tabela 9).

Tabela 9: Resultado do *admixture mapping* caso-controle da ancestralidade europeia dos dados de mortalidade dos pacientes com IC.

Chr	Primeiro SNP	Último SNP	OR	95% IC	p-valor*
4	rs62308307	rs17426309	1.92	1.45- 2.53	$8,294 \times 10^{-3}$
5	rs9282594	rs4703457	1.37	1.44 - 2.48	$5,992 \times 10^{-3}$
6	rs80234897	rs80251478	2.37	1.48 - 2.58	$6,805 \times 10^{-5}$
6	rs1737844	rs723054	1.72	1.44 - 2.50	$9,031 \times 10^{-3}$

Chr = cromossomo; OR= *odds ratio*; 95% IC = Intervalo de confiança de 95%.

* Nível de significância valor de $p \leq 9,78 \times 10^{-4}$.

Para confirmar essa associação foi realizada uma análise de *admixture mapping* somente dos casos. O *admixture mapping* dos casos foi realizado pela comparação da ancestralidade local dos 248 pacientes com IC em óbito com a média da ancestralidade global dos mesmos pacientes.

O resultado dessa análise foi plotado juntamente com os dados da análise caso-controle, para a ancestralidade europeia (Gráfico 14). O motivo de realizar dois tipos diferentes de testes (estudo de caso e caso-controle) baseou-se no entendimento de que coincidências de sinais de ambos os testes em uma dada região proporcionaria maior probabilidade de identificar um resultado positivo verdadeiro, eliminando sinais falso-positivos. Os resultados do *admixture mapping* caso-controle e somente de casos estão apresentados no Gráfico 14, onde se pode observar a presença do sinal de super-representação da ancestralidade europeia na região chr6p22.3, tanto na análise somente de casos quanto na análise caso-controle. No pico chr6p22.3, os casos apresentam maior contribuição de ancestralidade europeia do que os controles (caso-controle Z score = -4.26, $p=6,805 \times 10^{-5}$).

Gráfico 14: *Admixture mapping* das análises caso-controle e somente de casos da ancestralidade europeia para a mortalidade em pacientes com IC.

O eixo X representa as posições genômicas por cromossomo. No eixo Y à esquerda estão os valores de $\log_{10}$ para análise somente de casos para cada região dos cromossomos (LOD score, linha vermelha). A análise de caso-controle é plotada no eixo Y secundário (linha azul) e é representada pelo Z score, valores do lado direito. Foi utilizado o nível de significância no LOD score valores de $-\log_{10} \geq 4$ e no Z score valores de ≤ -4 , indicados pelas linhas pretas.

Como mencionado anteriormente, cada janela possui 25 SNPs, e a lista dos SNPs da janela que apresentou diferença estatisticamente significativa para a mortalidade em pacientes com IC está presente na Tabela 10.

Tabela 10: Lista dos 25 SNPs presentes na janela que apresentou diferença estatística.

SNPs				
rs80234897	rs4716059	rs116599639	rs116325143	rs421172
rs72833450	rs56404385	rs179990	rs179959	rs9464904
rs7760359	rs16877822	rs112188181	rs7738608	rs35210572
rs75358491	rs6901150	rs13207978	rs1042391	rs2282840
rs6914805	rs179994	rs9383153	rs12194346	rs80251478

Essa janela apresenta 244.512 pb de tamanho e está localizada entre os nucleotídeos 16.196.746 e 16.441.258 do cromossomo 6, mapeada na região p22.3 (Figura 6).

Figura 6: Localização no cromossomo 6 da janela que apresentou diferença estatisticamente significativa no *admixture mapping* da ancestralidade europeia para mortalidade em pacientes com insuficiência cardíaca. Modificado de <https://ghr.nlm.nih.gov/chromosome/6#ideogram>.

Nessa região estão presentes 30,7% do gene *Ataxina 1* (*ATXN1*) e o gene *Guanosina Monofosfato Redutase* (*GMPR*, do inglês *Guanosine Monophosphate Reductase*).

O gene *ATXN1*, está localizado no segmento que compreende do nucleotídeo 16.299.112 ao 16.761.490. Dos 462.379 pb do gene *ATXN1*, somente 142.377 pb (30,7%) estão presentes na região que nosso estudo encontrou associação.

Esse gene codifica uma proteína com o mesmo nome do gene (*ATXN1*). A proteína *ATXN1* possui 815 aminoácidos e está presente de forma ubíqua nos tecidos, principalmente expressa no endométrio e cérebro (Fagerberg et al., 2014). Além disso, o gene *ATXN1* foi encontrado na gordura ectópica e apresentou um efeito funcional durante a diferenciação de adipócitos (Chu et al., 2017).

O outro gene presente na região que apresentou o pico no *admixture mapping* foi o gene *GMPR*.

O gene *GMPR* está localizado do nucleotídeo 16.238.580 ao 16.295.549 e codifica a enzima guanosina monofosfato (GMP) redutase que apresenta 345 aminoácidos (Kondoh et al., 1991; Spector; Jones; Miller, 1979).

Após encontrar a região chr6p22.3 com super-representação na ancestralidade europeia, foi realizado um mapeamento fino, com o objetivo de avaliar se algum SNP presente nessa região seria responsável pelo pico apresentado no *admixture mapping*.

Primeiramente foi imputado os genótipos de 7.891 SNPs do projeto 1000 Genomas das populações europeias (CEU e TSI), com MAF >1%, juntamente com os 25 SNPs genotipados pelo nosso estudo. Um total de 7.916 SNPs foram utilizados para o estudo de mapeamento fino. A associação dos SNPs na ancestralidade europeia foi avaliada em uma regressão logística com ajuste de sexo e idade (Gráfico 15).

Dos SNPs utilizados para o estudo de mapeamento fino da região chr6p22.3, dois apresentaram diferença estatisticamente significativa, sendo um genotipado (rs1042391), e o outro imputado (rs2142672) ($p=1,140 \times 10^{-4}$, $p=3,921 \times 10^{-4}$, respectivamente) (Gráfico 15, Tabela 11).

Gráfico 15: Mapeamento fino com 7.916 SNPs da região chr6p22.3 na ancestralidade europeia.

Cada SNP é representado como um ponto no contexto de sua localização genômica e é colorido com base no fato de ser diretamente genotipado (vermelho) ou imputado (azul). A posição genômica é dada de acordo com as coordenadas do projeto Genoma humano *Genome Reference Consortium Human Build 37* (GRCh37). A associação foi definida com $-\log_{10} \geq 4$ indicada por uma linha vermelha tracejada. Os picos em preto representam a taxa de recombinação na região de acordo com a posição no cromossomo.

Tabela 11: SNPs que apresentaram diferenças estatisticamente significante no mapeamento fino da região chr6p22.3 na ancestralidade europeia em pacientes com IC do dado de mortalidade.

SNP	OR	95% IC	p-valor*	Alelo de Risco	FAR	FAR		
						EUR	AFR	AME
rs1042391	2.11	1.82- 2.33	1,140x10⁻⁴	T	0,72	0,60 ^a	0,12 ^a	0,13 ^a
rs2142672	2.31	1.86- 2.48	3,921x10⁻⁴	C	0,81	0,73 ^b	0,26 ^b	0,61 ^b

FAR: frequência do alelo de risco; EUR: europeu; AFR: africano; AME: ameríndio.

a= http://www.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=6:16290030-16291030;v=rs1042391;vdb=variation;vf=754967

b= http://www.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=6:16196463-16197463;v=rs2142672;vdb=variation;vf=1553938

* Nível de significância valor de $p \leq 1,78 \times 10^{-2}$.

Em seguida foi testada a existência de associação entre esses dois SNPs e o risco de morte por IC (Tabela 12).

Tabela 12: Análise associação entre SNPs e o risco de morte por IC.

SNP	Genótipo	HR	95% IC	p-valor*
rs1042391	TA	1.10	0.93-1.37	0,069
	TT	1.21	1.15-1.87	0,0013
rs2142672	CT	1.46	1.01-1.91	0,056
	CC	1.51	1.21-2.31	0,0004

HR= *hazard ratio*.* Nível de significância valor de $p \leq 0,05$.

No SNP rs1042391, o alelo de risco (T) apresentou diferença estatisticamente significativa no mapeamento fino da região chr6p22.3 (OR: 2.11; $p=1,140 \times 10^{-4}$) e esse alelo em homozigose (TT) foi associado a um aumento de 21% ao risco de morte por IC (HR=1.21, $p=0,0013$) (Tabelas 11 e 12). O alelo de risco (T) em nosso estudo apresentou uma frequência maior (72%) do que nas populações europeias presentes no projeto 1000 Genomas (60%) (Tabela 11).

O SNP rs1042391 está mapeado no nucleotídeo 16.290.530 do cromossomo 6. Nesse SNP ocorre uma substituição da base A por T na posição 766 do gene *GMPR* (766A>T) resultando em uma substituição de isoleucina por fenilalanina no aminoácido 256 na proteína (Ile256Phe). A enzima GMPR é altamente expressa nos testículos, próstata e coração (Fagerberg et al., 2014).

A GMPR catalisa a desaminação redutiva irreversível de guanosina monofosfato (GMP) para formação de inosina monofosfato (IMP), de forma dependente de nicotinamida adenina dinucleotídeo fosfatado (NADPH, do inglês *Nicotinamide Adenine Dinucleotide Phosphate*) como doador de elétron (Kondoh et al., 1991; Li et al., 2006; Spector; Jones; Miller, 1979). Os nucleotídeos adenosina monofostato (AMP) e guanosina monofosfato (GMP) podem ser posteriormente sintetizados a partir do IMP.

A proteína também participa da via de salvamento de purinas reutilizando bases intracelulares livres, mantendo o equilíbrio entre as bases guanina e adenina. Esta enzima é inibida pela adenosina trifosfato (ATP, do inglês *Adenosine Triphosphate*) em concentrações fisiológicas, porém essa inibição é completamente reversível pela guanosina trifosfato (GTP, do inglês *Guanosine Triphosphate*), além da enzima ser inativada sob condições de escassez de bases púricas (Li et al., 2006; Mackenzie; Sorensen, 1973; Mager, Magasanik, 1960).

Estudos de GWAS associaram esse gene com a presença de lipoproteínas de baixa densidade colesterol (LDL, do inglês *Low Density Lipoprotein*) em níveis elevados, em populações de ancestralidade europeia (Chasman et al., 2009; Waterworth et al., 2010). O LDL colesterol em níveis elevados está associado a complicações cardiovasculares, como aumento da pressão arterial, infarto e insuficiência cardíaca (Rader; Hovingh, 2014).

Em relação ao SNP rs1042391, o único estudo até o momento que avaliou esse polimorfismo foi realizado em 1991 por Kondoh e colaboradores. Nesse estudo foram avaliadas amostras de DNA obtidas de 4 caucasianos e 5 indivíduos orientais, saudáveis e não relacionados, e foi observado que a variante T apresentou frequência de 30%. Ambas as enzimas, GMPR de tipo selvagem e a enzima GMPR variante (Ile256Phe), apresentaram atividade normal de GMP-redutase (Kondoh et al., 1991). O detalhamento das propriedades cinéticas da enzima GMPR variante (Ile256Phe) ainda não foi realizado.

Portanto, seria interessante uma melhor caracterização enzimática da GMP redutase, uma vez que ela está envolvida na interconversão de bases púricas, sendo responsável pela conversão e o balanceamento entre os nucleotídeos de guanina e adenina. A elucidação dos aspectos cinéticos, químicos e termodinâmicos da reação catalisada pela enzima ajudaria a compreender seu papel e sua importância no balanceamento das bases púricas e então sua relação com a mortalidade por IC.

No mapeamento fino na região chr6p22.3 o SNP rs2142672, alelo de risco (C), também apresentou diferença estatisticamente significativa (OR: 2.31; $p=3,921 \times 10^{-4}$). Esse alelo em homozigose (CC) foi associado a um aumento de 51% no risco de morte por IC (HR=1.51, $p=0,0004$) (Tabelas 11 e 12). A frequência do alelo de risco (C) em nosso estudo foi maior (81%) do que nas populações europeias presentes no projeto 1000 Genomas (73%) (Tabela 11).

O SNP rs2142672 está presente no nucleotídeo 16.196.963 do cromossomo 6. Nesse SNP ocorre uma substituição da base T por C, e está presente em uma região intergênica entre os genes *GMPR* e *MYLIP* (*myosin regulatory light chain interacting protein*).

Em estudos de GWAS já foram encontradas associações desse SNP com níveis aumentados de LDL colesterol e colesterol total, em populações de ancestralidade europeia, população de descendentes de asiáticos e indivíduos

mexicanos, sugerindo que essa variante poderia influenciar na concentração plasmática de lipídeos devido a função regulatória que essa variante poderia apresentar (Waterworth et al., 2010; Weissglas-Volkov et al., 2011).

Por ser um SNP presente em uma região intergênica, sua função no metabolismo de lipídeos ainda é desconhecida, bem como sua interação com a mortalidade em pacientes brasileiros com IC. Estudos funcionais ajudariam a identificar as funções dessa variante, ou pelo menos esclarecer o papel dessa variante na funcionalidade das proteínas que estão próximas a ela.

A IC é uma condição crônica relativamente comum, caracterizada pela incapacidade do coração de bombear o sangue eficientemente e representar um ônus significativo na saúde pública, afetando quase 6 milhões de brasileiros (Albuquerque et al., 2015). No Brasil, a insuficiência cardíaca é a terceira causa de morte entre as doenças cardiovasculares (Albuquerque et al., 2015). O prognóstico e a sobrevida são uma preocupação significativa: após a IC ser diagnosticada, 1 em cada 5 pacientes morre no prazo de um ano (Lloyd-Jones et al., 2009; Morrison et al., 2010).

Nesse estudo, executou-se uma abordagem de associação genômica por meio da técnica de *admixture mapping* para identificar ancestralidade e região genômica associadas com a doença e com a mortalidade em pacientes com IC. Alguns estudos que utilizaram a técnica de *admixture mapping* têm identificado diversas variantes genéticas relacionadas com diferentes doenças cardiovasculares, mas até o momento nenhum mostrou resultados relacionados à mortalidade (Divers et al., 2013; Gomez et al., 2015; Shendre et al., 2017a; 2017b; Shetty et al., 2015; Sofer et al., 2017; Tandon et al., 2015).

Por outro lado, alguns estudos de GWAS já identificaram genes e/ou variantes genéticas associado com mortalidade por IC.

Já foram identificados, em pacientes de ancestralidade europeia da população dos Estados Unidos, um SNP (rs12638540) intrônico presente na região chr3p22 associado a um aumento de 53% no risco de morte por IC; um SNP intergênico (rs2207418) da região chr20p12 foi associado com risco aumentado de 51% para a mortalidade por IC, em pacientes caucasianos pertencentes ao um grupo Amish; e a variante genética intergênica (rs9885413) na região chr5q22 foi associada com 36% de risco aumentado de morte por IC em pacientes da população americana (Morrison et al., 2010; Parsa et al., 2011; Smith et al., 2016).

Embora esses resultados tenham tido êxito, podemos observar que esses estudos realizaram o GWAS em população de ancestralidade europeia, com exceção do estudo de Morrison et al., 2010 que avaliou seus resultados também em indivíduos de ancestralidade africana, não encontrando nenhuma região ou gene associado à mortalidade nessa população. No entanto, nenhum desses estudos testou se a ancestralidade estaria associada à mortalidade por IC.

Percebe-se que em todos esses estudos, diferentes variantes foram associadas à doença, refletindo uma dificuldade em obter resultados expressivos e concordantes. Esse fato pode ser consequência de características da doença.

Claramente a IC é uma doença multifatorial, com diversos fatores etiológicos influentes e que apresenta herança poligênica (Rau; Lusi; Wang, 2015).

A contribuição significativa de fatores socioambientais no desenvolvimento e gravidade da IC também é evidente. A IC não familiar é tipicamente uma doença muito presente nos idosos e é resultado não apenas de fatores genéticos, mas também de fatores socioambientais, como tabagismo, dieta e estilo de vida sedentário. Essa heterogeneidade complica as tentativas de determinar os fatores genéticos envolvidos na etiologia da IC. Consequentemente, estudos com GWAS ou *admixture mapping* podem exigir coortes muito grandes de casos para sobrepujar o ruído produzido por esses fatores (Rau; Lusi; Wang, 2015).

Existe também a dificuldade na obtenção de características quantitativas da IC. A falta de medidas quantitativas específicas para a IC resulta na dificuldade para mensurar com precisão a gravidade da doença. A avaliação atual da gravidade da IC depende da medida em grande parte não-invasiva dos parâmetros da função ventricular, como a fração de ejeção ou a distância de caminhada de 6 minutos, ou das classificações funcionais da *New York Heart Association*, que são altamente variáveis e principalmente de natureza qualitativa. Tem sido sugerida a utilização de parâmetros ecocardiográficos como um caminho possível para estudos de GWAS, e consequentemente *admixture mapping*, na IC (Fox et al., 2013; Rau; Lusi; Wang, 2015; Vasan et al., 2009).

Até o momento, nosso estudo é o primeiro a relatar a associação de uma região (chr6p22.3) na ancestralidade europeia com a mortalidade em pacientes com IC, usando uma abordagem de *admixture mapping*, seguida do mapeamento fino nessa região. Com isso, foi possível observar duas variantes associadas (rs1042391,

rs2142672) com a mortalidade por IC na ancestralidade europeia. Estes são achados iniciais e, como tal, devem ser considerados como geradores de hipóteses. Entretanto, essas descobertas nos proporcionam a oportunidade de explorar algumas novas regiões e genes. Assim, estudos futuros são necessários para replicar as associações encontradas nesse estudo e explorar aplicações clínicas para os genes consistentemente associados a mortalidade em pacientes com IC.

5 CONCLUSÃO

5 CONCLUSÃO

As análises de *admixture mapping* em 492 pacientes com IC, com o uso de 182.090 SNPs, foram utilizadas na tentativa de identificar regiões e ancestralidades genômicas associadas à IC, seus fatores de risco (diabetes e hipertensão), e a mortalidade por IC. Em seguida, a análise de mapeamento fino foi utilizada para avaliar a possibilidade de cada variante presente na região onde houve associação ser a responsável pelo pico apresentado no *admixture mapping*.

Em relação à ancestralidade global, os 492 pacientes apresentaram maior proporção de ancestralidade europeia, seguida pela ancestralidade africana e ameríndia.

Até o momento, nosso estudo é o primeiro a relatar associação de uma região no cromossomo 2 (p25.3) na ancestralidade africana com a hipertensão em pacientes com IC, usando a abordagem de *admixture mapping*.

Também é inédito o relato de associação de uma região no cromossomo 6 (p22.3), na ancestralidade europeia com a mortalidade por IC. Por meio do mapeamento fino foi evidenciado que o alelo em homozigose TT do SNP rs1042391 está associado a um aumento de 21% ao risco de morte por IC, e que o alelo em homozigose CC do SNP rs2142672 está associado a um aumento de 51% ao risco de morte por IC.

Estes são achados iniciais e, como tal, devem ser considerados como geradores de hipóteses. Apesar disso, ficou demonstrada a utilidade do estudo de *admixture mapping* para identificar regiões com diferentes ancestralidades genômicas que podem contribuir para o risco de doenças complexas em populações geneticamente miscigenadas.

Esses dados podem contribuir para a compreensão da etiologia genética da hipertensão e da mortalidade em pacientes com IC, relacionados à ancestralidade, podendo servir como ponto de partida para futuros estudos funcionais na tentativa de aprofundar os conhecimentos dos processos biológicos que levam à hipertensão e a IC. Estudos futuros são necessários para replicar as associações encontradas, e explorar aplicações clínicas para as regiões de genes consistentemente associadas a mortalidade em pacientes com IC e à hipertensão.

Essas identificações poderiam ser úteis para prevenção, definição de grupos de alto risco e guiar o tratamento na IC com objetivo de fornecer melhor qualidade de vida a esses indivíduos.

6 REFERÊNCIAS

6 REFERÊNCIAS¹

- Abadi A, Peralta-Romero J, Suarez F, Gomez-Zamudio J, Burguete-Garcia AI, Cruz M, et al. Assessing the effects of 35 European-derived BMI-associated SNPs in Mexican children. *Obesity (Silver Spring)*. 2016;24(9):1989-95.
- Albuquerque DC, Souza NJD, Bacal F, Rohde LEP, Bernardez-Pereira S, Berwanger O, et al. I Registro Brasileiro de Insuficiência Cardíaca – Aspectos Clínicos, Qualidade Assistencial e Desfechos Hospitalares. *Arq. Bras. Cardiol*. 2015;104(6):433-442.
- Almén MS, Jacobsson JA, Shaik JH, Olszewski PK, Cedernaes J, Alsiö J, et al. The obesity gene, TMEM18, is of ancient origin, found in majority of neuronal cells in all major brain regions and associated with obesity in severely obese children. *BMC Med Genet*. 2010;11:58.
- Alves-Silva J, da Silva Santos M, Guimarães PE, Ferreira AC, Bandelt HJ, Pena SD, et al. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 2000;67(2):444-61.
- Bahrami H, Kronmal R, Bluemke DA, Olson J, Shea S, Liu K, et al. Differences in the incidence of congestive heart failure by ethnicity: the multi-ethnic study of atherosclerosis. *Arch Intern Med*. 2008;168(19):2138-45.
- Bansal V, Libiger O. Fast individual ancestry inference from DNA sequence data leveraging allele frequencies for multiple populations. *BMC Bioinformatics*. 2015;16(1):4.
- Basu A, Tang H, Arnett D, Gu CC, Mosley T, Kardia S, et al. Admixture mapping of quantitative trait loci for BMI in African Americans: evidence for loci on chromosomes 3q, 5q, and 15q. *Obesity (Silver Spring)*. 2009;17(6):1226-31.
- Bensen JT, Xu Z, McKeigue PM, Smith GJ, Fontham ET, Mohler JL, et al. Admixture mapping of prostate cancer in African Americans participating in the North Carolina-Louisiana Prostate Cancer Project (PCaP). *Prostate*. 2014;74(1):1-9.
- Bernardez-Pereira S, Gioli-Pereira L, Marcondes-Braga FG, Santos PC, Spina JM, Horimoto AR, et al. Genomic ancestry as a predictor of haemodynamic profile in heart failure. *Open Heart*. 2016;3(2):e000434.
- Bernhard F, Landgraf K, Klötting N, Berthold A, Büttner P, Friebe D, et al. Functional relevance of genes implicated by obesity genome-wide association study signals for human adipocyte biology. *Diabetologia*. 2013;56(2):311-22.
- Bibbins-Domingo K, Pletcher MJ, Lin F, Vittinghoff E, Gardin JM, Arynchyn A, et al. Racial differences in incident heart failure among young adults. *N Engl J Med*. 2009;360(12):1179-90.
- Bocchi EA, Marcondes-Braga FG, Ayub-Ferreira SM, Rohde LE, Oliveira WA, Almeida, DR et al. Sociedade Brasileira de Cardiologia. III Diretriz Brasileira de Insuficiência Cardíaca Crônica. *Arq Bras Cardiol*. 2009;93(1):1-71.

¹ De acordo com Estilo Vancouver

Bocchi EA. Heart Failure in South America. *Curr Cardiol Rev.* 2013;9(2):147-56.

Bustamante CD, Burchard EG, De la Veja, FM. Genomics for the world. *Nature.* 2011;475(7355):163-5.

Cappola TP, Li M, He J, Ky B, Gilmore J, Qu L, et al. Common variants in HSPB7 and FRMD4B associated with advanced heart failure. *Circ Cardiovasc Genet.* 2010;3(2):147-54.

Cardena MMSGC. Avaliação da relação entre haplogrupo mitocondrial e ancestralidade genômica no desenvolvimento de insuficiência cardíaca em amostra brasileira [dissertação]. São Paulo: Faculdade de Medicina da Universidade de São Paulo; 2013.

Cardena MM, Ribeiro-Dos-Santos A, Santos S, Mansur AJ, Pereira AC, Fridman C. Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. *PLoS One.* 2013;8(4):e62005.

Cardena MM, Ribeiro-dos-Santos A, Santos S, Mansur AJ, Pereira AC, Fridman C. Amerindian genetic ancestry is associated with higher survival rates compared to African and European ancestry in Brazilian patients with heart failure. *Int J Cardiol.* 2014;176(2):527-8.

Cardena MMSG, Mansur AJ, Pereira AC, Fridman C. Y haplotype variability in a population of SÃO Paulo state, Brazil. *Forensic Sci Int Genet.* 2016;23:e22-e23.

Cardena MM, Ribeiro-Dos-Santos AK, Santos SE, Mansur AJ, Bernardez-Pereira S, Santos PC, et al. Mitochondrial and genomic ancestry are associated with etiology of heart failure in Brazilian patients. *J Hum Hypertens.* 2016;30(2):120-3.

Carvalho-Silva DR, Santos FR, Rocha J, Pena SD. The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet.* 2001;68(1):281-6.

Carvalho-Silva DR, Tarazona-Santos E, Rocha J, Pena SD, Santos FR. Y chromosome diversity in Brazilians: switching perspectives from slow to fast evolving markers. *Genetica.* 2006;126(1-2):251-60.

Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE. Caucasian genes in American blacks: new data. *Am J Hum Genet.* 1992;50(1):145-55.

Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet.* 2009;5(11):e1000730.

Cheng CY, Kao WH, Patterson N, Tandon A, Haiman CA, Harris TB, et al. Admixture mapping of 15,280 African Americans identifies obesity susceptibility loci on chromosomes 5 and X. *PLoS Genet.* 2009;5(5):e1000490.

Cheng CY, Reich D, Wong TY, Klein R, Klein BE, Patterson N, et al. Admixture mapping scans identify a locus affecting retinal vascular caliber in hypertensive African Americans: the Atherosclerosis Risk in Communities (ARIC) study. *PLoS Genet.* 2010 ;6(4):e1000908.

Cheng CY, Reich D, Haiman CA, Tandon A, Patterson N, Selvin E, et al. African ancestry and its correlation to type 2 diabetes in African Americans: a genetic admixture analysis in three U.S. population cohorts. *PLoS One.* 2012;7(3):e32840.

Chen M, Yang C, Li C, Hou L, Chen X, Zhao H. Admixture mapping analysis in the context of GWAS with GAW18 data. *BMC Proc.* 2014;8(Suppl 1):S3.

Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA.* 2003;289(19):2560-72.

Chu AY, Deng X, Fisher VA, Drong A, Zhang Y, Feitosa MF, et al. Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation. *Nat Genet.* 2017;49(1):125-130.

Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. *Nat Protoc.* 2011;6(2):121-33.

Darvasi A, Shifman S. The beauty of admixture. *Nat Genet.* 2005;37(2):118-9.

Datasus, 2017. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/niuf.def>. Acessado em 05/04/2018.

Davenport C, Cheng EY, Kwok YT, Lai AH, Wakabayashi T, Hyde C, et al. Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis. *Br J Gen Pract.* 2006;56(522):48-56.

Deo RC, Patterson N, Tandon A, McDonald GJ, Haiman CA, Ardlie K, et al. A high-density admixture scan in 1,670 African Americans with hypertension. *PLoS Genet.* 2007;3(11):e196.

Deniz E, Erman B. Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Funct Integr Genomics.* 2017;17:135-143.

Ding L, Wiener H, Abebe T, Altaye M, Go RC, Kercsmar C, et al. Comparison of measures of marker informativeness for ancestry and admixture mapping. *BMC Genomics.* 2011;12:622.

Divers J, Palmer ND, Lu L, Register TC, Carr JJ, Hicks PJ, et al. Admixture mapping of coronary artery calcified plaque in African Americans with type 2 diabetes mellitus. *Circ Cardiovasc Genet.* 2013;6(1):97-105.

Elbein SC, Das SK, Hallman DM, Hanis CL, Hasstedt SJ. Genome-wide linkage and admixture mapping of type 2 diabetes in African American families from the American Diabetes Association GENNID (Genetics of NIDDM) Study Cohort. *Diabetes*. 2009;58(1):268-74.

ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007;447(7146):799-816.

Enoch MA, Shen PH, Xu K, Hodgkinson C, Goldman D. Using ancestry-informative markers to define populations and detect population stratification. *J Psychopharmacol*. 2006;20(4 Suppl):19-26.

Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics*. 2014;13(2):397-406.

Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337-43.

Fejerman L, Chen GK, Eng C, Huntsman S, Hu D, Williams A, et al. Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Hum Mol Genet*. 2012;21(8):1907-17.

Fox ER, Musani SK, Barbalic M, Lin H, Yu B, Ogunyankin KO, et al. Genome-wide association study of cardiac structure and systolic function in African Americans: the Candidate Gene Association Resource (CARE) study. *Circ Cardiovasc Genet*. 2013;6(1):37-46.

Franceschini N, Reiner AP, Heiss G. Recent findings in the genetics of blood pressure and hypertension traits. *Am J Hypertens*. 2011;24(4):392-400.

Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet*. 2007;8:657–662.

Fridman C, Gonzalez RS, Pereira AC, Cardena MM. Haplotype diversity in mitochondrial DNA hypervariable region in a population of southeastern Brazil. *Int J Legal Med*. 2014;128(4):589-93.

Galanter JM, Gignoux CR, Torgerson DG, Roth LA, Eng C, Oh SS, et al. Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. *J Allergy Clin Immunol*. 2014;134(2):295-305.

Gallardo ME, García-Pavía P, Chamorro R, Vázquez ME, Gómez-Bueno M, Millán I, et al. Mitochondrial haplogroups associated with end-stage heart failure and coronary allograft vasculopathy in heart transplant patients. *Eur Heart J*. 2012;33(3):346-53.

Gauí EN, Oliveira GMM, Klein CH. Mortalidade por Insuficiência Cardíaca e Doença Isquêmica do Coração no Brasil de 1996 a 2011. *Arq Bras Cardiol.* 2014;102(6):557-565.

Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, et al. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation.* 2014;129(3):e28-e292.

Gomez F, Wang L, Abel H, Zhang Q, Province MA, Borecki IB. Admixture mapping of coronary artery calcification in African Americans from the NHLBI family heart study. *BMC Genet.* 2015;16:42.

Gong J, Schumacher F, Lim U, Hindorff LA, Haessler J, Buyske S, et al. Fine Mapping and Identification of BMI Loci in African Americans. *Am J Hum Genet.* 2013;93(4):661-71.

Gordon HS, Nowlin PR, Maynard D, Berbaum ML, Deswal A. Mortality after hospitalization for heart failure in blacks compared to whites. *Am J Cardiol.* 2010;105(5):694-700.

Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet.* 2006;38(3):320-3.

Graff M, North KE, Richardson AS, Young KL, Mazul AL, Highland HM, et al. BMI loci and longitudinal BMI from adolescence to young adulthood in an ethnically diverse cohort. *Int J Obes (Lond).* 2017;41(5):759-768.

Hellenthal G, Busby GBJ, Band G, Wilson JF, Capelli C, Falush D, et al. A genetic atlas of human admixture history. *Science.* 2014;343(6172):747-751.

Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, et al. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet.* 2003;72(6):1492-1504.

Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM. Design and analysis of admixture mapping studies. *Am J Hum Genet.* 2004;74(5):965-78.

Hohman TJ, Cooke-Bailey JN, Reitz C, Jun G, Naj A, Beecham GW, et al. Global and local ancestry in African-Americans: Implications for Alzheimer's disease risk. *Alzheimers Dement.* 2016;12(3):233-43.

Howie B, Marchini J, Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda).* 2011;1(6): 457-470.

Hormozdiari F, Kostem E, Kang EY, Pasaniuc B, Eskin E. Identifying causal variants at loci with multiple signals of association. *Genetics.* 2014;198(2):497-508.

Hunt SA; American College of Cardiology; American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure). ACC/AHA 2005 Guideline Update for the Diagnosis and Management of Chronic Heart Failure in the Adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. *J Am Coll Cardiol.* 2005;46(6):e1-82

Instituto Brasileiro de Estatística (IBGE). 2007. Brasil: 500 anos de povoamento. Rio de Janeiro: IBGE. p 237.

Jeff JM, Armstrong LL, Ritchie MD, Denny JC, Kho AN, Basford MA, et al. Admixture mapping and subsequent fine-mapping suggests a biologically relevant and novel association on chromosome 11 for type 2 diabetes in African Americans. *PLoS One.* 2014;9(3):e86931.

Jurvansuu JM, Goldman A. Obesity risk gene TMEM18 encodes a sequence-specific DNA-binding protein. *PLoS One.* 2011;9:e25317,

Kalnina I, Zaharenko L, Vaivade I, Rovite V, Nikitina-Zake L, Peculis R, et al. Polymorphisms in FTO and near TMEM18 associate with type 2 diabetes and predispose to younger age at diagnosis of diabetes. *Gene.* 2013;527(2):462-8.

Kehdy FS, Gouveia MH, Machado M, Magalhães WC, Horimoto A, Horta BL, 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):8696-701.

Kichaev G, Yang WY, Lindstrom S, Hormozdiari F, Eskin E, Price AL, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet.* 2014;10(10):e1004722.

Kondoh T, Kanno H, Chang LF, Yoshida A. Identification of common variant alleles of the human guanosine monophosphate reductase gene. *Hum Genet.* 1991;88(2):225-7.

Leite TK, Fonseca RM, de França NM, Parra EJ, Pereira RW. Genomic ancestry, self-reported "color" and quantitative measures of skin pigmentation in Brazilian admixed siblings. *PLoS One.* 2011;6(11):e27162.

Li J, Wei Z, Zheng M, Gu X, Deng Y, Qiu R, et al. Crystal structure of human guanosine monophosphate reductase 2 (GMPT2) in complex with GMP. *J Mol Biol.* 2006;355(5):980-8.

Lins TC, Vieira RG, Abreu BS, Gentil P, Moreno-Lima R, Oliveira RJ, et al. Genetic heterogeneity of self-reported ancestry groups in an admixed Brazilian population. *J Epidemiol.* 2011;21(4):240-5.

Lima-Costa MF, Mambrini JV, Leite ML, Peixoto SV, Firmo JO, Loyola Filho AI, et al. Socioeconomic Position, But Not African Genomic Ancestry, Is Associated With Blood Pressure in the Bambui-Epigen (Brazil) Cohort Study of Aging. *Hypertension*. 2016;67(2):349-55.

Liu S, Wilson JG, Jiang F, Griswold M, Correa A, Mei H. Multi-variant study of obesity risk genes in African Americans: The Jackson Heart Study. *Gene*. 2016;593(2):315-21.

Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, et al. Heart disease and stroke statistics--2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2009;119(3):480-6.

Mackenzie JJ, Sorensen LB. Guanosine 5'-phosphate reductase of human erythrocytes. *Biochim. Biophys. Acta*. 1973;327(2):282-294.

Mager J, Magasanik B. Guanosine 5'-phosphate reductase and its role in the interconversion of purine nucleotides. *J. Biol Chem*. 1960;235:1474-1478.

Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 2012;44(6):659-69.

Manta FS, Pereira R, Caiafa A, Silva DA, Gusmão L, Carvalho EF. Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. *Ann Hum Biol*. 2013;40(1):94-8.

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9(5):356-69.

McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. *N Engl J Med*. 1971;285(26):1441-6.

McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet*. 2000;64(Pt 2):171-86.

McKeigue PM. Prospects for admixture mapping of complex traits. *Am J Hum Genet*. 2005;76(1):1-7.

Mersha TB. Mapping asthma-associated variants in admixed populations. *Front Genet*. 2015;6:292.

Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics*. 2015;9:1.

- Moreno-Estrada A, Gignoux CR, Fernández-López JC, Zakharia F, Sikora M, et al. Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science*. 2014;344(6189):1280-5.
- Morris AP. Fine mapping of type 2 diabetes susceptibility loci. *Curr Diab Rep*. 2014;14(11):549.
- Morrison AC, Felix JF, Cupples LA, Glazer NL, Loehr LR, Dehghan A, et al. Genomic variation associated with mortality among adults of European and African ancestry with heart failure: the cohorts for heart and aging research in genomic epidemiology consortium. *Circ Cardiovasc Genet*. 2010;3(3):248-55.
- Morton NE. Sequential tests for the detection of linkage. *Am J Hum Genet*. 1955;7(3):277-318.
- Mosterd A, Hoes AW. Clinical epidemiology of heart failure. *Heart*. 2007;93(9):137-146.
- Moura RR, Coelho AV, Balbino Vde Q, Crovella S, Brandão LA. Meta-analysis of Brazilian genetic admixture and comparison with other Latin America countries. *Am J Hum Biol*. 2015;27(5):674-80.
- Ng MC. Genetics of Type 2 Diabetes in African Americans. *Curr Diab Rep*. 2015;15(10):74.
- Okin PM, Kjeldsen SE, Dahlöf B, Devereux RB. Racial differences in incident heart failure during antihypertensive therapy. *Circ Cardiovasc Qual Outcomes*. 2011;4(2):157-64.
- Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet*. 2015;6(5):275–284.
- Padhukasahasram B. Inferring ancestry from population genomic data and its applications. *Front Genet*. 2014;5:204.
- Palha T, Ribeiro-Rodrigues E, Ribeiro-dos-Santos A, Santos S. Fourteen short tandem repeat loci Y chromosome haplotypes: Genetic analysis in populations from northern Brazil. *Forensic Sci Int Genet*. 2012;6(3):413-8.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci U S A*. 2003;100(1):177-82.
- Parsa A, Chang YP, Kelly RJ, Corretti MC, Ryan KA, Robinson SW, et al. Hypertrophy-associated polymorphisms ascertained in a founder cohort applied to heart failure risk and mortality. *Clin Transl Sci*. 2011;4(1):17-23.
- Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, et al. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet*. 2004;74(5):979-1000.

Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy Fde S, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One*. 2011;6(2):e17063.

Pereira SB, Velloso MW, Chermont S, Quintão MM, Nunes Abdhala R, Giro C, et al. β -adrenergic receptor polymorphisms in susceptibility, response to treatment and prognosis in heart failure: implication of ethnicity. *Mol Med Rep*. 2013;7(1):259-65.

Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Hooshier Kashani B, et al. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. *Genome Res*. 2010;20(9):1174-9.

Picon RV, Fuchs FD, Moreira LB, Riegel G, Fuchs SC. Trends in prevalence of hypertension in Brazil: a systematic review with meta-analysis. *PLoS One*. 2012;7(10):e48255.

Piluso G, Mirabella M, Ricci E, Belsito A, Abbondanza C, Servidei S, et al. Gamma1- and gamma2-syntrophins, two novel dystrophin-binding proteins localized in neuronal cells. *J Biol Chem*. 2000;275(21):15851-60.

Qu Z, Adelson DL. Evolutionary conservation and functional roles of ncRNA. *Front Genet*. 2012;9(3):205.

Rader DJ, Hovingh, KG. HDL and cardiovascular disease. *The Lancet*. 2014;384(9943):618-625.

Rask-Andersen M, Jacobsson JA, Moschonis G, Chavan RA, Sikder MA, Allzén E, et al. Association of TMEM18 variants with BMI and waist circumference in children and correlation of mRNA expression in the PFC with body weight in rats. *Eur J Hum Genet*. 2012;20(2):192-7.

Rask-Andersen M, Sällman Almén M, Jacobsson JA, Ameer A, Moschonis G, Dedoussis G, et al. Determination of obesity associated gene variants related to TMEM18 through ultra-deep targeted re-sequencing in a case-control cohort for pediatric obesity. *Genet Res (Camb)*. 2015;97:e16.

Rau CD, Lusis AJ, Wang Y. Genetics of common forms of heart failure: challenges and potential solutions. *Curr Opin Cardiol*. 2015;30(3):222-227.

Reed TE. Caucasian genes in American Negroes. *Science*. 1969;165(3895):762-768.

Resque R, Gusmão L, Geppert M, Roewer L, Palha T, Alvarez L, et al. Male Lineages in Brazil: Intercontinental Admixture and Stratification of the European Background. *PLoS One*. 2016;11(4):e0152573.

Richardson P, McKenna W, Bristow M, Maisch B, Mautner B, O'Connell J, et al. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. *Circulation*. 1996;93(5):841-2.

Rio M, Royer G, Gobin S, de Blois MC, Ozilou C, Bernheim A, et al. Monozygotic twins discordant for submicroscopic chromosomal anomalies in 2p25.3 region detected by array CGH. *Clin Genet*. 2013;84(1):31-6.

Rohde K, Keller M, Klös M, Schleinitz D, Dietrich A, Schön MR, et al. Adipose tissue depot specific promoter methylation of TMEM18. *J Mol Med (Berl)*. 2014;92(8):881-8.

Salzano FM, Sans M. Interethnic admixture and the evolution of Latin American populations. *Genet Mol Biol*. 2014;37(1):151-170.

Schaan AP, Costa L, Santos D, Modesto A, Amador M, Lopes C, et al. mtDNA structure: the women who formed the Brazilian Northeast. *BMC Evol Biol*. 2017 9;17(1):185.

Scherer ML, Nalls MA, Pawlikowska L, Ziv E, Mitchell G, Huntsman S, et al. Admixture mapping of ankle-arm index: identification of a candidate locus associated with peripheral arterial disease. *J Med Genet*. 2010;47(1):1-7.

Schmidt MI, Hoffmann JF, de Fátima Sander Diniz M, Lotufo PA, Griep RH, Bensenor IM, et al. High prevalence of diabetes and intermediate hyperglycemia - The Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *Diabetol Metab Syndr*. 2014;6:123.

Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. 2007;316(5829):1341-5.

Shendre A, Wiener H, Irvin MR, Zhi D, Limdi NA, Overton ET, et al. Admixture Mapping of Subclinical Atherosclerosis and Subsequent Clinical Events Among African Americans in 2 Large Cohort Studies. *Circ Cardiovasc Genet*. 2017a;10(2). pii: e001569.

Shendre A, Irvin MR, Wiener H, Zhi D, Limdi NA, Overton ET, et al. Local Ancestry and Clinical Cardiovascular Events Among African Americans From the Atherosclerosis Risk in Communities Study. *J Am Heart Assoc*. 2017b;6(4).pii: e004739.

Shetty PB, Tang H, Tayo BO, Morrison AC, Hanis CL, Rao DC, et al. Variants in CXADR and F2RL1 are associated with blood pressure and obesity in African-Americans in regions identified through admixture mapping. *J Hypertens*. 2012;30(10):1970-6.

Shetty PB, Tang H, Feng T, Tayo B, Morrison AC, Kardina SL, et al. Variants for HDL-C, LDL-C, and triglycerides identified from admixture mapping and fine-mapping analysis in African American families. *Circ Cardiovasc Genet*. 2015;8(1):106-13.

Shriner D, Adeyemo A, Ramos E, Chen G, Rotimi CN. Mapping of disease-associated variants in admixed populations. *Genome Biol*. 2011;12(5):223.

Shriner D. Overview of admixture mapping. *Curr Protoc Hum Genet*. 2013;Chapter 1:Unit 1.23.

Shriner D, Kumkhaek C, Doumatey AP, Chen G, Bentley AR, Charles BA, et al. Evolutionary context for the association of γ -globin, serum uric acid, and hypertension in African Americans. *BMC Med Genet*. 2015;16:103.

Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, et al. Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet*. 1997;60(4):957-64.

Simão F, Ferreira AP, de Carvalho EF, Parson W, Gusmão L. Defining mtDNA origins and population stratification in Rio de Janeiro. *Forensic Sci Int Genet*. 2018;34:97-104.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007;445:881–885.

Smith N, Felix JF, Morrison AC, Demissie S, Glazer NL, Loehr LR, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet*. 2010;3(3):256-66.

Smith JG, Felix JF, Morrison AC, Kalogeropoulos A, Trompet S, Wilk JB, et al. Discovery of Genetic Variation on Chromosome 5q22 Associated with Mortality in Heart Failure. *PLoS Genet*. 2016;12(5):e1006034.

Sociedade Brasileira de Diabetes. Conduta Terapêutica no Diabetes Tipo 2 – Algoritmo SBD 2014. Posicionamento Oficial SBD nº 01/2014. <http://www.diabetes.org.br/publico/images/pdf/conduta-terapeutica-no-dm2-algoritmo-sbd-2014-versao-final-impressao.pdf>

Sofer T, Baier LJ, Browning SR, Thornton TA, Talavera GA, Wassertheil-Smoller S, et al. Admixture mapping in the Hispanic Community Health Study/Study of Latinos reveals regions of genetic associations with blood pressure traits. *PLoS One*. 2017;12(11):e0188400.

Spain SL, Barrett JC. Strategies for fine-mapping complex traits. *Hum Mol Genet*. 2015;24(R1):R111-119.

Spector T, Jones TE, Miller RL. Reaction mechanism and specificity of human GMP reductase. Substrates, inhibitors, activators, and inactivators. *J Biol. Chem*. 1979;254(7): 2308-2315.

Srivastava A, Mittal B, Prakash J, Srivastava P, Srivastava N, Srivastava N. A multianalytical approach to evaluate the association of 55 SNPs in 28 genes with obesity risk in North Indian adults. *Am J Hum Biol*. 2017;29(2).

Susic D, Varagic J. Obesity: A Perspective from Hypertension. *Med Clin North Am*. 2017;101(1):139-157.

Tandon A, Chen CJ, Penman A, Hancock H, James M, Husain D, et al. African Ancestry Analysis and Admixture Genetic Mapping for Proliferative Diabetic Retinopathy in African Americans. *Invest Ophthalmol Vis Sci*. 2015;56(6):3999-4005.

Tang H, Jorgenson E, Gadde M, Kardia SL, Rao DC, Zhu X, et al. Racial admixture and its impact on BMI and blood pressure in African and Mexican Americans. *Hum Genet*. 2006;119(6):624-33.

Taylor MR, Sun AY, Davis G, Fiuzat M, Liggett SB, Bristow MR. Race, common genetic variation, and therapeutic response disparities in heart failure. *JACC Heart Fail*. 2014;2(6):561-72.

The Criteria Committee of the New York Heart Association. Diseases of the heart and blood vessels: nomenclature and criteria for diagnosis. Boston: Little, Brown and Co; 1964.

Thornton TA, Bermejo JL. Local and global ancestry inference and applications to genetic association analysis for admixed populations. *Genet Epidemiol*. 2014;38(1): S5-S12.

van Riet EE, Hoes AW, Wagenaar KP, Limburg A, Landman MA, Rutten FH. Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review. *Eur J Heart Fail*. 2016;18(3):242-52.

Vasan RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *JAMA*. 2009;302(2):168-78.

Waterworth DM, Ricketts SL, Song K, Chen L, Zhao JH, Ripatti S, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2010;30(11):2264-76.

Weissglas-Volkov D, Calkin AC, Tusie-Luna T, Sinsheimer JS, Zelcer N, Riba L, et al. The N342S MYLIP polymorphism is associated with high total cholesterol and increased LDL receptor degradation in humans. *J Clin Invest*. 2011;121(8):3062-71.

Winkler CA, Nelson GW, Smith MW. Admixture mapping comes of age. *Annu Rev Genomics Hum Genet*. 2010;11:65-89.

Yancy CW, Fowler MB, Colucci WS, Gilbert EM, Bristow MR, Cohn JN, et al. Race and the response to adrenergic blockade with carvedilol in patients with chronic heart failure. *N Engl J Med*. 2001;344(18):1358-65.

Yamakawa H, Oyama S, Mitsuhashi H, Sasagawa N, Uchino S, Kohsaka S, et al. Neuroligins 3 and 4X interact with syntrophin-gamma2, and the interactions are affected by autism-related mutations. *Biochem Biophys Res Commun*. 2007;355(1):41-6.

Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet.* 2008;40(9):1092-7.

Zandoná MR, Sangalli CN, Campagnolo PD, Vitolo MR, Almeida S, Mattevi VS. Validation of obesity susceptibility loci identified by genome-wide association studies in early childhood in South Brazilian children. *Pediatr Obes.* 2017;12(1):85-92.

Zhu B, Ashley-Koch AE, Dunson DB. Generalized admixture mapping for complex traits. *G3 (Bethesda).* 2013;3(7):1165-1175.

Zhu X, Cooper RS, Elston RC. Linkage analysis of a complex disease through use of admixed populations. *Am J Hum Genet.* 2004;74:1136–1153.

Zhu X, Cooper RS. Admixture mapping provides evidence of association of the VNN1 gene with hypertension. *PLoS One.* 2007;2(11):e1244.

Zhu X, Luke A, Cooper RS, Quertermous T, Hanis C, Mosley T, Gu CC, et al. Admixture mapping for hypertension loci with genome-scan markers. *Nat Genet.* 2005;37(2):177-81.

Zhu X, Zhang S, Tang H, Cooper R. A classical likelihood based approach for admixture mapping using EM algorithm. *Hum Genet.* 2006;120(3):431-45.

ANEXOS

ANEXO 1

Ancestralidade global dos 492 pacientes com IC.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM001	0,613	0,098	0,289
ICCM003	0,682	0,101	0,217
ICCM004	0,615	0,325	0,060
ICCM005	0,398	0,521	0,081
ICCM006	0,252	0,536	0,212
ICCM007	0,298	0,647	0,055
ICCM008	0,891	0,081	0,028
ICCM009	0,892	0,074	0,034
ICCM010	0,414	0,409	0,177
ICCM011	0,717	0,186	0,096
ICCM012	0,742	0,211	0,047
ICCM013	0,446	0,407	0,147
ICCM014	0,751	0,152	0,097
ICCM015	0,853	0,032	0,115
ICCM016	0,355	0,156	0,489
ICCM017	0,586	0,313	0,101
ICCM018	0,240	0,642	0,118
ICCM019	0,641	0,094	0,265
ICCM020	0,768	0,011	0,221
ICCM021	0,577	0,072	0,351
ICCM022	0,402	0,144	0,454
ICCM023	0,571	0,292	0,137
ICCM024	0,552	0,009	0,439
ICCM025	0,601	0,362	0,037
ICCM026	0,578	0,332	0,090
ICCM027	0,798	0,001	0,201
ICCM028	0,837	0,060	0,103
ICCM029	0,496	0,448	0,056
ICCM030	0,566	0,333	0,101
ICCM031	0,684	0,034	0,282
ICCM032	0,711	0,004	0,285
ICCM033	0,750	0,202	0,048
ICCM035	0,279	0,293	0,428
ICCM036	0,749	0,087	0,164
ICCM037	0,723	0,218	0,059
ICCM038	0,078	0,771	0,151
ICCM039	0,801	0,024	0,175
ICCM040	0,587	0,330	0,083
ICCM041	0,528	0,351	0,121
ICCM042	0,138	0,505	0,357
ICCM043	0,681	0,226	0,093
ICCM044	0,772	0,187	0,041
ICCM045	0,807	0,026	0,167
ICCM046	0,700	0,242	0,058
ICCM047	0,775	0,137	0,087
ICCM048	0,645	0,224	0,131
ICCM049	0,697	0,224	0,079

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM050	0,649	0,113	0,238
ICCM051	0,628	0,131	0,241
ICCM052	0,622	0,283	0,094
ICCM053	0,597	0,291	0,113
ICCM054	0,457	0,458	0,085
ICCM055	0,873	0,070	0,057
ICCM056	0,546	0,141	0,313
ICCM057	0,678	0,024	0,298
ICCM059	0,417	0,428	0,155
ICCM060	0,765	0,171	0,064
ICCM061	0,463	0,391	0,145
ICCM062	0,811	0,031	0,158
ICCM063	0,621	0,148	0,231
ICCM064	0,702	0,062	0,236
ICCM066	0,758	0,055	0,187
ICCM067	0,838	0,087	0,075
ICCM069	0,552	0,355	0,093
ICCM070	0,691	0,236	0,073
ICCM071	0,759	0,011	0,230
ICCM072	0,471	0,361	0,168
ICCM073	0,642	0,221	0,137
ICCM074	0,749	0,158	0,093
ICCM075	0,197	0,781	0,022
ICCM076	0,197	0,795	0,007
ICCM077	0,331	0,431	0,238
ICCM078	0,647	0,144	0,209
ICCM079	0,879	0,077	0,044
ICCM080	0,576	0,313	0,111
ICCM081	0,621	0,064	0,315
ICCM082	0,696	0,118	0,186
ICCM083	0,821	0,041	0,138
ICCM084	0,672	0,241	0,086
ICCM085	0,358	0,571	0,071
ICCM086	0,907	0,064	0,029
ICCM087	0,470	0,187	0,343
ICCM088	0,323	0,577	0,100
ICCM090	0,509	0,306	0,185
ICCM091	0,471	0,093	0,436
ICCM092	0,552	0,315	0,132
ICCM093	0,256	0,321	0,423
ICCM094	0,759	0,192	0,050
ICCM095	0,516	0,410	0,074
ICCM096	0,867	0,058	0,074
ICCM097	0,467	0,385	0,148
ICCM098	0,369	0,308	0,323
ICCM099	0,350	0,593	0,057
ICCM100	0,702	0,234	0,064
ICCM101	0,466	0,386	0,148
ICCM102	0,750	0,149	0,101
ICCM103	0,409	0,548	0,043
ICCM104	0,836	0,001	0,163

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM105	0,767	0,131	0,102
ICCM106	0,764	0,009	0,227
ICCM107	0,805	0,005	0,190
ICCM108	0,437	0,403	0,161
ICCM109	0,741	0,191	0,067
ICCM110	0,728	0,195	0,077
ICCM111	0,765	0,013	0,222
ICCM112	0,819	0,001	0,180
ICCM115	0,774	0,180	0,046
ICCM116	0,811	0,034	0,155
ICCM117	0,815	0,001	0,184
ICCM118	0,746	0,012	0,242
ICCM119	0,427	0,504	0,069
ICCM121	0,744	0,129	0,127
ICCM122	0,727	0,097	0,176
ICCM123	0,122	0,856	0,022
ICCM124	0,714	0,022	0,264
ICCM125	0,261	0,718	0,021
ICCM126	0,863	0,001	0,136
ICCM127	0,653	0,214	0,133
ICCM128	0,573	0,343	0,084
ICCM129	0,855	0,046	0,100
ICCM130	0,480	0,419	0,101
ICCM131	0,660	0,094	0,246
ICCM132	0,615	0,265	0,120
ICCM133	0,483	0,424	0,093
ICCM134	0,817	0,020	0,163
ICCM135	0,768	0,038	0,194
ICCM136	0,507	0,214	0,279
ICCM137	0,913	0,036	0,051
ICCM138	0,808	0,007	0,185
ICCM139	0,755	0,024	0,221
ICCM140	0,373	0,585	0,042
ICCM141	0,682	0,209	0,109
ICCM142	0,895	0,032	0,073
ICCM143	0,755	0,041	0,204
ICCM144	0,878	0,043	0,079
ICCM145	0,815	0,031	0,154
ICCM146	0,975	0,022	0,003
ICCM147	0,721	0,052	0,227
ICCM148	0,789	0,117	0,094
ICCM149	0,481	0,038	0,481
ICCM150	0,527	0,402	0,071
ICCM151	0,453	0,471	0,076
ICCM152	0,323	0,179	0,498
ICCM153	0,708	0,194	0,098
ICCM154	0,737	0,053	0,210
ICCM155	0,831	0,063	0,106
ICCM156	0,564	0,380	0,056
ICCM157	0,061	0,911	0,028
ICCM158	0,365	0,544	0,091

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM160	0,390	0,501	0,109
ICCM161	0,597	0,252	0,151
ICCM162	0,678	0,317	0,005
ICCM163	0,567	0,264	0,169
ICCM164	0,469	0,446	0,085
ICCM165	0,754	0,109	0,137
ICCM166	0,328	0,639	0,033
ICCM167	0,788	0,136	0,077
ICCM168	0,731	0,026	0,243
ICCM169	0,891	0,048	0,061
ICCM170	0,632	0,258	0,110
ICCM171	0,755	0,035	0,210
ICCM172	0,154	0,836	0,010
ICCM174	0,471	0,458	0,072
ICCM175	0,784	0,005	0,211
ICCM176	0,319	0,573	0,108
ICCM177	0,291	0,615	0,095
ICCM178	0,611	0,282	0,107
ICCM179	0,844	0,083	0,073
ICCM180	0,708	0,147	0,145
ICCM182	0,983	0,014	0,002
ICCM183	0,225	0,675	0,101
ICCM184	0,866	0,096	0,037
ICCM185	0,269	0,649	0,082
ICCM186	0,787	0,126	0,086
ICCM187	0,724	0,192	0,084
ICCM188	0,716	0,254	0,031
ICCM189	0,164	0,769	0,067
ICCM190	0,386	0,505	0,108
ICCM191	0,772	0,024	0,204
ICCM192	0,531	0,431	0,038
ICCM193	0,863	0,061	0,076
ICCM194	0,631	0,108	0,261
ICCM195	0,097	0,873	0,029
ICCM198	0,370	0,136	0,493
ICCM199	0,693	0,227	0,080
ICCM200	0,805	0,014	0,181
ICCM201	0,733	0,120	0,148
ICCM202	0,973	0,019	0,008
ICCM203	0,582	0,304	0,114
ICCM204	0,763	0,169	0,068
ICCM205	0,562	0,367	0,072
ICCM206	0,104	0,881	0,014
ICCM207	0,640	0,241	0,119
ICCM210	0,805	0,152	0,043
ICCM211	0,602	0,317	0,081
ICCM212	0,621	0,091	0,288
ICCM213	0,532	0,374	0,094
ICCM214	0,852	0,043	0,105
ICCM215	0,503	0,321	0,176
ICCM216	0,639	0,190	0,170

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM217	0,225	0,742	0,033
ICCM218	0,318	0,613	0,069
ICCM220	0,811	0,001	0,188
ICCM221	0,438	0,468	0,094
ICCM222	0,722	0,172	0,106
ICCM223	0,939	0,051	0,009
ICCM224	0,651	0,095	0,254
ICCM225	0,427	0,459	0,114
ICCM226	0,807	0,001	0,192
ICCM227	0,929	0,012	0,059
ICCM228	0,421	0,475	0,105
ICCM229	0,167	0,795	0,037
ICCM230	0,514	0,161	0,325
ICCM231	0,113	0,833	0,054
ICCM232	0,567	0,330	0,102
ICCM233	0,115	0,884	0,001
ICCM234	0,741	0,177	0,083
ICCM235	0,465	0,415	0,120
ICCM236	0,557	0,371	0,072
ICCM237	0,232	0,754	0,015
ICCM238	0,574	0,384	0,042
ICCM239	0,331	0,270	0,399
ICCM240	0,494	0,376	0,131
ICCM241	0,050	0,945	0,004
ICCM242	0,667	0,282	0,051
ICCM243	0,979	0,017	0,004
ICCM244	0,796	0,130	0,074
ICCM245	0,663	0,230	0,106
ICCM246	0,782	0,100	0,117
ICCM247	0,981	0,018	0,001
ICCM248	0,357	0,590	0,054
ICCM249	0,707	0,033	0,260
ICCM250	0,483	0,385	0,132
ICCM251	0,873	0,016	0,111
ICCM252	0,753	0,029	0,218
ICCM253	0,977	0,017	0,006
ICCM254	0,505	0,049	0,446
ICCM255	0,674	0,024	0,302
ICCM256	0,061	0,802	0,137
ICCM257	0,200	0,799	0,001
ICCM258	0,750	0,062	0,188
ICCM259	0,173	0,789	0,038
ICCM260	0,968	0,031	0,001
ICCM261	0,788	0,027	0,184
ICCM262	0,912	0,063	0,025
ICCM263	0,454	0,480	0,066
ICCM264	0,961	0,025	0,013
ICCM265	0,941	0,054	0,005
ICCM266	0,658	0,131	0,211
ICCM267	0,688	0,226	0,086
ICCM268	0,849	0,073	0,078

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM269	0,872	0,090	0,038
ICCM270	0,740	0,202	0,059
ICCM271	0,809	0,128	0,063
ICCM272	0,570	0,273	0,157
ICCM273	0,758	0,033	0,209
ICCM274	0,880	0,070	0,051
ICCM275	0,421	0,473	0,105
ICCM276	0,589	0,316	0,096
ICCM277	0,982	0,001	0,017
ICCM278	0,986	0,013	0,001
ICCM279	0,583	0,375	0,041
ICCM280	0,716	0,178	0,106
ICCM281	0,503	0,398	0,098
ICCM283	0,470	0,436	0,094
ICCM284	0,603	0,334	0,063
ICCM285	0,917	0,032	0,052
ICCM286	0,922	0,038	0,040
ICCM287	0,639	0,248	0,113
ICCM288	0,511	0,123	0,366
ICCM289	0,994	0,001	0,005
ICCM290	0,732	0,192	0,076
ICCM291	0,836	0,001	0,163
ICCM292	0,541	0,318	0,141
ICCM293	0,495	0,417	0,087
ICCM294	0,615	0,300	0,085
ICCM295	0,808	0,054	0,139
ICCM296	0,721	0,174	0,105
ICCM297	0,259	0,683	0,058
ICCM298	0,763	0,151	0,086
ICCM299	0,959	0,039	0,003
ICCM300	0,667	0,202	0,131
ICCM301	0,512	0,015	0,473
ICCM302	0,770	0,022	0,208
ICCM303	0,945	0,018	0,037
ICCM304	0,653	0,227	0,120
ICCM305	0,177	0,716	0,107
ICCM306	0,971	0,025	0,004
ICCM307	0,396	0,426	0,179
ICCM308	0,667	0,265	0,068
ICCM310	0,553	0,366	0,081
ICCM311	0,435	0,176	0,389
ICCM312	0,816	0,071	0,113
ICCM313	0,695	0,258	0,047
ICCM314	0,626	0,111	0,263
ICCM315	0,266	0,371	0,364
ICCM316	0,718	0,163	0,119
ICCM317	0,215	0,733	0,052
ICCM318	0,913	0,029	0,058
ICCM319	0,529	0,385	0,086
ICCM320	0,619	0,132	0,249
ICCM321	0,401	0,111	0,488

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM322	0,832	0,027	0,141
ICCM323	0,831	0,097	0,072
ICCM324	0,725	0,009	0,266
ICCM325	0,708	0,177	0,115
ICCM326	0,072	0,853	0,075
ICCM327	0,727	0,069	0,204
ICCM328	0,185	0,328	0,487
ICCM329	0,914	0,020	0,066
ICCM330	0,913	0,037	0,050
ICCM331	0,570	0,078	0,352
ICCM332	0,502	0,218	0,280
ICCM333	0,270	0,642	0,089
ICCM334	0,862	0,045	0,093
ICCM335	0,597	0,293	0,110
ICCM336	0,599	0,138	0,263
ICCM337	0,542	0,083	0,375
ICCM338	0,667	0,168	0,166
ICCM339	0,504	0,226	0,270
ICCM341	0,673	0,281	0,045
ICCM342	0,665	0,270	0,065
ICCM343	0,662	0,261	0,077
ICCM344	0,973	0,007	0,020
ICCM345	0,764	0,150	0,086
ICCM346	0,707	0,051	0,242
ICCM347	0,025	0,966	0,009
ICCM348	0,790	0,055	0,155
ICCM349	0,525	0,378	0,097
ICCM350	0,415	0,233	0,352
ICCM351	0,444	0,520	0,036
ICCM352	0,976	0,001	0,023
ICCM353	0,349	0,625	0,026
ICCM354	0,644	0,230	0,126
ICCM355	0,561	0,208	0,231
ICCM356	0,697	0,183	0,120
ICCM357	0,661	0,063	0,276
ICCM358	0,465	0,355	0,180
ICCM359	0,765	0,218	0,017
ICCM360	0,721	0,060	0,220
ICCM361	0,596	0,299	0,104
ICCM362	0,445	0,464	0,091
ICCM363	0,431	0,348	0,221
ICCM364	0,195	0,242	0,563
ICCM365	0,563	0,369	0,068
ICCM366	0,356	0,635	0,010
ICCM367	0,707	0,257	0,036
ICCM368	0,724	0,180	0,097
ICCM369	0,721	0,174	0,104
ICCM370	0,392	0,525	0,083
ICCM371	0,641	0,311	0,048
ICCM372	0,439	0,429	0,132
ICCM373	0,118	0,753	0,130

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM374	0,482	0,444	0,074
ICCM375	0,918	0,055	0,027
ICCM376	0,863	0,103	0,035
ICCM377	0,599	0,270	0,131
ICCM378	0,901	0,003	0,096
ICCM379	0,675	0,024	0,301
ICCM380	0,971	0,028	0,001
ICCM385	0,593	0,219	0,188
ICCM386	0,813	0,034	0,153
ICCM387	0,645	0,266	0,089
ICCM388	0,982	0,012	0,006
ICCM389	0,422	0,345	0,233
ICCM390	0,968	0,018	0,015
ICCM391	0,575	0,091	0,334
ICCM392	0,567	0,112	0,321
ICCM393	0,801	0,114	0,085
ICCM394	0,814	0,164	0,022
ICCM395	0,403	0,490	0,107
ICCM396	0,641	0,208	0,150
ICCM397	0,553	0,098	0,349
ICCM398	0,594	0,164	0,242
ICCM399	0,814	0,047	0,139
ICCM400	0,422	0,176	0,402
ICCM401	0,757	0,024	0,219
ICCM402	0,581	0,359	0,060
ICCM403	0,821	0,051	0,128
ICCM404	0,420	0,493	0,087
ICCM405	0,427	0,483	0,090
ICCM406	0,712	0,057	0,231
ICCM407	0,241	0,662	0,096
ICCM408	0,750	0,142	0,108
ICCM409	0,543	0,091	0,366
ICCM410	0,603	0,382	0,015
ICCM411	0,646	0,186	0,168
ICCM412	0,181	0,786	0,033
ICCM413	0,520	0,424	0,056
ICCM414	0,769	0,204	0,027
ICCM415	0,595	0,222	0,183
ICCM416	0,801	0,001	0,198
ICCM417	0,481	0,442	0,077
ICCM418	0,851	0,028	0,121
ICCM419	0,936	0,053	0,011
ICCM420	0,926	0,057	0,017
ICCM421	0,373	0,301	0,326
ICCM422	0,990	0,001	0,009
ICCM423	0,850	0,079	0,071
ICCM424	0,528	0,373	0,099
ICCM425	0,541	0,261	0,199
ICCM426	0,459	0,301	0,240
ICCM427	0,894	0,070	0,036
ICCM428	0,236	0,679	0,085

Continuação da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM429	0,480	0,435	0,084
ICCM430	0,413	0,371	0,216
ICCM431	0,831	0,063	0,106
ICCM432	0,810	0,097	0,093
ICCM433	0,401	0,321	0,278
ICCM434	0,897	0,040	0,063
ICCM435	0,414	0,569	0,017
ICCM436	0,981	0,008	0,011
ICCM437	0,803	0,055	0,142
ICCM438	0,195	0,754	0,050
ICCM439	0,424	0,434	0,142
ICCM440	0,395	0,538	0,067
ICCM441	0,748	0,179	0,073
ICCM442	0,499	0,343	0,158
ICCM443	0,778	0,169	0,054
ICCM444	0,686	0,278	0,035
ICCM445	0,294	0,647	0,059
ICCM446	0,584	0,203	0,213
ICCM447	0,753	0,057	0,189
ICCM448	0,823	0,013	0,164
ICCM449	0,970	0,016	0,014
ICCM450	0,477	0,494	0,030
ICCM451	0,893	0,055	0,053
ICCM452	0,739	0,089	0,172
ICCM453	0,737	0,154	0,109
ICCM454	0,988	0,001	0,011
ICCM455	0,856	0,052	0,091
ICCM456	0,877	0,076	0,047
ICCM457	0,598	0,312	0,090
ICCM458	0,754	0,140	0,106
ICCM459	0,100	0,865	0,034
ICCM460	0,373	0,609	0,018
ICCM461	0,645	0,286	0,068
ICCM462	0,524	0,394	0,082
ICCM463	0,701	0,211	0,089
ICCM464	0,371	0,248	0,381
ICCM465	0,567	0,350	0,083
ICCM466	0,471	0,344	0,185
ICCM467	0,805	0,053	0,142
ICCM468	0,635	0,212	0,153
ICCM469	0,629	0,004	0,367
ICCM470	0,661	0,252	0,087
ICCM471	0,493	0,381	0,127
ICCM472	0,634	0,301	0,065
ICCM473	0,904	0,038	0,058
ICCM474	0,545	0,384	0,070
ICCM475	0,648	0,270	0,082
ICCM476	0,778	0,049	0,173
ICCM477	0,707	0,163	0,130
ICCM478	0,528	0,361	0,111
ICCM479	0,174	0,760	0,066

Finalização da tabela do Anexo 1.

Pacientes	Ancestralidade Europeia	Ancestralidade Africana	Ancestralidade Ameríndia
ICCM480	0,294	0,527	0,179
ICCM481	0,599	0,341	0,060
ICCM482	0,976	0,008	0,016
ICCM483	0,401	0,486	0,113
ICCM484	0,630	0,227	0,143
ICCM485	0,077	0,919	0,004
ICCM486	0,675	0,256	0,069
ICCM487	0,512	0,074	0,414
ICCM488	0,366	0,552	0,082
ICCM489	0,553	0,314	0,134
ICCM490	0,367	0,268	0,364
ICCM491	0,540	0,161	0,299
ICCM492	0,628	0,268	0,104
ICCM493	0,790	0,181	0,029
ICCM494	0,423	0,426	0,151
ICCM495	0,429	0,448	0,122
ICCM496	0,462	0,311	0,227
ICCM497	0,685	0,225	0,091
ICCM498	0,651	0,188	0,161
ICCM499	0,721	0,098	0,181
ICCM500	0,399	0,244	0,357
ICCM501	0,967	0,021	0,012
ICCM502	0,733	0,152	0,115
ICCM503	0,361	0,501	0,138
ICCM504	0,727	0,227	0,047
ICCM505	0,842	0,053	0,104
ICCM506	0,816	0,157	0,027
ICCM507	0,742	0,209	0,049
ICCM508	0,704	0,189	0,107
ICCM509	0,791	0,086	0,123
ICCM510	0,450	0,480	0,070
ICCM511	0,694	0,011	0,295
ICCM512	0,799	0,015	0,186
ICCM513	0,328	0,504	0,168
ICCM514	0,247	0,689	0,064
ICCM515	0,856	0,025	0,119
ICCM516	0,849	0,026	0,125

ANEXO 2

Proporções de cada ancestralidade nos 22 cromossomos dos 492 pacientes.

Cromossomos	Ancestralidade		
	Europeia	Ameríndia	Africana
chr 1	0,608±0,239	0,131±0,077	0,261±0,230
chr 2	0,612±0,236	0,130±0,081	0,258±0,225
chr 3	0,605±0,241	0,130±0,082	0,265±0,230
chr 4	0,614±0,240	0,129±0,084	0,257±0,230
chr 5	0,665±0,238	0,131±0,078	0,204±0,231
chr 6	0,617±0,241	0,130±0,078	0,253±0,234
chr 7	0,624±0,275	0,132±0,115	0,244±0,266
chr 8	0,612±0,239	0,129±0,093	0,259±0,229
chr 9	0,607±0,247	0,131±0,084	0,262±0,238
chr 10	0,613±0,253	0,129±0,085	0,258±0,243
chr 11	0,602±0,256	0,131±0,085	0,267±0,248
chr 12	0,617±0,236	0,130±0,086	0,253±0,230
chr 13	0,618±0,251	0,129±0,084	0,253±0,246
chr 14	0,651±0,249	0,131±0,092	0,218±0,239
chr 15	0,599±0,258	0,129±0,099	0,272±0,239
chr 16	0,619±0,246	0,131±0,086	0,250±0,240
chr 17	0,597±0,253	0,130±0,085	0,273±0,248
chr 18	0,611±0,249	0,131±0,092	0,258±0,236
chr 19	0,598±0,251	0,130±0,097	0,272±0,242
chr 20	0,626±0,247	0,129±0,100	0,245±0,234
chr 21	0,611±0,236	0,130±0,086	0,259±0,230
chr 22	0,629±0,261	0,130±0,109	0,241±0,251