

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO

CAÍQUE MANÓCHIO NUNES DA SILVA

**ANCESTRALIDADE GENÔMICA SUBCONTINENTAL E VARIANTES
DE INTERESSE FARMACOGENÔMICO**

UBERABA

2023

CAÍQUE MANÓCHIO NUNES DA SILVA

**ANCESTRALIDADE GENÔMICA SUBCONTINENTAL E VARIANTES
DE INTERESSE FARMACOGENÔMICO**

Dissertação apresentada ao curso de Pós-Graduação em Ciências da Saúde, área de concentração “Medicina Translacional” da Universidade Federal do Triângulo Mineiro, como requisito parcial para obtenção do título de Mestre em Ciências da Saúde

Orientador: Prof^a. Dra. Fernanda Rodrigues Soares

Uberaba-MG

2023

**Catálogo na fonte: Biblioteca da Universidade Federal do
Triângulo Mineiro**

S579a Silva, Caique Manóchio Nunes da
Ancestralidade genômica subcontinental e variantes de interesse
farmacogenômico / Caique Manóchio Nunes da Silva. – 2023.
72 f.: il., graf., tab.

Dissertação (Mestrado em Ciências da Saúde) -- Universidade Federal
do Triângulo Mineiro, Uberaba, MG, 2023
Orientadora: Profa. Dra. Fernanda Rodrigues Soares

1. Genômica. 2. Citocromo P-450 CYP2C9. 3. Citocromo P-450
CYP2C19. 4. Farmacogenética. I. Soares, Fernanda Rodrigues. II. Univer-
sidade Federal do Triângulo Mineiro. III. Título.

CDU 575.111

CAÍQUE MANÓCHIO NUNES DA SILVA

ANCESTRALIDADE GENÔMICA SUBCONTINENTAL E VARIANTES DE
INTERESSE FARMACOGENÔMICO

Dissertação apresentada ao curso de Pós-Graduação
em Ciências da Saúde, área de concentração
“Medicina Translacional” da Universidade Federal
do Triângulo Mineiro, como requisito parcial para
obtenção do título de Mestre em Ciências da Saúde

27 de novembro de 2023

Banca Examinadora:

Prof.º Dr. Gilderlânio Santana de Araújo

Profª. Drª. Mariângela Torreglosa Ruiz Cintra

Profª. Dra. Fernanda Rodrigues Soares

AGRADECIMENTOS

Está terminando mais uma jornada, o mestrado!!!! Foram alguns anos de muito estudo, esforço e empenho e por isso gostaria de escrever algumas palavras de gratidão e agradecimentos para as pessoas que foram importantes para que eu conseguisse concluir mais esse sonho. Primeiramente queria agradecer aos meus pais, Denize e José; ao meu irmão João e minha cunhada Camila; minhas avós Dirce e Romilda, por sempre apoiar e incentivar a eu procurar e encontrar meu caminho. Minha gratidão especial a Prof. Dr^a. Fernanda Rodrigues Soares, minha orientadora e amiga. Obrigado por sua dedicação, que por muitas vezes deixou de lado momentos de descanso para me ajudar e me orientar. E, um obrigado especial, pelos ensinamentos acadêmicos, e os não acadêmicos também, enfim obrigado pela confiança. Quero expressar minha gratidão, agradecimentos, carinho e muito amor por amigos que fazem parte do grupo de pesquisa, que contribuíram muito com meu aprendizado, em especial meu sincero obrigado a Mariana, Maria Laura, Isadora, Henrique, Bruno e Guilherme por se tornaram minha família aqui na UFTM. Ainda queria dizer obrigado há mais alguns amigos, Vanessa, Kamila, Lara, Larissa, Maicon, Herisy e Larissa, muito obrigado por me acompanharem e permitir que eu os acompanhe durante a vida sempre, celebrando nos momentos felizes e apoiando nos momentos difíceis.

RESUMO

Alterações na sequência de DNA podem ocasionar mudanças na função de enzimas, podendo causar uma variabilidade de resposta a fármacos. A farmacogenômica busca compreender essa variabilidade de resposta e assim propor um tratamento mais individualizado. Os CYPs são uma família de genes responsáveis por codificar enzimas que atuam no complexo do citocromo P450, nas vias de metabolização de compostos endógenos e xenobióticos. Dentre os genes dessa família, se destacam os genes *CYP2C19* e *CYP2C9* pela alta variabilidade alélica, além de participarem da metabolização de grande parte dos fármacos disponíveis no mercado. Também é importante considerar que os alelos, inclusive os de interesse farmacogenético, possuem variação na frequência entre as populações e etnias, essa diferença de frequência pode ser usada para inferir a ancestralidade genômica. A ancestralidade genômica é a informação genética de um indivíduo que revela de qual população parental ele descende, e quando associada com questões de saúde pode auxiliar no diagnóstico e no tratamento. Tendo isso em vista, esse trabalho tem o objetivo de correlacionar ancestralidade genômica subcontinental de componentes parentais com a frequência de alelos de interesse farmacogenômico dos genes *CYP2C19* e *CYP2C9*. Utilizamos o software Admixture para inferir a ancestralidade genômica com 8 componentes parentais de 2504 indivíduos pertencentes a 26 populações ao redor do mundo no banco de dados públicos *The 1000 Genomes Project*, também inferimos, utilizando o Stargazer, os alelos estrela e o fenótipo predito e depois calculamos frequência alélica e do fenótipo nas populações e por último realizamos a correlação entre a frequência alélica e frequência do fenótipo predito com os componentes parentais. Os componentes parentais foram separados em Leste Asiático – Oeste (EAS-W), Leste Asiático – Leste (EAS-E), Norte Europeu (NEUR), Sul Europeu (SEUR), Nativo Americano (NAT), Oeste Africano (WAFR), Leste Africano (EAFR), e Sul Asiático (SAS), e os alelos estrela encontrados foram para *CYP2C19*, 15 alelos e para *CYP2C9*, 18 alelos. Os *CYP2C19*2* e *CYP2C19*3* apresentaram uma frequência maior em populações asiáticas, enquanto que os *CYP2C19*5*, *CYP2C19*8*, *CYP2C19*9*, *CYP2C19*17* tiveram sua frequência maior em populações africanas. Para o *CYP2C9*2* a sua frequência foi maior em populações europeias, já o alelo *CYP2C9*7* foi encontrado apenas em populações sul asiáticas e o alelo *CYP2C9*13* apenas no leste asiáticos. Foi encontrada correlação positiva entre o *CYP2C19*2* e os componentes EAS-E e SAS e negativa com o componente SEUR, enquanto o *CYP2C19*3* apresentou correlação positiva com EAS-W e EAS-E e o *CYP2C19*17* apresentou correlação negativa com EAS-W e EAS-E e positivas com os WAFR e EAFR. Já o *CYP2C9*2* teve uma correlação positiva com o componente SEUR e o alelo *CYP2C9*3* teve negativa com os componentes WAFR e EAFR e positiva com SAS. Em relação ao fenótipo predito encontramos correlação positiva de metabolizadores lentos (PMs) do *CYP2C19* com SAS e metabolizadores intermediários (IMs) com os EAS-W e EAS-E e para o gene *CYP2C9* encontramos correlação positiva do IMs com SEUR e negativa com EAS-W. Sendo assim, é importante que no desenho de painéis de testes farmacogenéticos considere qual população fará uso e quais alelos estão correlacionados com aquela população.

Palavra-chave: Ancestralidade genômica; *CYP2C9*; *CYP2C19*; farmacogenômica

ABSTRACT

Variation in the DNA sequence may cause changes in the function of enzymes, which can cause variability in drug response. Pharmacogenomics seeks to understand this response variability and thus propose a more individualized treatment. CYPs are a family of genes responsible for encoding enzymes that act on cytochrome P450, enzymes that are part of the metabolism pathways of endogenous and xenobiotic compounds. Among these genes, the *CYP2C19* and *CYP2C9* genes stand out for their high allelic variability, in addition to participating in the metabolism of most drugs available on the market. It is also important to consider that alleles vary in frequency between populations and ethnicities; this frequency difference can be used to infer genomic ancestry and be associated with health issues. Genomic ancestry is the genetic information of an individual that reveals which parental population he descends from, and when associated with health issues, it can help in the diagnosis of treatment. This work aims to correlate subcontinental genomic ancestry of parental components with the frequency of alleles of pharmacogenomic interest of the *CYP2C19* and *CYP2C9* genes. We used the Admixture software to infer the genomic ancestry with 8 parental components of 2504 individuals belonging to 26 populations around the world, we also inferred, using Stargazer, the star alleles and predicted phenotype then we calculated allele and phenotype frequency in the populations and finally we performed the correlation between the allele and phenotype frequency with the parental components. The parental components were separated into Native American, Northern Europe, Southern Europe, West African, East African, South Asian, East Asian -West and East Asian -East and the star alleles found were for *CYP2C19*, 15 alleles and for *CYP2C9*, 18 alleles. *CYP2C19*2* and *CYP2C19*3* showed a higher frequency in Asian populations, while *CYP2C19*5*, *CYP2C19*8*, *CYP2C19*9* and *CYP2C19*17* had a higher frequency in African populations. Regarding *CYP2C9*2*, its frequency was higher in Europeans, *CYP2C9*7* was found only in South Asians and *CYP2C9*13* only East Asians. A positive correlation was found between *CYP2C19*2* and the EAS2 and SAS and a negative correlation with the EUR2 and NAT components, while *CYP2C19*3* showed a positive correlation with EAS and EAS2 and *CYP2C19*17* demonstrated positive correlation with AFR and AFR2 and a negative correlation with EAS and EAS2. The *CYP2C9*2* had a positive correlation with the EUR2 component, and the *CYP2C9*3* had a negative correlation with AFR and AFR2 components and a positive correlation with the SAS. Therefore, it is important that, when including or not alleles in pharmacogenetic test panels, consider which population will use them and which alleles are correlated with that population.

Keywords: Genomic ancestry; *CYP2C9*; *CYP2C19*; pharmacogenomics

LISTA DE ILUSTRAÇÕES

Figura 1 - Representação via de metabolização da Varfarina	09
Figura 2 - Representação via de metabolização do Clopidogrel	10
Figura 3 - Gráfico de barras da ancestralidade continental individual	21
Figura 4 - Gráfico de barras da ancestralidade subcontinental individual	22
Figura 5 - Mapa com localização geográfica de todas as populações	26
Figura 6 - Gráfico de barras da ancestralidade subcontinental individual	30
Figura 7. Mapa com as populações e barras representando a ancestralidade de todos os indivíduos de uma população	31
Figura 8 - Mapa localizando cada população europeia e gráficos de ancestralidade individual	32
Figura 9 - Mapa localizando cada população do leste asiático e gráficos de ancestralidade individual	33
Figura 10 - Mapa localizando cada população do sul asiático e gráficos de ancestralidade individual	34
Figura 11 - Mapa localizando cada população do africano e gráficos de ancestralidade individual	36
Figura 12 - Mapa localizando cada população americana e gráficos de ancestralidade individual	37
Figura 13 - Gráficos Ancestralidade pela frequência <i>CYP2C19*2</i>	55
Figura 14 - Gráficos Ancestralidade pela frequência <i>CYP2C19*3</i>	56
Figura 15 - Gráficos Ancestralidade pela frequência <i>CYP2C19*17</i>	57
Figura 16 - Gráficos Ancestralidade pela frequência <i>CYP2C9*2</i>	58
Figura 17 - Gráficos Ancestralidade pela frequência <i>CYP2C9*3</i>	59

LISTA DE TABELAS

Tabela 1 - Função enzimática dos alelos do gene <i>CYP2C19</i>	13
Tabela 2 - Classificação dos fenótipos <i>CYP2C19</i>	13
Tabela 3 - Função enzimática dos alelos do gene <i>CYP2C9</i>	15
Tabela 4 - Classificação dos fenótipos <i>CYP2C9</i>	16
Tabela 5 - Porcentagem de ancestralidade $k = 8$ para cada população	28
Tabela 6 - Frequência dos alelos em cada população para <i>CYP2C19</i>	40
Tabela 6.1 - Frequência dos alelos em cada população para <i>CYP2C19</i>	41
Tabela 7 - Frequência dos alelos em cada população do gene <i>CYP2C9</i>	44
Tabela 7.1 - Continuação da frequência dos alelos em cada população do gene <i>CYP2C9</i>	45
Tabela 8 - Frequência dos fenótipos preditos em cada população do gene	47
Tabela 9 - Frequência dos fenótipos preditos em cada população do gene <i>CYP2C19</i>	49
Tabela 10 - Valores de coeficiente de correlação e valores de p da correlação entre alelos e ancestralidade do gene <i>CYP2C19</i>	51
Tabela 11 - Valores de coeficiente de correlação e valores de p da correlação entre alelos e ancestralidade do gene <i>CYP2C9</i>	52
Tabela 12 - Valores de coeficiente de correlação e valores de p da correlação entre fenótipos preditos e ancestralidade do gene <i>CYP2C19</i>	61
Tabela 13 - Valores de coeficiente de correlação e valores de p da correlação entre fenótipos preditos e ancestralidade do gene <i>CYP2C9</i>	62

Sumário

1. Introdução	7
2. Revisão da Literatura	7
2.1 Farmacogenômica geral.	7
2.2 CYP2C19	12
2.3 CYP2C9	15
2.4 Variabilidade interétnica e Ancestralidade Genômica	18
3. Hipótese	23
4. Justificativa	23
5. Objetivo	24
5.1 Objetivo Geral	24
5.2 Objetivos específicos	24
6. Metodologia	25
7. Resultados e Discussão	27
7.1 Ancestralidade	27
7.2 Farmacogenômica	38
7.3 Correlação entre ancestralidade e frequência de alelos farmacogenômicos	49
8. Conclusão	62
9. Referências Bibliográficas	64

1. Introdução

A sensibilidade e a resposta a fármacos são diferentes para cada indivíduo. A variabilidade de resposta pode ser causada por vários fatores, como por exemplo doenças, fatores ambientais e genéticos (Metzger et al., 2006; Lam, 2019). Variações na sequência de DNA quando estão presentes em uma frequência estável maior ou igual a 1% na população são chamadas de polimorfismos (REGINA; WANYCE; ROBINSON, 2013).

Alguns polimorfismos podem influenciar a cinética ou dinâmica de um fármaco, impactando a resposta de um indivíduo a medicamentos; esses polimorfismos são o foco da farmacogenética (SUAREZ-KURTZ, 2010). A farmacogenética busca identificar, estudar a frequência e a função desses polimorfismos nas populações e utilizar essas informações para propor um tratamento individualizado (SCUDELER; RODRIGUES-SOARES, 2020a; STANEK et al., 2012)

As enzimas do citocromo P450 são responsáveis por metabolizar compostos endógenos (ácidos graxos, colesterol, esteróides) e xenobióticos (poluentes, produtos naturais de plantas e fármacos) (Nelson; Nebert, 2018; Scudeler & Rodrigues-Soares, 2020; Zanger & Schwab, 2013). As enzimas do complexo do citocromo P450 são codificadas pelos genes das CYPs, que são altamente polimórficos, e alguns polimorfismos alteram a atividade enzimática. Por fazer parte de várias vias importantes na metabolização de fármacos e possuir uma alta variabilidade alélica, os CYPs são genes relevantes nos estudos farmacogenéticos (NELSON; NEBERT, 2018; SOSA-MACÍAS; LLERENA, 2013). Dentre os genes da família CYP, os genes *CYP2C19* e *CYP2C9* estão entre os mais relevantes e com mais dados disponíveis (LAM, 2019).

Dentro dos genes CYPs, destaca-se o gene *CYP2C19*, com 39 alelos identificados no Pharmacogene Variation Consortium (Pharmvar), (<https://www.pharmvar.org/>), os quais podem alterar a metabolização de fármacos, especialmente em tratamentos de doenças gastroesofágicas e na metabolização do clopidogrel. O *CYP2C9* é um gene que de acordo com Pharmvar possui 85 alelos identificados e podem influenciar no metabolismo da varfarina, fenitoína e anti-inflamatórios não esteroides.

É necessário considerar que a frequência dos alelos possui uma variação entre as populações e etnias. Segundo Auton et al., (2015), as variantes mais comuns são compartilhadas entre a maioria das populações, enquanto as variações mais raras estão restritas a algumas populações. As variantes raras são mais prováveis de serem compartilhadas entre indivíduos da mesma

população ou por um grupo continental, devido a isolamentos geográficos e efeito de gargalo de garrafa (AUTON et al., 2015).

O isolamento geográfico de populações ao longo do tempo, juntamente com a ação de mecanismos evolutivos, auxilia a diferenciação genética entre essas populações atuando no controle da frequência de variantes novas ou variantes já existentes. Após um longo tempo de isolamento geográfico as populações passam a ter variantes exclusivas e são essas populações que são chamadas de populações parentais (HENN; CAVALLI-SFORZA; FELDMAN, 2012). As populações possuem muitas variantes geograficamente e geneticamente distintas entre os grupos continentais, o que torna possível a inferência da ancestralidade genômica de um indivíduo ou de uma população (AUTON et al., 2015). A ancestralidade genômica é a informação genética de um indivíduo que revela de qual população parental o indivíduo descende, pois durante os séculos mais recentes houve um grande fluxo migratório entre as populações humanas, proporcionando uma mistura entre os grupos, como exemplo as populações miscigenadas da América Latina (ZHANG et al., 2019).

Nos países latino-americanos, as proporções de cada raiz étnica são diferentes não só entre os países, mas também entre as diferentes regiões de um país, e o mesmo acontece com a ancestralidade genômica. Etnia é uma classificação sociológica feita a humanos que compartilham características biológicas e a autodeclaração étnica, além de considerar características biológicas, envolve fatores socioeconômicos, como falta de acesso a tratamentos médicos e racismo institucional, o que influencia a autodeclaração étnica em populações miscigenadas, que pode não corresponder à ancestralidade genômica (ZHANG et al., 2019).

A ancestralidade genômica pode ser inferida por meio dos Marcadores Informativos de Ancestralidade (AIMs - variantes selecionadas para inferir a ancestralidade) ou por uma abordagem *genome-wide* (milhares de marcadores aleatórios). Através dos AIMs, conseguimos fazer um perfil da porcentagem de ancestralidade genômica individual para cada componente parental continental (p. ex. um indivíduo pode ser 60% europeu, 30% nativo-americano e 10% africano) (Padhukasahasram, 2014; Scudeler; Rodrigues-Soares, 2020).

Esse perfil de ancestralidade pode ser feito com componentes parentais continentais, que consiste em relacionar a ancestralidade de um indivíduo com grandes grupos continentais, como: europeus, africanos, asiáticos e ameríndios. No entanto, também podemos separar componentes subcontinentais, relacionando a ancestralidade de um indivíduo com populações mais específicas como, por exemplo, se há uma descendência de grupos do norte ou sul europeu, leste ou sul asiático e leste ou oeste africano, utilizando abordagens *genome-wide*. (KEHDY et al., 2015; PADHUKASAHASRAM, 2014)

A variação das frequências alélicas entre as diferentes populações do mundo pode estar correlacionada também com questões de saúde. De acordo com Rodrigues-Soares et al. (2020) determinadas variações alélicas do gene *CYP2D6*, que interferem na resposta a vários fármacos, estão correlacionadas positivamente com a ancestralidade europeia e outros alelos, como *CYP2C19*17*, possuem uma correlação negativa com a ancestralidade nativo-americana, por ter uma frequência reduzida neste grupo.

Apesar de existirem estudos sobre ancestralidade e farmacogenômica, muitas perguntas ainda carecem de esclarecimento científico. A correlação entre ancestralidade genômica subcontinental e variantes farmacogenômicas, por exemplo, poderá produzir informações úteis para estudos clínicos e novos testes de farmacogenômica com as variantes específicas de cada população e seus descendentes, e, no caso das populações miscigenadas, testar com mais constância variantes “raras” em europeus, mas que podem estar associadas com outras ancestralidades parentais dos miscigenados.

2. Revisão da Literatura

2.1 Farmacogenômica geral.

A resposta de um indivíduo a um medicamento é um aspecto essencial para o sucesso do tratamento. Entretanto, a resposta terapêutica do paciente ao fármaco pode não ter resultados desejáveis, podendo ocorrer falha terapêutica ou desencadeamento de eventos adversos a medicamentos (LI et al., 2011a). A variabilidade de resposta a fármacos entre os pacientes é atribuída a diferenças inatas entre os indivíduos (LAM, 2019). Existem vários fatores que podem ser atribuídos a essas diferenças inatas entre os indivíduos, como por exemplo idade, patologias, farmacodinâmica, farmacocinética e o genótipo do indivíduo (METZGER; SOUZA-COSTA; TANUS-SANTOS, 2006). A farmacogenômica (PGx) busca compreender a influência dos genótipos nas respostas a medicamentos e incorporar esses dados na hora de decidir qual terapia usar, propondo um tratamento individualizado para o paciente e com isso aumentar a eficácia e diminuir os riscos de respostas adversas (LAM, 2019).

A PGx da farmacocinética tem como foco os genes envolvidos na absorção, distribuição, metabolismo e excreção de fármacos (ADME). Alterações na sequência de DNA, como deleções, micro e minissatélites, duplicações e substituições nucleotídeo, podem produzir

novas variantes genéticas, que são capazes de modificar o metabolismo de fármacos, alterando a resposta do indivíduo (METZGER; SOUZA-COSTA; TANUS-SANTOS, 2006).

O genoma humano possui aproximadamente 20 mil genes que são codificadores de proteínas e é formado por cerca de 3,12 bilhões de nucleotídeos, os quais apresentam mais de 22 milhões de polimorfismos de nucleotídeo único (SNPs) identificados (variantes com frequência maior ou igual a 1% na população) (LAM, 2019; METZGER; SOUZA-COSTA; TANUS-SANTOS, 2006).

Os polimorfismos que podem afetar o metabolismo de fármacos são geralmente em genes que irão codificar as proteínas que auxiliam na metabolização, realizar o transporte de fármacos ou de proteínas-alvo. Geralmente os polimorfismos causam alterações nas reações de Fase I e II, que consistem em reações envolvidas na metabolização de fármacos, e influenciando a resposta do indivíduo (LAM, 2019).

As reações de fase I e II são processos de biotransformação que representam a principal via que organismos utilizam para inativar e eliminar fármacos. As reações de fase I são catabólicas (oxidação, redução, hidrólise), geralmente transformando o metabólito em um composto mais polar que o fármaco original e proporcionando sua eliminação mais rápida pela urina, como por exemplo a metabolização da varfarina pela enzima CYP2C9, na qual a enzima transforma o composto S-Varfarina em 7-Hidroxi S-Varfarina (FIGURA 1.) (RANG et al., 2016)

Já as reações de fase II são sintéticas ou anabólicas, envolvendo a conjugação com o que normalmente resulta em produtos inativos. Quando a molécula do fármaco possui um gancho certo (grupo hidroxila, amino ou tiol), ela fica mais suscetível à ligação de um grupo substituinte, ou seja, a conjugação, entretanto, durante metabolização em ambas as fases algumas substâncias podem produzir metabólitos com atividade farmacológica ou tóxica, como é caso do clopidogrel que por meio de oxidações pela enzima CYP2C19 é transformado em um metabólito ativo, chamado clopidogrel tiol H4 (FIGURA 2) (Rang et al., 2016; PharmGKB (<https://www.pharmgkb.org/>), 2023).

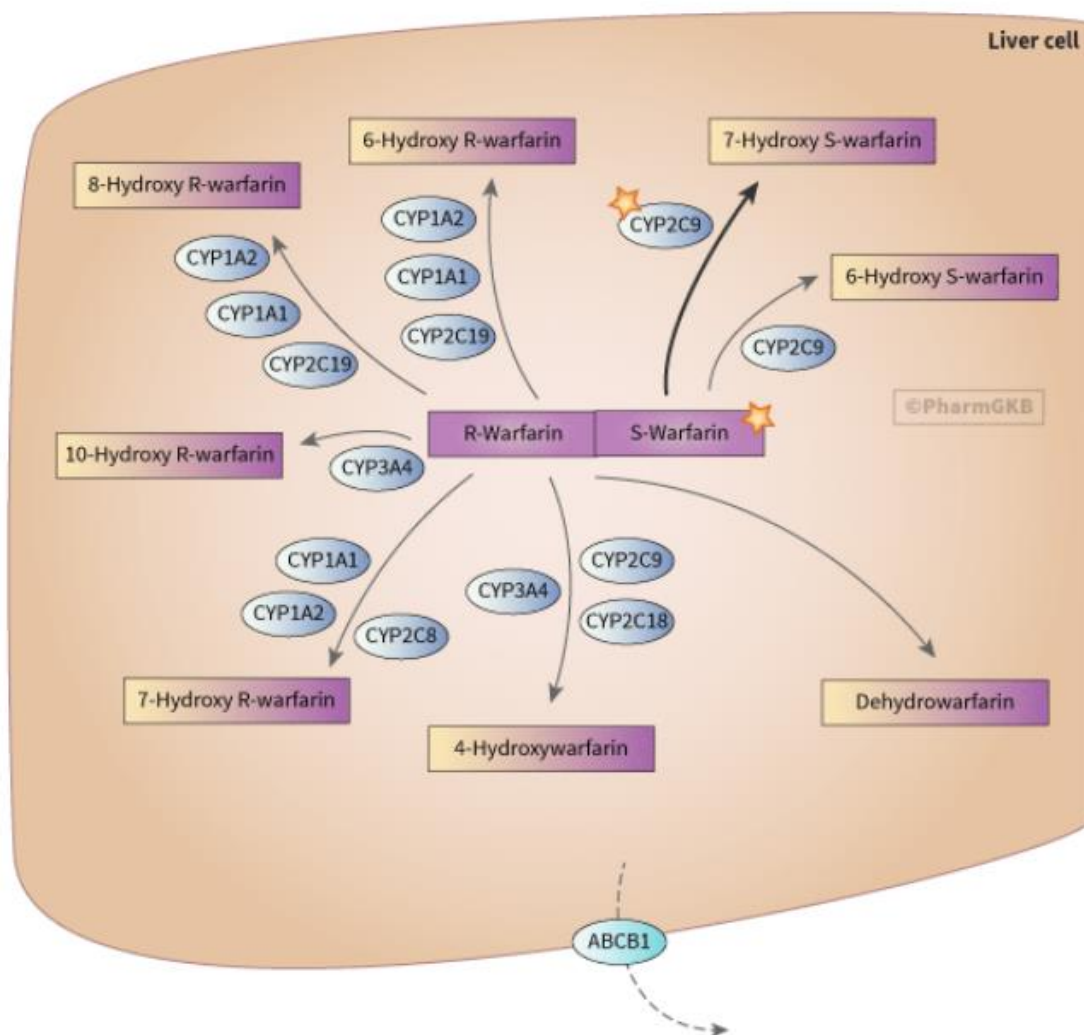


Figura 1. Representação via de metabolização da Varfarina Fonte: (PharmGKB (<https://www.pharmgkb.org/>), 2023)

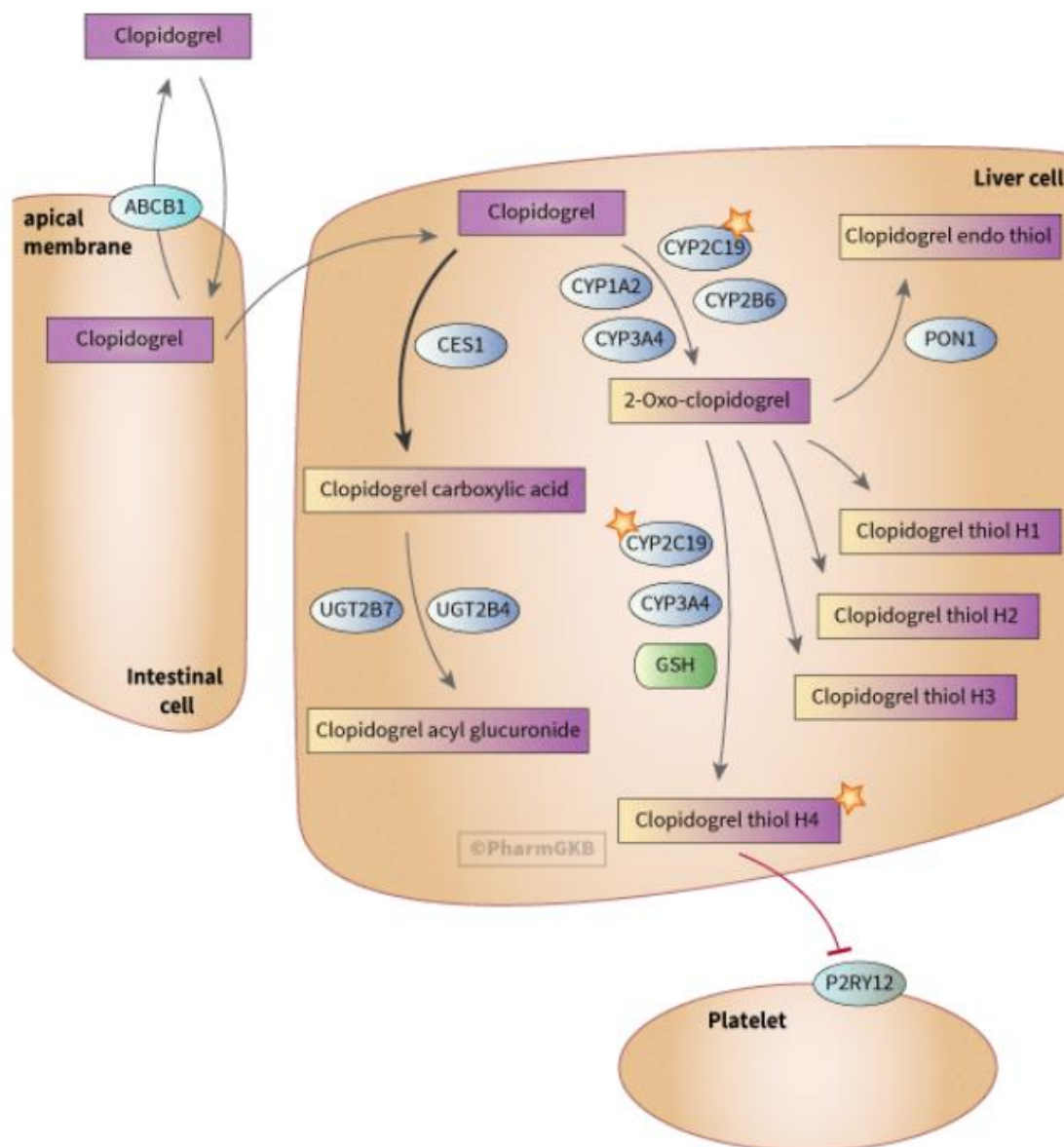


Figura 2. Representação via de metabolização do Clopidogrel. Fonte: (PharmGKB (<https://www.pharmgkb.org/>), 2023).

As enzimas do citocromo P450 são bastante importantes nessas reações, especialmente as da superfamília CYP que atuam nas reações de fase I, que ocorrem em sua maioria no fígado (LAM, 2019; RANG et al., 2016).

O complexo do citocromo P450 é composto por enzimas que participam catalisando a biotransformação oxidativa da maioria dos fármacos disponíveis no mercado atualmente, e possui as enzimas com alguns dos polimorfismos genéticos de maior relevância clínica e PGx

(Lam, 2019; Zanger; Schwab, 2013). Em humanos foram descritas 57 enzimas CYP funcionais e distintas, que são codificadas por genes de mesmo nome ((NUSSBAUM; WILLARD; HAMOSH, 2008)). Além dos 57 genes, são descritos 58 pseudogenes, que são divididos, baseando na similaridade em suas sequências de aminoácidos, em 18 famílias e 44 subfamílias. Desses genes e pseudogenes, 42 estão envolvidos no metabolismo de substâncias endógenas e xenobióticas e 15 são conhecidos por estar envolvidos na metabolização de fármacos em humanos (LAM, 2019).

Os genes codificadores das CYPs apresentam muitos polimorfismos, podendo ser SNPs, deleções, multiplicação de cópias ativas do gene, entre outros. Tais alterações podem alterar a atividade da enzima e consequentemente o metabolismo do fármaco. As variantes com atividade enzimática individual alterada são denominadas como alelos estrela (*), sendo geralmente representados como *CYP2D6*1*, onde o número pode representar um polimorfismo. O polimorfismo representado pelo número, pode ser um SNP, que consiste em na substituição de um único nucleotídeo, um indels, que são deleção ou inserção de bases no genoma, e haplótipo, podendo ser denominado como um grupo de alelos que tendem a ser herdados juntos (LAM, 2019; NUSSBAUM; WILLARD; HAMOSH, 2008; ONI-ORISAN et al., 2023).

Além de definir o genótipo do indivíduo, os alelos estrela são usados para prever o fenótipo metabólico apresentado pela pessoa, o que é chamado de fenótipo predito. O genótipo é representado da seguinte maneira: *CYP2D6*1/*2*. Tendo em vista que cada alelo estrela pode causar uma atividade enzimática diferente, o genótipo de uma pessoa pode ser usado para dizer como será a atividade enzimática de um indivíduo. Baseando-se no escore da atividade enzimática de cada alelo, pode-se classificar o metabolismo de um indivíduo como: metabolizados lentos (PMs - *poor metabolizers*), não possui uma atividade enzimática, metabolizadores intermediários (IMs - *intermediate metabolizers*), a atividade enzimática é reduzida, metabolizadores normais (EMs/NMs - *extensive/normal metabolizers*), tem uma atividade enzimática normal, e metabolizadores ultrarrápidos (UMs - *ultrarapid metabolizers*), possuem uma atividade enzimática aumentada (LAM, 2019).

O score de atividade (em inglês *activity score*), é baseado na atribuição de um score para a atividade enzimática daquele alelo estrela, onde os alelos que possuem uma atividade enzimática nula recebem o valor 0 (alelos nulos), os alelos estrela com atividade reduzida recebem um valor de 0,25 ou 0,5 (alelos lentos), os alelos com uma atividade normal recebem valor 1 (alelos extensivos) e os alelos com atividade enzimática aumentada recebem um valor >1 (alelos rápidos) (Caudle et al., 2020).

Para realizar o fenótipo predito, é feita uma soma dos scores de cada alelo que o indivíduo carrega, classificando-o assim em um dos tipos de metabolizadores. No caso do gene *CYP2D6*, por exemplo, quando o indivíduo tem o resultado da soma igual a 0 é classificado como um metabolizador lento, quando a resultado é entre 0,5 até 1,5 a pessoa é um metabolizador intermediário, o resultado sendo igual a 2 é um metabolizador normal e o resultado sendo maior que dois é um metabolizador ultrarrápido (Caudle et al., 2020).

Dentre as CYPs, os genes *CYP2C19* e *CYP2C9* são altamente polimórficos e participam da metabolização de grande parte dos fármacos disponíveis no mercado hoje.

2.2 CYP2C19

O *CYP2C19* é um gene da subfamília CYP2C, que está localizado no braço longo do cromossomo 10 (10q24.1-q24.3) contendo 90.209 nucleotídeos que codificam uma proteína composta por 490 aminoácidos (LAM, 2019; MRAZEK, 2010). Segundo Nussbaum et al., (2008), a enzima de fase I codificada pelo gene *CYP2C19* está presente na metabolização de mais de 50 fármacos disponíveis no mercado atualmente, incluindo antiplaquetários como o clopidogrel, inibidores de bomba de prótons, como por exemplo omeprazol, ansiolíticos, como o diazepam, e antidepressivos tricíclicos como por exemplo a amitriptilina e clomipramina (Lam, 2019; Mrazek, 2010).

Segundo o PharmGKB (<https://www.pharmgkb.org/>), a enzima *CYP2C19* está presente em 8 vias de agentes cardiovasculares e hematológicos, em 20 vias de agentes neurológicos, 6 de imunomoduladores, 4 relacionadas com agentes anticâncer, 2 de agentes respiratórios, 4 relacionadas com agentes gastrointestinais e 1 via relacionada a agentes anti-infecção.

O papel da *CYP2C19* nas vias de metabolização de fármacos demonstra a importância da mesma. De acordo com o PharmGKB, a *CYP2C19* atua juntamente com outras enzimas do citocromo P450 na conversão do pró-fármaco clopidogrel em seu metabólito ativo; além disso, a *CYP2C19* é a principal enzima responsável por metabolizar o voriconazol em variconazol N-óxido, e na metabolização da amitriptilina junto com a *CYP2D6*.

Segundo o Pharmavar, o *CYP2C19* possui 39 alelos identificados, que afetam a atividade enzimática de diferentes formas como mostrado na (Tabela 1).

Tabela 1. Função enzimática dos alelos do gene *CYP2C19*.

Função Enzimática	Alelos
Função Normal	*1, *11, *13, *15, *18, *28 e *38
Função Reduzida	*9, *10, *16, *19, *25, e *26
Sem função	*2, *3, *4, *5, *6, *7, *8, *22, *24, *35, *36 e *37
Função aumentada	*17
Desconhecido	*12, *14, *23, *29, *30, *31, *32, *33, *34 e *39

Fonte: (Pharmvar, 2023)

Genótipos com dois alelos com função aumentada (*17/*17) são classificados como UMs, indivíduos com dois alelos com função normal (*1/*1 ou *13/*13) são considerados NMs e genótipos com dois alelos sem função (*2/*2, *3/*3 ou *2/*3) ou com um alelo sem função e outro com atividade reduzida (*2/*9 ou *3/*16) são considerados PMs. Indivíduos carregando um alelo com função normal e um sem função (*1/*3) ou um com função aumentada e um sem função (*2/*17), são classificados como IMs, e pessoas que possuem dois alelos com atividade reduzida (*9/*9 ou *10/*19) ou um com função reduzida e outro com função normal (*10/*13) ou um com aumento de função e outro com função reduzida (*17/*9) são provavelmente IMs e quando possui um ou dois alelos cuja função não foi descrita ainda não conseguimos determinar o fenótipo predito (Tabela 2) (LEE et al., 2022).

Tabela 2. Classificação dos fenótipos *CYP2C19*.

<i>CYP2C19</i>	
Indivíduo com dois alelos de função enzimática aumentada (*17/*17)	Metabolizadores ultrarrápidos (UMs - ultrarapid metabolizers)
Indivíduos com dois alelos com função normal (*1/*1)	Metabolizadores normais (EMs/NMs - extensive/normal metabolizers)

Indivíduo com dois alelos sem função ou com um alelo sem função e outra atividade enzimática reduzida (*2/*2 ou *2/*9).	Metabolizados lentos (PMs - poor metabolizers)
Indivíduos com um alelo com função normal e um sem função ou função aumentada e um sem função (*1/*3 ou *17/*2)	Metabolizadores intermediários (IMs - intermediate metabolizers)
Indivíduos com dois alelos de atividade reduzida ou um com função reduzida e outro função normal ou função aumentada e outro reduzida (*9/*9 ou *9/*1 ou *17/*9)	Prováveis metabolizadores intermediários (IMs - intermediate metabolizers)

Fonte: (Lee et al., 2022).

De acordo com o Clinical Pharmacogenetics Implementation Consortium (CPIC) (<https://cpicpgx.org/>), que disponibiliza guidelines com elevados níveis de confiança para vários genes, incluindo o *CYP2C19*, há recomendações para os seguintes fármacos: clopidogrel, voriconazol, inibidores de bomba de prótons e para antidepressivos tricíclicos, sendo este último importante considerar o *CYP2D6* também.

Em relação ao clopidogrel, estudos demonstram que pacientes com alelos sem função enzimática (*2, *3, *4, *5 ou *6) possuem um risco maior de eventos cardiovasculares do que pacientes com outros alelos (SIMON et al., 2009). Os indivíduos com alelos de função reduzida apresentam menos metabólitos ativos e menor inibição plaquetária, além de apresentar maior índice de eventos cardiovasculares adversos (MEGA et al., 2009). Em contrapartida, quando possuem alelos com aumento da função há um maior risco de sangramentos (SIBBING et al., 2010). Tendo em vista os possíveis efeitos adversos, a CPIC recomenda o uso desse fármaco para NMs com dose de 75mg por dia, enquanto para PMs e IMs a recomendação é evitar o uso do clopidogrel e fazer uso de outros antiplaquetários, como o prasugrel ou ticagrelor.

Já em relação ao voriconazol, um estudo com pacientes que realizaram transplante de pulmão e possuíam genótipo *1/*17 ou *17/*17 que receberam voriconazol demonstraram um risco maior, estatisticamente significativo, de aumento de carcinoma de células escamosas (MARSH et al., 2016; WILLIAMS; ARRON, 2016). Em sua guia a CPIC recomenda para UMs e PMs que seja feito o uso de outro fármaco, não metabolizado pela enzima *CYP2C19*, e para IMs e NMs é recomendada a dosagem padrão.

Para tratamento com omeprazol, lansoprazol, pantoprazol e dexlansoprazol, a *guideline* da CPIC recomenda que UMs aumente a dose diária em 100%, podendo ser dividida durante o

dia, e monitorar a eficácia; NMs iniciar com a dose padrão, mas para o tratamento de *Helicobacter pylori* considerar aumentar a dose 50% a 100% e monitorar a eficácia. Indivíduos PMs e IMs: iniciar com dose padrão, mas para tratamento crônico quando alcançar a eficácia considerar diminuir em 50% a dose e monitorar a eficácia (LIMA et al., 2021).

Os antidepressivos tricíclicos, tais como amitriptilina e nortriptilina, também possuem uma *guideline* disponibilizada pela CPIC. A recomendação é para NMs e IMs começar a terapia com a dose inicial recomendada e para os UMs e PMs é evitar os antidepressivos tricíclicos, e usar fármacos que utilizem outras vias metabólicas. Entretanto, para PMs que forem usar esse fármaco por não ter uma alternativa, realizar a redução de 50% da dose recomendada e fazer monitoramento (HICKS et al., 2017).

2.3 CYP2C9

A enzima *CYP2C9*, codificada pelo gene de mesmo nome e pertencente à família do citocromo P450, é uma enzima que atua na metabolização de fase I, desempenhando um papel importante na oxidação de compostos endógenos e xenobióticos. Identificada na região 10q24 do cromossomo, sendo uma das enzimas CYP com maior expressão no fígado humano, com estimativas de ser responsável pela depuração de 15% a 20% dos fármacos metabolizados nas reações de fase I (LEE; GOLDSTEIN; PIEPER, 2002; RETTIE; JONES, 2005; VISSER et al., 2005).

De acordo com o PharmGKB, o *CYP2C9* está envolvido em mais de 20 vias de metabolização diferentes e possui *guideline* na CPIC para fármacos anti-inflamatórios não esteróides (AINEs), fenitoína e varfarina (JOHNSON et al., 2017; KARNES et al., 2021; THEKEN et al., 2020).

O papel do *CYP2C9* nas vias de metabolização dos fármacos pode alterar a biotransformação, dependendo do polimorfismo. Segundo o Pharmvar, o gene *CYP2C9* possui 85 alelos estrela identificados, como mostrado na (Tabela 3).

Tabela 3. Função enzimática dos alelos do gene *CYP2C9*.

Função Enzimática	Alelo
Função Normal	*1 e *9

Função Reduzida	*2, *4, *5, *8, *11, *12, *14, *16, *23, *26, *28, *29, *30, *31, *37, *38, *44, *46, *50, *55 e *61
Sem Função	*3, *6, *13, *15, *24, *25, *33, *35, *39, *42, *43, *45 e *52
Desconhecido	*7, *10, *17, *18, *19, *20, *21, *22, *27, *32, *34, *36, *40, *41, *47, *48, *49, *51, *53, *54, *56, *58, *59, *60, *2, *63, *64, *65, *66, *67, *68, *69, *70, *71, *57, *72, *73, *74, *75, *76, *77, *78, *79, *80, *82, *83, *84 e *85

Fonte: (Pharmvar, 2023).

De acordo com a CPIC, genótipos com dois alelos de função normal (*1/*1) são considerados NMs, genótipos com um alelo de função enzimática normal e outro com atividade enzimática reduzida ou um alelo sem função e com outro de função reduzida ou dois alelos com função reduzida (*1/*2 ou *1/*3 ou *2/*2) são considerados IMs e indivíduos com um alelo de função reduzida e outro sem função ou dois alelos sem função enzimática (*2/*3 ou *3/*3) são chamados de PMs (Tabela 4).

Tabela 4. Classificação dos fenótipos CYP2C9

CYP2C9	
Indivíduos com dois alelos com função normal (*1/*1)	Metabolizadores normais (EMs/NMs - extensive/normal metabolizers)
Indivíduo com um alelo de função normal e outro com função reduzida ou um sem função e um com função reduzida ou dois com função reduzida (*1/*2 ou *1/*3 ou *2/*2)	Metabolizadores intermediários (IMs - intermediate metabolizers)
Indivíduos com um alelo de função reduzida e outro sem função ou dois sem função (*2/*3 ou *3/*3)	Metabolizados lentos (PMs - poor metabolizers)

Fonte: (CPIC, 2023)

Em estudos com voluntários brancos saudáveis, foi observado que a depuração do ibuprofeno foi reduzida em indivíduos com os alelos *CYP2C9*3* e *CYP2C8*3* comparando com os alelos *CYP2C9*1* e *CYP2C8*1* (KARANIEWICZ-ADA; UCZAK; GLWKA, 2009). Segundo Pilotto et al. (2007), os indivíduos com os genótipos *CYP2C9*1/*2* e **1/*3* tiveram uma frequência, estatisticamente significativa, maior de sangramento gastroduodenal em casos *versus* controles. Em outro estudo de casos e controles, o alelo **2* teve uma associação com risco de sangramentos. Nesse mesmo estudo, o haplótipo *CYP2C8*3-CYP2C9*2* foi associado com maior risco de sangramentos relacionados com AINEs (BLANCO et al., 2008).

De acordo com a *guideline* da CPIC, para AINEs é recomendado que NMs iniciem com a dosagem padrão, enquanto para IMs com *activity score* de 1, iniciem com a menor dosagem recomendada pelo menor período possível, além de acompanhar possíveis efeitos adversos. Nos PMs é recomendado iniciar a terapia com redução de 25% a 50% da dose padrão e considerar terapias alternativas que não são influenciadas pelo *CYP2C9* (THEKEN et al., 2020).

Adicionalmente aos AINEs, segundo Ramasamy et al. (2010), o alelo **2* aumenta a concentração de fenitoína livre no plasma, e, quando presente em homozigose aumenta o risco de toxicidade com o fármaco. A *guideline* da CPIC recomenda para a fenitoína considerar também os genótipos para *HLA-B* junto com o *CYP2C9*. No caso de *HLA-B*15:02 negativo*, considerar para NMs de *CYP2C9*, nenhum ajuste na dosagem padrão recomendada, para IMs com *activity score* igual a 1, é recomendado iniciar a terapia com a dose padrão e nas próximas doses uma redução de 25% da dosagem padrão, enquanto nos PMs a redução deve ser 50%. Os indivíduos com *HLA-B*15:02 positivo*, independente do genótipo em *CYP2C9*, não fazer uso da fenitoína (KARNES et al., 2021).

Já considerando a varfarina, de acordo com Johnson et al. (2017), o alelo **2* pode prejudicar o metabolismo da varfarina em até 40%, enquanto o **3* pode chegar a 90%, por isso o paciente homozigoto para esses genótipos tem risco aumentado para sangramentos durante a terapia com varfarina e precisam de doses menores para atingir níveis semelhantes de anticoagulação sem efeitos adversos.

Tendo em vista todos esses polimorfismos, é importante considerar que a frequência deles varia em diferentes populações devido a alguns fatores como isolamento geográfico e entre outros (LLERENA et al., 2014). As diferentes frequências dos alelos podem ser usadas para correlacionar, como por exemplo com questões de saúde e ancestralidade. Segundo Rodrigues-Soares et al. (2020), o alelo *CYP2D6*41* foi positivamente associado com ancestralidade

européia, enquanto o alelo *CYP2D6*17* foi associado positivamente com a ancestralidade africana, indicando que quando maior a porcentagem da ancestralidade europeia maior a probabilidade de o indivíduo ter alelo **41* e o mesmo vale para o alelo **17*, para com ancestralidade africana.

2.4 Variabilidade interétnica e Ancestralidade Genômica

A diversidade genética é a variação existente no DNA das populações humanas. Essas variantes, inclusive as de relevância farmacogenômica, sofrem uma variação na frequência entre os grupos populacionais e étnicos. Conhecer essa diversidade e distribuição da genética humana nos possibilita uma melhor compreensão de doenças genéticas, da evolução humana, das respostas à fármacos entre outras áreas (1000 GENOMES PROJECT CONSORTIUM et al., 2010; BERGSTRÖM et al., 2020; JORDE et al., 2000; WATKINS et al., 2003). O genoma humano típico consegue diferenciar de um genoma humano de referência entre 4,1 e 5,0 milhões de *locus* (AUTON et al., 2015) e segundo Mallick e seus colaboradores (2016), na média dois indivíduos irão se diferenciar por cerca de 3 milhões de pares de base, o que equivale a 0.1% da nossa composição genética total.

Quando agrupamos os indivíduos em populações observamos que as variantes mudam sua frequência nas diferentes populações do mundo, por meio dessa diferença é possível destacar alguns padrões na distribuição alélica (JORDE et al., 2000; MALLICK et al., 2016). De acordo com Xing et al., (2010), 78% dos alelos comuns (alelos que frequência é igual ou maior que 5%) são compartilhados entre populações de quatro grandes regiões (África, Europa, Leste Asiático e Índia), 92% desses alelos são compartilhados entre populações de duas das grandes regiões, 7.4% estão apenas em populações na região da África e 0.5% são encontrados apenas em grupos não africanos. Em outros estudos, também é possível observar que os alelos comuns são mais prováveis de serem compartilhados entre as grandes regiões e os alelos mais raros são mais prováveis de serem encontrados em uma população específica (AUTON et al., 2015; BIDDANDA; RICE; NOVEMBRE, 2020).

O nível de diversidade genética não é considerado o mesmo em todas as populações, ou seja, algumas populações possuem uma maior diversidade genética que outras. As populações do continente africano possuem uma maior diversidade genética quando comparadas com populações não africanas (AUTON et al., 2015; JORDE et al., 2000). A

variação na diversidade genética e na frequência dos alelos entre as populações é a consequência da combinação de diversos fatores evolutivos como a recombinação gênica, a mutação, o fluxo gênico e à deriva genética (CHEVITARESE, 2009).

A causa das variações na molécula de DNA acontece devido modificações na sequência dessa molécula, essas modificações são chamadas de mutações. As mutações ocorrem de forma natural e aleatória, podendo ser causadas por erros químicos no processamento da molécula e/ou a exposição a agentes ambientais e podendo ser transmitida para descendentes quando ocorrem em células germinativas. As mutações podem ajudar, prejudicar ou não desempenhar nenhum papel, aos indivíduos, que estão sofrendo pressões ambientais em seus habitats para sobreviver e reproduzir (Griffiths et al., 2013).

A recombinação de acordo com Griffiths et al., (2013), é a formação de novas rearranjos alélicos que geralmente acontece durante o processo de meiose na formação de células germinativas, ou seja, no processo de meiose acontece uma recombinação entre os alelos e há formação de novos rearranjos entre esses alelos, logo a recombinação é responsável por reagrupar os alelos criados pela mutação, ajudando na criação de uma maior variabilidade genética nas populações (Hartl; Clark, 2010).

Uma outra forma de aumento da diversidade genética em uma população é por meio do fluxo gênico ou migração de indivíduos. A migração de indivíduos propicia que o fluxo gênico ocorra, permitindo que haja uma troca de alelos entre as populações podendo ter a inserção de alelos novos ou extinção de alelos em determinada população. Segundo Bergström et al. (2020), houve um fluxo gênico ao longo do tempo entre as populações africanas, sendo possível estimar o fluxo gênico entre os Mbuti, Biaka e Yoruba até os dias atuais.

Alguns alelos podem aumentar ou diminuir as chances de sobrevivência e reprodução do indivíduo; quando auxiliam em aumentar o sucesso reprodutivo pode haver gradualmente o aumento da frequência desses alelos ao passar cada geração e nos casos em que levarem à redução do sucesso reprodutivo a consequência seria diminuição ou a extinção desses alelos com o passar das gerações, a seleção natural é um dos mecanismos que pode atuar nesses processos (Hartl; Clark, 2010). A seleção natural é um mecanismo evolutivo que atua por meio de pressões ambientais para selecionar fenótipos, e conseqüentemente os genótipos que possibilitam a expressão daquele fenótipo, promovendo o sucesso reprodutivo naquele ambiente (RIDLY, 2006).

Além da seleção natural, outro mecanismo evolutivo é à deriva genética, que por aleatoriedade escolhe quais indivíduos, e seus genótipos, irão ou não sobreviver e reproduzir, e assim através do acaso à deriva genética altera as frequências dos alelos nas populações. Esse

mecanismo evolutivo é responsável por diminuir a variabilidade genéticas nas populações e tem uma influência maior em populações menores (Hartl; Clark, 2010; Ridley, 2006).

A diminuição aleatória da diversidade genética pode ocorrer por vários motivos, quando um desses motivos reduz uma população por pelo menos uma geração é dito que essa população passou por um gargalo ou pelo efeito de gargalo de garrafa. Junto com redução da população, há também a diminuição drástica da variabilidade genética daquela população (Hartl; Clark, 2010; Ridley, 2006). De acordo com Auton et al., (2015), as populações europeia, asiática e nativo americana compartilham ter passado por um severo efeito de gargalo entre 15 mil e 20 mil anos atrás enquanto nesse mesmo período as populações africanas sofreram um impacto menor desse efeito.

A diminuição variabilidade genética aleatória também pode ser causa pelo efeito de fundador, que consiste em um grupo de indivíduos de uma população migram e começam uma nova população. Os indivíduos fundadores dessa nova população podem não representar geneticamente toda a população original e nova população que será formada terá uma menor diversidade genética comparado com a população de origem (Hartl; Clark, 2010; Ridley, 2006). Esse efeito também pode ter influenciado a menor diversidade genética das populações não africanas, pois de acordo com teoria de saída da África, pequenos grupos de humanos realizam a migração e ocupação dos outros continentes ao redor do mundo (HENN; CAVALLI-SFORZA; FELDMAN, 2012; HENRY, 2019).

Em algumas regiões do mundo a diversidade genética pode acompanhar padrões geográficos, ou seja, a diferenciação genética aumenta com a distância geográfica (BRADBURY; RALPH; COOP, 2016; HANDLEY et al., 2007; PETER; PETKOVA; NOVEMBRE, 2020).

Quando um alelo está presente em um grupo numa frequência alta e em outros grupos com uma frequência baixa, este pode ser correlacionado com a ancestralidade da população que apresenta a alta frequência (RODRIGUES-SOARES et al., 2020; URBAN, 2010). Dados como esse tornam possível inferir a ancestralidade genômica e correlacioná-la com outros polimorfismos, como os de interesse farmacogenômico.

A ancestralidade genômica é a informação genética de um indivíduo que revela de qual população parental o indivíduo descende. Essa ancestralidade pode ser inferida de duas maneiras, por meio dos Marcadores Informativos de Ancestralidade (AIMs), que são variantes selecionadas para inferir a ancestralidade, ou pelo método *genome-wide*, utilizando milhares/milhões de marcadores aleatórios (PADHUKASAHASRAM, 2014; SCUDELER; RODRIGUES-SOARES, 2020a).

Nos últimos séculos, com o advento de novas tecnologias, houve uma crescente onda migratória entre as populações, proporcionando a formação de populações miscigenadas, como por exemplo a da América Latina. Devido ao processo de miscigenação das populações ter ocorrido de forma não uniforme, os países com populações miscigenadas podem apresentar ancestralidade genômica diferente da etnia autodeclarada (TAN-KOI et al., 2018; ZHANG et al., 2019). Segundo Céspedes-Garro et al., (2014), países da América Central possuem uma variabilidade interétnica dentro da América Central e com outras populações latino-americanas.

A autodeclaração étnica envolve mais fatores do que apenas a ancestralidade genômica. Para Tan-Koi et al. (2019), etnia e raça são misturas de construções biológicas e sociais, sendo altamente complexas e dinâmicas. Zhang et al. (2019), define etnia como uma classificação sociológica de seres humanos que compartilham características biológicas, não levando em conta a ancestralidade genômica. A identificação com alguma etnia não depende exclusivamente dos componentes genéticos da pessoa, mas sim, de fatores socioeconômicos do indivíduo (ZHANG et al., 2019).

Por meio da ancestralidade genômica é possível realizar o perfil de ancestralidade individual, mostrando a porcentagem de cada componente parental que o indivíduo descende (p. ex. um indivíduo pode ser 10% europeu, 40% nativo-americano e 50% africano (Fig. 3) (Scudeler & Rodrigues-Soares, 2020).

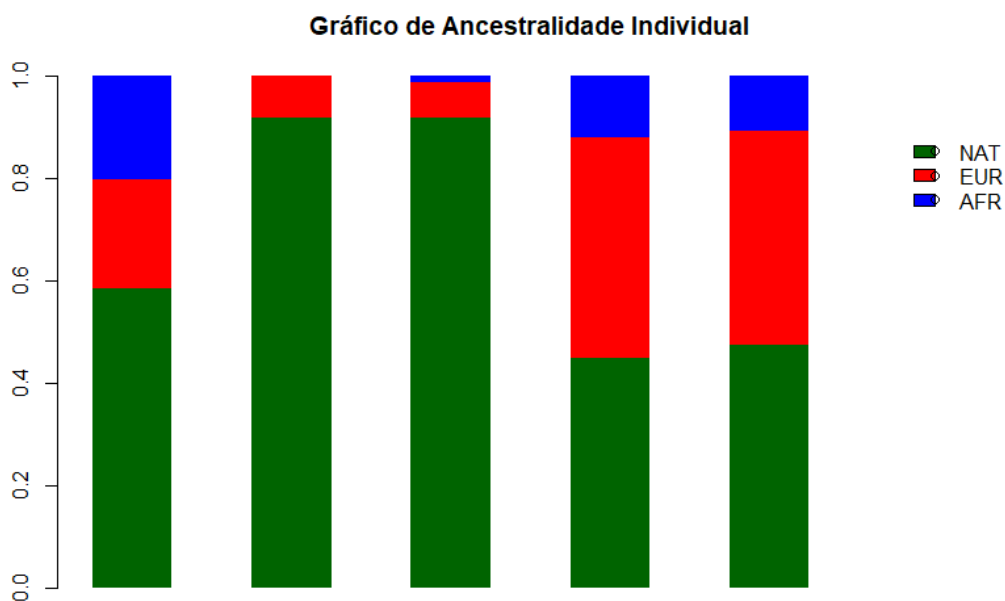


Figura 3. Gráfico de barras da ancestralidade continental individual. Representando 5 indivíduos de uma população. (AFR, África; EUR, Europa; NAT, Nativo Americano). Fonte: Elaborado pelo autor, 2023.

O perfil de ancestralidade pode ser inferido para componentes continentais, separando os componentes das populações continentais como: europeia, africana, nativo-americana e asiática (Fig 3). Esse perfil também pode ser mais específico, inferindo a ancestralidade em componentes subcontinentais, onde podemos diferenciar a ancestralidade em populações mais específicas, como: oeste africano, leste africano, sul asiático, sudeste asiático, norte europeu, sul europeu (KEHDY et al., 2015; PADHUKASAHASRAM, 2014) (Fig. 4). Um estudo desenvolvido no Brasil por Kehdy et al. (2015), demonstra que a população da cidade de Salvador-BA, possui uma ancestralidade da África Ocidental, do povo Yoruba/Mandenka, enquanto a ancestralidade africana das populações de Bambuí-MG e Pelotas-RS são da África Oriental, do povo Bantu.

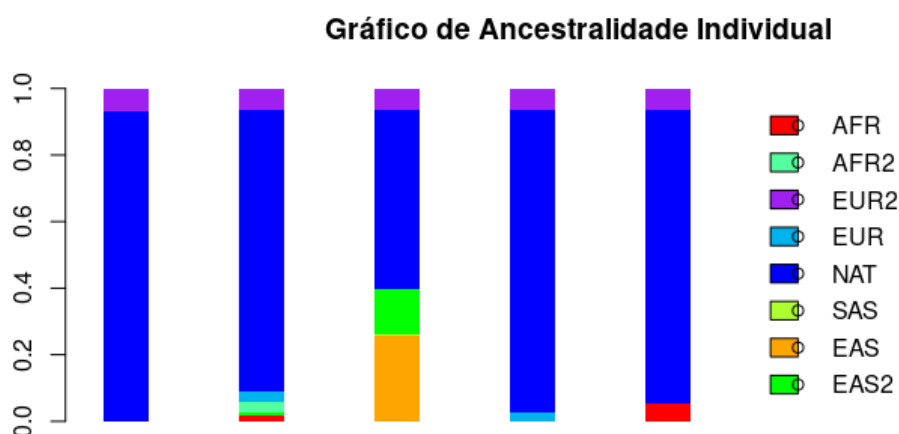


Figura 4. Gráfico de barras da ancestralidade subcontinental individual. Representando 5 indivíduos de uma população. (AFR, África; AFR2, África 2; EUR2, Europa 2; EUR, Europa; NAT, Nativo Americano; SAS, Sul Asiático; EAS, Leste Asiático; EAS2, Leste Asiático 2). Fonte: Elaborado pelo autor, 2023

Urban (2010), afirma que a ancestralidade pode influenciar no desfecho de tratamentos, sendo assim importante considerar a ancestralidade genômica em estudos de saúde, além de ser importante em iniciativas de estudo da variabilidade genética entre as populações como o 1000 Genomes Project Consortium (AUTON et al., 2015).

O 1000 Genomes é um banco de dados público que possui cerca de 2.500 genomas humanos completos e seu objetivo é garantir acessibilidade e usabilidade desses dados, como

também inclusão de novos dados que surgirem seguindo os princípios de amostragem estabelecidos pela The International Genome Sample Resource (IGSR). O banco de dados possui amostras de 26 populações da África, Leste Asiático, Sul Asiático, Europa e Américas. As análises feitas com os dados do projeto comparam variantes genéticas compartilhadas entre populações e podem ser relacionadas com a história delas (AUTON et al., 2015).

Apesar de haver estudos sobre a ancestralidade genômica e farmacogenômica, ainda existem lacunas que precisam ser respondidas. Levando em consideração o banco de dados do 1000 Genomes, ainda não há estudos relacionando ancestralidade genômica subcontinental com alelos de interesse farmacogenômico.

3. Hipótese

Com os dados de alta densidade do banco de dados 1000 Genomes, temos a hipótese de conseguir correlações entre frequências de alelos dos genes *CYP2C19* e *CYP2C9* e ancestralidade subcontinental, uma análise pioneira na área de genômica humana com grande potencial de publicação.

4. Justificativa

Os avanços que estão sendo alcançados na área da genômica, estão evidenciando a importância de vários biomarcadores que influenciam na toxicidade ou na resposta em geral a fármacos (RAMOS et al., 2014). A farmacogenômica é responsável por levar em consideração as diferenças individuais na escolha do tratamento, permitindo um aprimoramento na eficácia desse tratamento reduzindo possíveis efeitos adversos ou falhas na terapia e reduzindo também os custos com os tratamentos (VELLEKOOP et al., 2022).

Existe uma variabilidade de respostas a tratamentos entre os indivíduos e o sucesso de um tratamento depende diretamente da resposta desse indivíduo (LI et al., 2011b). Atualmente na maioria dos sistemas de saúde no mundo, quando uma pessoa é diagnosticada com alguma patologia, esse indivíduo receberá um tratamento padrão, contudo nem todos os indivíduos apresentam a resposta esperada para essa terapia, alguns indivíduos irão apresentar reações adversas e outros irão apresentar falha no tratamento.

A farmacogenômica propõe alterações nesse sistema, onde para escolher o tratamento é levado em consideração essa variabilidade de respostas entre os indivíduos, sendo que essa diversidade de respostas a medicamentos pode ser influenciada pela genética (MA; LU, 2011). A genética auxilia na

previsão do tipo de resposta que esse indivíduo terá ao tratamento e através dessa informação propor uma terapia mais eficiente, e com uma redução de casos de reações adversas e falhas ao tratamento ((SCUDELER; RODRIGUES-SOARES, 2020b).

Outra vantagem da farmacogenômica é apresentar um custo-benefício maior quando comparado o modelo de terapia atual. Morris e colaboradores (2022), identificou que tratamentos guiados por teste farmacogenômicos obtiveram um melhor custo-benefício ou tiveram uma economia no custo, assim como Koleva-Kolarova et al., (2023) e Verbelen et al., (2017), observou que o custo-benefício de tratamentos guiados medicina personalizada é maior.

A ancestralidade genômica é informação genética que um indivíduo possui sobre a população parental que ele desce. Essa informação é usada para rastreamento de indivíduos em investigações forenses, corrigir estratificação populacional em estudos de associação genômica e no desenvolvimento de abordagens personalizadas para diagnóstico e tratamento (PADHUKASAHASRAM, 2014). Então a utilização da ancestralidade genômica pode ser um fator importante na hora das escolhas de terapias mais individualizada.

A associação da ancestralidade genômica com a farmacogenômica pode direcionar a escolha de terapias mais assertivas para o indivíduo ou para populações mais específicas, possibilitando a criação de painéis de testes farmacogenômicos mais específicos para cada população com menores custos (RODRIGUES-SOARES et al., 2020). Sendo então possível colocarmos em pauta a importância das escolhas dos alelos a serem usados em testes farmacogenéticos na rotina clínica em vários locais do mundo. Além disso, cientificamente, a metodologia será delineada pela primeira vez neste estudo, sendo reprodutível para outras populações posteriormente.

O banco de dados 1000 Genomes foi estudado para vários desfechos na área de saúde, no entanto não há até o momento um estudo de correlação entre ancestralidade genômica e alelos de interesse farmacogenético com estes dados. Além disso, cientificamente, a metodologia será delineada pela primeira vez neste estudo, sendo reprodutível para outras populações posteriormente.

5. Objetivo

4.1 Objetivos Geral

Correlacionar a ancestralidade genômica subcontinental de oito componentes parentais com frequências de alelos de interesse farmacogenômico dos genes *CYP2C9* e *CYP2C19*.

4.2 Objetivos específicos

- Inferir a ancestralidade genômica subcontinental de 8 componentes das 26 populações do 1000 Genomes;
- Filtrar os dados de variantes dos genes *CYP2C9* e *CYP2C19* e inferir seus alelos estrela;
- Calcular a frequência alélica dos alelos estrela dos genes *CYP2C9* e *CYP2C19* nas populações do 1000 Genomes;
- Realizar a correlação entre os oito componentes de ancestralidade subcontinental e a frequência de variantes de interesse farmacogenômico nos dois genes.

5. Metodologia

Amostras

Foram utilizados dados das frequências de variantes de interesse farmacogenômico dos genes *CYP2C9* e *CYP2C19* e de ancestralidade *genome-wide* presentes em populações do 1000 Genomes. Utilizamos o genoma de 2504 indivíduos presentes em 26 populações diferentes, sendo todos retirados do banco de dados público 1000 Genomes (AUTON et al., 2015).

Os indivíduos estão distribuídos em 26 populações ao redor do mundo, sendo 91 indivíduos britânicos vivendo na Inglaterra e Escócia (GBR), 99 finlandeses (FIN), 99 descendentes de europeus residentes em Utah (CEU), 107 indivíduos da Toscana na Itália (TSI), 107 ibéricos da Espanha (IBS), 104 japoneses em Tóquio (JPT), 105 chineses Han da região sul (CHS), 103 chineses Han de Beijing (CHB), 93 chineses Dai em Xishuangbanna, China (CDX), 99 vietnamita Kinh (KHV), 103 indianos Gujarati em Houston, TX (GIH), 102 indianos Telugu vivendo na Inglaterra (ITU), 102 indivíduos Tamil do Sri Lanka vivendo na Inglaterra (STU), 86 indivíduos de Bengali em Bangladesh (BEB), 96 Punjabi em Lahore no Paquistão (PJI), 113 Mandinka em Gâmbia (GWD), 85 indivíduos Mande de Serra Leoa (MSL), 108 Yorubas da Nigéria (YRI), 99 Esan da Nigéria (ESN), 99 indivíduos Luhya no Quênia (LWK), 96 afro-caribenhos em Barbados (ACB), 61 afro-americanos sudoeste dos Estados Unidos (ASW), 85 peruanos de Lima, Peru (PEL), 64 descendentes de mexicanos em Los Angeles (MXL), 94 colombianos de Medellín (CLM) e 104 porto-riquenhos (PUR) (Fig. 5) (AUTON et al., 2015).

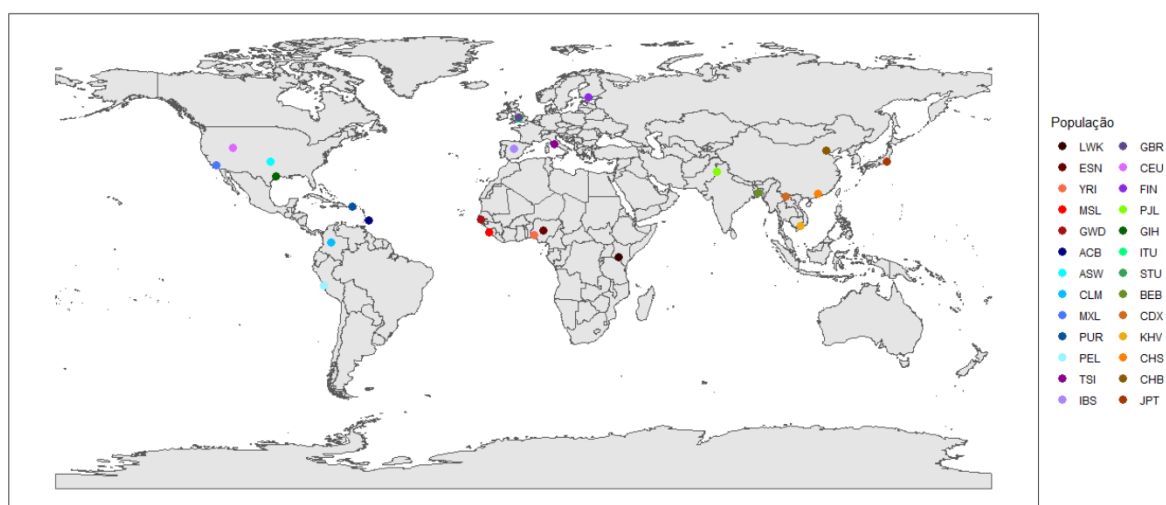


Figura 5. Mapa com localização geográfica de todas as populações. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandinka de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana

vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio). Fonte: Elaborado pelo autor, 2023

O download e os filtros de controle de qualidade dos dados do 1000 Genomes Project foi realizado pelo Laboratório de Diversidade de Genética Humana (LDGH) e em seguida os seguintes filtros para controle de qualidade foram aplicados: a remoção do chr 0, de dados duplicados, de dados faltantes, remoção de variantes A|T e C|G e remoção de 100 de variantes heterozigose, utilizando o script MosaiQC (<https://github.com/ldgh/MosaiQC-public>).

Análise de ancestralidade genômica

A ancestralidade foi inferida pelo software Admixture (Alexander et al., 2009), para escolha dos SNPs foram aplicados filtros para exclusão de variantes bi alélicas, que possuíam algum dado faltante e com uma frequência menor (MAF) que 1%. Também excluimos variantes cujo desequilíbrio de ligação (LD) era maior ou igual a $r^2 = 0.8$, resultando em 3.489.554 SNPs. Para a realização da análise no software Admixture, utilizamos o modo não-supervisionado, assumindo oito ($k=8$) componentes de ancestralidade.

O output do Admixture foi um arquivo com a porcentagem de ancestralidade de cada indivíduo e através da função *mean* na Plataforma R (www.r-project.org), que é um ambiente de software livre para computação estatística e gráficos, calculamos a média de cada ancestralidade para cada população (https://github.com/ManochioCaique/Genomic-ancestry-and-Star-allele/blob/main/Calculo_media_ancestralidade.R).

Inferência dos alelos estrela e cálculo das frequências

Os alelos estrela e os fenótipos preditos (importantes para interpretação clínica dos resultados) dos dois genes foram inferidos pela ferramenta Stargazer (Lee et al., 2019), que é um software que consegue identificar alelo estrela de dados *Next Generation Sequencing* (NGS) e depois já atribuir o score de atividade e realiza o fenótipo predito. Após inferidos os alelos estrela e o fenótipo, o *output* foi importado para a plataforma R.

Na plataforma R, realizamos o cálculo de frequências alélicas e dos fenótipos para os dois genes. Para o cálculo das frequências alélicas e do fenótipo em cada população, utilizamos a função *makefreq* do pacote *adegenet* (Jombart; Ahmed, 2011), que consiste em um pacote dedicado a realização de análise de dados genéticos (https://github.com/ManochioCaique/Genomic-ancestry-and-Star-allele/blob/main/Calculo_frequencia_alelica_CYP2C9.R).

Correlação entre ancestralidade genômica e frequência dos alelos estrela

No cálculo de correlação entre a frequência alélica e fenótipo predito com os componentes de ancestralidades, utilizamos a função *cor.test*, (https://github.com/ManochioCaique/Genomic-ancestry-and-Star-allele/blob/main/Correla%C3%A7%C3%A3o_alelica_CYP2C9.R)

Os resultados foram confirmados quanto à anotação haplotípica, predição fenotípica e anotações clínicas com os bancos de dados PharmVar (<https://www.pharmvar.org/>) e PharmGKB (<https://www.pharmgkb.org/>).

6. Resultados e Discussão

6.1 Ancestralidade

A ancestralidade foi inferida em uma análise não supervisionada no ADMIXTURE, utilizando uma lista com 3.489.554 SNPs para $k = 8$. A inferência da ancestralidade foi feita usando dados de 2504 indivíduos, pertencente a 26 diferentes populações do 1000 Genomes Project, conseguindo então fazer uma separação subcontinental dos componentes (Tabela 5) (Fig. 6).

Tabela 5. Porcentagem de ancestralidade $k = 8$ para cada população. **NEUR**, Norte Europeu; **WAFR**, Oeste Africano; **NAT**, Nativo Americano; **EAS-W**, Leste Asiático – Oeste; **SEUR**, Sul Europeu; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático - Leste; **SAS**, Sul Asiático. **POP**, população; **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandinka de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de

Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japoneses de Tóquio;.

POP	EUR	AFR	NAT	EAS	EUR2	AFR2	EAS2	SAS
LWK	0,0010	0,0443	0,0005	0,0018	0,0353	0,9003	0,0021	0,0143
ESN	0,0000	0,3852	0,0000	0,0000	0,0000	0,6146	0,0000	0,0000
YRI	0,0001	0,4370	0,0001	0,0000	0,0002	0,5624	0,0000	0,0000
MSL	0,0000	0,7742	0,0001	0,0000	0,0002	0,2250	0,0000	0,0001
GWD	0,0012	0,9424	0,0001	0,0002	0,0145	0,0387	0,0003	0,0022
ACB	0,0410	0,4273	0,0013	0,0024	0,0681	0,4530	0,0008	0,0057
ASW	0,0732	0,3685	0,0346	0,0037	0,1187	0,3911	0,0042	0,0056
CLM	0,0368	0,0396	0,2792	0,0047	0,5885	0,0377	0,0071	0,0060
MXL	0,0443	0,0226	0,4883	0,0059	0,3881	0,0131	0,0270	0,0102
PUR	0,0245	0,0798	0,1487	0,0042	0,6741	0,0556	0,0068	0,0060
PEL	0,0201	0,0110	0,8056	0,0069	0,1373	0,0123	0,0036	0,0029
TSI	0,0469	0,0000	0,0008	0,0028	0,9163	0,0000	0,0024	0,0306
IBS	0,1116	0,0063	0,0023	0,0029	0,8655	0,0040	0,0028	0,0041
GBR	0,3777	0,0000	0,0008	0,0000	0,6207	0,0000	0,0000	0,0004
CEU	0,3878	0,0000	0,0009	0,0000	0,6093	0,0000	0,0002	0,0013
FIN	0,9114	0,0000	0,0020	0,0028	0,0640	0,0000	0,0182	0,0012
PJL	0,0956	0,0000	0,0059	0,0087	0,1252	0,0001	0,0089	0,7552
GIH	0,0898	0,0000	0,0032	0,0012	0,0619	0,0000	0,0019	0,8418
ITU	0,0288	0,0000	0,0029	0,0153	0,0393	0,0000	0,0095	0,9038
STU	0,0293	0,0000	0,0027	0,0201	0,0308	0,0000	0,0149	0,9018
BEB	0,0570	0,0000	0,0058	0,0905	0,0206	0,0000	0,0370	0,7887
CDX	0,0000	0,0000	0,0011	0,9771	0,0000	0,0000	0,0189	0,0027
KHV	0,0012	0,0000	0,0023	0,8698	0,0015	0,0000	0,1027	0,0223
CHS	0,0000	0,0000	0,0004	0,6029	0,0000	0,0000	0,3964	0,0001
CHB	0,0029	0,0000	0,0027	0,4190	0,0016	0,0000	0,5730	0,0004
JPT	0,0000	0,0000	0,0002	0,0162	0,0000	0,0000	0,9832	0,0001

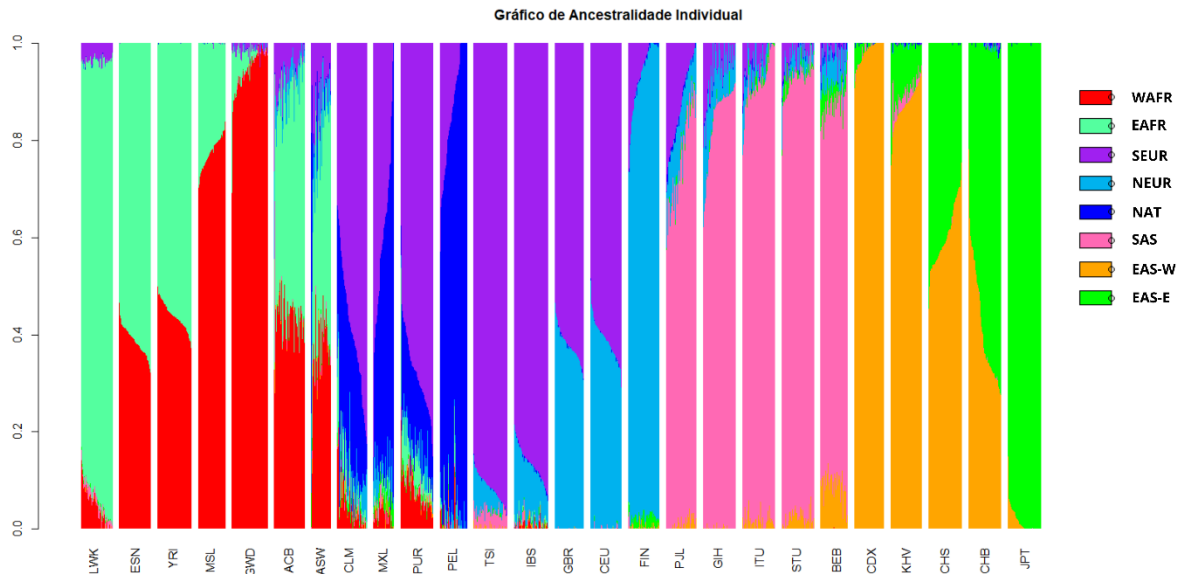


Figura 6. Gráfico de barras da ancestralidade subcontinental individual. (**EAS-W**, Leste Asiático – Oeste; **SEUR**, Sul Europeu; **NEUR**, Norte Europeu; **NAT**, Nativo Americano; **EAFR**, Leste Africano; **EAS - E**, Leste Asiático- Leste; **SAS**, Sul Asiático; **WAFR**, Oeste Africano. **POP**: população. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJI**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio). Fonte: Elaborado pelo autor, 2023.

A separação dos componentes se deu da seguinte maneira: componente norte europeu (NEUR), componente sul europeu (SEUR), componente leste asiático – oeste (EAS-W), componente leste asiático – leste (EAS-E), sul asiático (SAS), componente africano oeste (WAFR), componente africano leste (EAFR) e nativo americano (NAT). As populações parentais africana, asiática e europeia tiveram seus componentes separados em subcontinentais (Fig. 7), assim como em trabalhos anteriores utilizando os indivíduos do 1000 Genomes Project. Na análise feita por Auton et al., (2015) utilizando um diferente set de SNPs, mas com o mesmo valor de k houve a separação dessas três populações parentais em componentes subcontinentais. Já nas análises de Kehdy et al., (2015), para o mesmo valor de k , também com um set diferente de 331.790 SNPs e utilizando dados públicos e particulares, obteve a separação apenas das populações europeia e africana em componentes subcontinentais.

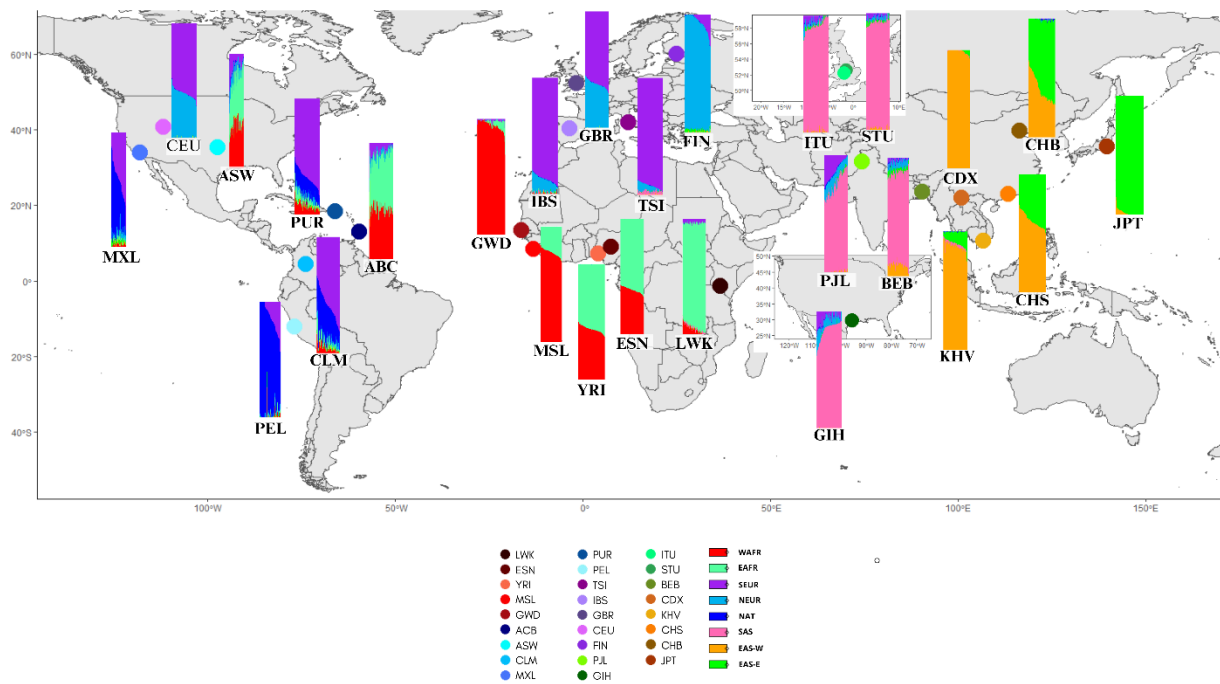


Figura 7. Mapa com as populações e barras representando a ancestralidade de todos os indivíduos de uma população. (**EAS-W**, Leste Asiático – Oeste; **SEUR**, Sul Europeu; **NEUR**, Norte Europeu; **NAT**, Nativo Americano; **EAFR**, Leste Africano; **EAS - E**, Leste Asiático- Leste; **SAS**, Sul Asiático; **WAFR**, Oeste Africano. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandinka de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio). Fonte Elaborado pelo autor, 2023

Em nosso estudo houve a separação do componente continental EUR em dois componentes subcontinentais (**NEUR** e **SEUR**), podendo ser visto nas populações europeias (**GBR**, **FIN**, **CEU**, **TSI** e **IBS**). O componente **NEUR** mais presente em populações do norte europeu (**FIN**), e o componente **SEUR** mais evidente em populações do sul europeu (**TSI** e **IBS**), enquanto nas populações **GBR** e **CEU**, é possível observar uma miscigenação com esses dois componentes (Fig. 8). Essa separação é corroborada pelas análises de Auton et al., (2015) e nas análises de Kehdy et al. (2015), há separação em dois componentes europeus, onde um componente representando as populações norte e outro do sul europeu.

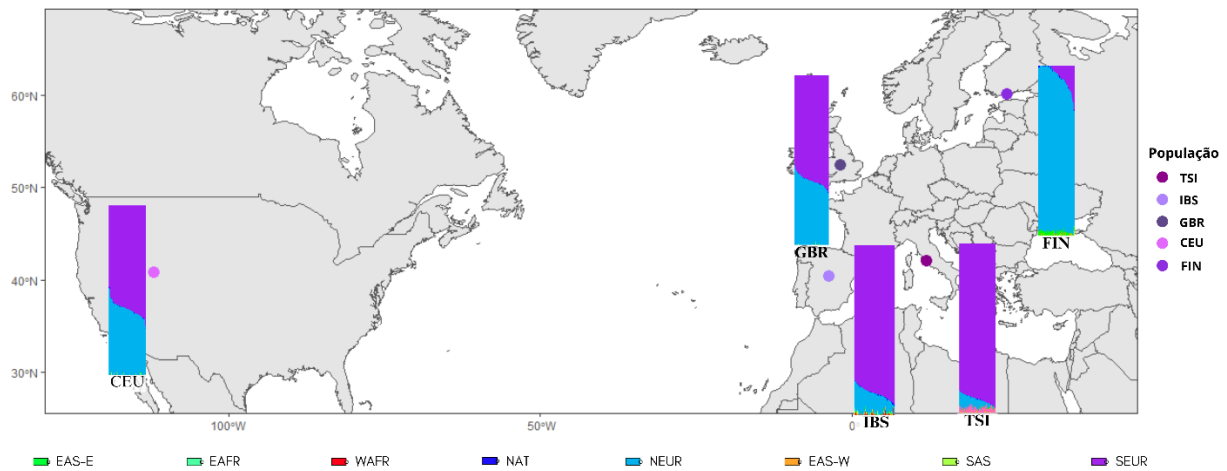


Figura 8. Mapa localizando cada população europeia e gráficos de ancestralidade individual. (**EAS-W**, Leste Asiático – Oeste; **SEUR**, Sul Europeu; **NEUR**, Norte Europeu; **NAT**, Nativo Americano; **EAFR**, Leste Africano; **EAS - E**, Leste Asiático- Leste; **SAS**, Sul Asiático; **WAFR**, Oeste Africano. **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia. Fonte: elaborado pelo autor, 2023.

Em outro estudo, Lazaridis et al., (2014), para $k=3$ até $k=20$, observaram que até $k=8$ não houve separação do componente europeu em subcontinental, já nas análises para k igual ou maior que 9 há separação em componentes subcontinentais. Nas populações norte-europeias, incluindo os finlandeses, há uma predominância do mesmo componente subcontinental, já para as populações italianas a uma miscigenação entre dois componentes e o mesmo acontece com populações espanholas e britânicas. A divisão dos componentes nas populações italianas, finlandesas e britânicas, mostram resultados parecidos com os de Kehdy et al. (2015). No entanto, quando comparamos os resultados do nosso trabalho com os de Auton et al., (2015), para $k=8$, temos uma diferença, onde a população italiana (TSI) e ibéricos (IBS), possuem uma predominância de apenas um dos componentes subcontinentais. No trabalho de Auton, quando o k é 11, há uma divisão subcontinental do componente europeu em 3, isolando um para os FIN, outro com uma predominância em GBR e CEU e um aparecendo em maior porcentagem no IBS e TSI.

As diferenças nos resultados em relação ao componente europeu podem ocorrer devido ao pouco aprofundamento nos estudos da ancestralidade genômica subcontinental nas populações europeias. Os componentes subcontinentais de populações como africana e asiática são bem documentados, mas para as populações europeias isso não acontece. Contudo, estudos mais recentes comprovam que as populações europeias possuem uma diferença notável a ponto de uma separação do componente continental em nível subcontinental (GILBERT; SHANMUGAM; CAVALLERI, 2022; GOUVEIA et al., 2022).

De acordo com Gouveia et al., (2022), o componente europeu pode ser separado em 3 componentes subcontinentais; sendo eles norte, sudeste e sudoeste europeu. Além dessa divisão do componente europeu, os resultados mostram que o componente norte europeu diminui conforme seguimos geograficamente do norte a sul da Europa e esse mesmo padrão é observado em nossas análises (Fig. 8).

O componente continental asiático foi separado em três componentes subcontinentais (EAS-W, EAS-E e SAS), podendo ser observado nas populações CHS, CHB, CDX, KHV, JPT, GIH, ITU, STU, BEB e PJI (Fig. 9 e 10). O componente leste asiático foi dividido em dois (EAS-W, EAS-E), onde o EAS-W pode ser visto com predominância nas populações CDX e KHV, e o EAS-E é observado principalmente na população JPT (Fig. 9). As populações CHS e CHB aparecem com uma miscigenação dos dois componentes do leste asiático. O componente sul-asiático (SAS) é visto nas populações GIH, ITU, STU, BEB e PJI (Fig. 10).

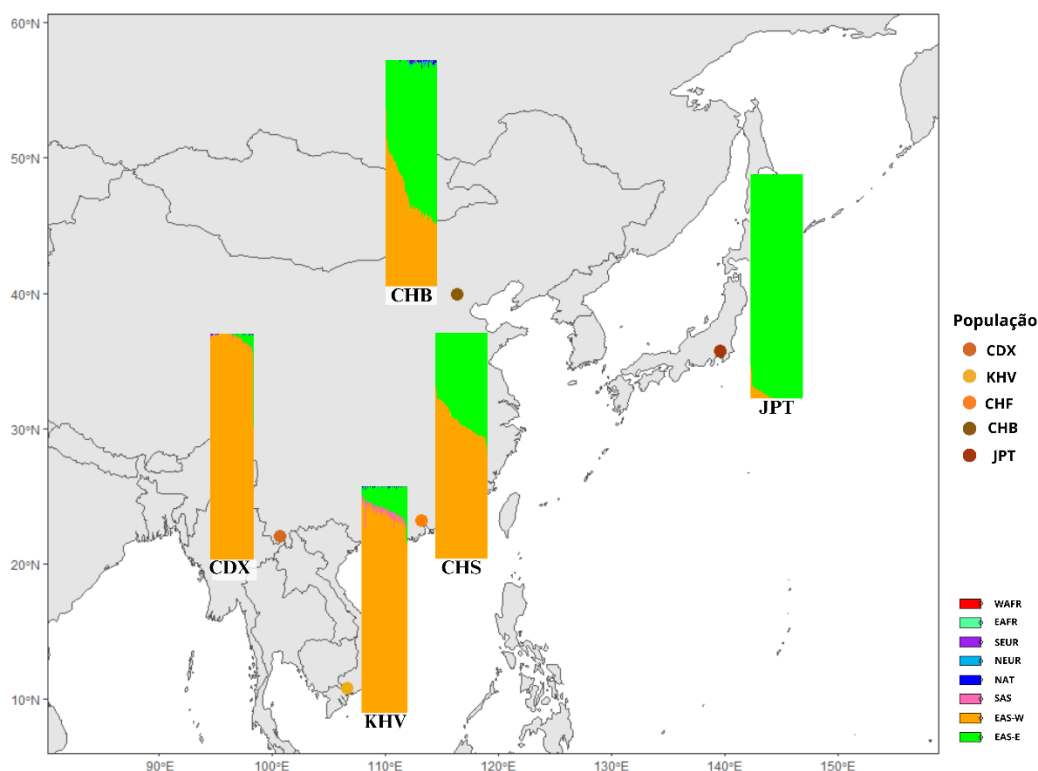


Figura 9. Mapa localizando cada população do leste asiático e gráficos de ancestralidade individual. (EAS-W, Leste Asiático – Oeste; SEUR, Sul Europeu; NEUR, Norte Europeu; NAT, Nativo Americano; EAFR, Leste Africano; EAS - E, Leste Asiático - Leste; SAS, Sul Asiático; WAFR, Oeste Africano. CDX, Chineses Dai de Xishuangbanna; KHV, Vietnamitas Kinh da cidade Ho Chi Minh; CHS, Chineses Han do Sul; CHB, Chineses Han de Beijinn; JPT, japonês de Tóquio). Fonte: Elaborado pelo autor, 2023

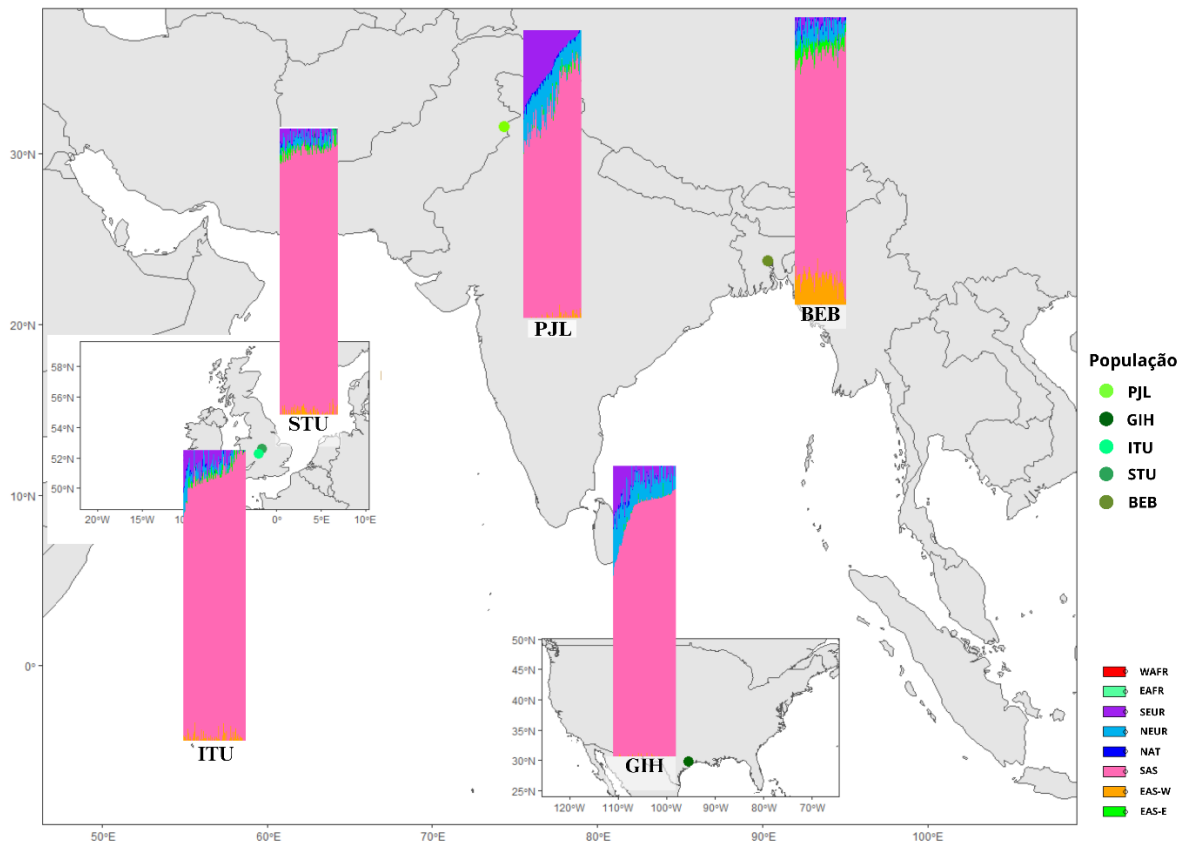


Figura 10. Mapa localizando cada população do sul asiático e gráficos de ancestralidade individual. (**EAS-W**, Leste Asiático – Oeste; **SEUR**, Sul Europeu; **NEUR**, Norte Europeu; **NAT**, Nativo Americano; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático – Leste; **SAS**, Sul Asiático; **WAFR**, Oeste Africano. **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh). Elaborado pelo autor, 2023.

Na separação dos componentes realizada no trabalho de Auton et al., (2015), com $k=8$, também houve essa mesma separação, enquanto Jun et al., (2019), em seu trabalho por meio de AIMs, utilizou 100 SNPs para diferenciar as populações chinesa, coreana, japonesa, indonésia, filipina e tailandesa entre si utilizando $k=8$. Nesse mesmo trabalho há uma miscigenação da população chinesa com as outras populações do leste asiático (coreana e japonesa) e populações do sudeste asiático (tailandesa).

Wu et al., (2019), em seu estudo usando as populações do 1000 Genomes Project e de chineses, indianos e malaios vivendo em Singapura, para $k=8$, separou a população JPT com prevalência de um componente de leste asiático, já as populações CHS e CHB tiveram uma miscigenação entre três componentes, um componente leste asiático com maior presença nos JPT, e os outros dois componentes de leste asiático com predominância em amostra de uma população chinesa vivendo em Singapura. Nas populações CDX e KHV, além de uma miscigenação com os três

componentes do leste asiático, há uma pequena porcentagem de componente presente em uma população malaia vivendo em Singapura. Em outro estudo as populações chinesas e vietnamita, apresentam um padrão de separação semelhante ao do nosso trabalho (RAMANI et al., 2020).

A população japonesa nas análises realizadas é a população mais homogênea, não mostrando uma miscigenação com outras populações do leste asiático, enquanto a população chinesa demonstra uma maior miscigenação com outras populações do leste e sudeste asiático (AUTON et al., 2015; JUNG et al., 2019; RAMANI et al., 2020; WU et al., 2019). A formação primária das populações do leste asiático acontece com junção de duas populações, os caçadores-coletores do pré-neolítico e norte asiáticos do início do neolítico, paralelamente contribuindo com essa formação rotas regionais de migração acrescentam novas variantes na formação dessas populações do leste asiático (Pan; Xu, 2020). Sendo possível presumir a separação dessas populações, baseando em análise de diferenciação genética e tamanho efetivo de população, que os chineses da etnia Han e os japoneses divergiram por volta 3600-3000 anos atrás e os coreanos e chineses da etnia Han por volta de 1200 anos atrás. Entretanto mesmo com formação de grupos homogêneos, há uma evidência de um fluxo gênico entre essas populações (Pan; Xu, 2020). O fluxo gênico e uma formação heterogênea dessas populações chinesas podem explicar a miscigenação nessas populações.

Nas populações GIH, ITU, STU, BEB e PJJ, o componente com maior porcentagem é SAS (Fig. 10), entretanto é possível observar uma pequena porcentagem de ancestralidade europeia, assim como no trabalho de Auton et al., (2015) para $k=8$. Já para Ramani et al., (2020), que em seu estudo tinha como objetivo diferenciar as populações asiáticas, também observou um componente que diferenciava a população indiana e de Bangladesh das demais do estudo. Entretanto para análises com k igual ou maior que 9 Auton et al., (2015) conseguiu uma separação de dois componentes para as populações sul asiáticas. A Índia é formada por muitas etnias e possui uma grande diversidade genética, além de possuir uma história complexa de eventos de miscigenação (Das; Upadhyai, 2018; Majumder, 2010). Em relação a porcentagem do componente europeu nas populações sul asiática, de acordo com Xing et al., (2010), a formação genética indiana foi influenciada por eurásianos do oeste, sendo um possível motivo para a presença desse componente nessas populações.

O componente continental africano foi dividido em dois, o WAFR e EAFR. As populações GWD, MSL, YRI e ESN localizadas na parte ocidental do continente africano mostraram 97%, 76%, 43% e 37% respectivamente do componente WAFR. A população GWD foi a que

apresentou uma porcentagem mais alta do componente WAFR, enquanto nas outras populações a porcentagem do componente EAFR mostrou um gradiente com MSL apresentado uma porcentagem baixa desse componente e nas duas populações nigerianas (YRI e ENS) chega a ser maior que o WAFR. A população LWK apresentou quase que exclusivamente o componente EAFR (90%) (Fig. 11).

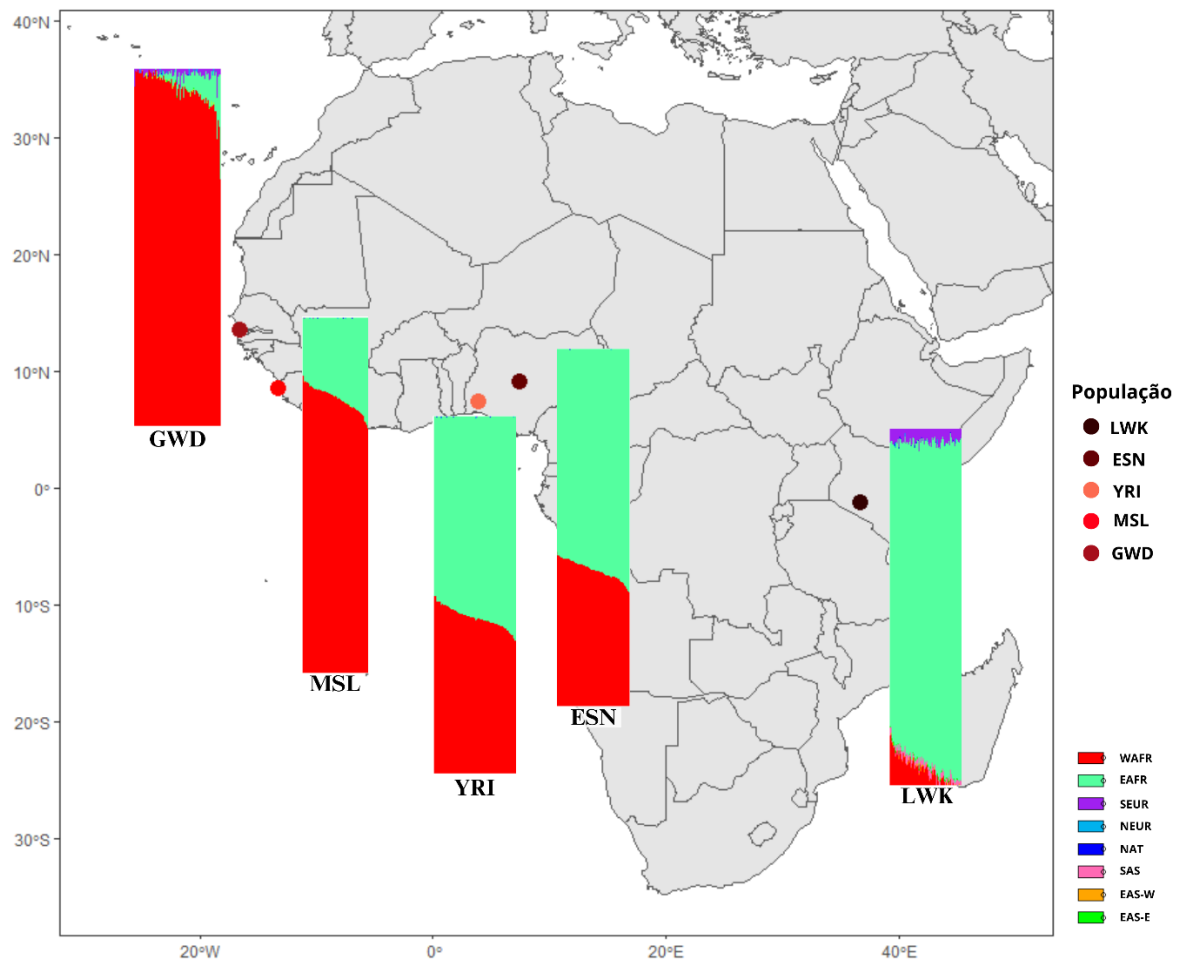


Figura 11. Mapa localizando cada população do africano e gráficos de ancestralidade individual. (EAS-W, Sul do Leste Asiático – Oeste; SEUR, Sul Europeu; NEUR, Norte Europeu; NAT, Nativo Americano; EAFR, Leste Africano; EAS-E, Leste Asiático Leste; SAS, Sul Asiático; WAFR, Oeste Africano. LWK, Luhya do Quênia; ESN, Esan da Nigéria; YRI, Yoruba da Nigéria; MSL, Mende de Serra Leoa; GWD, Mandika de Gâmbia).

Os dois componentes africanos apareceram em populações tanto no oeste africano como no leste, à medida que três populações do oeste africano (MSL, YRI e ESN) apresentam porcentagens mais elevadas dos dois componentes, principalmente as duas populações nigerianas. Em contrapartida, LWK apresenta praticamente o componente EAFR e a GWD apresenta quase exclusivamente o WAFR (Fig. 11).

Auton et al., (2015), para $k=8$, foi apresentada uma separação do componente continental africano na mesma forma que em nosso trabalho e em $k=10$ há separação do WAFR em dois componentes, diferenciando as populações GWD de YRI e ESN. No estudo de Kehdy et al., (2015), também há separação do componente africano em subcontinentais, sendo um componente com prevalência em população do oeste africano e outro em populações do leste africano. A miscigenação entre as populações africanas pode ser causada por um fluxo gênico entre elas de tempo em tempo. Algumas populações africanas mesmo após divergirem entre si, em determinados momentos ocorria o fluxo gênico entre elas contribuindo para uma miscigenação (BERGSTRÖM et al., 2020).

As populações no continente americano (ACB, AWS, PEL, MXL, CLM e PUR), possuem uma miscigenação com principalmente três populações continentais parentais; são elas africana, nativo americana e europeia. A miscigenação nessas populações iniciou recentemente, nos últimos 500 anos, com o avanço de novos meios de transporte e de maneira não uniforme, permitindo uma composição genética bem diferente em cada população (Fig. 12).

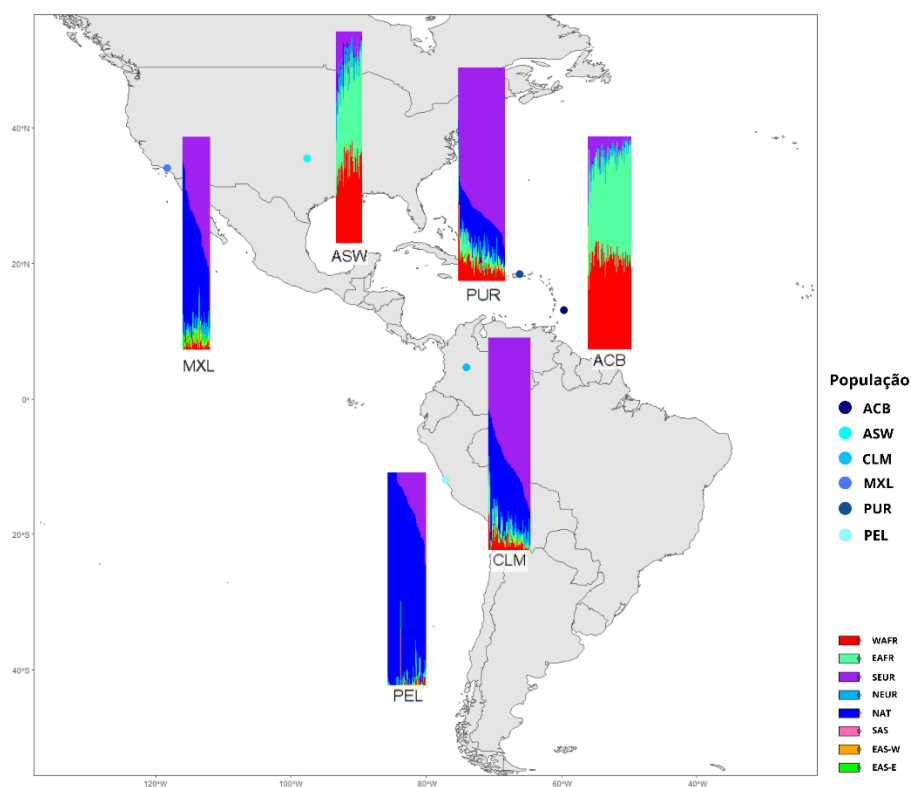


Figura 12. Mapa localizando cada população americana e gráficos de ancestralidade individual. (**EAS-W**, Sul do Leste Asiático – Oeste; **SEUR**, Sul Europeu; **NEUR**, Norte Europeu; **NAT**, Nativo Americano; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático – Leste; **SAS**, Sul Asiático; **WAFR**, Oeste Africano. **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no

sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima.

As populações ACB e AWS são populações de africanos morando em países da América. Na análise do software *admixture* as duas populações apresentaram alta porcentagem de componente africano. A população ACB possui uma ancestralidade bem miscigenada entre os componentes subcontinentes WAFR e EAFR e uma pequena porcentagem do componente SEUR, enquanto população AWS segue apresentando miscigenação entre os dois componentes subcontinentais WAFR e EAFR e uma porcentagem um pouco maior do componente SEUR quando comparado a população ACB.

A mesma separação é feita por Auton et al., (2015) em seu estudo e no trabalho de Kehdy et al., (2015), para $k=8$, na população AWS observa uma maior ancestralidade africana.

As populações CLM e PUR apresentam principalmente dois componentes predominante EUR2 e NAT e uma pequena porcentagem do componente WAFR. O componente com maior porcentagem nessas populações é o SEUR, enquanto o NAT aparece como a segundo maior porcentagem. Essa mesma composição é corroborada por Auton et al., (2015), já em Kehdy et al., (2015) a uma mistura de dois componentes subcontinentais europeu, norte e sul.

Nas populações PEL e MXL, é onde observa maior porcentagem do componente NAT. Na população PEL, o componente NAT possui a maior porcentagem dividindo apenas uma pequena porcentagem com o componente SEUR. Já a população MXL, possui uma porcentagem semelhantes dos componentes NAT e SEUR. Auton et al., (2015) e Kehdy et al., (2015) também apresentaram separação semelhante para essas populações. A maior presença do componente SEUR nas populações miscigenadas pode ser explicado pelo período das colonizações europeia, levando em conta que as populações mencionadas sofreram uma colonização maior de países ibéricos, que em nossas análises também apresentam uma predominância do componente SEUR. Já nos componentes africanos, a alta porcentagem dos dois componentes nas populações pode ser um reflexo do modo como as populações africanas foram escravizadas e retiradas de várias regiões da África e trazidas para América.

6.2 Farmacogenômica

Os alelos estrela dos genes *CYP2C19* e *CYP2C9* foram inferidos pelo software *Stargazer*. Para o *CYP2C19* foram inferidos os seguintes alelos: *2, *3, *4, *8, *9, *10, *13, *15, *16, *17, *22, *24, *30, *34 e *35. *CYP2C9* teve inferidos os alelos *2, *3, *5, *6, *7, *8, *9, *11, *12, *13, *14, *16, *29, *31, *33, *36, *44 e *45.

Segundo o Pharmvar, os alelos *CYP2C19**13 e *15 estão associados com uma atividade enzimática normal, os alelos *9, *10 e *16, com uma atividade reduzida, já os *2, *3, *4, *8, *22, *24 e *35 são considerados alelos nulos, o alelo *17 possui função aumentada e os *30 e *34 não possui sua função descrita. No gene *CYP2C9*, o Pharmvar classifica com função normal o alelo *9, sem função *3, *6, *13, *33 e *45, enquanto os alelos *2, *5, *8, *11, *12, *14, *16, *29, *31 e *44 são considerados com a função reduzida e os alelos *7 e *36 não possuem uma função descrita.

Do gene *CYP2C19*, os alelos *2 e *17 apareceram em todas as populações, o alelo *CYP2C19**2 apresentou frequência maior em populações sul asiáticas e em algumas do leste asiático. A população STU=0,41 apresentou maior frequência do alelo *CYP2C19**2, enquanto a frequência em outras populações sul asiáticas foi: PJI=0,34, GIH=0,33, ITU=0,37 e BEB=0,32 e as populações do leste asiático CHS=0,35, CHB=0,33 e JPT=0,32. Já as populações europeias, exceto os FIN, populações do oeste africano e miscigenadas apresentaram a menor frequência.

Em revisão publicada anteriormente, o alelo *CYP2C19**2 foi encontrado em maior frequência em nativos oceânicos, seguido por sul e leste asiáticos (FRICKE-GALINDO et al., 2016). Este resultado difere do nosso resultado que encontrou uma frequência maior em sul asiáticos, entretanto em nossas análises não incluímos populações nativas oceânicas.

O alelo *CYP2C19**17 possui uma frequência maior em populações africanas, exceto a LWK, e em miscigenadas, com porcentagem alta de componentes de ancestralidade africanos e europeus. A população ACB=0,27, apresenta maior frequência para esse alelo, seguida por populações africanas ESN=0,24, YRI=0,24, MSL=0,25 e GWD=0,23 e as populações europeias TSI=0,22, IBS=0,21, GBR=0,24, CEU=0,22 e FIN=0,21. Em relação a menor frequência, foi observada nas populações do leste asiático e peruana. Já era esperado uma menor frequência do alelo *CYP2C19**17 em populações com uma alta porcentagem de ancestralidade nativa América, devido a correlação negativa entre o alelo e componente NAT que Rodrigues-Soares et al. (2020), observou em trabalho.

Na revisão feita por Fricke-Galindo et al., (2016), a maior frequência do alelo *CYP2C19*17* foi encontrada em populações do sul europeu e em menor frequência em lestes asiáticos. Já em nosso trabalho a maior frequência foi encontrada em ACB, uma população miscigenada. Enquanto a menor frequência foi observada em leste asiáticos, sendo semelhante com a revisão de Fricke-Galindo et al., (2016).

O alelo *CYP2C19*3* é observado em populações asiáticas, LWK e ACB, sendo que nas últimas duas populações mencionadas sua frequência é menor ou igual a 0,01. Logo sua maior frequência é em populações do leste asiático [CDX=0,07, KHV=0,04, CHS=0,04, CHB=0,04 e JPT=0,07], enquanto as populações sul asiáticas possuem frequência menor quando comparada com as do leste asiático.

Fricke-Galindo et al., (2016), observaram em sua revisão que o alelo *CYP2C19*3* possui uma maior frequência em populações de nativos oceânicos e leste asiáticos. Em nossas análises as populações do leste asiático apresentaram uma maior frequência comparado com outras populações, entretanto não incluímos nas análises populações nativas oceânicas.

Os alelos *CYP2C19*9*, **10*, **13*, **15* e **35* são observados em praticamente todas as populações africanas e apresentam sua maior frequência nessas populações. Outra população onde é observada a ocorrência desses alelos é população PUR, grupo miscigenado com baixa porcentagem dos componentes de ancestralidade subcontinentais africanos, e os alelos *CYP2C19*15* e **35*, também são vistos nas populações PJJ e PEL respectivamente.

Já o alelo *CYP2C19*34* é encontrado apenas em populações sul asiáticas, enquanto o alelo *CYP2C19*16* é encontrado apenas na população LWK, o *CYP2C19*24* apenas na população AWS e o *CYP2C19*30* apenas ESN. E para os alelos *CYP2C19*4*, **8* e **22*, são encontrados em populações com diferentes componentes subcontinental, como por exemplo o *CYP2C19*4* é observado em populações miscigenadas (MXL e PUR), europeia (IBS) e leste asiática (CHB).

Tabela 6. Frequência dos alelos em cada população para *CYP2C19*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJJ**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio.

POP	*2	*3	*4	*8	*9	*10	*13
LWK	0,212	0,010	0	0	0,010	0	0,040
ESN	0,207	0	0	0	0,010	0	0
YRI	0,167	0	0	0	0,005	0	0,019
MSL	0,176	0	0	0	0,018	0,006	0,006
GWD	0,133	0	0	0	0,013	0,004	0,027
ACB	0,151	0,005	0	0	0,005	0	0,031
ASW	0,139	0	0	0,008	0,008	0	0,008
CLM	0,106	0	0	0	0	0	0
MXL	0,125	0	0,008	0	0	0	0
PUR	0,125	0	0,005	0	0,005	0,005	0,010
PEL	0,059	0	0	0	0	0	0
TSI	0,093	0	0	0	0	0	0
IBS	0,146	0	0,005	0	0	0	0
GBR	0,143	0	0	0	0	0	0
CEU	0,131	0	0	0,015	0	0	0
FIN	0,217	0	0	0	0	0	0
PJL	0,344	0,016	0	0	0	0	0
GIH	0,330	0,005	0	0	0	0	0
ITU	0,373	0,005	0	0,005	0	0	0
STU	0,412	0,015	0	0	0	0	0
BEB	0,326	0,023	0	0	0	0	0
CDX	0,263	0,075	0	0	0	0	0
KHV	0,283	0,040	0	0	0	0	0
CHS	0,352	0,048	0	0	0	0	0
CHB	0,335	0,044	0,005	0	0	0	0
JPT	0,322	0,072	0	0	0	0	0

Tabela 6.1. Frequência dos alelos em cada população para *CYP2C19*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio.

POP	*15	*16	*17	*22	*24	*30	*34	*35
LWK	0,005	0,005	0,172	0	0	0	0	0,025
ESN	0,025	0	0,247	0	0	0,005	0	0,015
YRI	0,019	0	0,245	0	0	0	0	0,056
MSL	0	0	0,259	0	0	0	0	0,018

GWD	0,035	0	0,239	0	0	0	0	0,031
ACB	0,010	0	0,271	0,005	0	0	0	0,021
ASW	0,016	0	0,197	0	0,008	0	0	0,008
CLM	0	0	0,128	0,005	0	0	0	0
MXL	0	0	0,117	0	0	0	0	0
PUR	0	0	0,178	0	0	0	0	0
PEL	0	0	0,041	0	0	0	0	0
TSI	0	0	0,224	0	0	0	0	0
IBS	0	0	0,217	0	0	0	0	0
GBR	0	0	0,242	0	0	0	0	0
CEU	0	0	0,222	0	0	0	0	0
FIN	0	0	0,217	0	0	0	0	0
PJL	0	0	0,146	0	0	0	0	0
GIH	0,005	0	0,136	0	0	0	0,015	0
ITU	0	0	0,142	0	0	0	0,015	0
STU	0	0	0,142	0	0	0	0,015	0
BEB	0	0	0,110	0	0	0	0,023	0
CDX	0	0	0,016	0	0	0	0	0
KHV	0	0	0,020	0	0	0	0	0
CHS	0	0	0,010	0	0	0	0	0
CHB	0	0	0,024	0	0	0	0	0
JPT	0	0	0,005	0	0	0	0	0

No gene *CYP2C9*, nenhum alelo foi encontrado em todas as populações, entretanto os alelos *CYP2C9**2 e *3 foram observados na maioria das populações. O alelo *CYP2C9**2 foi observado em maior frequência em populações europeias; TSI=0,15, IBS=0,14, GBR=0,08, CEU=0,15 e FIN=0,08 e nas populações americanas CLM=0,12, MXL=0,10 e PUR=0,13 com uma porcentagem maior de componentes europeus. As populações sul asiáticas apresentaram uma frequência menor quando comparadas com a europeia, e as populações africanas e do leste asiático não apresentaram esse alelo, exceto GWD=0,004 e CHS=0,004 respectivamente.

Em concordância com nosso estudo, na revisão de Céspedes-Garro et al. (2015), foi encontrada a maior frequência do alelo *CYP2C9**2 em populações europeias, onde a maior frequência eram europeus americanos. Além desse estudo, Zhou et al. (2023) também observaram uma maior frequência do alelo em populações europeias e do oriente médio. Céspedes-Garro et al. (2015) observaram uma frequência menor de *CYP2C9**2 nas populações africanas e do leste asiático, enquanto Zhou et al. (2023) encontraram uma frequência menor nas populações sul asiáticas e ausência nas populações do leste asiático, corroborando assim, o que foi encontrado em nosso trabalho.

Outro estudo realizado por Gaikwad et al. (2014), que revisa a literatura para identificar a distribuição dos alelos do gene *CYP2C9* na população asiática, também não observa a presença do alelo *CYP2C9*2* em populações do leste asiático, observou a presença em apenas populações sul asiática, com uma frequência menor, e em populações da Asia Central, com uma frequência maior quando comparado entre as populações asiáticas.

O alelo *CYP2C9*3* é observado em todas as populações europeias, americanas, sul asiática e leste asiática, mas sua maior frequência é nas populações GHI=0,13, ITU=0,10 e BEB=0,11, que são sul asiáticas. As populações com a menor frequência são os lestes asiáticos e nos africanos não encontramos esse alelo. Em corroboração ao nosso estudo, o alelo *CYP2C9*3* também foi encontrado em maior frequência em populações sul asiáticas (CÉSPEDES-GARRO et al., 2015). Na revisão da literatura de Gaikwad et al. (2014), o alelo *CYP2C9*3* também foi encontrado em menor frequência nas populações do leste asiático e em uma frequência maior nas populações sul asiática. E em um outro de Zhou et al. (2023), também foi encontrado a maior frequência do *CYP2C9*3* em populações sul asiática e sua menor frequência nas populações do leste asiático.

Em todas as populações africanas foram observados os alelos *CYP2C9*5*, *8, *9, e *11, onde esses mesmos alelos são mais frequentes nas populações africanas, enquanto o alelo *CYP2C9*6* encontramos apenas nas populações ESN, YRI e GWD. Já o alelo *CYP2C9*31* foi observado apenas na população LWK=0,01 e o alelo *CYP2C9*36* estava presente apenas nos YRI=0,004. Concordando com nosso estudo, a revisão de Céspedes-Garro et al., (2015) também encontrou os alelos *CYP2C9*5*, *8 e *11 em maior frequência em africanos.

Em relação às populações europeias, encontramos também em algumas populações os alelos *CYP2C9*8* (IBS e CEU), *CYP2C9*9* (CEU), *CYP2C9*11* (FIN), *CYP2C9*12* (TSI e CEU) e *CYP2C9*29* (FIN), entretanto os alelos *CYP2C9*8*, *9 e *11 estavam menor frequência quando comparado com populações africanas. E o alelo *CYP2C9*45* observamos apenas na população FIN.

As populações miscigenadas também apresentaram os alelos *CYP2C9*5*, *6, *8, *9, *11, *12 e *44, nas populações com maior porcentagem de componentes subcontinentais africanos a frequência dos alelos são semelhantes a populações africanas, como por exemplo os alelos *CYP2C9*5*, *8, *9 e *11.

Nas populações sul asiáticas, o alelo *CYP2C9*14* apresentou uma maior frequência, enquanto o alelo *CYP2C9*7* foi observado apenas nessas populações. E nas populações do leste asiáticos

observamos também os alelos *CYP2C9*13*, **14*, **16*, **29*, **33* e **44*, enquanto o alelo *CYP2C9*13* foi encontrado apenas nessas populações, os *CYP2C9*16* e **33* apenas na população CHB. Em estudo anterior, o alelo *CYP2C9*13* também foi encontrado apenas populações do leste asiático (CÉSPEDES-GARRO et al., 2015).

Tabela 7. Frequência dos alelos em cada população do gene *CYP2C9*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio.

POP	*2	*3	*5	*6	*7	*8	*9	*11	*12
LWK	0	0	0,010	0	0	0,071	0,146	0,015	0
ESN	0	0	0,025	0,020	0	0,076	0,111	0,010	0
YRI	0	0	0,023	0,019	0	0,051	0,093	0,051	0
MSL	0	0	0,006	0	0	0,035	0,035	0,024	0
GWD	0,004	0	0,009	0,009	0	0,018	0,031	0,035	0
ACB	0,026	0,005	0,016	0,005	0	0,083	0,063	0,016	0
ASW	0,041	0,016	0,025	0	0	0,033	0,082	0,008	0
CLM	0,122	0,064	0	0	0	0	0	0	0
MXL	0,102	0,023	0	0	0	0	0	0	0
PUR	0,139	0,043	0,005	0	0	0,005	0	0,005	0,005
PEL	0,024	0,012	0	0	0	0	0	0	0
TSI	0,154	0,084	0	0	0	0	0	0	0,009
IBS	0,142	0,085	0	0	0	0,005	0	0	0
GBR	0,088	0,071	0	0	0	0	0	0	0
CEU	0,152	0,066	0	0	0	0,005	0,005	0	0,005
FIN	0,081	0,056	0	0	0	0	0	0,010	0
PJL	0,052	0,099	0	0	0	0,005	0	0	0
GIH	0,049	0,131	0	0	0,015	0	0	0	0
ITU	0,025	0,103	0	0	0,010	0	0	0,005	0
STU	0,029	0,098	0	0	0,005	0	0	0	0
BEB	0,017	0,116	0	0	0,006	0	0	0	0
CDX	0	0,027	0	0	0	0	0	0	0
KHV	0	0,035	0	0	0	0	0	0	0
CHS	0,005	0,043	0	0	0	0	0	0	0
CHB	0	0,039	0	0	0	0	0	0	0
JPT	0	0,019	0	0	0	0	0	0	0

As populações do 1000 Genomes Project para o gene *CYP2C19* apresentaram 16 alelos, enquanto para o gene *CYP2C9* foram inferidos 19 alelos. A quantidade de alelos identificados nos dois genes mostra uma variabilidade genética nas populações. A maioria dos alelos estão presentes apenas em algumas populações, como os alelos *16, (LWK), *24(AWS) e *30 (ESN) do gene *CYP2C19* e os *31 (LWK), *33 (CHB), *36 (YRI), e *45 (FIN) do *CYP2C9*. Outros alelos são compartilhados entre populações de regiões próximas. Entretanto, a maioria desses alelos são raros.

Populações que compartilham componentes continentais também compartilham alguns alelos entre si. As populações africanas (GWD, MSL, YRI, ESN e LWK) apresentam alelos em comum. No gene *CYP2C9* a frequência dos alelos *5, *6, *8, *9 e *11 é maior que em outras populações, assim como os alelos *9 e *35 do *CYP2C19*. Nas populações asiáticas (JPT, CHS, CHB, CDX, KHV, BEB, STU, P JL, ITU e GIH), os alelos como *7, *13 e *14, do *CYP2C9* e o alelo *34 do *CYP2C19* são compartilhados entre elas. Os alelos *7 e *14 do *CYP2C9* e *34 do *CYP2C19* são compartilhados apenas entre populações sul asiáticas (BEB, STU, P JL, ITU e GIH).

Além da variabilidade interétnica dos alelos, há uma variabilidade interétnica entre os scores de atividade e os fenótipos preditos. A frequência do score de atividade 0 para o gene *CYP2C19* foi maior nas populações STU=0,19, P JL=0,17 e JPT=0,16 e não possuindo uma frequência nas populações IBS e CLM. Os TSI=0,009 foram os que apresentaram a menor frequência.

Indivíduos com score de atividade 0,5 foram encontrados apenas nas populações LWK, MSL e YRI, já a frequência do score de atividade 1 foi maior nas populações CHS=0,57, CHB=0,51, JPT=0,45 e CDX=0,45 e baixa nas populações PEL=0,09, TSI=0,14 e GWD=0,15. Para o score de atividade 1,5 as populações que apresentaram uma maior frequência foram as ESN=0,16, FIN=0,14 e MSL=0,11, enquanto as menores frequência ficaram com as populações do leste asiático.

O score de atividade 2 apresentou uma frequência maior nas populações PEL=0,81, CLM=0,58 e MXL0,57, seguidas por populações como PUR=0,44 e TSI=0,44 e uma menor frequência nas populações STU=0,17, ITU=0,19= e P JL=0,25.

E em relação ao score de atividade 2,5 apresentou maior frequência nas populações CEU=0,33 e TSI=0,31, e menor frequência nas populações do leste asiático. Já o score de atividade de

valor 3 possuem uma frequência maior nos YRI=0,07, ESN=0,06 e MSL=0,05 e as populações do leste asiático não apresentam indivíduos com esse score de atividade.

No gene *CYP2C19* a frequência do fenótipo predito PMs foi maior nas populações STU=0,19, PJI=0,17 e JPT=0,16 respectivamente, e a população TSI=0,009 apresentou a menor frequência, enquanto as populações IBS e CLM não possuem indivíduos metabolizadores lentos (Tabela 8). Já no trabalho de Fricke-Galindo et al (2016), a maior frequência de PMs foi encontrada em populações oceânicas, seguido por populações do leste asiático e menor frequência foi encontrada em populações do norte da África, diferente do nosso trabalho, onde encontramos maior frequência em populações do sul asiático e menor frequência TSI, contudo, não estudamos populações da Oceania e do Norte da África.

Tabela 8. Frequência dos fenótipos preditos em cada população do gene *CYP2C19*. **POP**, População; **NMs**, Metabolizadores Normais; **PMs**, Metabolizadores Lentos; **IMs**, Metabolizadores Intermediários; **RMs**, Metabolizadores Rápidos; **UMs**, Metabolizadores Ultrarrápidos; **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJI**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio.

POP	NMs	PMs	IMs	RMs	UMs
LWK	0,424	0,061	0,293	0,192	0,030
ESN	0,455	0,061	0,182	0,232	0,061
YRI	0,324	0,056	0,269	0,278	0,074
MSL	0,400	0,047	0,200	0,294	0,059
GWD	0,460	0,044	0,159	0,283	0,053
ACB	0,406	0,021	0,208	0,313	0,052
ASW	0,475	0,016	0,213	0,279	0,016
CLM	0,638	0,000	0,170	0,181	0,011
MXL	0,609	0,016	0,203	0,141	0,031
PUR	0,490	0,019	0,192	0,269	0,029
PEL	0,824	0,012	0,094	0,071	0,000
TSI	0,467	0,009	0,150	0,318	0,056
IBS	0,415	0,000	0,245	0,302	0,038
GBR	0,429	0,011	0,198	0,308	0,055
CEU	0,414	0,020	0,202	0,333	0,030
FIN	0,475	0,020	0,253	0,212	0,040
PJI	0,302	0,177	0,313	0,177	0,031

GIH	0,320	0,155	0,301	0,175	0,019
ITU	0,275	0,147	0,382	0,137	0,029
STU	0,255	0,196	0,363	0,118	0,039
BEB	0,349	0,140	0,326	0,140	0,000
CDX	0,441	0,108	0,441	0,011	0,000
KHV	0,424	0,091	0,455	0,030	0,000
CHS	0,295	0,114	0,571	0,019	0,000
CHB	0,369	0,107	0,515	0,010	0,000
JPT	0,385	0,163	0,452	0,000	0,000

As populações do leste asiático foram as que apresentaram a maior frequência de IMs para o gene *CYP2C19*, com frequência acima de 40%, seguidos pelas populações do sul asiáticos, com uma frequência entre 38% e 30%. As populações PEL=0,09 e GWD=0,15 apresentaram uma frequência menor de metabolizadores intermediários.

Em relação aos RMs, as populações que apresentaram maior frequência são CEU=0,33, TSI=0,31 e ACB=0,31, enquanto a menor frequência é encontrada CHB=0,009 e os JPT que não apresentaram indivíduos com esse fenótipo. Já os UMs possuem uma frequência maior nos YRI=0,07, ESN=0,06, MSL=0,058 e TSI=0,056 e as populações do leste asiático e PEL não apresentaram nenhum indivíduo com esse fenótipo. No estudo realizado por Fricke-Galindo (2016), mostra que a frequência de UMs é maior em populações Oriente Médio, Europeus americanos e Americanos miscigenados, diferenciando da nossa análise, que encontramos uma frequência maior dos UMs em populações africanas. Em nosso trabalho consideramos que genótipo dos UMs é o homozigoto para o alelo *CYP2C19*17*, e os indivíduos com genótipo contendo um alelo de função enzimática normal e outro de função aumentada (*CYP2C19*1/*17*) são considerados RMs. Fricke-Galindo (2016), considerou esses dois grupos de metabolizadores como apenas um grupo de UMs, causando assim a diferença entre as frequências encontradas em nossos trabalhos.

Já o score de atividade do gene *CYP2C9*, temos para o valor 0,5 apenas duas populações: ESN=0,01 e YRI=0,009. O score de atividade 1 apresenta uma maior variabilidade, sendo as populações CEU=0,007, IBS=0,06 e YRI=0,04 com maior frequência desse score de atividade, e as populações do leste asiático não apresentaram nenhum indivíduo com esse score de atividade, exceto os CHS=0,009.

Em relação ao score de atividade de valor 1,5 temos as populações TSI=0,42, IBS=0,33 e PUR=0,33 apresentando uma frequência maior e as populações do leste asiático

apresentando a menor frequência. E o score de atividade com valor 2, apresentou as populações JPT=0,95, CDX=0,92 e PEL=0,92 com a maior frequência e as populações TSI=0,54, IBS=0,60 e CEU=0,61 com a menor frequência.

Os fenótipos preditos encontrados para o gene *CYP2C9* foram IMs e NMs. As populações JPT, PEL, MXL e MSL apresentaram uma frequência de 100% para NMs, enquanto a população IBS=0,92 apresentou a menor porcentagem para esse fenótipo. Os IMs apresentaram uma frequência maior nas populações CEU=0,07, IBS=0,06 e ESN=0,05 e populações do leste asiáticos não apresentaram indivíduos com esse fenótipo ou a frequência é muito baixa como por exemplo CHS=0,009 (Tabela 9).

Tabela 9. Frequência dos fenótipos preditos em cada população do gene *CYP2C19*. **POP**, População; **NMs**, Metabolizadores Normais; **IMs**, Metabolizadores Intermediários; **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio

POP	IMs	NMs
LWK	0,010	0,990
ESN	0,051	0,949
YRI	0,056	0,944
MSL	0,000	1,000
GWD	0,027	0,973
ACB	0,031	0,969
ASW	0,016	0,984
CLM	0,021	0,979
MXL	0,000	1,000
PUR	0,029	0,962
PEL	0,000	1,000
TSI	0,037	0,963
IBS	0,065	0,935

GBR	0,033	0,967
CEU	0,071	0,929
FIN	0,010	0,970
PJL	0,031	0,948
GIH	0,029	0,922
ITU	0,020	0,941
STU	0,020	0,931
BEB	0,000	0,942
CDX	0,000	0,989
KHV	0,000	0,990
CHS	0,010	0,971
CHB	0,000	0,971
JPT	0,000	1,000

A prevalência de PMs no gene *CYP2C9* mundialmente é relativamente baixa, sendo no sudeste europeu e leste do mediterrâneo em torno de 3-4%, caindo para menos de 1% em populações asiáticas e africanas (ZHOU et al., 2023). Em um outro estudo conduzido por Céspedes-Garro et al. (2016), também foi observado que frequência PMs é maior em europeus e em algumas populações do sul asiático e não encontrou PMs em populações africanas e nem americanas. O que pode explicar a ausência de PMs nos nossos resultados.

No estudo de Zhou e seus colaboradores (2023), os IMs segue um padrão similar com os PMs, com algumas exceções como os Emirados Árabes, Croácia e Irã, onde a frequência de IMs é mais alta. Na África a frequência também é muito baixa, em alguns países sendo quase inexistente a presença de indivíduos com metabolismo reduzidos, como em Gana e na Nigéria por exemplo.

6.3 Correlação entre ancestralidade e frequência de alelos farmacogenômicos

Foi realizada uma correlação entre a frequência de todos os alelos estrela inferidos para os dois genes e os oito componentes parentais, utilizando a plataforma R, com scripts desenvolvidos durante este trabalho (https://github.com/ManochioCaique/Genomic-ancestry-and-Star-allele/blob/main/Correlacao_alelica_CYP2C19.R). Os valores do coeficiente de correlação e valores de p podem ser vistos na Tabela 10 para o *CYP2C19* e Tabela 11 para o *CYP2C9*.

Tabela 10. Valores de coeficiente de correlação e valores de p da correlação entre alelos e ancestralidade do gene *CYP2C19*. (**EAS-W**, Leste Asiático – Oeste; **NEUR**, Norte Europeu; **WAFR**, Oeste Africano; **NAT**, Nativo Americano; **SEUR**, Sul Europeu; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático – Leste; **SAS**, Sul Asiático). Negrito: valores significativos.

Alelos	EAS-W	NEUR	WAFR	NAT	SEUR	EAFR	EAS-E	SAS
*2								
Valor p	0,081	0,556	0,108	0,018	0,002	0,317	0,040	<0,001
Valor r	0,349	-	-0,323	0,460	-0,575	-0,204	0,405	0,677
*3								
Valor p	<0,001	0,205	0,120	0,264	0,038	0,228	<0,001	0,878
Valor r	0,749	-	-0,313	0,227	-0,410	-0,245	0,709	-0,032
*4								
Valor p	0,868	0,627	0,384	0,125	0,067	0,356	0,641	0,326
Valor r	-0,034	0,100	-0,178	0,309	0,365	-0,189	0,096	-0,200
*8								
Valor p	0,504	0,233	0,856	0,616	0,384	0,928	0,556	0,975
Valor r	-0,137	0,242	-0,037	0,103	0,178	-0,019	-0,121	-0,006
*9								
Valor p	0,216	0,239	< 0,001	0,389	0,177	0,001	0,277	0,156
Valor r	-0,251	0,239	0,872	0,176	-0,273	0,617	-0,221	-0,286
*10								
Valor p	0,452	0,446	<0,001	0,837	0,945	0,939	0,512	0,386
Valor r	-0,154	0,156	0,658	0,042	0,014	-0,016	-0,135	-0,177
*13								
Valor p	0,297	0,336	0,006	0,456	0,299	<0,001	0,359	0,239
Valor r	-0,213	0,197	0,521	0,153	-0,212	0,722	-0,188	-0,239
*15								
Valor p	0,298	0,357	<0,001	0,402	0,146	0,010	0,359	0,351
Valor r	-0,212	0,188	0,770	0,171	-0,293	0,498	-0,187	-0,191
*16								
Valor p	0,677	0,646	0,726	0,716	0,570	<0,001	0,715	0,655
Valor r	-0,086	0,095	-0,072	0,075	-0,117	0,644	-0,075	-0,092
*17								

Valor p	0,000	0,135	0,004	0,184	0,114	0,033	0,001	0,637
Valor r	-0,656	0,301	0,544	-	0,317	0,420	-0,594	-0,097
*22								
Valor p	0,553	0,699	0,598	0,575	0,553	0,497	0,606	0,495
Valor r	-0,122	-	0,108	0,115	0,122	0,139	-0,106	-0,140
*24								
Valor p	0,682	0,926	0,370	0,855	0,771	0,281	0,722	0,637
Valor r	-0,084	-	0,183	-	-0,060	0,220	-0,073	-0,097
*30								
Valor p	0,672	0,642	0,336	0,714	0,493	0,040	0,708	0,625
Valor r	-0,087	-	0,196	-	-0,141	0,406	-0,077	-0,101
*34								
Valor p	0,559	0,676	0,275	0,471	0,235	0,281	0,542	<0,001
Valor r	-0,120	-	-0,222	-	-0,241	-0,220	-0,125	0,862
*34								
Valor p	0,267	0,268	<0,001	0,570	0,107	<0,001	0,327	0,203
Valor r	-0,226	-	0,709	-	-0,323	0,709	-0,200	-0,258

Tabela 11. Valores de coeficiente de correlação e valores de p da correlação entre alelos e ancestralidade do gene *CYP2C9*. (**EAS-W**, Leste Asiático – Oeste; **NEUR**, Norte Europeu; **WAFR**, Oeste Africano; **NAT**, Nativo Americano; **SEUR**, Sul Europeu; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático – Leste; **SAS**, Sul Asiático). Negrito: valores significativos.

Alelos	EAS-W	NEUR	WAFR	NAT	SEUR	EAFR	EAS-E	SAS
*2								
Valor p	0,067	0,053	0,090	0,443	< 0,001	0,082	0,124	0,555
Valor r	-0,365	0,384	-0,339	0,157	0,933	-0,347	-0,310	-0,121
*3								
Valor p	0,609	0,286	0,002	0,345	0,132	0,002	0,490	< 0,001
Valor r	-0,105	0,218	-0,581	-0,193	0,303	-0,574	-0,142	0,757
*5								
Valor p	0,237	0,318	0,001	0,407	0,183	< 0,001	0,299	0,174
Valor r	-0,240	-0,204	0,618	-0,170	-0,269	0,811	-0,212	-0,275
*6								

Valor p	0,423	0,400	0,004	0,487	0,210	0,003	0,478	0,355
Valor r	-0,164	-0,172	0,544	-0,143	-0,254	0,561	-0,146	-0,189
<i>*7</i>								
Valor p	0,509	0,749	0,311	0,501	0,296	0,317	0,532	< 0,001
Valor r	-0,136	-0,066	-0,207	-0,138	-0,213	-0,204	-0,128	0,811
<i>*8</i>								
Valor p	0,222	0,314	0,003	0,325	0,155	< 0,001	0,282	0,192
Valor r	-0,248	-0,206	0,553	-0,201	-0,287	0,927	-0,219	-0,264
<i>*9</i>								
Valor p	0,257	0,319	0,015	0,361	0,131	< 0,001	0,318	0,198
Valor r	-0,231	-0,203	0,471	-0,187	-0,304	0,985	-0,204	-0,261
<i>*11</i>								
Valor p	0,238	0,658	< 0,001	0,347	0,116	0,003	0,304	0,250
Valor r	-0,240	-0,091	0,787	-0,192	-0,316	0,557	-0,210	-0,234
<i>*12</i>								
Valor p	0,478	0,734	0,446	0,778	< 0,001	0,432	0,536	0,451
Valor r	-0,146	0,070	-0,156	-0,058	0,659	-0,161	-0,127	-0,154
<i>*13</i>								
Valor p	< 0,001	0,401	0,344	0,514	0,212	0,349	0,027	0,400
Valor r	0,695	-0,172	-0,193	-0,134	-0,253	-0,191	0,432	-0,172
<i>*14</i>								
Valor p	0,895	0,628	0,204	0,404	0,186	0,209	0,480	< 0,001
Valor r	-0,027	-0,100	-0,258	-0,171	-0,268	-0,255	-0,145	0,896
<i>*16</i>								
Valor p	0,274	0,652	0,602	0,725	0,496	0,606	0,024	0,626
Valor r	0,223	-0,093	-0,107	-0,073	-0,140	-0,106	0,440	-0,100
<i>*29</i>								
Valor p	0,003	0,114	0,261	0,444	0,174	0,267	0,079	0,312
Valor r	0,559	0,318	-0,229	-0,157	-0,275	-0,226	0,351	-0,206
<i>*31</i>								
Valor p	0,677	0,646	0,726	0,716	0,570	< 0,001	0,715	0,655
Valor r	-0,086	-0,095	-0,072	-0,075	-0,117	0,644	-0,075	-0,092
<i>*33</i>								
Valor p	0,274	0,652	0,602	0,725	0,496	0,606	0,024	0,626
Valor r	0,223	-0,093	-0,107	-0,073	-0,140	-0,106	0,440	-0,100
<i>*36</i>								
Valor p	0,672	0,642	0,243	0,714	0,493	0,069	0,708	0,625
Valor r	-0,087	-0,096	0,237	-0,076	-0,141	0,362	-0,077	-0,101

*44								
Valor p	0,335	0,561	0,595	0,964	0,529	0,561	0,464	0,487
Valor r	0,197	-0,120	-0,109	0,009	0,129	-0,119	0,150	-0,143
*45								
Valor p	0,680	0,001	0,602	0,722	0,636	0,606	0,768	0,627
Valor r	-0,085	0,856	-0,107	-0,073	-0,097	-0,106	-0,061	-0,100

Em relação ao gene *CYP2C19*, foram encontradas 5 correlações negativas e 18 correlações positivas. O componente SEUR obteve correlação negativa com alelos *CYP2C19*2* e *CYP2C19*3*, o NAT apresentou correlação negativa com o alelo *CYP2C19*2*. O componente subcontinental WAFR, obteve correlações positivas com os alelos *CYP2C19*9*, **10*, **13*, **15*, **17* e **35* e nenhuma correlação negativa, o mesmo acontece com o componente EAFR, onde os alelos *CYP2C19*9*, **13*, **15*, **16*, **17*, **30* e **35* apresentam correlação positiva e nenhuma negativa.

Em relação aos componentes subcontinentais asiáticos, para o EAS-W obtivemos correlação positiva com o alelo *CYP2C19*3*, e negativa com alelo *CYP2C19*17*. Já o componente EAS-E apresentou correlações positivas com os alelos *CYP2C19*2* e **3* e negativa com alelo *CYP2C19*17*. E o componente SAS obteve correlações positivas com os alelos *CYP2C19*2* e **34*.

Para o gene *CYP2C9*, encontramos 22 correlações positivas e 2 negativas. No componente NEUR, foi observada correlação positiva com o alelo *CYP2C9*45* e para o componente SEUR correlações positivas foram encontradas para os alelos *CYP2C9*2* e **12*. Em relação aos componentes subcontinentais africanos, para o WAFR encontramos correlações positivas com os alelos *CYP2C9*5*, **6*, **8*, **9* e **11* e negativa com o alelo *CYP2C9*3* e componente EAFR observamos correlações positivas com os alelos *CYP2C9*5*, **6*, **8*, **9*, **11* e **31* e negativa com alelo *CYP2C9*3*.

Para os componentes subcontinentais asiáticos, observamos para EAS-W correlações positivas com os alelos *CYP2C9*13* e **29*, para o componente EAS-E correlações positivas com os alelos *CYP2C9*13*, **16* e **33* e o componente SAS encontramos correlações positivas com os alelos *CYP2C9*3*, **7* e **14*.

Dentre os alelos inferidos, alguns possuem uma maior relevância clínica. Esses alelos se destacam entre os outros por terem anotações com um nível de confiabilidade alto, além de possuírem uma frequência maior nas populações em relação aos outros alelos. Os **2*, **3* e **17*

para o *CYP2C19* e o *2 e *3 do *CYP2C9*, são alelos com uma alta relevância clínica, além de estarem presentes em *guidelines* do CPIC (www.cpicpgx.org).

Os alelos de relevância clínica correlacionados com componentes parentais foram usados para a criação de gráficos de frequência do alelo em relação ao componente parental (Fig. 13 a 17). O *CYP2C19**2, alelo sem função enzimática, foi correlacionado positivamente com EAS-E e SAS: quanto maior a ancestralidade leste e sul asiática, maior a probabilidade de o indivíduo possuir o *2, e negativamente com o SEUR e NAT, ou seja, quando maior a ancestralidade sul europeia ou nativa americana menor a probabilidade de o indivíduo possuir o *2 (Fig. 13). O alelo *CYP2C19**3 também é sem função enzimática foi correlacionado positivamente com os componentes EAS-W e EAS-E e negativamente com SEUR (Fig. 14). Já o *CYP2C19**17 com função enzimática aumentada foi correlacionado positivamente AFR e AFR2 e negativamente com os componentes EAS e EAS2 (Fig. 15).

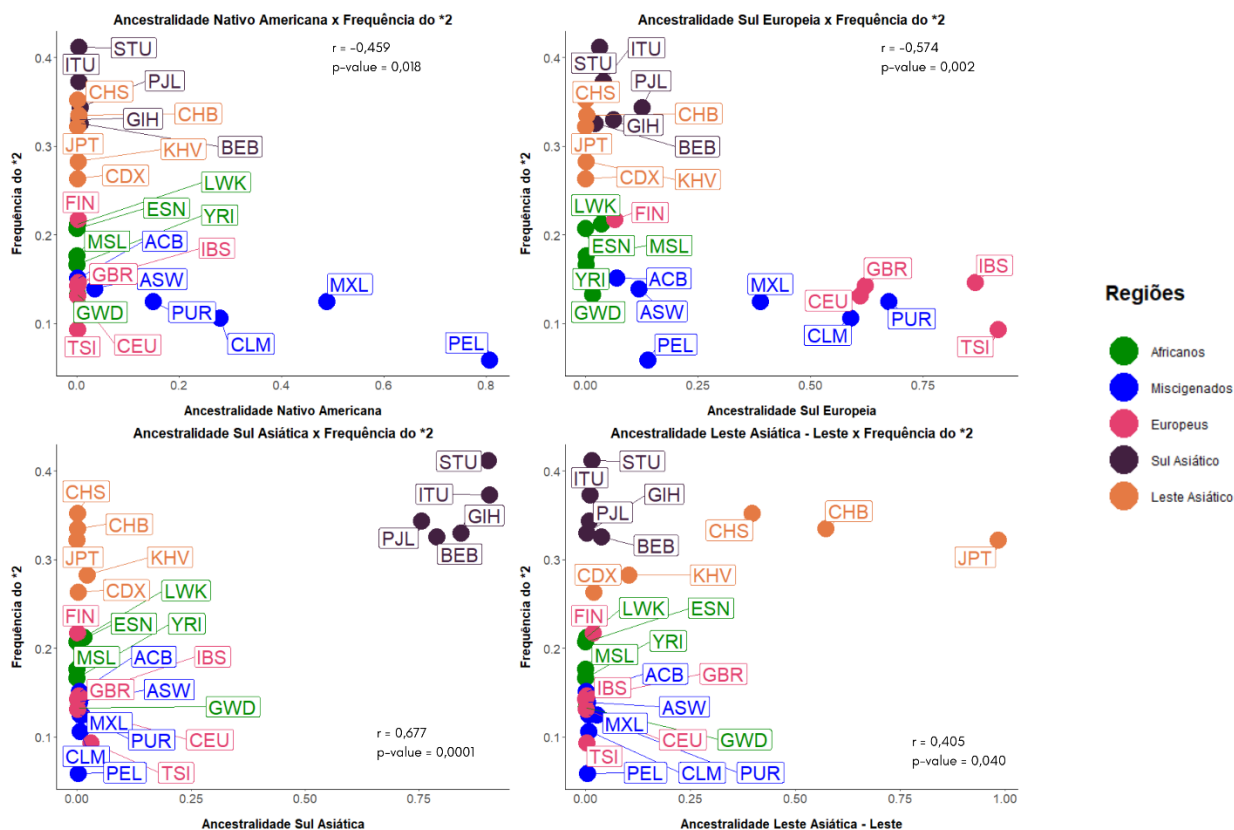


Figura 13. Gráficos Ancestralidade pela frequência *CYP2C19**2. LWK, Luhya do Quênia; ESN, Esan da Nigéria; YRI, Yoruba da Nigéria; MSL, Mende de Serra Leoa; GWD, Mandika de Gâmbia; ACB, Afro-Caribenhos de Barbados; ASW, Indivíduos com ancestralidade africana no sudeste dos EUA;

CLM, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio).

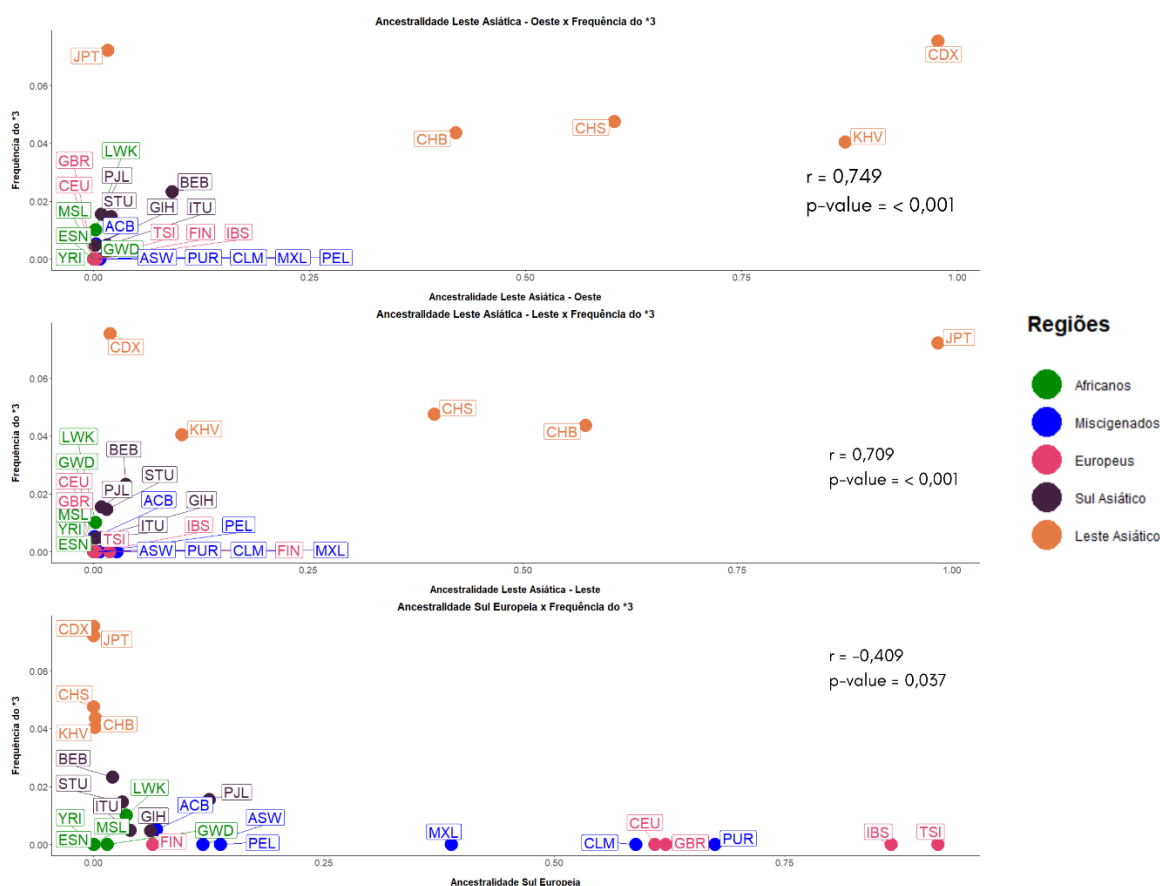


Figura 14. Gráficos Ancestralidade pela frequência *CYP2C19*3*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio).

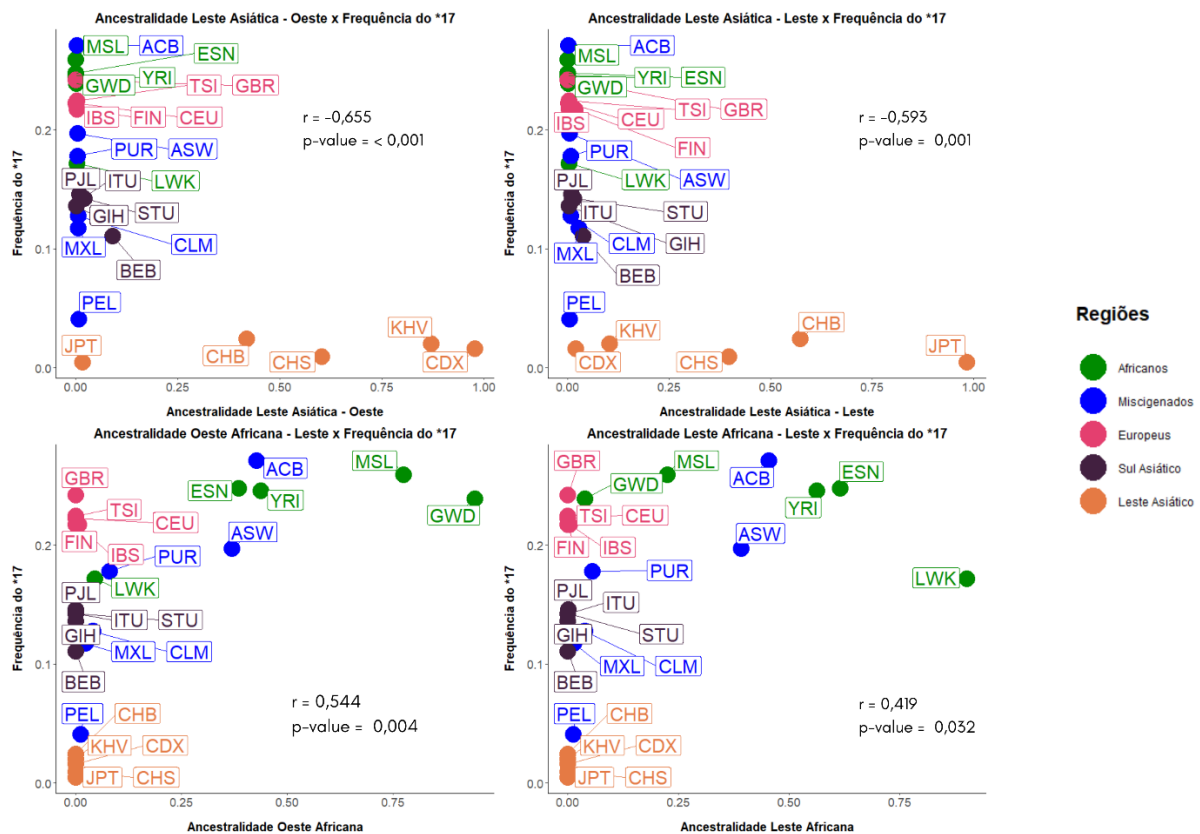


Figura 15. Gráficos Ancestralidade pela frequência *CYP2C19*17*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio).

Em um estudo realizado por Rodrigues-Soares et al., (2020), associando alelos dos genes *CYP2C19* e *CYP2C9* com ancestralidade continental constatou que o alelo *CYP2C19*2* é negativamente associado com a ancestralidade nativo americana, e positivamente associado com ancestralidade africana. Já o alelo *CYP2C19*17* também foi associado negativamente com ancestralidade nativo americana. Observamos uma concordância em nossos estudos em relação ao componente nativo americano e o alelo *CYP2C19*2* que estão correlacionados negativamente, entretanto não encontramos associação positiva do alelo *CYP2C19*2* com componente africano, além de não observamos uma correlação negativa entre o alelo *CYP2C19*17* e componente NAT. Pontos que podem ser levantados como prováveis causas nas diferença entre os dois trabalhos são, as populações estudadas, no trabalho de Rodrigues-Soares et al., (2020), os autores utilizaram populações nativas americanas e latino americanas

e não utilizou populações asiáticas, outro ponto é ancestralidade continental, enquanto em nosso trabalho realizamos a subcontinental.

Além desses alelos mencionados, outros alelos apresentaram correlações. Os alelos de função normal *CYP2C19**13 e *15 tiveram uma correlação positiva com os dois componentes subcontinentais africanos. O alelo *CYP2C19**9, com função enzimática reduzida, também apresentou correlação positiva com os WAFR e EAFR, já o alelo *CYP2C19**10 apenas com o componente WAFR. O *CYP2C19**35, alelo sem função enzimática, que apresentou uma correlação positiva com os dois componentes subcontinentais africanos. Enquanto os alelos *CYP2C19**30 e *34 apresentaram correlações positivas com EAFR e SAS, respectivamente.

Em relação ao *CYP2C9**2, com função enzimática reduzida, foi correlacionado positivamente com componente SEUR (Fig. 16) e o *CYP2C9**3, que é sem função enzimática, correlacionado positivamente com SAS e negativamente WAFR e EAFR (Figura 17).

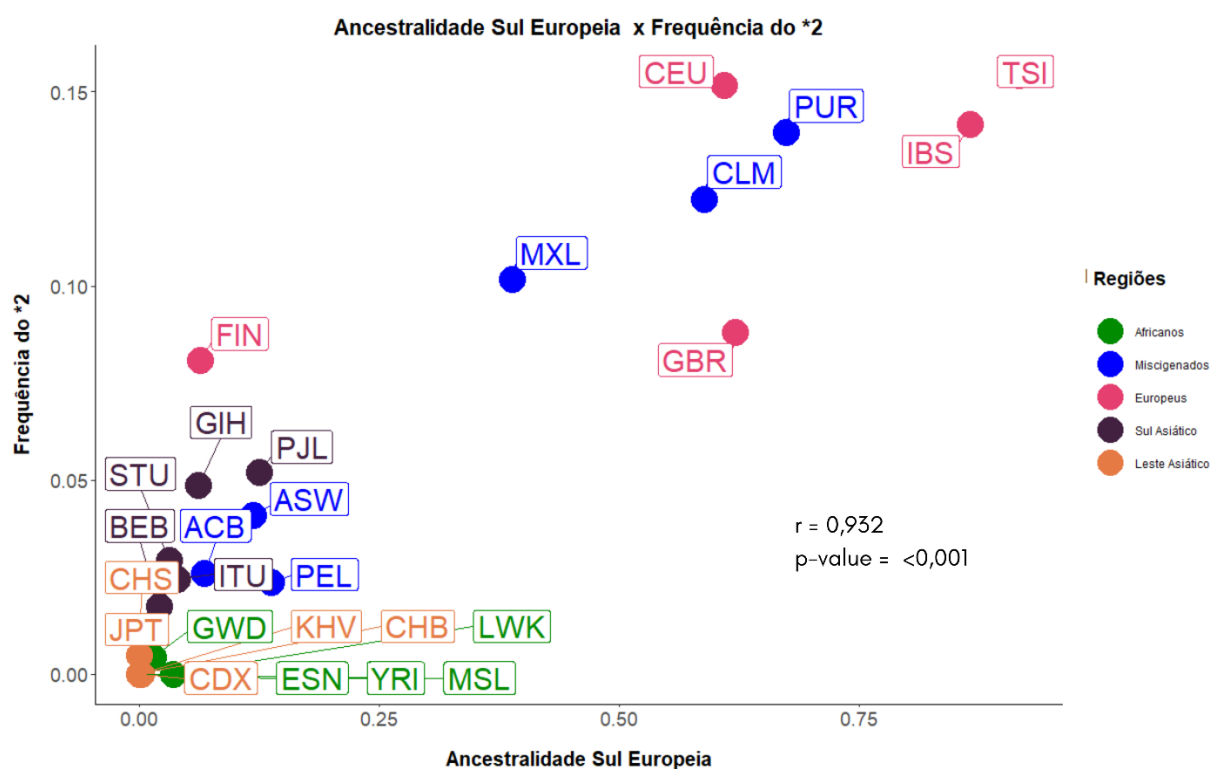


Figura 16. Gráficos Ancestralidade pela frequência *CYP2C9**2. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**,

Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio).

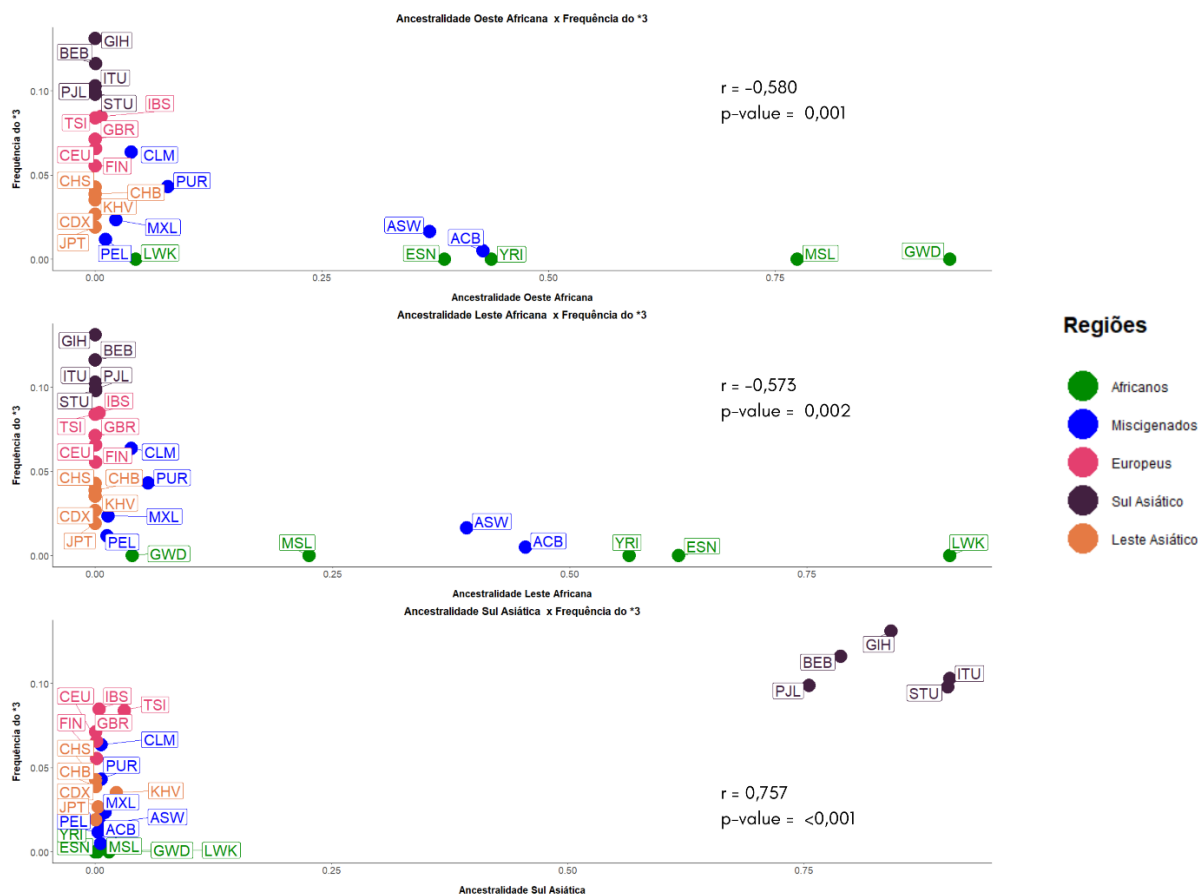


Figura 17. Gráficos Ancestralidade pela frequência *CYP2C9*3*. **LWK**, Luhya do Quênia; **ESN**, Esan da Nigéria; **YRI**, Yoruba da Nigéria; **MSL**, Mende de Serra Leoa; **GWD**, Mandika de Gâmbia; **ACB**, Afro-Caribenhos de Barbados; **ASW**, Indivíduos com ancestralidade africana no sudeste dos EUA; **CLM**, Colombianos de Medellín na Colômbia; **MXL**, Indivíduos com ancestralidade Mexicana vivendo em Los Angeles; **PUR**, Porto-Riquenhos de Porto Rico; **PEL**, Peruanos de Lima; **TSI**, Italianos da Toscana; **IBS**, Ibéricos da Espanha; **GBR**, Britânicos da Inglaterra e Escócia; **CEU**, Descendentes de europeus de Utah; **FIN**, Finlandeses da Finlândia; **PJL**, Paquistaneses; **GIH**, Indianos Gujarati em Houston EUA; **ITU**, Indianos Telugu no Reino Unido; **STU**, Sri Lanka de Tamil no Reino Unido; **BEB**, Bengali de Bangladesh; **CDX**, Chineses Dai de Xishuangbanna; **KHV**, Vietnamitas Kinh da cidade Ho Chi Minh; **CHS**, Chineses Han do sul; **CHB**, Chineses Han de Beijinn; **JPT**, Japonês de Tóquio).

No estudo de Rodrigues-Soares et al., (2020), foi encontrada uma correlação positiva do alelo *CYP2C9*2* com ancestralidade europeia, contudo em nosso trabalho, encontramos essa correlação com o componente subcontinental SEUR, que está presente em maior porcentagem

nas populações do sul europeu. Para o alelo *CYP2C9**3 não encontramos a mesma associação, Rodrigues-Soares et al., (2020), observaram correlação positiva com europeus e negativa com nativos americanos, enquanto em nossos estudos encontramos correlação positiva com o componente SAS e negativa com os componentes WAFR e EAFR. Essa diferença pode ser atribuídas as populações estudadas, pois no trabalho de Rodrigues-Soares et al., (2020), foram estudadas diversas populações nativas americana, enquanto no nosso apenas populações miscigenadas com ancestralidade nativo americana. Um outro motivo são as análises de ancestralidade, em nosso trabalho dividimos os componentes em subcontinentais e no trabalho de Rodrigues-Soares et al., (2020) os componentes foram divididos em continentais.

Outros alelos do gene *CYP2C9* também obtiveram correlação. Os componentes WAFR e EAFR apresentaram correlação positiva com os alelos *CYP2C9**5, *6, *8, *9 e *11, os componentes EAS-W e EAS-E apresentaram correlação positiva com *CYP2C9**13, enquanto o SAS demonstrou correlação positiva com os alelos *CYP2C9**7 e *14. O componente EAS-E apresentou correlação positiva com os alelos *CYP2C9**16 e *33 e o componente EAS-W exibiu correlação com os alelos *CYP2C9**29. Já o componente subcontinental NEUR apresentou correlação positiva com o alelo *CYP2C9**45, enquanto o SEUR apresentou correlação positiva com o alelo *CYP2C9**12. Já o componente EAFR demonstrou correlação com alelo *CYP2C9**31.

A ancestralidade genômica também foi correlacionada com o score de atividade e o fenótipo predito para os dois genes. No *CYP2C19* obtivemos 22 correlações entre componentes de ancestralidade e score de atividade, sendo 12 correlações positivas e 10 negativas, e para fenótipo predito temos, 16 correlações, onde 8 foram correlações positivas e 8 foram correlações negativas.

O score de atividade de valor 0, no gene *CYP2C19* observamos uma correlação positiva com o componente SAS, e uma correlação negativa com o componente SEUR. Já score de atividade 0,5 encontramos correlações positivas com os dois componentes africanos (WAFR e EAFR), enquanto o score de valor 1 observamos correlações negativas com os componentes NAT e SEUR, e positivas com os dois componentes do leste asiático (EAS-W e EAS-E).

Correlações também foram encontradas entre o score de atividade 1,5 e os componentes e os dois componentes do leste asiático (EAS-W e EAS-E), sendo correlações negativas, e correlações positiva com esse score de atividade observamos nos dois componentes africanos

(WAFR e EAFR). O valor de score de atividade 2, observamos uma correlação positiva com NAT e negativa com o SAS.

Já o score de atividade de valor 2,5 também apresentou uma correlação positiva com os dois componentes africanos (WAFR e EAFR) e negativas com os dois componentes do leste asiáticos (EAS-W e EAS-E). E valor 3 de score de atividade também apresenta o mesmo padrão que o 2,5, apresentando correlações positivas com WAFR e EAFR e negativas EAS-W e EAS-E.

Em relação às correlações com fenótipo predito do gene *CYP2C19*, para os PMs encontramos correlação positiva com o SAS e correlação negativa com SEUR, o que difere do trabalho de Rodrigues-Soares et al., (2020), que não encontrou correlação desse fenótipo com nenhum componente continental, contudo nesse trabalho os autores não utilizaram em suas análises populações asiáticas (Tabela 12).

Tabela 12. Valores de coeficiente de correlação e valores de p da correlação entre fenótipos preditos e ancestralidade do gene *CYP2C19*. (NMs, Metabolizadores Normais; PMs, Metabolizadores Lentos; IMs, Metabolizadores Intermediários; RMs, Metabolizadores Rápidos; UMs, Metabolizadores Ultrarrápidos; EAS-W, Leste Asiático – Oeste; NEUR, Norte Europeu; WAFR, Oeste Africano; NAT, Nativo Americano; SEUR, Sul Europeu; EAFR, Leste Africano; EAS-E, Leste Asiático – Leste; SAS, Sul Asiático). Negrito: valores significativos.

Fenótipos	EAS-W	NEUR	WAFR	NAT	SEUR	EAFR	EAS-E	SAS
PMs								
Valor p	0,213	0,181	0,217	0,084	0,002	0,318	0,050	< 0,001
Valor r	0,253	-0,270	-0,251	-0,346	-0,583	-0,204	0,388	0,743
IMs								
Valor p	<0,001	0,371	0,060	0,036	0,013	0,305	0,001	0,273
Valor r	0,686	-0,183	-0,374	-0,412	-0,481	-0,209	0,622	0,223
NMs								
Valor p	0,544	0,803	0,980	<0,001	0,125	0,884	0,362	0,006
Valor r	-0,125	0,051	0,005	0,841	0,309	-0,030	-0,186	-0,526
RMs								
Valor p	<0,001	0,173	0,018	0,336	0,007	0,119	0,001	0,417
Valor r	-0,642	0,275	0,460	-0,197	0,513	0,314	-0,595	-0,166
UMs								
Valor p	0,004	0,393	0,004	0,194	0,312	0,037	0,015	0,620
Valor r	-0,544	0,175	0,550	-0,263	0,206	0,412	-0,470	-0,102

Já com IMs encontramos correlações positivas com dois componentes do leste asiático (EAS-W e EAS-E) e correlações negativas com NAT e SEUR. E os NMs observamos correlação positiva com componente NAT e negativa com o SAS.

A correlação positiva com PMs com o componente SAS pode exemplificar a presença de casamentos endogâmicos nas sociedades sul asiáticas. De acordo com Majumder (2010), a atual sociedade indiana é formada por uma grande variedade de etnias que são casadas entre si, o que gera praticamente nenhum fluxo gênico entre essas etnias, isso ocorre devido a rígidas regras sociais e ao sistema de casamento entre as castas. Este fator cultural pode explicar o excesso de homozigose (que dá origem ao fenótipo PM) nessas populações.

Os RMs apresentaram correlação positiva com os componentes WAFR e SEUR e correlação negativa com os dois componentes do leste asiático (EAS-W e EAS-E). E para UMs observamos que correlações positivas ocorreram com dois componentes africanos (WAFR e EAFR) e correlações negativa os dois componentes do leste asiáticos (EAS-W e EAS-E). No trabalho de Rodrigues-Soares et al., (2020), que correlacionou fenótipos preditos com componentes continental, foi encontrado uma correlação negativa de UMs com o componente Nativo Americano. A diferença com nosso trabalho pode ser atribuída pelo fato que não utilizamos populações de nativos americanos, apenas populações miscigenadas.

O gene *CYP2C9* apresentou 9 correlações, sendo que 5 foram correlações positivas e 4 negativas entre os componentes de ancestralidade e score de atividade. Já os fenótipos preditos foram encontrados 4 correlações, 2 correlações positivas e 2 negativas.

Uma correlação positiva foi observada entre o score de atividade 0,5 e componente EAFR, com score de atividade de valor 1, encontramos correlação positiva com SEUR e negativa com EAS-W. Já o score 1,5 foi correlacionado positivamente com SEUR e negativamente com dois componentes do leste asiático (EAS-W e EAS-E). Enquanto o score atividade de valor 2 observamos correlação positiva com os dois componentes do leste asiático (EAS-W e EAS-E) e uma correlação negativa com SEUR.

Em relação a correlação com fenótipos preditos, os NMs encontramos correlação positiva com o componente NAT e uma correlação negativa SAS. Já nos IMs encontramos correlação positiva com SEUR e uma correlação negativa com componente EAS-W (Tabela 13).

Tabela 13. Valores de coeficiente de correlação e valores de p da correlação entre fenótipos preditos e ancestralidade do gene *CYP2C9*. (NMs, Metabolizadores Normais; IMs, Metabolizadores

Intermediários; **EAS-W**, Leste Asiático – Oeste; **NEUR**, Norte Europeu; **WAFR**, Oeste Africano; **NAT**, Nativo Americano; **SEUR**, Sul Europeu; **EAFR**, Leste Africano; **EAS-E**, Leste Asiático – Leste; **SAS**, Sul Asiático). Negrito: valores significativos.

Fenótipos	EAS-W	NEUR	WAFR	NAT	SEUR	EAFR	EAS-E	SAS
IMs								
Valor p	0,040	0,421	0,592	0,167	0,007	0,373	0,072	0,828
Valor r	-0,406	0,165	0,110	-0,279	0,515	0,182	-0,359	-0,045
NMs								
Valor p	0,172	0,428	0,401	0,035	0,354	0,602	0,144	0,001
Valor r	0,276	-0,162	0,172	0,415	-0,189	0,107	0,295	-0,602

7. Conclusão

Dada a escassez de literatura sobre a correlação entre ancestralidade genômica e alelos de interesse farmacogenômico, mais pesquisas ainda são necessárias para aprimorar nosso conhecimento sobre farmacogenética populacional, especialmente considerando a importância das diferenças subcontinentais. Assim, na nossa iniciativa para preencher esta lacuna, descrevemos correlações, que podem ser valiosas para a seleção de alelos para painéis farmacogenéticos e, portanto, contribuindo para o futuro da implementação da farmacogenética em regiões específicas do mundo.

8. Referências Bibliográficas

1000 GENOMES PROJECT CONSORTIUM et al. A map of human genome variation from population-scale sequencing. **Nature**, v. 467, n. 7319, p. 1061–1073, 28 out. 2010.

ALEXANDER, D. H.; NOVEMBRE, J.; LANGE, K. Fast model-based estimation of ancestry in unrelated individuals. **Genome research**, v. 19, n. 9, p. 1655–1664, set. 2009.

AUTON, A. et al. A global reference for human genetic variation. **Nature** 2015 **526:7571**, v. 526, n. 7571, p. 68–74, 30 set. 2015.

BERGSTRÖM, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. **Science**, v. 367, n. 6484, 20 mar. 2020.

BIDDANDA, A.; RICE, D. P.; NOVEMBRE, J. A variant-centric perspective on geographic patterns of human allele frequency variation. **eLife**, v. 9, p. 1–23, 1 dez. 2020.

BLANCO, G. et al. Interaction of CYP2C8 and CYP2C9 genotypes modifies the risk for nonsteroidal anti-inflammatory drugs-related acute gastrointestinal bleeding. **Pharmacogenetics and genomics**, v. 18, n. 1, p. 37–43, jan. 2008.

BRADBURD, G. S.; RALPH, P. L.; COOP, G. M. A Spatial Framework for Understanding Population Structure and Admixture. **PLoS genetics**, v. 12, n. 1, 2016.

CAUDLE, K. E. et al. Standardizing CYP 2D6 Genotype to Phenotype Translation: Consensus Recommendations from the Clinical Pharmacogenetics Implementation Consortium and Dutch Pharmacogenetics Working Group. **Clinical and Translational Science**, v. 13, n. 1, p. 116–124, 24 jan. 2020.

CÉSPEDES-GARRO, C. et al. Worldwide interethnic variability and geographical distribution of CYP2C9 genotypes and phenotypes. **Expert Opinion on Drug Metabolism and Toxicology**, v. 11, n. 12, p. 1893–1905, 2015.

CHEVITARESE, J. **DETERMINAÇÃO DA ESTRUTURA GENÉTICA DAS POPULAÇÕES HUMANAS E INFERÊNCIA DOS FATORES EVOLUTIVOS QUE CONTRIBUÍRAM PARA A SUA FORMAÇÃO**. [s.l: s.n.].

CPIC. Disponível em: <<https://cpicpgx.org/>>. Acesso em: 23 ago. 2023.

DAS, R.; UPADHYAI, P. An Ancestry Informative Marker Set Which Recapitulates the Known Fine Structure of Populations in South Asia. **Genome biology and evolution**, v. 10, n. 9, p. 2408–2416, 1 set. 2018.

FRICKE-GALINDO, I. et al. Interethnic variation of CYP2C19 alleles, “predicted” phenotypes and “measured” metabolic phenotypes across world populations. **Pharmacogenomics Journal**, v. 16, n. 2, p. 113–123, 2016.

GAIKWAD, T.; GHOSH, K.; SHETTY, S. VKORC1 and CYP2C9 genotype distribution in Asian countries. **Thrombosis Research**, v. 134, n. 3, p. 537–544, 1 set. 2014.

GILBERT, E.; SHANMUGAM, A.; CAVALLERI, G. L. Revealing the recent demographic history of Europe via haplotype sharing in the UK Biobank. **Proceedings of the National Academy of Sciences of the United States of America**, v. 119, n. 25, 21 jun. 2022.

GOUVEIA, M. H. et al. Unappreciated Subcontinental Admixture in Europeans and European Americans: Implications for Genetic Epidemiology Studies. **bioRxiv**, 2022.

HANDLEY, L. J. L. et al. Going the distance: human population genetics in a clinal world. **Trends in Genetics**, v. 23, n. 9, p. 432–439, set. 2007.

HARTL, D. L.; CLARK, A. G. *Princípios de Genética de Populações*. 2010.

HENN, B. M.; CAVALLI-SFORZA, L. L.; FELDMAN, M. W. The great human expansion. **Proceedings of the National Academy of Sciences of the United States of America**, v. 109, n. 44, p. 17758–17764, 30 out. 2012.

HENRY, J. P. [Genetics and origin of Homo sapiens]. **Medecine sciences : M/S**, v. 35, n. 1, p. 39–45, 1 jan. 2019.

HICKS, J. K. et al. Clinical pharmacogenetics implementation consortium guideline (CPIC) for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants: 2016 update. **Clinical pharmacology and therapeutics**, v. 102, n. 1, p. 37–44, 1 jul. 2017.

JOHNSON, J. A. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. 2017.

JOMBART, T.; AHMED, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. **Bioinformatics (Oxford, England)**, v. 27, n. 21, p. 3070–3071, nov. 2011.

JORDE, L. B. et al. The Distribution of Human Genetic Diversity: A Comparison of Mitochondrial, Autosomal, and Y-Chromosome Data. **The American Journal of Human Genetics**, v. 66, n. 3, p. 979–988, mar. 2000.

JUNG, J. Y. et al. Ancestry informative markers (AIMs) for Korean and other East Asian and South East Asian populations. **International journal of legal medicine**, v. 133, n. 6, p. 1711–1719, 1 nov. 2019.

KARANIEWICZ-ADA, M.; UCZAK, M.; GŁWKA, F. Pharmacokinetic studies of enantiomers of ibuprofen and its chiral metabolites in humans with different variants of genes coding CYP2C8 and CYP2C9 isoenzymes. **Xenobiotica; the fate of foreign compounds in biological systems**, v. 39, n. 6, p. 476–485, 2009.

KARNES, J. H. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C9 and HLA-B Genotypes and Phenytoin Dosing: 2020 Update. v. 109, p. 302, 2021.

KEHDY, F. S. G. et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 28, p. 8696–8701, 14 jul. 2015.

KOLEVA-KOLAROVA, R. et al. Financial incentives to promote personalized medicine in Europe: an overview and guidance for implementation. **Personalized medicine**, v. 20, n. 4, p. 305–319, 1 jul. 2023.

LAM, Y. W. F. Principles of Pharmacogenomics. Em: **Pharmacogenomics**. Second Edi ed. [s.l.] Elsevier, 2019. p. 1–53.

LAZARIDIS, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. **Nature**, v. 513, n. 7518, p. 409–413, 18 set. 2014.

LEE, C. R. et al. Clinical Pharmacogenetics Implementation Consortium Guideline for CYP2C19 Genotype and Clopidogrel Therapy: 2022 Update. **Clinical Pharmacology and Therapeutics**, v. 112, n. 5, p. 959–967, 2022.

LEE, C. R.; GOLDSTEIN, J. A.; PIEPER, J. A. Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data. **Pharmacogenetics**, v. 12, n. 3, p. 251–263, 2002.

LEE, S. BEEN et al. Calling Star Alleles With Stargazer in 28 Pharmacogenes With Whole Genome Sequences. **Clinical Pharmacology and Therapeutics**, v. 106, n. 6, p. 1328–1337, 2019.

LI, J. et al. Global patterns of genetic diversity and signals of natural selection for human ADME genes. **Human molecular genetics**, v. 20, n. 3, p. 528–540, fev. 2011a.

LI, J. et al. Global patterns of genetic diversity and signals of natural selection for human ADME genes. **Human molecular genetics**, v. 20, n. 3, p. 528–540, fev. 2011b.

LIMA, J. J. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and Proton Pump Inhibitor Dosing. **Clinical pharmacology and therapeutics**, v. 109, n. 6, p. 1417–1423, 1 jun. 2021.

LLERENA, A. et al. Interethnic variability of CYP2D6 alleles and of predicted and measured metabolic phenotypes across world populations. **Expert opinion on drug metabolism & toxicology**, v. 10, n. 11, p. 1569–1583, 1 nov. 2014.

MA, Q.; LU, A. Y. H. Pharmacogenetics, pharmacogenomics, and individualized medicine. **Pharmacological reviews**, v. 63, n. 2, p. 437–459, jun. 2011.

MAJUMDER, P. P. The human genetic history of South Asia. **Current biology : CB**, v. 20, n. 4, 23 fev. 2010.

MALLICK, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. **Nature**, v. 538, n. 7624, p. 201–206, 2016.

MARSH, S. E. et al. The adaptive immune system restrains Alzheimer's disease pathogenesis by modulating microglial function. **Proceedings of the National Academy of Sciences**, v. 113, n. 9, p. E1316–E1325, 16 mar. 2016.

MEGA, J. L. et al. Cytochrome p-450 polymorphisms and response to clopidogrel. **The New England journal of medicine**, v. 360, n. 4, p. 354–362, 22 jan. 2009.

METZGER, I. F.; SOUZA-COSTA, D. C.; TANUS-SANTOS, J. E. FARMACOGENÉTICA: PRINCÍPIOS, APLICAÇÕES E PERSPECTIVAS. **Medicina (Ribeirão Preto)**, v. 39, n. 4, p. 515–521, 30 dez. 2006.

MORRIS, S. A. et al. Cost Effectiveness of Pharmacogenetic Testing for Drugs with Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines: A Systematic Review. **Clinical pharmacology and therapeutics**, v. 112, n. 6, p. 1318–1328, 1 dez. 2022.

MRAZEK, D. A. Psychiatric pharmacogenomic testing in clinical practice. **Dialogues in clinical neuroscience**, v. 12, n. 1, p. 69–76, 2010.

NELSON, D. R.; NEBERT, D. W. Cytochrome *P450* (*CYP*) Gene Superfamily. Em: **eLS**. [s.l.] Wiley, 2018. p. 1–19.

NUSSBAUM, R. L.; WILLARD, H. F.; HAMOSH, A. **THOMPSON & THOMPSON GENÉTICA MÉDICA**. [s.l: s.n.].

ONI-ORISAN, A. et al. An Introductory Tutorial on Cardiovascular Pharmacogenetics for Healthcare Providers. **Clinical Pharmacology and Therapeutics**, 1 ago. 2023.

PADHUKASAHASRAM, B. Inferring ancestry from population genomic data and its applications. **Frontiers in genetics**, v. 5, n. JUL, 2014.

PAN, Z.; XU, S. Population genomics of East Asian ethnic groups. **Hereditas**, v. 157, n. 1, p. 1–10, 1 dez. 2020.

PETER, B. M.; PETKOVA, D.; NOVEMBRE, J. Genetic Landscapes Reveal How Human Genetic Diversity Aligns with Geography. **Molecular biology and evolution**, v. 37, n. 4, p. 943–951, 1 abr. 2020.

PharmGKB. Disponível em: <<https://www.pharmgkb.org/>>. Acesso em: 10 dez. 2022.

PharmVar. Disponível em: <<https://www.pharmvar.org/>>. Acesso em: 10 dez. 2022.

RAMANI, A. et al. Differentiation of Asian population samples using the Illumina ForenSeq kit. **Forensic Science International: Genetics**, v. 48, p. 102318, 1 set. 2020.

RAMASAMY, K. et al. Influence of CYP2C9 genetic polymorphism and undernourishment on plasma-free phenytoin concentrations in epileptic patients. **Therapeutic Drug Monitoring**, v. 32, n. 6, p. 762–766, dez. 2010.

RAMOS, E. et al. Pharmacogenomics, ancestry and clinical decision making for global populations. **The pharmacogenomics journal**, v. 14, n. 3, p. 217–222, 2014.

RANG, H. P. et al. **Rang & Dale: Farmacologia**. [s.l: s.n.].

REGINA, M.; WANYCE, B.-O.; ROBINSON, M. **Genética Humana**. Porto Alegre: Artmed, 2013.

RETTIE, A. E.; JONES, J. P. Clinical and toxicological relevance of CYP2C9: drug-drug interactions and pharmacogenetics. **Annual review of pharmacology and toxicology**, v. 45, p. 477–494, 2005.

RIDLY, M. **EVOLUÇÃO**. [s.l: s.n.].

RODRIGUES-SOARES, F. et al. Genomic Ancestry, CYP2D6, CYP2C9, and CYP2C19 Among Latin Americans. **Clinical Pharmacology and Therapeutics**, v. 107, n. 1, p. 257–268, 2020.

SCUDELER, M. M.; RODRIGUES-SOARES, F. FARMACOGENÉTICA NA AMÉRICA LATINA PHARMACOGENETICS IN LATIN AMERICA. **Acta Biologica Brasiliensia**, n. 2, 2020a.

SCUDELER, M. M.; RODRIGUES-SOARES, F. FARMACOGENÉTICA NA AMÉRICA LATINA PHARMACOGENETICS IN LATIN AMERICA. **Acta Biologica Brasiliensia**, n. 2, 2020b.

SIBBING, D. et al. Cytochrome 2C19*17 Allelic Variant, Platelet Aggregation, Bleeding Events, and Stent Thrombosis in Clopidogrel-Treated Patients With Coronary Stent Placement. **Circulation**, v. 121, n. 4, p. 512–518, 2 fev. 2010.

SIMON, T. et al. Genetic Determinants of Response to Clopidogrel and Cardiovascular Events. **New England Journal of Medicine**, v. 360, n. 4, p. 363–375, 22 jan. 2009.

SOSA-MACÍAS, M.; LLERENA, A. Cytochrome P450 genetic polymorphisms of Mexican indigenous populations. **Drug Metabolism and Drug Interactions**, v. 28, n. 4, p. 193–208, 1 jan. 2013.

STANEK, E. J. et al. Adoption of Pharmacogenomic Testing by US Physicians: Results of a Nationwide Survey. **Clinical Pharmacology & Therapeutics**, v. 91, n. 3, p. 450–458, 25 mar. 2012.

SUAREZ-KURTZ, G. Pharmacogenetics in the Brazilian population. **Frontiers in Pharmacology**, v. OCT, p. 118, 2010.

TAN-KOI, W. C. et al. The importance of ethnicity definitions and pharmacogenomics in ethnobridging and pharmacovigilance. **Pharmacogenomics: Challenges and Opportunities in Therapeutic Implementation**, p. 289–327, 1 jan. 2018.

THEKEN, K. N. et al. Clinical Pharmacogenetics Implementation Consortium Guideline (CPIC) for CYP2C9 and Nonsteroidal Anti-Inflammatory Drugs. **Clinical pharmacology and therapeutics**, v. 108, n. 2, p. 191–200, 1 ago. 2020.

URBAN, T. J. Race, Ethnicity, Ancestry, and Pharmacogenetics. **Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine**, v. 77, n. 2, p. 133–139, mar. 2010.

VELLEKOOP, H. et al. The Net Benefit of Personalized Medicine: A Systematic Literature Review and Regression Analysis. **Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research**, v. 25, n. 8, p. 1428–1438, 1 ago. 2022.

VERBELEN, M.; WEALE, M. E.; LEWIS, C. M. Cost-effectiveness of pharmacogenetic-guided treatment: are we there yet? **The pharmacogenomics journal**, v. 17, n. 5, p. 395–402, 1 out. 2017.

VISSER, L. E. et al. Allelic variants of cytochrome P450 2C9 modify the interaction between nonsteroidal anti-inflammatory drugs and coumarin anticoagulants. **Clinical Pharmacology and Therapeutics**, v. 77, n. 6, p. 479–485, 2005.

WATKINS, W. S. et al. Genetic Variation Among World Populations: Inferences From 100 *Alu* Insertion Polymorphisms. **Genome Research**, v. 13, n. 7, p. 1607–1618, 1 jul. 2003.

WILLIAMS, K.; ARRON, S. T. Association of CYP2C19 *17/*17 Genotype With the Risk of Voriconazole-Associated Squamous Cell Carcinoma. **JAMA dermatology**, v. 152, n. 6, p. 719–720, 1 jun. 2016.

WU, D. et al. Large-Scale Whole-Genome Sequencing of Three Diverse Asian Populations in Singapore. **Cell**, v. 179, n. 3, p. 736–749.e15, 17 out. 2019.

XING, J. et al. Mobile elements create structural variation: Analysis of a complete human genome. **Genome Research**, v. 19, n. 9, p. 1516–1526, set. 2009.

ZANGER, U. M.; SCHWAB, M. Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. **Pharmacology & Therapeutics**, v. 138, n. 1, p. 103–141, abr. 2013.

ZHANG, H. et al. The Advantages and Challenges of Diversity in Pharmacogenomics: Can Minority Populations Bring Us Closer to Implementation? **Clinical pharmacology and therapeutics**, v. 106, n. 2, p. 338–349, 1 ago. 2019.

ZHOU, Y. et al. Global distribution of functionally important CYP2C9 alleles and their inferred metabolic consequences. **Human genomics**, v. 17, n. 1, 1 dez. 2023.