

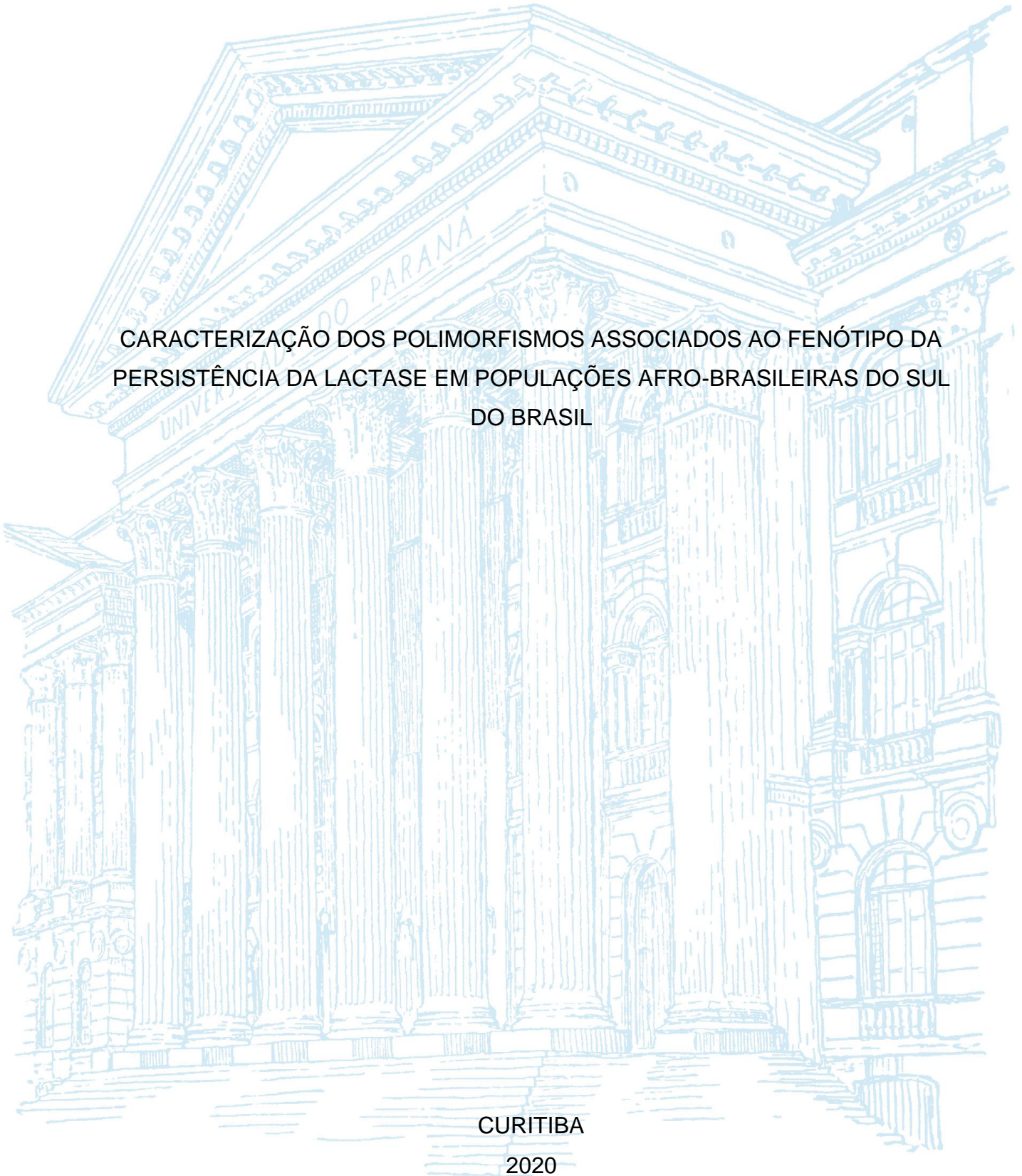
UNIVERSIDADE FEDERAL DO PARANÁ

NATALIE MARY SUKOW

CARACTERIZAÇÃO DOS POLIMORFISMOS ASSOCIADOS AO FENÓTIPO DA
PERSISTÊNCIA DA LACTASE EM POPULAÇÕES AFRO-BRASILEIRAS DO SUL
DO BRASIL

CURITIBA

2020



NATALIE MARY SUKOW

CARACTERIZAÇÃO DOS POLIMORFISMOS ASSOCIADOS AO FENÓTIPO DA
PERSISTÊNCIA DA LACTASE EM POPULAÇÕES AFRO-BRASILEIRAS DO SUL
DO BRASIL

Monografia apresentada ao Curso de Ciências Biológicas, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Bacharela em Ciências Biológicas.

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

Coorientadora: Ana Cecília Guimarães Alves.

CURITIBA

2020

AGRADECIMENTOS

À professora Márcia, pela orientação atenciosa, apoio constante e por todos os ensinamentos e sugestões que proporcionaram não só a construção deste trabalho, mas também meu desenvolvimento científico e pessoal.

À Ana Cecília, minha co-orientadora e companheira de projeto que, desde meu ingresso no LGMH, auxiliou-me das mais diversas formas e dividiu comigo as alegrias e frustrações vividas no desenvolvimento desta pesquisa. Agradeço também suas correções minuciosas que contribuíram imensamente para a finalização do presente trabalho.

Aos demais colegas de LGMH, pelos eventuais auxílios técnicos e científicos e pelo convívio.

À professora Ilíada Rainha de Souza e ao professor Ricardo Lehtonen Rodrigues de Souza, por terem gentilmente cedido as amostras que propiciaram o desenvolvimento desta monografia.

A todos os brasileiros que, voluntariamente, doaram amostras de sangue e, assim, propiciaram o desenvolvimento deste e de outros trabalhos.

Aos meus amigos da Biologia, Ana Clara, Gabriel, Laura, Milena, Natália e Victor, e aos amigos não-biólogos, Ariadne, Guilherme, Suelen e Vinícius, agradeço pelo companheirismo, amizade, carinho e incentivo em todos os momentos – bons ou ruins. Vocês são imprescindíveis!

À minha família, pelo amor, carinho e visão crítica de mundo que compartilham comigo constantemente. Aos meus pais, Zélia e José Augusto, por sempre estarem ao meu lado, me apoiando incondicionalmente, por serem meus grandes exemplos e por me tornarem quem eu sou. Às minhas irmãs, Maria Augusta e Nikita, por todas as conversas, risadas, brincadeiras e, acima de tudo, pelo companheirismo cotidiano, sempre permeado de muita ternura. Agradeço, em especial, minha irmã Nikita pela sua “consultoria histórica” em alguns capítulos deste trabalho. À minha avó, Estela, agradeço pelo cuidado, carinho e pelas inúmeras histórias já compartilhadas. Sem o apoio de vocês, nada seria possível. Obrigada!

“Gosto de ser gente porque, inacabado, sei que sou um ser condicionado mas, consciente do inacabamento, sei que posso ir mais além dele”.

Paulo Freire

RESUMO

A lactase é uma enzima localizada no intestino delgado, responsável pela hidrólise da lactose (dissacarídeo) a glicose e galactose – monossacarídeos passíveis de absorção pela parede intestinal. Sua expressão é elevada em mamíferos recém-nascidos e diminui progressivamente após o desmame (fenótipo lactase não persistente ou LNP). Contudo, algumas populações humanas com histórico de pastoralismo e ordenha, associado ao consumo de leite e derivados, têm a atividade enzimática mantida ao longo da vida (fenótipo lactase persistente ou LP), sendo um dos exemplos mais conhecidos de coevolução gene-cultura. A manutenção da expressão da enzima, nesses casos, está associada a polimorfismos de nucleotídeo único (SNPs) na região intensificadora do gene *LCT* (responsável por codificar a lactase), que sobrepõem ao gene *MCM6*. As substituições que originaram estes polimorfismos ocorreram por convergência evolutiva em populações pastoralistas de diferentes continentes, como as mutações originadas na Europa -13910 C>T e -22018 G>A (localizadas nos íntrons 13 e 9, respectivamente); as originadas na África -13907 C>G, -14009 T>G, -14010 G>C e a originada no Oriente Médio -13915 T>G (localizadas no íntron 13). Considerando o caráter miscigenado da população brasileira, que possui ancestralidade ameríndia, africana, europeia e asiática (em menor proporção), é esperado que esses diferentes alelos ocorram no país. No entanto, os estudos acerca de polimorfismos associados à LP no Brasil referem-se majoritariamente ao alelo -13910*T, de origem europeia, e nunca incluíram populações de Comunidades Quilombolas – cuja ancestralidade tende a ser predominantemente africana. Sendo assim, o objetivo do presente trabalho foi identificar polimorfismos associados ao fenótipo da persistência da lactase em populações afro-brasileiras de Curitiba e Região Metropolitana (PR) e da Comunidade Quilombola de Sertão do Valongo (SC), no Sul do Brasil. Para tanto, foram incluídas amostras coletadas previamente de 241 e 18 indivíduos de Curitiba e Região e de Sertão do Valongo, respectivamente, as quais passaram por reações de amplificação e sequenciamento de uma região do íntron 13 do gene *MCM6* e análises bioinformáticas. Foram observadas mutações associadas ao fenótipo LP somente nos afro-brasileiros de Curitiba e Região, sendo elas as variantes -13910*T (18,7%), -13915*G (0,2%), originadas na Europa e Oriente Médio, respectivamente, e a variante -14011*T (0,2%), de origem não especificada. O polimorfismo -13730 T>G, que possui associação não confirmada à LP, foi o único observado em ambas as populações, nas frequências de 2,5% em Curitiba e Região e de 16,7% em Sertão do Valongo – a maior já reportada para este alelo. Consequentemente, através de inferência, foi constatada predominância de indivíduos LNP em ambas as populações, sendo que em Sertão do Valongo este foi o único fenótipo observado. A ausência de outros alelos de origem africana em ambos os grupos investigados provavelmente decorre de eventos de deriva genética. A prevalência do SNP de origem europeia nos afro-brasileiros de Curitiba se deve a elevada proporção de ancestralidade europeia encontrada em brasileiros – sobretudo na região Sul do país. Na população de Curitiba investigada foram identificados sete haplótipos, dos quais apenas dois estão também presentes em Sertão do Valongo. A menor diversidade genética observada na Comunidade de Sertão do Valongo concorda com a ausência de fluxo gênico recente e ação da deriva genética devido ao efeito fundador na formação da comunidade. Tais fatores contribuem para a diferenciação genética intermediária constatada com relação a outras populações afro-brasileiras.

Palavras-Chave: Persistência da lactase; afro-brasileiros; Comunidade Quilombola.

ABSTRACT

Lactase is an enzyme located in the small intestine. It is responsible for the hydrolysis of lactose (disaccharide) into glucose and galactose – monosaccharides absorbed by the intestinal mucosa. Lactase expression is high in newborn mammals and decreases progressively after weaning (lactase non-persistence phenotype or LNP). However, some human populations with pastoralism and milking history, associated with milk and dairy products consumption, show high enzymatic activity throughout adulthood (lactase persistence phenotype or LP) – one of the most popular examples of gene-culture coevolution. The maintenance of lactase activity, in these cases, is associated with single nucleotide polymorphisms (SNPs) in the enhancer region of the *LCT* gene (accountable for codifying the lactase enzyme), overlapping the *MCM6* gene. The nucleotide substitutions that originated these polymorphisms have arisen due to convergent evolution in pastoralist populations from different continents, like the -13910 C>T and -22018 G>A (located in introns 13 and 9, respectively) polymorphisms that originated in Europe; the -13907 C>G, -14009 T>G, -14010 G>C polymorphisms that originated in Africa, and the -13915 T>G (residing in intron 13) polymorphisms that originated in the Middle East. Considering the Brazilian population's admixture, which has Amerindian, African, European, and Asian ancestry (in lower proportion), it is expected that these different alleles be found in the country. Nevertheless, the studies regarding polymorphisms associated with LP in Brazil refer mostly to the -13910*T allele of European origin and never included data from the Quilombolas Communities – whose ancestry tends to be predominantly African. Therefore, the present study aimed to identify polymorphisms associated with lactase persistence phenotype in Afro-Brazilian populations from Curitiba and Metropolitan Region (Paraná state), and Sertão do Valongo's Quilombola Community (Santa Catarina state), both located in Southern Brazil. Collected samples were included, from 241 and 18 individuals from Curitiba and Sertão do Valongo, respectively, which were submitted to amplification reactions and sequencing of a region from the intron 13 from the *MCM6* gene, as well as bioinformatics analysis. Polymorphisms associated with LP phenotype were observed only in Afro-Brazilians from Curitiba, including the -13910*T (18,7%) and -13915*G (0,2%) alleles, which originated in Europe and the Middle East, respectively, and the -14011*T (0,2%), unspecified origin. The -13730*G variant, which has an unconfirmed association with LP phenotype, was the only polymorphism observed in both populations, with 2,5% in Curitiba and 16,7% in Sertão do Valongo – the highest frequency ever reported for this allele. Hence, through phenotype inference, a predominance of LNP individuals in both populations was verified, as this phenotype was the only one observed in Sertão do Valongo. The absence of other African origin alleles in both groups assessed comes, presumably, from genetic drift events. In contrast, the prevalence of the European origin SNP in Curitiba's Afro-Brazilians results from the high European ancestry proportion found in southern Brazilians. In the examined population of Curitiba, seven haplotypes were identified, of which only two are also present in Sertão do Valongo. The lowest values of genetic diversity observed in the Sertão do Valongo' Community correlate with the recent absence of genetic flow and the genetic drift activity caused by the founding effect on the community's

development. These factors contribute to the intermediate genetic differentiation found with other Afro-Brazilian populations.

Keywords: Lactase persistence; Afro-Brazilians; Quilombola Community.

LISTA DE FIGURAS

FIGURA 1 -	ESTRUTURA DA LACTOSE.....	19
FIGURA 2 -	METABOLISMO DA LACTOSE	20
FIGURA 3 -	DIGESTÃO DA LACTOSE NO TRATO INTESTINAL	22
FIGURA 4 -	DISTRIBUIÇÃO MUNDIAL DO FENÓTIPO LP	29
FIGURA 5 -	LOCALIZAÇÃO DOS GENES <i>LCT</i> E <i>MCM6</i> NO CROMOSSOMO 2..	31
FIGURA 6 -	POLIMORFISMOS DO GENE <i>MCM6</i> ASSOCIADOS À PERSISTÊNCIA DA LACTASE.....	33
FIGURA 7 -	LOCALIZAÇÃO DO SÍTIO DE LIGAÇÃO DO FATOR DE TRANSCRIÇÃO Oct-1	34
FIGURA 8 -	DISTRIBUIÇÃO GEOGRÁFICA DAS FREQUÊNCIAS DOS SNPS ASSOCIADOS À PERSISTÊNCIA DA LACTASE NA ÁFRICA E EURÁSIA	36
FIGURA 9 -	POSIÇÕES NUCLEOTÍDICAS UTILIZADAS PARA DETERMINAR OS HAPLÓTIPOS DA LACTASE	42
FIGURA 10 -	HAPLÓTIPOS DA LACTASE MAIS COMUNS MUNDIALMENTE	42
FIGURA 11 -	PRINCIPAIS ROTAS DO COMÉRCIO TRANSATLÂNTICO DE ESCRAVIZADOS PARA O BRASIL DURANTE OS SÉCULOS XVI AO XIX	59
FIGURA 12 -	DISTRIBUIÇÃO DA POPULAÇÃO NEGRA NOS MUNICÍPIOS DO ESTADO DO PARANÁ	64
FIGURA 13 -	LOCALIDADES DAS POPULAÇÕES AMOSTRADAS	71

LISTA DE TABELAS

TABELA 1 -	HAPLÓTIPOS ASSOCIADOS À PERSISTÊNCIA DA LACTASE	45
TABELA 2 -	INICIADORES ESPECÍFICOS UTILIZADOS NA REAÇÃO DE PCR DO GENE <i>MCM6</i>	72
TABELA 3 -	PROTOCOLO DE PCR UTILIZADO PARA AS AMOSTRAS PROVENIENTES DO LABORATÓRIO DE GENÉTICA MOLECULAR HUMANA (UFPR).....	72
TABELA 4 -	PROTOCOLO DE PCR UTILIZADO PARA AS AMOSTRAS PROVENIENTES DO LABORATÓRIO DE LIGAÇÃO E POLIMORFISMOS (UFPR) E DO LABORATÓRIO DE POLIMORFISMOS GENÉTICOS (UFSC).....	72
TABELA 5 -	PROGRAMA DO TERMOCICLADOR UTILIZADO NA PCR DO GENE <i>MCM6</i>	73
TABELA 6 -	CONCENTRAÇÃO DOS REAGENTES PARA A REAÇÃO DE PURIFICAÇÃO.....	73
TABELA 7 -	CONCENTRAÇÃO DOS REAGENTES PARA A REAÇÃO DE SEQUENCIAMENTO	74
TABELA 8 -	PROGRAMA DO TERMOCICLADOR PARA A REAÇÃO DE SEQUENCIAMENTO	74
TABELA 9 -	ALELOS UTILIZADOS PARA ATRIBUIÇÃO DO FENÓTIPO LP.....	76
TABELA 10 -	FREQUÊNCIAS ALÉLICAS DAS VARIANTES DO ÍNTRON 13 DO GENE <i>MCM6</i>	78
TABELA 11 -	FREQUÊNCIAS GENOTÍPICAS DOS LOCI DO ÍNTRON 13 DO GENE <i>MCM6</i> ANALISADOS.....	79
TABELA 12 -	FREQUÊNCIAS FENOTÍPICAS DAS POPULAÇÕES AFRO-BRASILEIRAS DE CURITIBA E REGIÃO METROPOLITANA E DE SERTÃO DO VALONGO	80
TABELA 13 -	HAPLÓTIPOS CONTENDO AS VARIANTES ASSOCIADAS AO FENÓTIPO LP NA POPULAÇÃO DA REGIÃO DE CURITIBA.....	80
TABELA 14 -	HAPLÓTIPOS PRESENTES NAS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS	81

TABELA 15 - HETEROZIGOSIDADE ESPERADA E OBSERVADA NAS POPULAÇÕES DA REGIÃO DE CURITIBA E DE SERTÃO DO VALONGO	82
TABELA 16 - COMPARAÇÃO DAS FREQUÊNCIAS GENOTÍPICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS DE CURITIBA E REGIÃO E DE SERTÃO DO VALONGO	83
TABELA 17 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE <i>MCM6</i> EM POPULAÇÕES BRASILEIRAS.....	84
TABELA 18 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E DEMAIS POPULAÇÕES BRASILEIRAS.....	85
TABELA 19 - FREQUÊNCIAS ALÉLICAS DO POLIMORFISMO -13910 C>T EM POPULAÇÕES LATINO-AMERICANAS.....	87
TABELA 20 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E OUTRAS POPULAÇÕES LATINO-AMERICANAS.....	88
TABELA 21 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE <i>MCM6</i> EM POPULAÇÕES LATINO-AMERICANAS DO <i>1000 GENOMES PROJECT</i> E <i>gnomAD</i>	89
TABELA 22 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E POPULAÇÕES LATINO-AMERICANAS DO <i>1000 GENOMES PROJECT</i> E <i>gnomAD</i>	89
TABELA 23 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE <i>MCM6</i> EM POPULAÇÕES AFRICANAS	93
TABELA 24 - VALORES DE <i>P</i> COM RELAÇÃO À SIMILARIDADE DAS FREQUÊNCIAS ALÉLICAS ENTRE A POPULAÇÃO AFRO-BRASILEIRA DA REGIÃO DE CURITIBA E POPULAÇÕES AFRICANAS.....	94
TABELA 25 - VALORES DE <i>P</i> COM RELAÇÃO À SIMILARIDADE DAS FREQUÊNCIAS ALÉLICAS ENTRE A POPULAÇÃO AFRO-	

	BRASILEIRA DA COMUNIDADE QUILOMBOLA DE SERTÃO DO VALONGO E POPULAÇÕES AFRICANAS.....	95
TABELA 26 -	FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE <i>MCM6</i> EM POPULAÇÕES DE ORIGEM AFRICANA DO <i>1000 GENOMES PROJECT</i> E <i>gnomAD</i>	96
TABELA 27 -	DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E POPULAÇÕES DE ORIGEM AFRICANA DO <i>1000 GENOMES PROJECT</i> E <i>gnomAD</i>	97

SUMÁRIO

1	INTRODUÇÃO	15
2	REVISÃO BIBLIOGRÁFICA	18
2.1	LEITE	18
2.2	LACTOSE E LACTASE	18
2.3	DEFICIÊNCIAS DA LACTASE.....	22
2.4	FENÓTIPO DA PERSISTÊNCIA DA LACTASE	25
2.4.1	Frequência do fenótipo de persistência da lactase	26
2.5	BASES MOLECULARES DA PERSISTÊNCIA DA LACTASE.....	29
2.5.1	Genes <i>LCT</i> e <i>MCM6</i>	30
2.5.2	Distribuição dos polimorfismos associados à persistência da lactase.....	35
2.5.3	Haplótipos associados à persistência da lactase	40
2.6	EVOLUÇÃO DA PERSISTÊNCIA DA LACTASE.....	45
2.6.1	Surgimento do consumo adulto de leite	49
2.6.2	Surgimento e seleção das variantes associadas à persistência da lactase ..	50
2.7	PERSISTÊNCIA DA LACTASE NO BRASIL	55
2.8	COMPOSIÇÃO GENÉTICA DA POPULAÇÃO BRASILEIRA.....	57
2.8.1	População afro-brasileira do Paraná	62
2.8.2	Comunidades Quilombolas	65
2.8.2.1	Sertão do Valongo.....	67
3	OBJETIVOS	69
3.1	OBJETIVO GERAL	69
3.2	OBJETIVOS ESPECÍFICOS.....	69
4	MATERIAIS E MÉTODOS	70
4.1	AMOSTRAS.....	70
4.2	PCR GENE <i>MCM6</i>	71
4.3	SEQUENCIAMENTO.....	73
4.4	ANÁLISES BIOINFORMÁTICAS	75
5	RESULTADOS	77
5.1	FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS	77
5.2	FREQUÊNCIAS FENOTÍPICAS E HAPLÓTIPOS LP.....	79
5.3	HAPLÓTIPOS DA REGIÃO INTENSIFICADORA DO GENE <i>LCT</i> E DESEQUILÍBRIO DE LIGAÇÃO	80

5.4	DIVERSIDADE E DISTÂNCIA GENÉTICA.....	82
5.5	COMPARAÇÕES POPULACIONAIS	83
5.5.1	Afro-brasileiros de Curitiba e Região e de Sertão do Valongo	83
5.5.2	Populações brasileiras	83
5.5.3	Populações latino-americanas	85
5.5.4	Populações africanas	89
6	DISCUSSÃO	98
6.1	POLIMORFISMOS ASSOCIADOS AO FENÓTIPO LP EM POPULAÇÕES AFRO-BRASILEIRAS DO SUL DO BRASIL.....	98
6.2	DESEQUILÍBRIO DE LIGAÇÃO	103
6.3	DIVERSIDADE E DISTÂNCIA GENÉTICA ENTRE OS AFRO-BRASILEIROS ANALISADOS.....	104
6.4	COMPARAÇÕES POPULACIONAIS	105
6.4.1	Entre populações latino-americanas	105
6.4.2	Entre populações africanas	108
7	CONCLUSÕES	111
	REFERÊNCIAS	114
	APÊNDICE 1 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA EUROPA	140
	APÊNDICE 2 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA ÁFRICA	140
	APÊNDICE 3 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DO ORIENTE MÉDIO	142
	APÊNDICE 4 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA ÁSIA	143
	APÊNDICE 5 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA AMÉRICA	144
	APÊNDICE 6 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA OCEANIA	145
	APÊNDICE 7 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA EUROPA E DE ANCESTRALIDADE EUROPEIA	146
	APÊNDICE 8 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA ÁFRICA	147

APÊNDICE 9 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA ÁSIA E ORIENTE MÉDIO	148
APÊNDICE 10 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA AMÉRICA E OCEANIA	149
APÊNDICE 11 – FREQUÊNCIAS ALÉLICAS DO SNP -13907 C>G	150
APÊNDICE 12 – FREQUÊNCIAS ALÉLICAS DO SNP -13915 T>G.....	151
APÊNDICE 13 – FREQUÊNCIAS ALÉLICAS DO SNP -14009 T>G.....	153
APÊNDICE 14 – FREQUÊNCIAS ALÉLICAS DO SNP -14010 G>C....	154
APÊNDICE 15 – FREQUÊNCIAS ALÉLICAS DO SNP -14011 C>T.....	156
ANEXO 1 – PARÂMETROS DE DESEQUILÍBRIO DE LIGAÇÃO	157

1 INTRODUÇÃO

A lactase, ou β -galactosidase (E.C. 3.2.1.23), é encontrada exclusivamente na mucosa do intestino delgado, mais precisamente nas bordas em escova dos enterócitos (SWALLOW, 2003). A enzima hidrolisa a lactose (dissacarídeo) em glicose e galactose (monossacarídeos), absorvíveis pela parede intestinal (CAMPBELL; WAUD; MATTHEWS, 2009). Em seres humanos, a lactase é codificada exclusivamente pelo gene *LCT*, localizado no braço longo do cromossomo 2, na posição 21 (HARVEY et al., 1993).

A expressão da lactase é elevada em mamíferos recém natos, incluindo seres humanos, e diminui significativamente após o desmame (ocasião em que o leite deixa de fazer parte da dieta desses animais), conferindo-lhes o fenótipo da lactase não persistente (LNP) (FLATZ, 1987; SWALLOW, 2003). Contudo, em seres humanos, a atividade enzimática pode ser mantida mesmo na vida adulta – caracterizando o fenótipo de persistência da lactase (LP) (FLATZ, 1987).

O declínio da expressão e, conseqüentemente, da atividade enzimática da lactase ao longo do desenvolvimento dos indivíduos resulta (de maneira assintomática) na má digestão e absorção da lactose (MATTAR; MAZO, 2010), podendo progredir para um quadro sintomático denominado hipolactasia primária, condição que possui manifestações clínicas como náusea, vômito, flatulência, borboríngos, cólica abdominal, distensão abdominal, esteatorreia (presença de gordura nas fezes), constipação e diarreia após o consumo de leite e laticínios (FLATZ, 1987; SAHI, 1994a). Há ainda outras patologias relacionadas a modificações na expressão da lactase, como a hipolactasia congênita, condição na qual recém-nascidos apresentam baixa atividade da lactase ou até mesmo sua completa ausência (ASP et al., 1973; HOLZEL; SCHWARZ; SUTCLIFFE, 1959; ROBAYO-TORRES; NICHOLS, 2007), e a hipolactasia secundária, quadro no qual a diminuição da atividade enzimática da lactase decorre de danos ocasionados na mucosa intestinal (VILLAKO; MAAROOS, 1994). No presente estudo, serão investigados alelos associados à não persistência da lactase decorrente exclusivamente da hipolactasia primária.

A alta prevalência da persistência da lactase é encontrada em populações com histórico pastoralista e de ordenha, englobando populações ao Leste e Oeste da África, Oriente Médio e Noroeste da Europa (ITAN et al., 2010; MCCRACKEN, 1971;

RANCIARO et al., 2014). Tal correlação indica o caráter adaptativo desta condição (JONES et al., 2013; LEONARDI et al., 2012).

A ocorrência do fenótipo LP nas populações das localidades mencionadas está associada a polimorfismos presentes na região intensificadora do gene *LCT*. As primeiras variantes descritas nesta região foram os SNPs (polimorfismos de nucleotídeo único, do inglês *single nucleotide polymorphisms*) -13910 C>T (rs4988235) e -22018 G>A (rs182549), localizados respectivamente nos íntrons 13 e 9 do gene *MCM6*, encontrados em indivíduos lactase persistentes do Noroeste da Europa, especificamente em finlandeses (ENATTAH et al., 2002). Estudos posteriores identificaram, em populações da África Oriental e Oriente Médio, outros polimorfismos associados ao fenótipo em questão, incluindo os SNPs -14010 G>C (rs145946881); -13907 C>G (rs41525747) e -13915 T>G (rs41380347) (INGRAM et al., 2007; TISHKOFF et al., 2007).

A existência de diferentes mutações, em altas frequências, culminando num mesmo fenótipo em populações de localidades distintas evidencia um processo denominado de evolução convergente. Além disso, trata-se de um exemplo de coevolução gene-cultura pois o aumento da frequência é decorrente da seleção positiva produzida pelo consumo de leite (TISHKOFF et al., 2007).

As populações que compõem o Brasil possuem ancestralidade bastante diversa, decorrente de seu histórico marcado principalmente pelo povoamento inicial por ameríndios; posterior colonização e invasão por europeus a partir do século XVI; intenso tráfico de africanos da região subsaariana para trabalho escravo entre os séculos XVI e XIX; e imigração acentuada de europeus para substituição de mão de obra escrava e “branqueamento” nos séculos XIX e XX (CALÓGERAS, 1938; SANTOS, 2002). Consequentemente a ancestralidade genômica dos brasileiros – apesar de variar de acordo com a região – é bastante heterogênea, com predominância europeia seguida de africana e ameríndia (PENA et al., 2011). É importante salientar que mesmo indivíduos que se autodeclaram afro-brasileiros apresentam porcentagens significativas de ancestralidade europeia (KIMURA et al., 2013; MUNIZ et al., 2008; PENA et al., 2011).

Os africanos trazidos para o Brasil, para desempenharem trabalho escravo, entre os séculos XVI e XIX eram provenientes da África Ocidental e Ocidental-Central (GONÇALVES et al., 2008; SOUZA, 2008). Essas áreas possuíam uma vasta extensão geográfica e continham grande diversidade de povos, incluindo alguns de

origem étnica Bantu (HÜNEMEIER et al., 2007), de forma que haplótipos associados a essa ancestralidade já foram identificados em populações afro-brasileiras (SALZANO; BORTOLINI, 2002). As localidades ao Leste do continente africano têm populações que, atualmente, apresentam alta prevalência do fenótipo de persistência da lactase e elevada frequência do alelo -13910^*T – inserido nesses povos a partir do fluxo gênico com populações não africanas (BERSAGLIERI et al., 2004; ITAN et al., 2010; MATTAR; MAZO; CARRILHO, 2012; RANCIARO et al., 2014; TISHKOFF et al., 2009). Contudo, é importante destacar a alta frequência desse fenótipo relacionada aos alelos -14010^*C ; -13907^*G ; -13915^*G em populações de origem Bantu (TISHKOFF et al., 2007).

Em populações brasileiras, foi reportada a prevalência (43,0%) do polimorfismo $-13910 C>T$ em brancos ($n = 399$) e pardos ($n = 65$), enquanto que em pretos ($n = 50$) e descendentes de japoneses ($n = 53$) as frequências são de 20,0% e 0,0%, respectivamente (MATTAR et al., 2009). Além do SNP mencionado, já foram identificados em brasileiros com ancestralidade africana os polimorfismos $-14010 G>C$ e $-14011 C>T$ associados a persistência da lactase, bem como as variantes -13779^*C e -13937^*A – cuja associação ao fenótipo não foi confirmada (FRIEDRICH et al., 2012a; LIEBERT et al., 2016).

Em síntese, embora haja alguns estudos que trazem dados sobre polimorfismos relacionados ao fenótipo LP na população afro-brasileira (FRIEDRICH et al., 2012b; MATTAR et al., 2009, 2013), eles se referem majoritariamente a investigação do SNP $-13910 C>T$, de origem norte-europeia, e não incluem populações do Paraná e Santa Catarina nem populações de Comunidades Quilombolas. Além disso, os testes genéticos para diagnóstico da hipolactasia primária no país baseiam-se predominantemente no polimorfismo de origem europeia $-13910 C>T$, diminuindo a efetividade do exame devido à composição multiétnica da população brasileira. Portanto, a caracterização das frequências das variantes $-13907 C>G$, $-13915 T>G$, $-14009 T>G$ (rs869051967), $-14010 G>C$, $-14011 C>T$ (rs4988233) – também associadas à persistência da lactase – permitirá a otimização do diagnóstico genético da hipolactasia primária em brasileiros.

2 REVISÃO BIBLIOGRÁFICA

2.1 LEITE

A lactação é um importante atributo de mamíferos, e caracteriza-se pelo suprimento nutritivo da prole através da secreção de leite pelas glândulas mamárias (LEFÈVRE; SHARP; NICHOLAS, 2010). A importância desta característica para o grupo é tão significativa que ela foi utilizada para nomeá-lo, por Carl Linnaeus, no século XVIII (GREGORY, 1910).

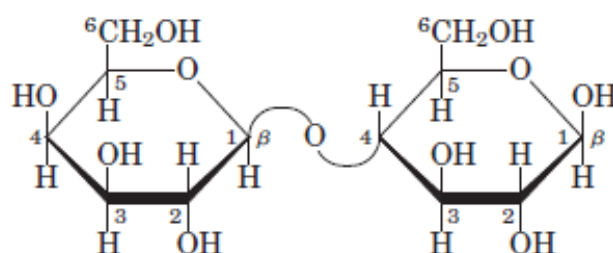
O leite assegura a nutrição e, conseqüentemente, a sobrevivência dos filhotes recém-nascidos, sendo assim os componentes e suas respectivas concentrações irão variar de acordo com as necessidades de cada espécie (POTTER; HOTCHKISS, 1995). Além disso, por tratar-se de um fluido corporal, sua composição também estará sujeita a alterações devido a diversos fatores como o estado de saúde do animal, idade, época do ano, dentre outros (FLATZ, 1987; JENSEN; BLANC; PATTON, 1995; PICCIANO, 2001; POTTER; HOTCHKISS, 1995). A composição do leite inclui 83 a 87,0% de água e 12,0 a 17,0% de componentes sólidos, os quais compreendem lipídeos, proteínas, sais minerais e carboidratos (porcentagens avaliadas em vacas, cabras, ovelhas, búfalas, zebus e humanos) (POTTER; HOTCHKISS, 1995). De todos os componentes sólidos mencionados, os mais abundantes são os carboidratos – sendo a lactose o principal deles. Sua concentração pode variar de 4,0 a 7,0% da composição total do leite, e está presente em todos os mamíferos, com exceção das famílias Otariidae e Odobenidae – pertencentes à subordem Pinnipedia – que correspondem às focas e leões marinhos e às morsas, respectivamente (POTTER; HOTCHKISS, 1995; REICH; ARNOULD, 2007). O leite humano é o que apresenta a maior concentração de lactose já encontrada, sendo equivalente a 6,98% do volume total, o que corresponde a 68 gramas por litro (g/L) em média (JENSEN; BLANC; PATTON, 1995; PICCIANO, 2001; POTTER; HOTCHKISS, 1995).

2.2 LACTOSE E LACTASE

A lactose (4-O- β -D-galactopyranosyl-D-gluco-pyranose) é um dissacarídeo presente exclusivamente no leite (FIGURA 1). É composta pelos monômeros glicose e galactose e sua síntese ocorre nas células das glândulas mamárias de mamíferos,

através da enzima lactose sintase (E.C 2.4.1.22) (KRETCHMER, 1972; KUHN; CARRICK; WILDE, 1980). Como já mencionado, este carboidrato possui um importante papel na nutrição de mamíferos e também atua na diminuição da carga osmótica (MUSTAPHA; HERTZLER; SAVAIANO, 1997).

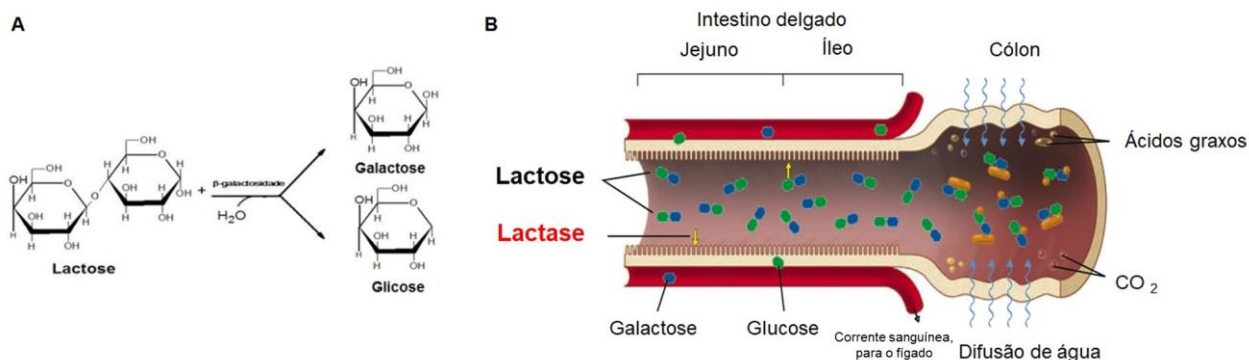
FIGURA 1 - ESTRUTURA DA LACTOSE



FONTE: NELSON; COX (2014)

A digestão da lactose demanda uma enzima específica denominada lactase-florizina hidrolase (E.C. 3.2.1.108) – comumente conhecida apenas como lactase. Trata-se de uma β -galactosidase localizada no jejuno (segunda porção do intestino delgado), na superfície apical das bordas em escova dos enterócitos (células do epitélio do intestino delgado) juntamente com outras hidrolases como a maltase-glicoamilase e a sucrase-isomaltase (KRETCHMER, 1972; NICHOLS et al., 2002; SWALLOW, 2003). A lactase promove a hidrólise da lactose em glicose e galactose, monossacarídeos absorvíveis pelos enterócitos e liberados na corrente sanguínea (LOMER; PARKES; SANDERSON, 2008). Parte da glicose liberada é aproveitada pelos enterócitos, enquanto o restante é convertido para a geração de energia. Da mesma forma, grande parte da galactose é absorvida no intestino delgado e é empregada na biossíntese de glicolipídeos e glicoproteínas, ou pode ser transportada para o fígado e armazenada na forma de glicogênio (KRETCHMER, 1972; LOMER; PARKES; SANDERSON, 2008; MATTAR; MAZO, 2010; WALKER; THOMAS, 2019). Estudos recentes sugerem que a galactose pode permanecer no cólon (localizado no intestino grosso) e servir como fonte de nutrição para bactérias lá presentes – possivelmente atuando como prebiótico (CEDERLUND et al., 2013; MACFARLANE; CUMMINGS, 1999; WALKER; THOMAS, 2019) (FIGURA 2).

FIGURA 2 - METABOLISMO DA LACTOSE



FONTE: MOREIRA et al. (2018); adaptado de comunicação pessoal de Alessia Ranciaro, 17 de março (2020).

LEGENDA: A = hidrólise enzimática da lactose catalisada pela lactase; B = metabolismo da lactose no intestino delgado.

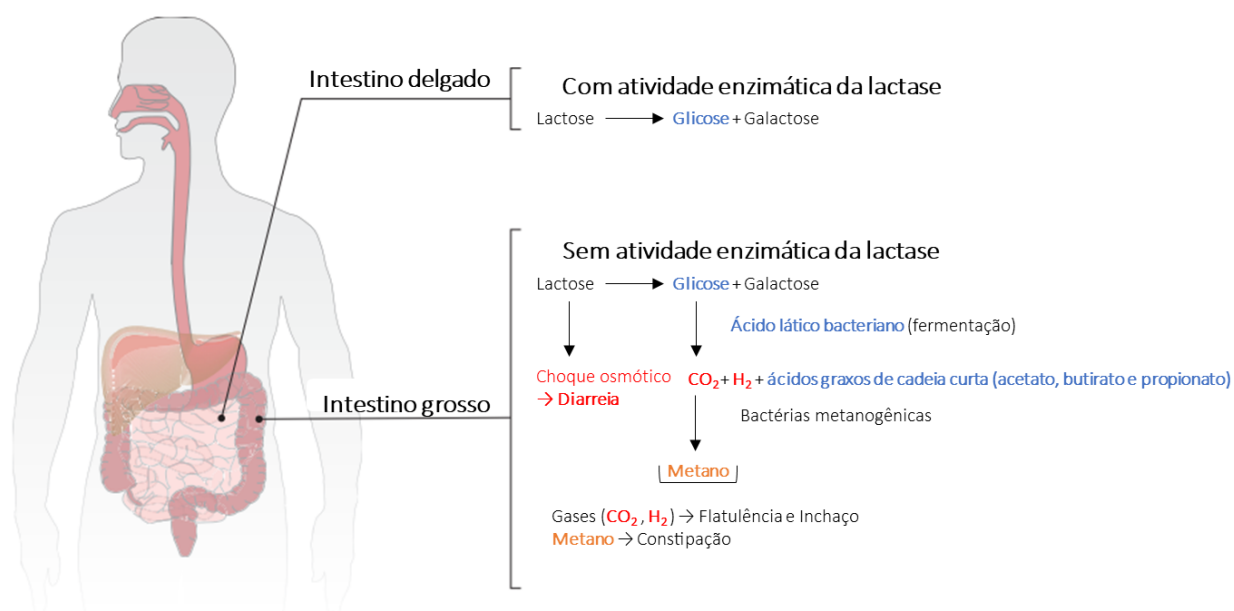
Por ser uma enzima empregada exclusivamente na digestão da lactose presente no leite e este, por sua vez, é disponível somente nas primeiras fases da vida do animal, a lactase tem sua atividade reduzida após o desmame, sendo tal redução mantida ao longo do desenvolvimento dos mamíferos (FLATZ, 1987). Apesar desse padrão geral ser o mesmo entre os animais, há algumas diferenças. Em mamíferos não humanos, a atividade enzimática só é detectada após o nascimento e decai no período pós-amamentação para níveis inferiores a 10,0% (MONTGOMERY et al., 2007). Em contrapartida, em humanos, a atividade enzimática (apesar de reduzida) já pode ser identificada na oitava semana de gestação, seus níveis aumentam até a 34ª semana, no terceiro trimestre, e alcançam o pico de atividade no momento do nascimento, enquanto o declínio da atividade ocorre na metade da infância – podendo variar (LOMER; PARKES; SANDERSON, 2008).

Lebenthal e colaboradores (1975) demonstraram que, em humanos, a atividade da lactase sofre poucas alterações até os três anos de idade e variações entre populações são imperceptíveis. A partir dos cinco anos a diferença no comportamento enzimático entre populações torna-se evidente, emergindo grupos com muita e pouca atividade enzimática. No Sri Lanka, por exemplo, foi observado que a redução da atividade enzimática ocorre entre os cinco e dez anos de idade e acentua-se a partir dos onze anos (THOMAS et al., 1990). Em populações da África do Sul, o declínio foi identificado em crianças indianas por volta dos oito anos de idade, enquanto que em crianças sul-africanas foi observado a partir dos três anos de idade (similar ao encontrado em crianças do Japão) (NOSE et al., 1979; WITTENBERG; MOOSA, 1990). Contrapondo-se a esses achados, Sahi e colaboradores (1983) identificaram

na Finlândia a redução ocorrendo em etapas mais avançadas do desenvolvimento, entre dez e vinte anos de idade. A atividade enzimática continua a declinar ao longo da vida, de modo que a má digestão da lactose é mais prevalente em idosos do que em adultos jovens (RAO et al., 1994; THOMAS et al., 1990).

A redução ou ausência da atividade da lactase impossibilita a digestão e absorção da lactose. O dissacarídeo, então, deixa o intestino delgado e segue para a porção proximal do cólon, onde, por meio da ação da microbiota intestinal, ocorre sua fermentação – produzindo ácidos graxos de cadeia curta (principalmente acetato, butirato e propionato) e gases (hidrogênio (H₂), gás carbônico (CO₂) e metano (CH₄)) como metabólitos (HE et al., 2008). A concentração em que cada um dos metabólitos é produzido e, conseqüentemente, os sintomas gerados variam de acordo com a composição da microbiota de cada indivíduo (ANGUITA-RUIZ; AGUILERA; GIL, 2020; LOMER; PARKES; SANDERSON, 2008). Os ácidos graxos podem ser removidos do cólon por diferentes vias: absorção ou utilização pelas células do cólon (colonócitos), eliminação pelas fezes ou incorporação na biomassa da microbiota. Os gases produzidos, por sua vez, são absorvidos do intestino pelo sangue e podem ser eliminados pelos pulmões ou através de flatos (HE et al., 2008). Este processo de fermentação pode, embora não necessariamente ocorra, gerar sintomas como dor abdominal e sensação de estufamento devido ao aumento do trânsito intestinal, (HE et al., 2008) e diarreia, devido ao gradiente osmótico estabelecido na parede intestinal (HE et al., 2008; INGRAM et al., 2009a) (FIGURA 3).

FIGURA 3 - DIGESTÃO DA LACTOSE NO TRATO INTESTINAL



FONTE: Adaptado de SÉGUREL e BON (2017).

2.3 DEFICIÊNCIAS DA LACTASE

Enfermidades associadas à má digestão da lactose em humanos começaram a ser estudadas no início do século XX, a partir da correlação com a diarreia (KRETCHMER, 1971; PORTER; DUNN, 1915). Na década de 1950, Holzel e colaboradores (1959) identificaram pela primeira vez, em crianças, a incapacidade de realizar a hidrólise da lactose, ocasionando sintomas como diarreia e desnutrição. Por incluir irmãos gêmeos com o mesmo fenótipo no estudo, os autores sugeriram um caráter hereditário para esta condição (HOLZEL; SCHWARZ; SUTCLIFFE, 1959). A ausência da atividade da lactase e a má digestão da lactose foram observadas também em adultos saudáveis (AURICCHIO et al., 1963; DAHLQVIST et al., 1964), havendo indícios de que seriam condições adquiridas na maioria dos casos, uma vez que os indivíduos não apresentavam tais sintomas quando crianças (DUNPHY et al., 1965; HAEMMERLI et al., 1965). Posteriormente, estudos demonstraram que a ocorrência desta condição seria mais frequente do que se supunha (CUATRECASAS; LOCKWOOD; CALDWELL, 1965; SIMOONS, 1969).

A alteração do comportamento enzimático da lactase ao longo da vida é consequência de modificações na expressão do gene *LCT*, responsável por codificá-la, sendo um processo programado no qual a diminuição da expressão gênica pode ser observada em crianças a partir dos dois anos de idade (WANG et al., 1998). Tal

condição caracteriza a hipolactasia primária (também chamada de hipolactasia do tipo adulto), traço recessivo autossômico que, frequentemente, possui sintomas como flatulência, dor abdominal e diarreia após o consumo de leite e laticínios – sendo esta a deficiência da lactase abordada no presente estudo (TROELSEN, 2005).

Além da hipolactasia primária, há outras duas patologias associadas a baixa expressão ou ausência da lactase e, por apresentarem etiologias distintas, é importante diferenciá-las. A hipolactasia secundária é um quadro clínico no qual a diminuição da atividade enzimática é adquirida através de eventos que danifiquem a mucosa intestinal de alguma forma, portanto o impacto na digestão da lactose irá depender da dimensão do dano – podendo ser completo ou parcial (VILLAKO; MAAROOS, 1994). Além disso, nos casos em que a integridade do tecido é reestabelecida, o quadro pode ser revertido – único caso no qual a má-digestão da lactose é reversível (MONTALTO et al., 2006). Dentre os eventos que proporcionam o desenvolvimento da hipolactasia secundária estão inclusas algumas enteroparasitoses (como giardíases e ascaridíases); gastroenterites agudas (bacterianas e virais); administração de alguns fármacos (como neomicina e colchicina) e patologias gastrointestinais como doença celíaca e doença de Crohn (CARRERA; NESHEIM; CROMPTON, 1984; KERLIN; WONG, 1987; VILLAKO; MAAROOS, 1994). A hipolactasia congênita (assim como a primária) é de origem genética, sendo um traço autossômico recessivo raro (identificado, na Finlândia, em 1 a cada 60.000 indivíduos), caracterizado pela baixa atividade da lactase ou até mesmo sua completa ausência em recém-nascidos – apesar da mucosa intestinal ser classificada como histologicamente saudável (ASP et al., 1973; NEVANLINNA, 1972; NORIO; NEVANLINNA; PERHEENTUPA, 1973; ROBAYO-TORRES; NICHOLS, 2007). Possui sintomas como diarreia líquida, desidratação severa, acidose e perda de peso após a ingestão de leite, havendo desaparecimento desses sinais clínicos quando a dieta é livre de lactose (ROBAYO-TORRES; NICHOLS, 2007; SAVILAHTI; LAUNIALA; KUITUNEN, 1983).

Deve-se destacar a diferença existente entre a má-digestão e a intolerância à lactose que, embora possam estar associadas, não possuem o mesmo significado (BAYLESS; BROWN; PAIGE, 2017; SAHI, 1978). A má-digestão da lactose, como o próprio nome sugere, refere-se à problemas no processo de hidrólise do dissacarídeo decorrente de alterações na atividade da lactase – o que não necessariamente resulta em manifestações clínicas (HE et al., 2008; LOMER; PARKES; SANDERSON, 2008).

Já a intolerância à lactose é a designação conferida aos sintomas gastrointestinais ocasionados pela digestão incompleta do carboidrato, que incluem dor, inchaço, flatulência, diarreia e borboríngos (HE et al., 2008; MATTHEWS et al., 2005; SAHI, 1994a). Desta forma, indivíduos assintomáticos podem ser classificados como má-digestores e tolerantes à lactose simultaneamente (HE et al., 2008).

Diversos métodos diagnósticos foram desenvolvidos e permitem identificar, direta ou indiretamente, a má-absorção e a intolerância à lactose (FLATZ, 1987). Um método direto invasivo e, portanto, pouco empregado, envolve a biópsia de uma amostra do jejuno aliada à adição de um composto que reage diretamente com a glicose liberada a partir da hidrólise da lactose (MATTAR; MAZO; CARRILHO, 2012).

Os métodos indiretos envolvem os chamados “testes de tolerância a lactose”, e consistem na ingestão de lactose pelo paciente seguida de uma avaliação dos efeitos deste consumo (que pode ser desempenhada por diferentes técnicas) (FLATZ, 1987; INGRAM et al., 2009a). Uma das técnicas mais utilizadas é a da curva glicêmica, na qual o sangue é coletado antes e após a ingestão da lactose (em intervalos pré-determinados) e observa-se a concentração de glicose sanguínea. Neste teste, indivíduos com alta capacidade de digestão da lactose apresentam níveis equivalentes ou superiores a 1,4 milimolar por litro (mmol/L), enquanto que naqueles com baixa capacidade de digestão os níveis não ultrapassam a 1,1 mmol/L (FLATZ, 1987; MATTAR; MAZO, 2010). Outra técnica emprega a administração de etanol e posterior avaliação da concentração de galactose e glicose no sangue. O etanol inibe a conversão de galactose a glicose no fígado, após a digestão da lactose, e, portanto, a detecção de galactose só ocorrerá em indivíduos que conseguem digerir o dissacarídeo (FLATZ, 1987; MATTAR; MAZO; CARRILHO, 2012).

O teste respiratório do hidrogênio expirado também é um método indireto e é considerado o mais recomendado para medir a digestão da lactose – com especificidade mensurada entre 85,0 e 100,0% e sensibilidade variando de 80,0 a 92,3% (LOMER; PARKES; SANDERSON, 2008; MARTON; XUE; SZILAGYI, 2012; MATTAR; MAZO, 2010; ROMAGNUOLO; SCHILLER; BAILEY, 2002). Este teste baseia-se na produção de hidrogênio (que é absorvido e eliminado pelos pulmões) através da fermentação da lactose não digerida no cólon. Como a fermentação é desempenhada pela microbiota intestinal, sua composição – variável entre os indivíduos – influencia o resultado do teste, podendo ocasionar conclusões errôneas (BABU et al., 2010; MYO-KHIN et al., 1999). Antes e após a ingestão de lactose, em

intervalos padronizados, a concentração de hidrogênio proveniente da expiração do paciente é examinada, sendo considerado intolerante quando ocorre um aumento (em relação aos valores pré-ingestão de lactose) do hidrogênio expirado em 20 partes por milhão (ppm) (MATTAR; MAZO, 2010; ROMAGNUOLO; SCHILLER; BAILEY, 2002).

A genotipagem através da identificação de polimorfismos associados à expressão do gene da lactase é o método diagnóstico para a má-digestão da lactose mais recente e menos invasivo (MATTAR; MAZO; CARRILHO, 2012). Apesar de sua acurácia já ser reconhecida em alguns locais, sua efetividade depende da identificação e compreensão da distribuição de variantes polimórficas associadas à persistência da lactase em diferentes populações (MARTON; XUE; SZILAGYI, 2012; MATTAR et al., 2010; MATTAR; MAZO; CARRILHO, 2012; TISHKOFF et al., 2007; TROELSEN, 2005).

2.4 FENÓTIPO DA PERSISTÊNCIA DA LACTASE

Apesar do estabelecimento no início do século XX de que em mamíferos, no geral, os recém-nascidos apresentam o fenótipo de persistência da lactase (LP) e os adultos o fenótipo de não persistência da lactase (LNP) (MENDEL; MITCHELL, 1907), a identificação de humanos LNP ocorreu somente em 1959 (HOLZEL; SCHWARZ; SUTCLIFFE, 1959). Desta forma, a concepção científica de que se tratava de um traço “anormal” permaneceu até que, a partir da metade da década de 1960, estudos começaram a observar sua ampla ocorrência em diferentes populações (FLATZ, 1987; SIMOONS, 1969). Cuatrecasas e colaboradores (1965) constataram a frequência de 55,0% para o fenótipo LNP em populações dos Estados Unidos, sendo a elevada frequência identificada principalmente em afro-americanos (com prevalência de 95,0%, em detrimento da frequência de apenas 10,0% em euro-americanos) (BAYLESS; ROSENSWEIG, 1966). Posteriormente, a alta frequência deste fenótipo foi também observada em populações de origem Bantu e não-Bantu da Uganda (COOK; KAJUBI, 1966); em asiáticos do Japão, Coreia e China (CHUNG; MCGILL, 1968); em populações da Tailândia (KEUSCH et al., 1969); em ameríndios da Colômbia (ALZATE; GONZÁLEZ; GUZMÁN, 1969) e em populações do Peru (FIGUEROA et al., 1971).

Através dos estudos populacionais, filogenia de haplótipos e de modelos matemáticos para explicar as frequências mundiais dos fenótipos, foi possível assumir

que, na verdade, o fenótipo LNP (associado à hipolactasia primária) seria o caráter ancestral – assim como o padrão existente nos demais mamíferos – e o fenótipo LP em adultos, o caráter derivado (MCBEAN; MILLER, 1998; SWALLOW, 2003).

2.4.1 Frequências do fenótipo de persistência da lactase

A partir da década de 1960, a identificação dos fenótipos LP e LNP começou a ser realizada em diferentes populações, demonstrando frequências discrepantes (SIMOONS, 1969). De acordo com a hipótese cultural-histórica, a mais aceita dentre as hipóteses que explicam a ocorrência dos fenótipos, essa variação é atribuída às diferenças em relação ao costume de consumir leite e laticínios entre populações (HOLDEN; MACE, 1997; SIMOONS, 1970). Deste modo, devido a seleção natural, o fenótipo LP é observado com maior frequência em populações com tradições pastoralistas e de ordenha, enquanto que o fenótipo LNP é prevalente em populações cujos hábitos não incluem (tradicionalmente) o consumo de leite (MCCRACKEN, 1971; SIMOONS, 1970). De maneira geral, estima-se que um terço da população mundial seja persistente para a lactase, em detrimento de dois terços que não o seriam (INGRAM et al., 2009a).

Na Europa, as frequências mais elevadas da persistência da lactase são encontradas na região Noroeste do continente – em especial na Escandinávia – com frequências variando de 90,0 a 100,0% (FLATZ, 1987; SÉGUREL; BON, 2017). Na Europa Central, a frequência da LP também é elevada, embora em menor proporção e com maiores variações: na Itália, por exemplo, já foram observadas frequências de 82,0% e 46,0% nas regiões Centro e Sul do país, respectivamente (BURGIO et al., 1984; CAVALLI-SFORZA et al., 1987; ITAN et al., 2010). Países como Áustria, República Tcheca, Alemanha e Suíça – também na Europa Central – possuem populações com frequências superiores a 75,0% (BEYERLEIN et al., 2008; FLATZ et al., 1982; LEICHTER, 1972; ROSENKRANZ et al., 1982; SAHI, 1994b; SÉGUREL; BON, 2017). Em síntese, a distribuição do fenótipo LP neste continente diminui gradualmente nos sentidos leste - oeste / norte - sul (FLATZ, 1987; SWALLOW, 2003) (APÊNDICE 1) .

Na África, as frequências da persistência da lactase não estão relacionadas unicamente com as regiões em que as populações habitam, como Bayoumi e colaboradores (1981) evidenciaram num estudo na República do Sudão, em que

povoados vizinhos – com hábitos pastoralistas e não pastoralistas – apresentavam frequências variando de 38,0 a 76,0%. O fenótipo LNP é predominante na região Sul do continente africano, apesar de estudos demonstrarem (através de inferência) a também ocorrência do fenótipo LP (BRETON et al., 2014; MACHOLDT et al., 2015; RANCIARO et al., 2014; TORNIAINEN et al., 2009). Na porção ocidental do continente, Níger e Senegal são os países em que foram identificadas as mais altas prevalências do fenótipo LP (com frequências iguais a 87,3% na população Tuareg do Níger e 71,0% numa população miscigenada do Senegal) (ARNOLD et al., 1980; FLATZ; SCHILDGE; SEKOU, 1986; SAHI, 1994b). Na África Oriental, as mais altas prevalências foram observadas no Quênia (88,0 e 89,0% nas populações Maasai e Samburu, respectivamente) (TISHKOFF et al., 2007), Ruanda (93,0% na população Tutsi) (COX; ELLIOTT, 1974), na Tanzânia (85,7% na população Iraqw) (TISHKOFF et al., 2007) e Sudão (83,2% na população pastoril Beja) (BAYOUMI et al., 1982) (APÊNDICE 2). No Norte da África e Oriente Médio, altas prevalências foram identificadas em Beduínos da Arábia Saudita (81,0%) e Jordânia (76,0%) (DISSANAYAKE et al., 1990; HIJAZI et al., 1983), bem como em populações da região de Sinai, no Egito (65,7%) (HUSSEIN; EZZILARAB, 1994). É importante destacar que todas as populações com elevadas frequências do fenótipo LP mencionadas apresentam hábitos pastoralistas (APÊNDICE 3) .

Na Ásia (com exceção da região do Oriente Médio), as frequências do fenótipo da persistência da lactase são inferiores às encontradas na Europa e África – apesar da existência de populações com tradições pastoralistas neste continente (SÉGUREL; BON, 2017). A elevada frequência deste fenótipo foi identificada em populações do Norte da Índia (72,6%) – contrapondo-se às baixas frequências encontradas no Sul do país (33,4%) – (TANDON; SINGH, 1981) e nas populações Sindhi, Punjabi, Pathan, Mohajir e Baloochi, do Sul do Paquistão (com frequências variando de 80,0 a 100,0%) (RAB; BASEER, 1976). Um estudo posterior observou frequências consideravelmente menores em populações de grupos similares do Norte do Paquistão (variando de 30,0 a 42%) (AHMAD; FLATZ, 1984; SÉGUREL; BON, 2017). A discrepância, de acordo com Flatz (1987), seria consequência das metodologias empregadas. O fenótipo de não persistência da lactase é predominante nas demais localidades do continente, e as maiores frequências do traço LP identificadas são provenientes das populações de Udmúrtia e Komi-Izhems, na Rússia (60,0% e 38,0%, respectivamente) (ITAN et al., 2010; KOZLOV, 1998); no Sri Lanka (27,5%)

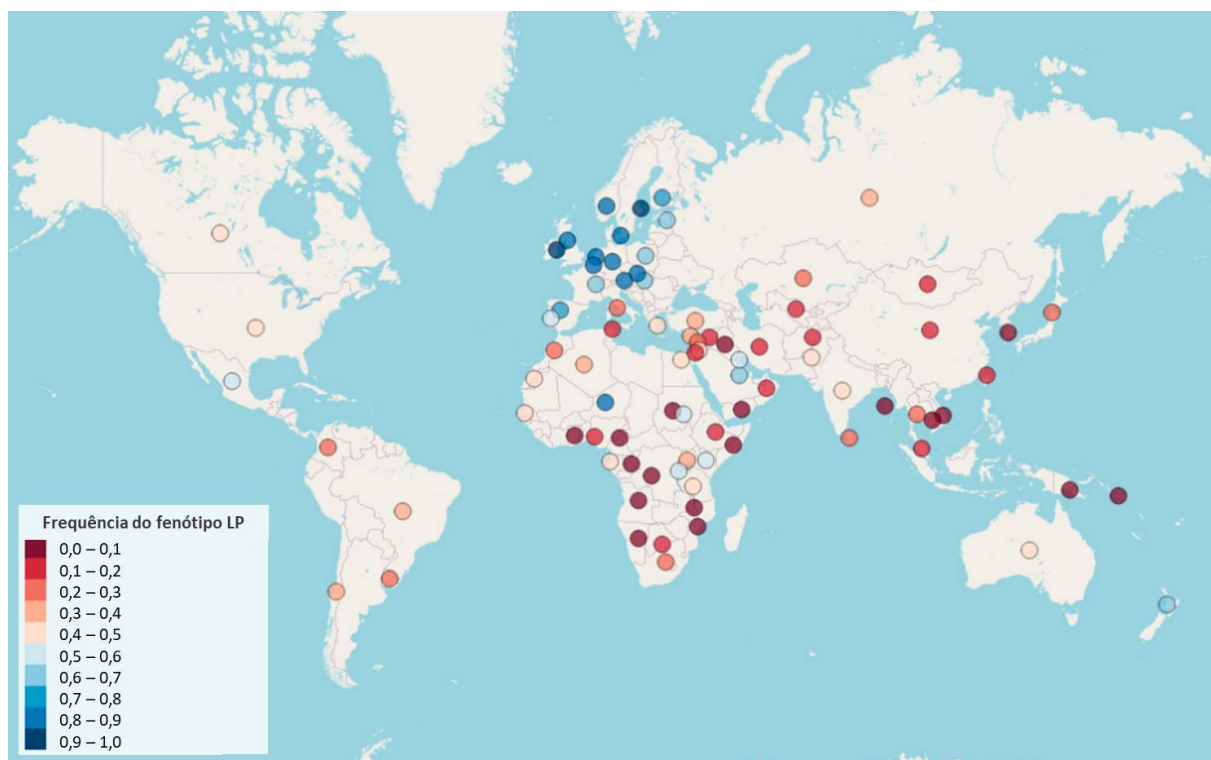
(SENEWIRATNE; THAMBIPIILLAI; PERERA, 1977) e nas populações Kazakh e Tajiko-Uzbek, no Uzbequistão (frequências de 21,0% e 11,0%, respectivamente) (HEYER et al., 2011) (APÊNDICE 4).

Na América, as populações tradicionais ameríndias possuem baixa frequência do fenótipo da persistência da lactase – provavelmente decorrente da origem asiática desses povos, através da migração via estreito de Bering (FLATZ, 1987; SALZANO; BORTOLINI, 2002). Em populações esquimós do Alaska (Estados Unidos), foi encontrada a frequência de 16,6% para o fenótipo LP (DUNCAN; SCOTT, 1972). Nesse mesmo país, estudos com ameríndios nativos de Oklahoma constataram a frequência de 17,5% para o fenótipo (CASKEY et al., 1977). No México, a ausência deste fenótipo e a frequência de apenas 2,3% foram observadas nas populações Huicholes e Nahuas, respectivamente, enquanto que em populações miscigenadas as frequências eram maiores – variando de 31,4% a 56,3% (OJEDA-GRANADOS et al., 2016). A baixa prevalência também é observada em populações tradicionais da América do Sul, sendo observadas as frequências de 12,5% em ameríndios do Equador (PAZ-Y-MIÑO et al., 2016) e de 10,0% em populações Mapuche, do Chile (FERNÁNDEZ et al., 2016) (APÊNDICE 5).

Na Oceania, assim como nas Américas, ocorre a predominância da não persistência da lactase nas populações tradicionais (FLATZ, 1987). Um estudo com populações aborígenes da Austrália observou a frequência de 16,0% para o fenótipo LP (BRAND et al., 1983). Na população de Sepik, de Papua-Nova Guiné, foi identificada a frequência de 22,0% (ARNHOLD; PERMAN; NURSE, 1981), enquanto que nas Ilhas Salomão a frequência não ultrapassa 1,0% (BERSAGLIERI et al., 2004; ITAN et al., 2010; STORHAUG; FOSSE; FADNES, 2017) (APÊNDICE 6).

É importante destacar, contudo, que países com histórico de colonização europeia (comumente associado à dizimação de populações nativas) e, conseqüentemente, populações com altas proporções de componentes de ancestralidade europeia, possuem elevadas frequências do fenótipo LP (SÉGUREL; BON, 2017; STORHAUG; FOSSE; FADNES, 2017). Isto pode ser observado na população geral da Nova Zelândia (frequência de aproximadamente 90,0%), Austrália (frequências variando de 56,0% a 90,0%) e Canadá (frequência de 41,0%) (CAMPBELL; WAUD; MATTHEWS, 2009; STORHAUG; FOSSE; FADNES, 2017).

FIGURA 4 - DISTRIBUIÇÃO MUNDIAL DO FENÓTIPO LP



FONTE: Adaptado de ANGUITA-RUIZ; AGUILERA; GIL (2020).

2.5 BASES MOLECULARES DA PERSISTÊNCIA DA LACTASE

A discrepância na distribuição dos fenótipos em diferentes povos e partes do mundo tornou-se conspícua com o avanço dos estudos populacionais, de modo que hipóteses de possíveis explicações para este fenômeno começaram a ser formuladas. Alguns autores passaram a sugerir uma origem genética para os fenótipos LP e LNP (BAYLESS; ROSENSWEIG, 1966; COOK, 1967; FERGUSON; MAXWELL, 1967; HAEMMERLI et al., 1965), embora ainda se cogitasse que essas características fossem decorrentes de adaptações ontogenéticas (BOLIN; DAVIS, 1970; CUATRECASAS; LOCKWOOD; CALDWELL, 1965; YOSHIDA et al., 1975). Simoons (1969) e Harrison (1975), através da comparação entre as diferentes hipóteses para a ocorrência dos fenótipos, atribuíram a origem genética como sendo a etiologia mais plausível.

Flatz e Saengudom (1969) investigaram a condição de duas famílias da Tailândia que possuíam integrantes tolerantes à lactose – embora o acesso a leite na região fosse bastante restrito e seu consumo não pertencer aos hábitos familiares. Nessas famílias, os autores identificaram que a distribuição da tolerância à lactose

nos indivíduos era condizente com a existência de um único gene autossômico como causador dos fenótipos LP e LNP – sugerido também por Ferguson e Maxwell (1967) num estudo semelhante que incluía integrantes de apenas uma família. A mesma conclusão foi obtida, nos anos seguintes, em outra pesquisa (com número amostral superior) incluindo famílias da Finlândia (SAHI et al., 1973). Estudos com gêmeos corroboraram esses achados, de modo que todos os gêmeos monozigóticos demonstraram concordância em relação ao fenótipo LP (MÉTNEKI et al., 1984).

Através de ensaio enzimático, Ho, Povey e Swallow (1982) observaram que as diferenças nos níveis de atividade da lactase em adultos apresentavam uma distribuição trimodal e estariam relacionadas a três genótipos: homocigoto persistente, heterocigoto e homocigoto não-persistente, apresentando alta, intermediária e baixa atividade enzimática, respectivamente. Em bebês, a distribuição observada segue um padrão unimodal, tornando-se trimodal ao longo do desenvolvimento (SWALLOW, 2003).

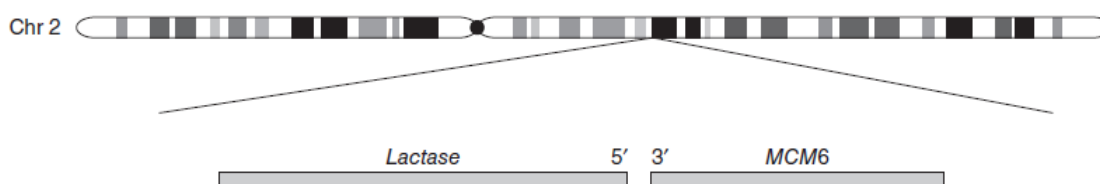
Apesar da hipótese da herança recessiva do fenótipo da não persistência da lactase ser compatível com as proporções identificadas em estudos populacionais, a suposição não explicava mudanças observadas no comportamento da lactase ao longo da vida – uma vez que os alelos deveriam controlar a produção da enzima da mesma maneira durante todo o desenvolvimento dos indivíduos (HARRISON, 1975). Sendo assim, Gilat e colaboradores (1973) sugeriram que alterações em genes regulatórios poderiam estar associadas à essa mudança de fenótipo.

2.5.1 Genes *LCT* e *MCM6*

O gene *LCT*, único responsável por codificar a enzima lactase em humanos, localiza-se no braço longo do cromossomo 2, região 2, banda 1, sub banda 3, posição 21 (HARVEY et al., 1993; KRUSE et al., 1988). Possui aproximadamente 55 kb (mil pares de bases) e 17 éxons (BOLL; WAGNER; MANTEI, 1991; SWALLOW, 2003). Pesquisas comparando indivíduos persistentes e não persistentes para a lactase indicaram que a sequência codificante deste gene e, conseqüentemente, seu transcrito são iguais, de modo que a enzima produzida é a mesma independentemente do fenótipo (BOLL; WAGNER; MANTEI, 1991). Além disso, evidências suportam a não ocorrência de edição de RNA e *splicing* alternativo no gene *LCT* (BOLL; WAGNER; MANTEI, 1991).

Wang e colaboradores (1995) evidenciaram que elementos *cis* atuam modificando a expressão do RNA mensageiro (mRNA) transcrito do gene *LCT*, aumentando ou diminuindo-a. Como os fatores *cis* incluem regiões promotoras e intensificadoras, alterações genéticas responsáveis pelos diferentes fenótipos da lactase poderiam ocorrer tanto no gene em questão quanto em localidades adjacentes à ele (LIN; BARBOSA, 2002; WANG et al., 1995). Aliado a isto, foi identificada uma região com cerca de 70 kb em alto desequilíbrio de ligação (LD, do inglês *linkage disequilibrium*) (HARVEY et al., 1995), que inclui todo o gene *LCT* e também o gene *MCM6* – envolvido no ciclo celular e altamente expresso em tecidos contendo abundância de células proliferativas, incluindo o intestino (HARVEY et al., 1996; HOLLOX et al., 2001) (FIGURA 5).

FIGURA 5 - LOCALIZAÇÃO DOS GENES *LCT* E *MCM6* NO CROMOSSOMO 2

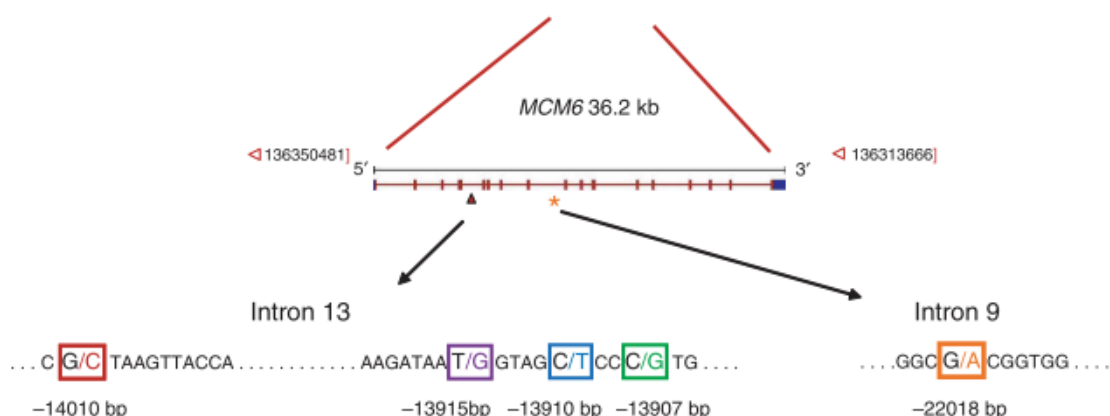


FONTE: MCINTOSH; SCHEINFELDT (2012).

Em 2002, foram identificados os primeiros polimorfismos associados aos fenótipos de persistência e não persistência da lactase em famílias da Finlândia (ENATTAH et al., 2002). Localizados nos íntrons 9 e 13 do gene *MCM6* (a 22 kb e 13,9 kb a montante do gene *LCT*, respectivamente), os SNPs -22018 (rs182549) e -13910 (rs4988235) possuem substituições nucleotídicas $G>A$ e $C>T$, nessa ordem, e estão em alto desequilíbrio de ligação na população europeia. Além disso, suas frequências alélicas mostraram-se compatíveis com as frequências fenotípicas da LP nas populações finlandesa, sul-coreana, alemã e italiana (correlação de 97,0% para a variante -22018*G e de 100,0% para a variante -13010*T) – indicando sua ligação com o fenótipo (ENATTAH et al., 2002). Corroborando as correlações mencionadas, Kuokkanen e colaboradores (2003) demonstraram que indivíduos com as variantes -13910*T e -22018*A apresentavam maior expressão de RNA mensageiro da lactase, bem como uma maior atividade enzimática (se comparados às amostras com os alelos -13910*C e -22018*G). Desta forma, indivíduos homocigotos *TT* e *AA* produzem maiores quantidades de enzima, heterocigotos produzem quantidades intermediárias

e homozigotos *CC* e *GG* as menores quantidades – concordando com o proposto previamente por Ho, Povey e Swallow (1982).

Apesar do polimorfismo *-13910 C>T* ser preditivo da persistência da lactase em populações europeias, sua ocorrência raramente é observada em populações que sabidamente apresentam o fenótipo LP no continente africano – sobretudo na região subsaariana – e Oriente Médio (MULCARE et al., 2004; MYLES et al., 2005). Isto culminou na necessidade de identificar outros alelos com a mesma capacidade preditiva. Através de novos sequenciamentos nas regiões a montante do gene *LCT*, localizadas no gene *MCM6*, em indivíduos das localidades mencionadas – incluindo populações de Camarões, Etiópia, Sudão, Tanzânia, Quênia, Jordânia, Arábia Saudita, Israel e Palestina – foram identificados novos SNPs (ainda não descritos até aquele momento) que demonstravam alguma relação com os fenótipos do metabolismo da lactose (INGRAM et al., 2007; TISHKOFF et al., 2007). As variantes referidas são: *-13907 C>G* (rs41525747), *-13913 T>C* (rs41456145), *-13915 T>G* (rs41380347), *-14010 G>C* (rs145946881) (FIGURA 6). Num estudo posterior, incluindo apenas populações da etnia Somali da Etiópia, foram constatadas outras posições polimórficas, incluindo os SNPs *-13730 T>G* (rs4954492), *-13779 G>C* (rs527991977), *-13806 A>G* (rs869122273) e *-14009 T>G*, dos quais apenas o último demonstrou relação com os fenótipos (INGRAM et al., 2009b).

FIGURA 6 - POLIMORFISMOS DO GENE *MCM6* ASSOCIADOS À PERSISTÊNCIA DA LACTASE

FONTE: TISHKOFF et al. (2007).

A região que inclui os *loci* -13910 e -22018 foi investigada, através de transfecção de gene repórter em linhagens de células intestinais Caco-2 avaliadas em teste de luciferase, quanto ao seu papel na determinação do fenótipo LP (TROEISEN et al., 2003). Foi identificado que a expressão do gene repórter é potencializada em quatro vezes na presença do alelo -13910*T, quando comparado ao alelo -13910*C (TROEISEN et al., 2003). Essa acentuação ocasionada pelo polimorfismo -13910 C>T é decorrente do aumento do reconhecimento (ocasionado pelo alelo T) do fator de transcrição nuclear Oct-1 (do inglês *Octamer-binding Transcription Factor 1*) pelo seu sítio de ligação, fazendo com que a atividade do promotor do gene *LCT* seja estimulada. Na ausência deste alelo, não ocorre intensificação (LEWINSKY et al., 2005). Os autores sugerem que o Oct-1 induz mudanças na cromatina das regiões próximas ao *LCT*, podendo recrutar cofatores que promovem o aumento da atividade gênica, no entanto isso só ocorre quando o fator de transcrição HNF1 α também é expressado (LEWINSKY et al., 2005). Outros fatores de transcrição também possuem sítios de ligação nessa mesma região intensificadora, dentre os quais estão inclusos o GATA-6 (sítio de ligação localizado na região entre as posições -13909 e -13934), HNF4 α e Fox/HNF3 α (entre as posições -13857 e -13817) e Cdx-2 (entre as posições -14022 e -14032) (ENATTAH et al., 2008; LEWINSKY et al., 2005). Os alelos da posição -22018 não pertencem a nenhum sítio de ligação, contudo especula-se que interfiram a atividade intensificadora dos polimorfismos da posição -13910 (TROEISEN et al., 2003). Em contrapartida, as posições dos polimorfismos originários da África e Oriente Médio (-13907, -13913, -13915, -14009, -14010)

localizam-se a poucos nucleotídeos ou nas próprias regiões que constituem os sítios de ligação mencionados – sugerindo efeitos similares aos da variante *-13910*T* na transcrição do gene *LCT* (ENATTAH et al., 2008; INGRAM et al., 2007) (FIGURA 7).

FIGURA 7 - LOCALIZAÇÃO DO SÍTIO DE LIGAÇÃO DO FATOR DE TRANSCRIÇÃO Oct-1

ANCESTRAL	TGGCAATACAGATAAGATAATGTAGCCCTGGCCTCAAAGGAACTCTCC
<i>-13915*G</i>	TGGCAATACAGATAAGATAA <u>G</u> GTAGCCCTGGCCTCAAAGGAACTCTCC
<i>-13913*C</i>	TGGCAATACAGATAAGATAATG <u>C</u> AGCCCTGGCCTCAAAGGAACTCTCC
<i>-13910*T</i>	TGGCAATACAGATAAGATAATGTAG <u>T</u> CCCTGGCCTCAAAGGAACTCTCC
<i>-13907*G</i>	TGGCAATACAGATAAGATAATGTAGCC <u>G</u> TGGCCTCAAAGGAACTCTCC

FONTE: INGRAM et al. (2007).

LEGENDA: Comparação entre as sequências ancestral e das variantes do íntron 13 do gene *MCM6*. A região que compreende o sítio de ligação do fator de transcrição Oct-1 está sombreada. As variantes associadas à LP estão mostradas em negrito e sublinhadas nas sequências.

Enattah e colaboradores (2008) demonstraram que a variante *-13907*G* possui uma forte ligação específica com o Oct-1 (assim como já descrito para a variante *-13910*T*), apesar de estar localizada a poucos nucleotídeos de seu sítio de ligação. O estudo também identificou a variante *-3712*C* em indivíduos diagnosticados com o fenótipo LP. Este *locus* inclui a região do sítio de ligação do HNF1 α , porém não foi possível diferenciar as fortes ligações que ambos os alelos (*T* e *C*, ancestral e derivado, respectivamente) exerciam sobre a proteína reguladora. Experimentalmente foi demonstrado que os SNPs *-13915 T>G* e *-3712 T>C*, conjuntamente, contribuíam para uma ligação mais intensa com o HNF1 α e Oct-1 o que, conseqüentemente, culminava numa maior ativação do promotor do gene *LCT* (quando comparada com a intensidade da ativação do promotor na presença das variantes *-13915*T* e *-3712*T*). Todavia, quando avaliado individualmente, o SNP *-13915 T>G* demonstrou fraca ligação com o fator de transcrição Oct-1, assim como o alelo *-13913*C* (ENATTAH et al., 2008). Resultados diferentes foram obtidos numa pesquisa posterior, na qual a variante *-13915*G* demonstrou maior interação com o Oct-1 do que o alelo ancestral *-13915*T* em ensaios *in vitro*, assegurando sua ação intensificadora no promotor do *LCT* (OLDS; AHN; SIBLEY, 2011). Jones e colaboradores (2013), por sua vez, identificaram que as variantes *-13913*G* e *-13730*G* presentes em populações

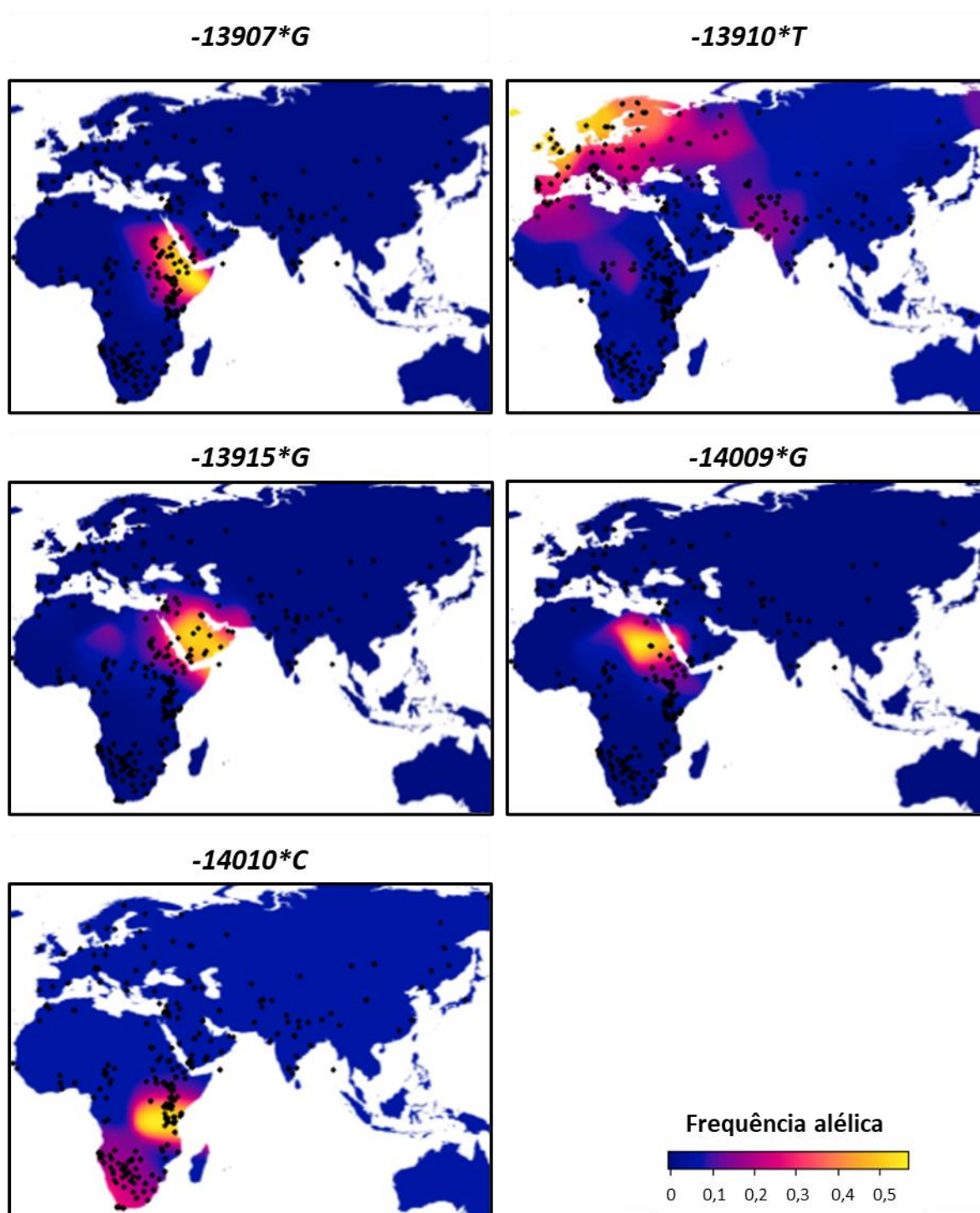
etíopes eram mais frequentes em indivíduos que não digeriam a lactose – sugerindo que tais alelos não estariam associados ao fenótipo LP.

A deleção experimental do nucleotídeo na posição -14010 demonstrou reduzir a atividade da região intensificadora do gene *LCT* (JENSEN et al., 2011). Além disso, em ensaios de transfecção, foi constatado que a variante -14010*C apresenta maior afinidade pelo fator de transcrição Oct-1 quando comparada ao seu alelo ancestral -14010*G (JENSEN et al., 2011). Os autores sugerem que outros polimorfismos localizados próximos ao *locus* em questão podem também influenciar a ligação do fator de transcrição – incluindo as variantes -14009*G e -14011*T (JENSEN et al., 2011). Tal hipótese foi corroborada por Liebert e colaboradores (2016), que demonstraram que os três SNPs (-14009 T>G, -14010 G>C e -14011 C>T) possuem efeitos similares na intensificação do promotor do gene *LCT*, porém através de mecanismos moleculares diferentes. As variantes -14010*C e -14011*T promovem uma ligação com o Oct-1 maior do que a produzida por seus respectivos alelos ancestrais, -14010*G e -14011*C (LIEBERT et al., 2016). Da mesma forma, a variante -14009*G proporciona uma maior ligação a outro fator de transcrição (possivelmente pertencente à família Ets, cujos membros são conhecidos por interagirem com outras proteínas a fim de regular processos de transcrição), quando comparada a seu alelo ancestral -14009*T (LIEBERT et al., 2016). Nesse mesmo estudo de 2016, não foram obtidas informações conclusivas acerca dos papéis dos polimorfismos -13779 G>C (rs527991977) e -14028 T>C (rs759157971) na transcrição do gene *LCT*, embora tenham sido demonstradas algumas alterações referentes à ligação diferencial com fatores de transcrição (LIEBERT et al., 2016).

2.5.2 Distribuição dos polimorfismos associados à persistência da lactase

Por apresentarem origens distintas, os polimorfismos associados à persistência da lactase estão distribuídos de maneira desigual entre as populações do mundo – mesmo entre as que possuem tradições pastoralistas (ITAN et al., 2010) (FIGURA 8). Vale ressaltar que o primeiro SNP descrito (-13910 C>T) possui a maior quantidade de estudos acerca de sua distribuição – incluindo a maior quantidade de localidades geográficas. As demais variantes não possuem quantias tão amplas de trabalhos referentes às suas frequências, de forma que a escassez de informações certamente influencia a compreensão de suas distribuições entre as populações.

FIGURA 8 - DISTRIBUIÇÃO GEOGRÁFICA DAS FREQUÊNCIAS DOS SNPS ASSOCIADOS À PERSISTÊNCIA DA LACTASE NA ÁFRICA E EURÁSIA



FONTE: Adaptado de LIEBERT et al. (2017).

NOTA: Os pontos pretos no mapa indicam as localidades com populações amostradas.

Como já mencionado, o SNP *-13910 C>T* foi identificado primeiramente em elevadas frequências em populações da Finlândia (ENATTAH et al., 2002). Na Europa, de maneira geral, as maiores frequências do polimorfismo são encontradas

principalmente nas porções Norte (populações da Suécia e Finlândia com frequências variando de 58,0% a 81,5%) (BERSAGLIERI et al., 2004; ENATTAH et al., 2008) e Noroeste (frequências variando de 68,8% a 81,9% na Escócia; frequência de 69,8% e 75,1% no Sudeste e Norte da Inglaterra, respectivamente) (BERSAGLIERI et al., 2004; DAVEY SMITH et al., 2009). No restante do continente, a prevalência do SNP -13910 C>T foi descrita também na República Tcheca (76,0%) (HUBÁČEK et al., 2017); em povos Bascos da França (66,0 a 67,0%) (BERSAGLIERI et al., 2004; ENATTAH et al., 2008); na Hungria (62,2%) (NAGY et al., 2009); na Áustria (58,0%) (BODLAJ et al., 2006; KRÜTTLI et al., 2014); na Alemanha (51,7%) (BÜNING et al., 2003) e Estônia (51,4%) (LEMBER et al., 2006) – todos pertencentes à localidades da Europa Central e Oriental. Países de outros continentes cujo fluxo migratório e/ou processos de colonização recentes foram desempenhados por europeus também demonstram elevadas frequências dessa variante (ao contrário dos povos nativos dessas localidades) como, por exemplo, populações da Nova Zelândia (frequências de 69,1% a 72,0%) (NOLAN et al., 2010; UPTON; GEORGE, 2010), populações euro-americanas dos Estados Unidos (frequência variando de 59,8% a 77,2%) (BERSAGLIERI et al., 2004; CHIN et al., 2019) e populações de ancestralidade europeia do Canadá (frequência de 48,1%) (ALHARBI; EL-SOHEMY, 2017) (APÊNDICE 7).

A rara ocorrência da variante -13910 C>T foi constatada em países africanos da região subsaariana – estando presente (em baixas frequências) nas populações das etnias Hausa (13,9%), Mambila (0,4%) e de outras etnias de Camarões (4,3%) e do Senegal (2,6%) (MULCARE et al., 2004). Posteriormente, foi observada a maior prevalência deste SNP na África em populações da etnia Fulani do Sudão e de Camarões, com frequências de 48,0% e 39,0%, respectivamente (ENATTAH et al., 2008; INGRAM et al., 2007). No Norte da África, este SNP apresenta maiores frequências do que no restante do continente, sendo identificado em populações Mozabite (Argélia) e Saharawi (Saara Ocidental) com frequências de 21,7 a 22,7% e 23,0 a 26,0%, respectivamente (BERSAGLIERI et al., 2004; ENATTAH et al., 2007, 2008; RANCIARO et al., 2014) (APÊNDICE 8).

Na Ásia e Oriente Médio, as frequências do SNP -13910 C>T raramente ultrapassam 10,0% (ITAN et al., 2010). Frequências maiores (porém não tão elevadas como as mencionadas anteriormente para o continente europeu) foram observadas no Paquistão (populações Sindi, Balooch, Pathan, Brahui e Mohammed com

frequências de 41,0%, 34,0 a 36,0%, 30,0%, 27,0 a 34,0% e 28,0%, respectivamente) (BERSAGLIERI et al., 2004; ENATTAH et al., 2008), Índia (frequências variando de 15,9% a 19,5% na região Norte, 21,3% na região Oeste e 18,4% em indo-europeus) (BABU et al., 2010; GALLEGO ROMERO et al., 2012) e Rússia (populações Udmurt, Mokshas e Erzas com frequências de 33,0%, 28,0% e 27,0%, respectivamente) (ENATTAH et al., 2007) (APÊNDICE 9).

Na Oceania, estudos com populações nativas demonstraram a ausência da variante -13910 C>T (BERSAGLIERI et al., 2004). Nas Américas, de maneira similar, a baixa frequência e inexistência desse SNP também foram observadas em povos ameríndios nativos do Chile (populações Mapuche com frequência de 0,5% (FERNÁNDEZ; FLORES, 2014), México (população Nahuatl com frequência de 1,2%) (OJEDA-GRANADOS et al., 2016) e Equador (7,4%) (PAZ-Y-MIÑO et al., 2016). Prevalências maiores foram identificadas em populações miscigenadas do continente (FERNÁNDEZ et al., 2016; MORALES et al., 2011; OJEDA-GRANADOS et al., 2016) (APÊNDICE 10).

A variante -13907 C>G, por sua vez, foi identificada somente em populações da África e Oriente Médio. Foi observada pela primeira vez no continente africano em populações Afar, da Etiópia, com frequências variando de 20,0 a 29,7% (INGRAM et al., 2007, 2009b; JONES et al., 2015). Neste mesmo continente, as maiores prevalências do SNP foram identificadas em populações Tigre, da Etiópia (19,3%) (JONES et al., 2013) e em povos Beja Beni Amer (25,0 a 13,0%) e Beja Hadandawa (18,2%) (HASSAN et al., 2016; JONES et al., 2015; RANCIARO et al., 2014) do Sudão. Frequências menores (inferiores a 15,0%) foram constatadas em outras populações também da Etiópia, Sudão e Quênia (ENATTAH et al., 2008; HASSAN et al., 2016; INGRAM et al., 2007, 2009b; JONES et al., 2013, 2015; PRIEHODOVÁ et al., 2014; RANCIARO et al., 2014). No Oriente Médio, a variante foi observada em populações Mutran do Kuwait (3,0%) (HILL; MOHAMMAD; TOOMASKIVISILD, 2013); em populações da região Hadramaut (1,2%) (LIEBERT et al., 2017) e da cidade de Al Hudaydah (1,5%) (PRIEHODOVÁ et al., 2014) – ambas no Iêmen – e em povos árabes da Arábia Saudita (0,8%) (ENATTAH et al., 2008) (APÊNDICE 11) .

O polimorfismo -13915 T>G é prevalente no Oriente Médio, com as maiores frequências observadas em populações árabes Dhofaris do Omã (72,4%) (AL-ABRI et al., 2012); no Iêmen, em populações das cidades de Al-Akhkum (65,1%), Hajjah (61,8%), Hudeida (59,1%) e da Ilha de Soqatra (58,5%) (PRIEHODOVÁ et al., 2014);

na Arábia Saudita, com frequências variando entre 52,4 a 64,5% de acordo com a região do país (IMTIAZ et al., 2007); e no Kuwait, nas populações Ajman (57,0%) e Mutran (55,0%) (HILL; MOHAMMAD; TOOMASKIVISILD, 2013). A frequência do SNP -13915 T>G também é elevada em populações de Beduínos, porém as frequências modificam-se de acordo com a localidade: 48,9% na Arábia Saudita (INGRAM et al., 2007), 39,1% na Jordânia (INGRAM et al., 2007) e 27,9% na Palestina (RAZ et al., 2013). Na África, foi observada a prevalência da variante nas populações Rashaayda (76,9%) do Sudão (PRIEHODOVÁ et al., 2014), Fulani (60,0%) de Mali (LOKKI et al., 2011) e Abbala (46,3%) do Chade (PRIEHODOVÁ et al., 2014). Frequências menores foram identificadas em outras populações africanas do Chade, Etiópia, Quênia, Marrocos, Nigéria e Sudão (ENATTAH et al., 2008; HASSAN et al., 2016; JONES et al., 2015; RANCIARO et al., 2014). Nos demais continentes, o SNP foi observado somente em indivíduos do Sul da Índia (0,8%) (GALLEGO ROMERO et al., 2012) e em povos miscigenados (*Mestízos*) do Equador (0,5%) (PAZ-Y-MIÑO et al., 2016) (APÊNDICE 12) .

As maiores frequências já identificadas da variante -14009 T>G foram encontradas na África, em populações provenientes do Sudão: 23,5% nos povos Beja Beni Amer e Beja Hadandawa (RANCIARO et al., 2014) e 16,7% na população Shaigi (INGRAM et al., 2009b; JONES et al., 2015; LIEBERT et al., 2017). Frequências inferiores a 10,0% foram observadas em populações de Camarões, Chade, Egito, Etiópia e Nigéria (INGRAM et al., 2009b; JONES et al., 2013, 2015; PRIEHODOVÁ et al., 2014; RANCIARO et al., 2014). No Oriente Médio, o SNP está presente também em baixas frequências, tendo sido identificado em populações palestinas e libanesas (na frequência de 1,0% em cada) (RANCIARO et al., 2014); em kuwaitianos (1,5%) (LIEBERT et al., 2017); em populações de origem asiática do Omã (1,5%) (AL-ABRI et al., 2012) e em populações da região de Al-Akhkum (1,2%) (PRIEHODOVÁ et al., 2014) e povos de origem judaica (1,0%) (RANCIARO et al., 2014) do Iêmen. No restante do continente asiático, o polimorfismo foi identificado somente nas populações paquistanesas Hazara e Kalash, na frequência de 1,0% em cada (RANCIARO et al., 2014) (APÊNDICE 13).

O SNP -14010 G>C também apresenta as maiores prevalências no continente africano (assim como as demais variantes mencionadas, com exceção da -13910 C>T), sendo a mais elevada encontrada na população pastoralista Il'gvesi Maasai (68,2%), que habita a região do Quênia (RANCIARO et al., 2014). Outros povos desse

mesmo país demonstraram frequências elevadas do polimorfismo em questão, dentre os quais estão inclusas as populações Ilchamus Maasai (55,0%); Yaaku (53,6%); Kikuyu (50,0%) e Mumonyot Maasai (50,0%) (RANCIARO et al., 2014). A elevada frequência do SNP é também observada na Tanzânia, nas populações Datoga (62,5%); Fiome (55,0%) e Iraqw (51,9%) (RANCIARO et al., 2014). No Sul do continente, a variante está presente em frequências reduzidas, tendo já sido identificada em populações da África do Sul, Angola, Botsuana, Moçambique, Namíbia e Zâmbia (BRETON et al., 2014; COELHO et al., 2009; MACHOLDT et al., 2014; PINTO et al., 2016; RANCIARO et al., 2014). Além do continente africano, este polimorfismo foi observado somente em povos do lêmén, com frequência de 1,2% em populações da cidade de Al-Akhkum (PRIEHODOVÁ et al., 2014); 0,8% e 0,6% em populações da região de Hodeida (PRIEHODOVÁ et al., 2014) e de Hadramaut (LIEBERT et al., 2017), respectivamente (APÊNDICE 14).

Por fim, o polimorfismo $-14011 C>T$ foi observado em baixas frequências nas populações estudadas até o momento, com a maior frequência sendo de 2,5% na população Kgalagadi, de Botsuana (MACHOLDT et al., 2014). Na África Ocidental, o SNP foi observado na frequência de 1,9% (RANCIARO et al., 2014). Na Etiópia, na África Oriental, foi constatada a variante $-14011 *T$ nas frequências de 1,1% em povos falantes de línguas cushíticas e de 0,6% em etíopes não classificados quanto a sua origem (JONES et al., 2013). Foi identificado também na Índia, com frequência de 2,1% na região Norte e 0,8% na região Oeste (GALLEGO ROMERO et al., 2012). No Omã, está presente em populações árabes da porção Norte do país na frequência de 1,0% (AL-ABRI et al., 2012) (APÊNDICE 15).

De acordo com Enattah e colaboradores (2007), a baixa frequência de algumas dessas variantes pode ser reflexo de que a adaptação para persistência da lactase ainda está ocorrendo em algumas populações.

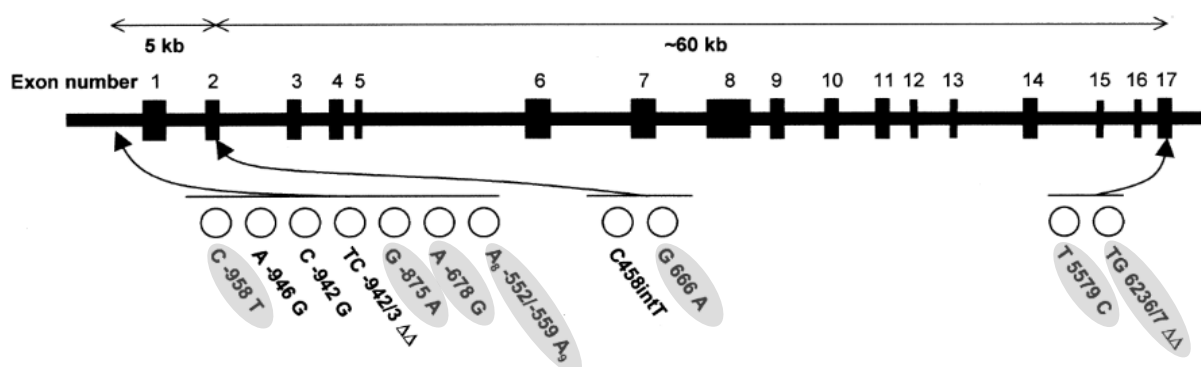
2.5.3 Haplótipos associados à persistência da lactase

Embora não associados ao fenótipo de persistência da lactase, o gene *LCT* e sua região promotora possuem diversos polimorfismos já descritos, culminando numa alta diversidade de haplótipos (BOLL; WAGNER; MANTEI, 1991; LLOYD et al., 1992). Harvey e colaboradores (1995) analisaram sete sítios polimórficos do *LCT* ao longo de 70 kb (que incluíam SNPs localizados em éxons e na região flanqueadora 5') e

determinaram as combinações haplotípicas existentes nos indivíduos amostrados – provenientes majoritariamente da França e de Utah, nos Estados Unidos. Ao todo, foram obtidos nove haplótipos (nomeados alfabeticamente de A a I) dos quais apenas três eram de comum ocorrência (A, B e C, observados nas frequências 75,4%, 13,8% e 5,8%, nessa ordem). Apesar de terem a mesma prevalência, os haplótipos demonstraram frequências diferentes entre as populações (58,3%, 27,0% e 6,3% para os haplótipos A, B e C, respectivamente, nos indivíduos franceses analisados, e frequência de 80,7%, 9,7% e 5,1% para os haplótipos A, B e C, respectivamente, nos indivíduos estadunidenses analisados). Posteriormente, através de análises alélicas das mesmas sete posições polimórficas, foram reportadas as frequências haplotípicas de populações da Índia e do Norte e Sul europeu (com amostras principalmente do Reino Unido e Itália, respectivamente). Os haplótipos A, B, C e E estavam presentes em todas as populações, contudo os haplótipos D e G foram encontrados exclusivamente em europeus e indianos, respectivamente. Entre os europeus, o haplótipo A era o mais frequente e os haplótipos B e C demonstraram maiores frequências em populações do Sul do continente. Nos asiáticos, os haplótipos A e C possuíam a mesma prevalência e o B era o menos frequente (HARVEY et al., 1998).

Alguns anos depois, Hollox e colaboradores (2001) analisaram onze populações da África, Ásia e Europa com diferentes fenótipos da lactase (incluindo povos da Namíbia, África do Sul, Reino Unido, Itália, Rússia, Eslováquia, Índia, Sri Lanka, Malásia, China, Japão e Papua-Nova Guiné) quanto aos nucleotídeos presentes em onze sítios polimórficos do gene *LCT* – os quais incluíam os sete *loci* utilizados por Harvey (1995, 1998) e quatro posições ainda não avaliadas (FIGURA 9). Foram identificadas 42 combinações haplotípicas no total, sendo que apenas quatro delas possuíam frequências maiores ou iguais a 5,0% (haplótipos A, B, C e U) (FIGURA 10). Tais haplótipos são derivados e não possuem relação direta entre si. É importante salientar que as populações africanas analisadas possuíam a maior diversidade haplotípica (incluindo combinações exclusivas) e os maiores valores de heterozigosidade (0,91 em populações San da Namíbia e 0,87 em populações falantes da língua Bantu da África do Sul) – características condizentes com o modelo “*Out of Africa*” de povoamento dos continentes (HOLLOX et al., 2001).

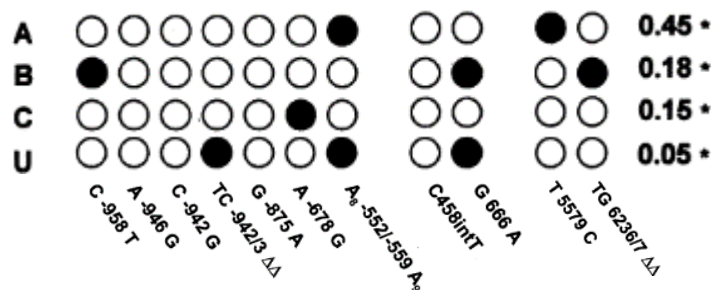
FIGURA 9 - POSIÇÕES NUCLEOTÍDICAS UTILIZADAS PARA DETERMINAR OS HAPLÓTIPOS DA LACTASE



FONTE: Modificado de HOLLOX et al. (2001).

NOTA: As posições destacadas em cinza referem-se aos *loci* também utilizados por Harvey e colaboradores (1995, 1998).

FIGURA 10 - HAPLÓTIPOS DA LACTASE MAIS COMUNS MUNDIALMENTE



FONTE: Modificado de Hollox et al. (2001).

LEGENDA: Círculos pretos indicam os alelos derivados e círculos brancos os alelos ancestrais (classificados através de comparação com sequências de chimpanzés). Os valores à direita indicam as frequências de cada haplótipo em todas as populações avaliadas.

Com exceção dos povos provenientes da África subsaariana, Hollox e colaboradores (2001) constataram que o haplótipo A (como já observado em trabalhos anteriores) é o mais comum em todas as populações analisadas – em especial no Norte da Europa, em que os baixos valores de heteroziguidade (0,26) enfatizam o predomínio deste haplótipo. O haplótipo B, por sua vez, também está presente em quase todas as populações (ausente somente nos indivíduos da África do Sul) e sua maior frequência foi observada na Papua-Nova Guiné. O haplótipo C foi o único identificado em todas as populações, porém baixas frequências são encontradas nos povos San da Namíbia e no Norte da Europa. O haplótipo U, apesar de estar entre os quatro mais frequentes, está presente somente em populações africanas e raramente é observado em povos indo-europeus. A perda de haplótipos em uma população onde

o haplótipo já existiu, assim como sua distribuição atual em populações não africanas, são resultantes de eventos de deriva genética (HOLLOX et al., 2001).

Através de análises de associação alélica visando identificar possíveis variantes relacionadas à persistência da lactase, Harvey e colaboradores (1998) demonstraram que o haplótipo A era o mais comum na maioria dos indivíduos persistentes para a enzima, enquanto que em não persistentes havia mais de um haplótipo associado. A elevada frequência do haplótipo A em populações do Norte da Europa, majoritariamente persistentes para a lactase, sugeriam uma associação existente com o fenótipo LP – ao menos nessa localidade (HOLLOX et al., 2001). Estudos posteriores observaram que o haplótipo A também estava presente em diversas populações cujo fenótipo LNP era predominante (HOLLOX et al., 2001) e, em raros casos, a LP demonstrou estar relacionada aos haplótipos B e I em populações europeias (HARVEY et al., 1998; WANG et al., 1995).

Poulter e colaboradores (2003) observaram que as variantes -13910^*T e -22018^*A são resultantes de mutações que ocorreram exclusivamente no haplótipo A estendido, evidenciando sua associação ao fenótipo LP (apesar de existirem cromossomos com esse haplótipo que não possuem os SNPs). No haplótipo B, foram identificados somente os alelos ancestrais dessas posições, -13910^*C e -22018^*G (POULTER et al., 2003). Ingram e colaboradores (2007) observaram que o SNP $-13910 C>T$ está associado ao haplótipo A também na população Fulani de Camarões, bem como nos poucos indivíduos que possuem o polimorfismo nas populações sudanesas Beja Beni Amer e Jaali e nos povos Drusos, Palestinos e Beduínos do Oriente Médio – indicando que este segmento teria a mesma origem que o encontrado em europeus que, através de fluxo gênico, o teriam inserido nas populações em questão. A correlação existente entre o referido SNP e o haplótipo A também foi observada em populações pastoris do Chile (MONTALVA et al., 2019) e em diversas populações da África, Ásia e Europa analisadas por Liebert e colaboradores (2017). Contrapondo-se, Myles e colaboradores (2005) identificaram, nas populações norte-africanas da etnia Berber, a ocorrência da variante -13910^*T em cromossomos com os haplótipos B, G, P, W e W* (sendo o último um haplótipo descrito pela primeira vez no trabalho em questão). Os autores sugerem que eventos de recombinação poderiam ter transferido o SNP do haplótipo A para os outros haplótipos mencionados (MYLES et al., 2005).

O haplótipo C demonstrou ser frequente em populações da Arábia Saudita, Jordânia e Sudão com alta frequência da persistência da lactase – o que está relacionado com a presença da variante -13915^*G , associada à LP, nos cromossomos que contêm tal haplótipo (INGRAM et al., 2007, 2009b). A variante foi identificada no haplótipo C também em populações da Etiópia (JONES et al., 2013) e de outras localidades da África, Ásia e Europa (LIEBERT et al., 2017). Este SNP foi identificado (embora raramente) nos haplótipos E, K e c presentes nas populações Beja Beni Amer (do Sudão), Beduínas (da Arábia Saudita e Jordânia) e em israelenses (INGRAM et al., 2007, 2009b).

Estudos preliminares demonstraram que as variantes -13907^*G e -13913^*C estariam associadas aos haplótipos A e B, respectivamente (apesar do polimorfismo $-13907 C>G$ também ter sido identificado nos haplótipos B, G, K e F) (ENATTAH et al., 2008; INGRAM et al., 2007, 2009b; JONES et al., 2013; LIEBERT et al., 2017). A ocorrência de um mesmo polimorfismo em haplótipos diferentes entre populações pode ser resultante tanto de eventos de recombinação quanto de novas mutações originando uma mesma variante genética (GABRIEL et al., 2002).

O polimorfismo $-14010 G>C$ foi identificado pela primeira vez no haplótipo F em populações do Quênia e da Tanzânia (TISHKOFF et al., 2007). Em povos da etnia Somali, da Etiópia, o alelo foi encontrado exclusivamente no haplótipo B (INGRAM et al., 2009b). Posteriormente, estudos reportaram a comum ocorrência desta variante também no haplótipo P em indivíduos de populações do Brasil, Etiópia e de outras localidades da Eurásia (FRIEDRICH et al., 2012a; JONES et al., 2013; LIEBERT et al., 2017).

A variante -14009^*G está presente mais comumente no haplótipo X, embora também tenha sido encontrada no haplótipo H e U (INGRAM et al., 2009b; JONES et al., 2013; LIEBERT et al., 2017). Vale ressaltar que no haplótipo H usualmente encontra-se o polimorfismo $-13730 T>G$ (LIEBERT et al., 2017).

O SNP $-14011 C>T$ foi identificado majoritariamente em cromossomos contendo o haplótipo A, porém, em alguns casos, foi observado também no haplótipo B, de modo que os autores sugerem que tal substituição nucleotídica poderia ter surgido mais de uma vez em diferentes localidades geográficas – embora não se possa desconsiderar que eventos de recombinação e fluxo gênico teriam ocasionado tal achado (LIEBERT et al., 2017).

TABELA 1 - HAPLÓTIPOS ASSOCIADOS À PERSISTÊNCIA DA LACTASE

Variante LP	Haplótipos	Referências
-13907*G	A , B, G, K, F	Enattah et al. (2008); Ingram et al. (2007, 2009b)
-13910*T	A , B, G, J, K, P, W e W*	Friedrich et al. (2012a, 2012b); Ingram et al. (2007, 2009b); Liebert et al. (2017); Montalva et al. (2019); Myles et al. (2005); Poulter et al. (2003)
-13913*C	B	Ingram et al. (2007, 2009b); Jones et al. (2013); Liebert et al. (2017)
-13915*G	C , E, K, c	Ingram et al. (2007, 2009b); Jones et al. (2013); Liebert et al. (2017)
-14009*G	H, U, X	Ingram et al. (2009b); Jones et al. (2013); Liebert et al. (2017)
-14010*C	B, F, P	Friedrich et al. (2012b); Ingram et al. (2009b); Liebert et al. (2017); Tishkoff et al. (2007)
-14011*T	A , B	Liebert et al. (2017)

FONTE: A autora (2020).

NOTA: Em negrito estão indicados os haplótipos nos quais os alelos referidos estão presentes mais comumente.

2.6 EVOLUÇÃO DA PERSISTÊNCIA DA LACTASE

Diferenças referentes às frequências mundiais do fenótipo de persistência da lactase (descritas no tópico 2.4.1 deste capítulo) promoveram a busca e desenvolvimento de explicações que justificassem tal distribuição. A primeira delas, denominada hipótese cultural-histórica, foi formulada por Simoons (1970) a partir de observações de que em populações nas quais o consumo de leite e derivados não era habitual (ou tornou-se um hábito somente recentemente), os indivíduos adultos majoritariamente apresentam a não persistência da enzima, sendo o oposto encontrado em populações que consomem tais alimentos. O fenótipo LP não poderia ter sido selecionado num contexto em que a lactose está ausente na dieta, pois a habilidade de digeri-la ao longo da vida não seria utilizada e, portanto, o indivíduo que possuísse tal fenótipo não apresentaria vantagem sobre indivíduos que não possuem a mesma habilidade (SIMOONS, 1978). A prevalência da condição LNP pode ser observada em parte da Ásia, na Oceania e nas Américas – que apresentam populações originárias sem tradições pastoralistas (tanto por carecerem de animais

lactantes – domésticos ou silvestres – quanto por não serem sociedades pecuárias) (SIMOONS, 1970). Além dessas regiões, o mesmo pode ser identificado em criadores de renas no extremo Norte da Eurásia e em alguns povos caçadores-coletores africanos – indicando que os hábitos da população seriam de fato mais influentes do que a região geográfica na qual ela está localizada (SIMOONS, 1970). Algumas populações da Índia, China, Sudeste da Ásia e nas porções Sul e na Costa Oeste da África subsaariana desenvolveram a criação de animais lactantes – incluindo búfalas, *mithan* (vaca selvagem), vacas, ovelhas e cabras – porém, como a ordenha não é um costume, os indivíduos possuem caracteristicamente o traço LNP (DIAMOND, 2019; SIMOONS, 1954, 1970). Desta forma, a persistência da lactase teria sido selecionada como um caráter adaptativo vantajoso apenas em adultos de populações que possuíam hábitos pastoralistas e de ordenha milenares, e que, além disso, consumissem leite e derivados (SIMOONS, 1970). A mesma explicação foi proposta por McCracken (1971), que considerou a elevada frequência do fenótipo LNP em afro-americanos dos Estados Unidos um indicativo de ausência de seleção (considerando que a ancestralidade desta população remonta ao Oeste africano que possui, tradicionalmente, povos não consumidores de leite).

Contra-pondo-se à hipótese cultural-histórica, a hipótese da causa-reversa preconiza que as práticas de pastoralismo e ordenha surgiram em populações que já apresentavam o fenótipo LP, sendo tais hábitos rejeitados por povos que não possuíam habilidade de digerir a lactose devido aos sintomas desagradáveis produzidos (MCCRACKEN, 1971).

A hipótese da assimilação de cálcio, formulada por Flatz e Rotthauwe (1973), propõe que, como a lactose aumenta a captação de cálcio (assim como a vitamina D), o consumo de leite promoveria o suplemento desse mineral mesmo em ambientes de alta latitude com pouca disponibilidade de radiação solar e, conseqüentemente, com populações deficientes para a vitamina D. Segundo os autores, isso justificaria a ampla ocorrência da persistência da lactase nos povos do Norte da Europa, visto que seria um traço adaptativo vantajoso nessa localidade, pois impediria o desenvolvimento de doenças decorrentes da desmineralização óssea como o raquitismo e a osteomalácia (FLATZ; ROTTHAUWE, 1973).

Uma quarta hipótese, preconizada nos trabalhos de Cook e Al-Torki (1975) e Cook (1978), sugere que o leite seria um suprimento hidroeletrólítico e calórico em ambientes áridos (que possuem escassez de água e de fontes nutritivas). Desta

forma, a prevalência da LP na Península Arábica seria decorrente da seleção positiva promovida pelas vantagens citadas, ocasionadas pelo consumo de leite de camelo por parte das populações que habitam tal região, como, por exemplo, os Beduínos da Arábia Saudita (COOK; AL-TORKI, 1975). Ingram e colaboradores (2009a) propõem uma extensão à esta teoria, na qual a seleção ao fenótipo LP ocorreria não somente em regiões desérticas, mas sim em localidades que apresentassem circunstâncias semelhantes como a seca severa e a fome. De acordo com os autores, a vantagem do fenótipo se tornaria conspícua nessas condições – especialmente em momentos de surtos de diarreia – já que o leite seria a única fonte de água disponível, porém vantajosa apenas para indivíduos lactase persistentes (INGRAM et al., 2009a; SVERRISDÓTTIR et al., 2014).

Contrariamente às hipóteses descritas, mais recentemente autores propuseram uma explicação para justificar a ampla ocorrência do fenótipo LNP em algumas populações africanas. A hipótese baseia-se na correlação existente entre o fenótipo e as áreas endêmicas para a malária no continente, sugerindo que a ausência de ingestão de leite privaria os parasitas do gênero *Plasmodium* (presentes em indivíduos infectados) de compostos orgânicos como o ácido para-aminobenzóico (PABA) e as riboflavinas – essenciais para rotas anabólicas desses protozoários – e, assim, reduziria a multiplicação parasitária nos eritrócitos, promovendo diminuição na mortalidade causada por malária (ANDERSON; VULLO, 1994; CORDAIN; HICKEY; KIM, 2012). Contudo, estudos com indivíduos da Sardenha, na Itália, comparando populações com histórico de endemismo de malária e sem tal histórico demonstraram que a hipótese não é válida – ao menos para esta localidade (MELONI et al., 1998). Desta forma, Cordain e colaboradores (2012) sugerem que essa explicação, de acordo com suas análises, poderia ser aplicada apenas para populações africanas.

Ranciaro e colaboradores (2014) sugerem que, possivelmente, a alta frequência da persistência da lactase na população de caçadores-coletores Hadza, da Tanzânia, se deve à seleção positiva do mecanismo de metabolismo da florizina que a enzima também apresenta. Este composto está presente, naturalmente, em plantas da família Rosaceae (que inclui espécies de peras, maçãs e cerejas) e é empregado medicinalmente no tratamento contra malária. Portanto, a manutenção da enzima na vida adulta, nessas populações, promoveria a diminuição dos sintomas provocados pela malária – sendo vantajoso em regiões nas quais a doença é endêmica (contrapondo-se à hipótese anterior).

Holden e Mace (1997), através de análises filogenéticas de populações avaliadas para o fenótipo LP, demonstraram que a distribuição desta condição por todo o mundo condiz com a distribuição de populações tradicionalmente pastoralistas. Em contrapartida, tal correlação é inexistente quando comparada com povos, de diferentes continentes, provenientes de locais com baixos níveis de radiação solar e com muitos meses de estiagem por ano (HOLDEN; MACE, 1997). A ausência de correlação entre regiões de altas latitudes com baixa radiação solar e a LP foi encontrada também por Itan e colaboradores (2009). Sendo assim, as autoras sugerem que somente a hipótese cultural-histórica demonstra-se apta para explicar a ocorrência da persistência da lactase nas diferentes populações (HOLDEN; MACE, 1997). Ainda, Bloom e Sherman (2005) demonstraram que a má-absorção da lactose está associada à povos de regiões de climas extremos (presentes tanto em altas quanto em baixas latitudes) e, mais fortemente, à ocorrência histórica de doenças severas que acometem gados, fazendo com que o fenótipo LNP predomine em populações impossibilitadas de criarem gado de ordenha – corroborando a hipótese cultural-histórica.

A hipótese cultural-histórica caracteriza-se por ser um exemplo de coevolução gene-cultura, pois o hábito de ordenha e consumo de leite (sem efeitos colaterais) favoreceu a seleção positiva dos indivíduos lactase persistentes – aumentando sua frequência e favorecendo a continuidade de tradições pastoralistas (BEJA-PEREIRA et al., 2003; DURHAM, 1991). É importante salientar, contudo, que esta hipótese não explica a baixa ocorrência da persistência da lactase em populações asiáticas pastoralistas – principalmente da Ásia Central (SÉGUREL et al., 2020). Em populações da Mongólia, por exemplo, (apesar da frequência do fenótipo LP ser de apenas 12,0%) (HEYER et al., 2011) foi identificado que laticínios compreendem cerca de 35,0% de toda energia fornecida pela dieta e que seu consumo remonta a aproximadamente 3000 AC (antes de Cristo) (JEONG et al., 2018; WILKIN et al., 2020). A mesma inconsistência foi encontrada em populações do Tibete, nas quais a criação de iaques para ordenha ocorre há mais de 5000 anos e o consumo de leite e derivados é parte fundamental da dieta, apesar da frequência da persistência da lactase não ultrapassar 30,0% (DONG; LONG; KANG, 2003; MIN-SHENG et al., 2012; RHODE et al., 2007). A baixa prevalência (entre 25,0 e 32,0%) também foi observada nas populações pastoralistas Kazakh do Uzbequistão – região com tradição de consumo de leite há mais de 5000 anos (HEYER et al., 2011; OUTRAM et al., 2009).

Ségurel e colaboradores (2020) sugerem que mecanismos de adaptação cultural, incluindo processos de fermentação do leite (culminando na diminuição na concentração de lactose do alimento) (BUTTRISS, 1997) e adaptação colônica (através da aclimatação da microbiota do cólon mediante consumo constante de lactose) (HERTZLER; SAVAIANO, 1996), poderiam justificar o desenvolvimento de tradições de ordenha juntamente com o consumo de leite mesmo em populações que apresentam a não persistência da lactase, como as mencionadas, pois evitariam os sintomas desagradáveis associados à má digestão do dissacarídeo. O mesmo teria ocorrido nas populações africanas da etnia Somali, nas quais muitos dos indivíduos que apresentam o fenótipo LNP consomem leite fresco sem resultar em manifestações clínicas (INGRAM et al., 2009a).

2.6.1 Surgimento do consumo adulto de leite

Evidências indicam que o aparecimento de sociedades pastoralistas ocorreu no Oriente Médio durante o período Neolítico, que corresponde ao intervalo de 7000 a 2500 AC (VIGNE; HELMER, 2007). Os primeiros indícios dessas populações datam de aproximadamente 11500 AC e baseiam-se na domesticação de espécies das quais são derivadas as ovelhas, cabras, vacas e porcos atuais (VIGNE, 2011). Eventos de domesticação independentes também ocorreram em 8500 AC e 5000 AC no Vale do Indo (Paquistão), em 8000 AC na China e em 7000 AC no Peru (VIGNE, 2011). Contudo, registros do uso de leite por essas populações foram encontrados somente milênios após a domesticação. Os indícios mais antigos foram obtidos de vasos de cerâmica provenientes de povos que habitavam a região da Anatólia (que corresponde à porção asiática da Turquia) em aproximadamente 7000 AC (EVERSHED et al., 2008). Evidências encontradas nos vasos analisados apontam para a ocorrência do processamento do leite (através da fermentação) já nessa época (EVERSHED et al., 2008). Análises de isótopos e osteoarqueológicas sugerem que a utilização de leite já ocorria no início do oitavo milênio AC no Oriente Médio e na metade do sexto milênio AC no Sul da Europa (VIGNE; HELMER, 2007). No restante do continente europeu, indícios do consumo da bebida datam de 6110 AC no Reino Unido (COPLEY et al., 2005); 5900-5500 AC na região da Romênia e Hungria (CRAIG et al., 2005) e aproximadamente 5000 AC na Península Escandinava (CRAIG et al., 2011).

No continente africano, os primeiros registros de utilização de leite são remanescentes de cerâmicas encontradas no deserto do Saara (na região pertencente ao território da Líbia) e datam de 7150-5750 AC (DUNNE et al., 2012). No Leste africano, sobretudo no Sudão e Chade, pesquisas linguísticas e de arte rupestre indicam o uso da bebida a partir de 3000 AC (JESSE et al., 2013). Na porção Sul do continente, evidências da utilização de leite (também identificadas através de resquícios de gordura nas cerâmicas) remontam ao primeiro milênio AC e são provenientes de populações de caçadores-coletores que habitavam a região que, atualmente, corresponde ao Lesoto (FEWLASS et al., 2020). Os autores sugerem que o uso de leite nesses povos, que comumente apresentam o fenótipo de não persistência da lactase, ocorria através de laticínios e não da bebida fresca (FEWLASS et al., 2020).

Na Ásia, as evidências mais antigas referentes ao uso de leite datam de 3500 AC e referem-se ao seu consumo por populações pastoralistas que habitavam a região que corresponde, atualmente, ao Norte do Cazaquistão (OUTRAM et al., 2009; SÉGUREL et al., 2020). Características das cerâmicas nas quais constatou-se a presença de leite apontam para seu processamento antes do consumo por parte desses povos (OUTRAM et al., 2009). Na Mongólia, análises de espectrometria de massa em cálculos dentais fossilizados identificaram o consumo a partir de 1456 AC (JEONG et al., 2018). Em fósseis desse mesmo país, Wilkin e colaboradores (2020) obtiveram, a partir da análise proteômica também de cálculos dentais, indícios do consumo lácteo a partir de 3000 AC – sendo a evidência mais antiga do consumo na Estepe Euroasiática. O mesmo tipo de análise, porém desempenhada em amostras de indivíduos mumificados das tumbas de Xiaohe, na China, permitiu a identificação de hábitos de consumo de leite e, principalmente, de seu processamento para produção de *kefir* entre 1980-1450 AC (YANG et al., 2014).

A domesticação de mamíferos nas Américas data de 5050-4050 AC e envolve os ancestrais das lhamas, guanacos, vicunhas e alpacas atuais, que se expandiram por toda a extensão dos Andes sul-americano (KADWELL et al., 2001; WHEELER, 1995). Apesar disso, a ausência de indícios arqueológicos e culturais sugere que o leite desses animais não teria sido utilizado pelas populações locais (GADE, 1999).

2.6.2 Surgimento e seleção das variantes associadas à persistência da lactase

Análises utilizando a metodologia de homozigose do haplótipo estendido (EHH) permitem estimar o tamanho da região genômica na qual determinado alelo está inserido e, a partir disso, prever sua idade e (possivelmente) identificar se a seleção natural está atuando sobre ele positiva ou negativamente (SABETI et al., 2002). Em populações do Norte da Europa, foi identificada a existência de uma região do gene *LCT* que se estende por mais de 1 Mb (megabases) em intenso desequilíbrio de ligação – constatada especialmente no haplótipo A, no qual a variante *-13910*T* está inserida – o que é consistente com a ocorrência de seleção positiva recente dessa região (POULTER et al., 2003). Bersaglieri e colaboradores (2004) demonstraram (através de diferentes metodologias, incluindo a EHH) que, de fato, esta seleção teria ocorrido, visto que o haplótipo contendo as variantes *-13910*T* e *-22018*A* é demasiadamente extenso – algo incomum, considerando sua elevada frequência, e que não se aplica aos demais blocos de haplótipos que o circundam. Os autores identificaram que o surgimento da variante *-13910*T* teria ocorrido entre 20650 e 2188 anos atrás (período em que a o continente europeu já seria povoado por humanos e que a separação entre a sua população e as populações africanas e asiáticas ancestrais estaria bem estabelecida). Além disso, o coeficiente de seleção da variante é equivalente a 0,014 até 0,15 na população euro-americana dos Estados Unidos e entre 0,09 a 0,19 na população escandinava, enfatizando a recente e grande intensidade na qual a seleção natural estaria atuando positivamente sobre o haplótipo que contém este alelo – um dos exemplos mais fortes encontrados no genoma humano até o momento (BERSAGLIERI et al., 2004; SÉGUREL; BON, 2017). Valores similares de coeficiente de seleção foram obtidos por Enattah e colaboradores (2008), incluindo também populações do Oriente Médio.

Apesar de, possivelmente, a variante *-13910*T* ter surgido antes ou simultaneamente às práticas de ordenha, estudos envolvendo DNA antigo apontam para sua prevalência somente milhares de anos depois (WALKER; THOMAS, 2019). De acordo com Itan e colaboradores (2009), o alelo *-13910*T* começou a ser selecionado em populações europeias a partir de 5500 AC, na região central da Península Balcânica e da Europa Central. Porém, a ocorrência mais antiga do SNP - *13910 C>T* foi encontrada em um indivíduo da Ucrânia datado de 4010 AC (SÉGUREL et al., 2020). Em amostras de 26 indivíduos, encontrados num sepultamento no País Basco, que data de aproximadamente 3050-2550 AC (correspondendo ao período Neolítico Tardio), foi identificada a frequência de 23,0% para o polimorfismo *-13910*

C>T – a maior encontrada, até o momento, em populações pré-históricas (PLANTINGA et al., 2012). Referente a esse mesmo intervalo de tempo, foi observada a frequência de 1,2% em populações da Europa Central, bem como a ausência do polimorfismo em populações asiáticas (SÉGUREL et al., 2020). A variante -13910*T foi encontrada também em um único indivíduo do Noroeste europeu – dentre um total de 230 – pertencente à Cultura do Vaso Campaniforme (*Bell-Beaker culture*, em inglês) datado entre 2450-2140 AC (MATHIESON et al., 2015; OLALDE et al., 2018). A ausência do SNP -13910 C>T foi constatada em oito indivíduos do Neolítico e em um indivíduo do Mesolítico (13000-9000 AC) – todos da Europa Central (BURGER et al., 2007). No Norte da Europa, através da análise de dez indivíduos de populações de caçadores-coletores do Neolítico, foi identificada a frequência de apenas 5,0% para o polimorfismo -13910 C>T, o que, de acordo com os autores, seria característica dessas populações e teria sido modificado para uma maior prevalência quando populações sedentárias de agricultores as sucederam, dando origem aos escandinavos atuais (MALMSTRÖM et al., 2010). Gamba e colaboradores (2014) utilizando amostras de 13 indivíduos do Leste da Hungria, datados desde o Neolítico Recente (5700 AC) até a Idade do Ferro (800 AC), constataram a presença da variante -13910*T somente a partir da Idade do Bronze (aproximadamente 1000 AC). Estudos posteriores observaram, em populações de diversas regiões da Eurásia, que a frequência do alelo -13910*T também era baixa na Idade do Bronze (equivalente a 10,0%) – sendo um pouco maior em povos da Escandinávia (ALLENTOFT et al., 2015). Similarmente, o primeiro registro da variante -13910*T no continente asiático é proveniente deste mesmo período, tendo sido identificado em um indivíduo do Cazaquistão datando de 1763 AC (SÉGUREL et al., 2020).

Estudos referentes a populações europeias da Idade Média sugerem que nesse período o SNP -13910 C>T já apresentaria a prevalência encontrada em populações atuais. Através de amostras de ossos de povos que viveram entre os séculos X e XI na região que corresponde a atual Hungria, foi encontrada a frequência alélica de 27,8% para a variante -13910*T, encontrada apenas em camponeses e ausente em indivíduos das camadas mais altas dessa sociedade (NAGY et al., 2011). O mesmo tipo de análise, porém contendo populações dos séculos X ao XII da Alemanha, identificaram a frequência alélica de 50,0% para o polimorfismo -13910 C>T (KRÜTTLI et al., 2014). Em populações medievais da Polônia Central, a frequência variou de 20,0 a 64,0% (WITAS et al., 2015). Ségurel e Bon (2017) propõem que a

heterogeneidade nas frequências observadas nessas populações da Idade Média pode decorrer tanto de erro amostral (devido à escassez de amostras disponíveis) quanto da forte estratificação social existente neste período, inibindo que ocorresse fluxo gênico entre diferentes camadas da sociedade. Contudo, apesar da discrepância de valores, o aumento na frequência é significativo (se comparado aos povos pré-históricos) e indica que o estabelecimento do alelo *-13910*T* no continente europeu teria ocorrido recentemente (FAN et al., 2016; KRÜTTLI et al., 2014).

Apesar de restrito a poucas populações, o alelo *-13910*T* está presente em algumas localidades do Norte e Centro da África. Diferentes estudos identificaram que a variante, nessas populações, está contida em haplótipos derivados do mesmo ancestral encontrado na Europa (no caso do haplótipo encontrado na população Fulani, de Camarões) e Oriente Médio (no caso do haplótipo encontrado no povo Mozabite, da Argélia), de modo que a ocorrência do SNP *-13910 C>T* no continente africano seria decorrente de fluxo gênico com populações de fora da África (ENATTAH et al., 2007; RANCIARO et al., 2014).

Tishkoff e colaboradores (2007) avaliaram o tamanho dos haplótipos estendidos que incluíam as variantes *-13907*G*, *-13915*G* e *-14010*C* em populações africanas e observaram desequilíbrio de ligação com extensões de aproximadamente 1,4 Mb; 1,1 Mb e mais de 2 Mb, respectivamente. Assim como identificado para a variante *-13910*T*, as amplas extensões são incomuns e sugerem ausência de recombinação o que, por sua vez, aponta para a ocorrência de seleção positiva local e recente desta região (BERSAGLIERI et al., 2004; POULTER et al., 2003; TISHKOFF et al., 2007).

A origem do alelo *-13907*G* foi estimada em aproximadamente 5000 anos. O SNP teria surgido em populações do Leste da Etiópia falantes das línguas cushíticas, que, posteriormente, teriam migrado para o Quênia e Sudão (RANCIARO et al., 2014). Enattah e colaboradores (2008), através de análises de rede de haplótipos, identificaram que as variantes *-13910*T* e *-13907*G* pertenceriam a haplótipos derivados de um mesmo bloco ancestral. Com isso, os autores sugeriram que ambos os alelos poderiam ter suas frequências aumentadas a partir do mesmo evento de domesticação de gado, contrapondo-se a proposta inicial de que ambos fariam parte de contextos evolutivos distintos devido às suas origens separadas (TISHKOFF et al., 2007). Posteriormente, Ranciaro et al. (2014) constataram ausência de correlação evolutiva entre os haplótipos contendo as variantes em questão.

Através da estimativa da idade do alelo ancestral comum mais recente, Enattah e colaboradores (2008) identificaram que o surgimento dos alelos -13915^*G e -3712^*C teria ocorrido há 4095 anos (aproximadamente 2000 AC) em populações da região que corresponde a atual Arábia Saudita. O aparecimento das variantes estaria associado a diferentes eventos de domesticação de camelídeos ocorridos nesta época (ENATTAH et al., 2008). O coeficiente de seleção estimado para o SNP $-13915 T>G$, em populações árabes, varia de 0,04 a 0,14 – indicando que a seleção positiva também estaria atuando fortemente sobre este alelo (ENATTAH et al., 2008). Vale ressaltar que a presença desta variante em populações do Norte, Leste e Centro da África é consequência do fluxo gênico ocorrido com populações do Oriente Médio nessas regiões (RANCIARO et al., 2014). Estima-se que a inclusão do alelo em questão no genoma das populações leste-africanas teria ocorrido a partir de 1400 anos atrás, por meio da expansão do Islamismo (ALEXANDER, 2001; RANCIARO et al., 2014).

A idade da variante -14010^*C foi estimada, em populações do Quênia e da Tanzânia, entre aproximadamente 2700 - 6800 anos (TISHKOFF et al., 2007). O surgimento do SNP é concomitante ao período em que ocorreram migrações de povos pastoralistas do Norte africano para o Leste do continente (local em que o polimorfismo teria surgido) (AMBROSE, 1998; SMITH, 1992). Ranciaro e colaboradores (2014) propõem que o alelo apareceu primeiramente em populações falantes das línguas afro-asiáticas, posteriormente sendo introduzido em povos falantes das línguas nilo-saarianas. Macholdt e colaboradores (2015), através de análises de variações de microssatélites, sugeriram que o surgimento do alelo seria entre 6000-7000 anos atrás no Leste da África, mas sua seleção positiva teria ocorrido somente a partir de 5000 anos, coincidindo com a primeira domesticação de gado no continente africano (PHILLIPSON, 2005). Esses povos do Leste migraram para o Sul do continente a partir de 2000 anos atrás, explicando a ocorrência do SNP (em haplótipos derivados) também em populações dessa região (MACHOLDT et al., 2014). Os valores de coeficiente de seleção deste alelo, em populações africanas, variam de 0,04 a 0,10 – consistentes com seleção positiva, assim como as demais variantes associadas à LP já descritas (MACHOLDT et al., 2015; TISHKOFF et al., 2007).

O surgimento e fixação (em alguns casos) independente de vários alelos que culminam no mesmo fenótipo de persistência da lactase em populações distintas

evidencia a ocorrência de um processo denominado convergência evolutiva (ENATTAH et al., 2007; FAN et al., 2016; TISHKOFF et al., 2007). Tal processo é caracterizado pela evolução independente de características similares a partir de pressões seletivas semelhantes que, nesse caso, seria a vantagem dos indivíduos LP sobre os LNP, decorrente da habilidade dos primeiros em consumir leite e laticínios sem desenvolver sintomas desagradáveis (CHRISTIN; WEINREICH; BESNARD, 2010).

2.7 PERSISTÊNCIA DA LACTASE NO BRASIL

Os primeiros estudos sobre variantes associadas ao fenótipo da persistência da lactase realizado em populações brasileiras identificaram, através de genotipagem, os alelos -13910^*T e -22018^*A na população euro-brasileira de Porto Alegre, no Rio Grande do Sul, (n = 18) e em pacientes com doença inflamatória intestinal (n = 75), diagnosticados como LP previamente através do teste do hidrogênio expirado (BERNARDES-SILVA et al., 2007; BULHÕES et al., 2007). Posteriormente, Mattar e colaboradores (2009) também analisaram o SNP $-13910 C>T$, através de genotipagem, em brancos (n = 399), pardos (n = 65) e pretos (n = 50), observando o genótipo homozigoto numa maior proporção em indivíduos brancos (6,0%, n = 24), e seu genótipo heterozigoto em maior proporção em indivíduos pardos (40,0%, n = 26). Esse mesmo estudo apontou a frequência reduzida da variante -13910^*T em indivíduos pretos (12,0%, n = 6), bem como sua ausência em descendentes de japoneses (n = 53) – população cujo SNP $-22018 G>A$ demonstrou ser o marcador mais adequado para diagnóstico da LP, devido à ausência de outros polimorfismos associados à persistência da lactase na população que a originou (MATTAR et al., 2010).

Nas populações ameríndias brasileiras Guarani-Kaiowá (n = 84), Guarani-Ñandeva (n = 59), Kaingang (n = 72) e Xavante (n = 101), quando genotipadas para a substituição $-13910 C>T$, também foi observado este alelo em baixa frequência, com porcentagens de 0,6%; 7,6%; 4,9% e 0,5% respectivamente (FRIEDRICH et al., 2012b). Os autores sugerem que a variante em questão tenha sido introduzida nessas populações a partir de fluxo gênico com populações europeias (FRIEDRICH et al., 2012b). Ainda nesse estudo, foram identificados 15 haplótipos – inferidos a partir de polimorfismos localizados no gene *LCT* e na sua região flanqueadora 5', utilizados por

Hollox e colaboradores (2001) – dos quais apenas cinco estavam presentes, em todas as populações avaliadas, em frequências iguais ou superiores a 5,0%, sendo eles os haplótipos C (41,2%), A (29,2%), E (10,1%), B (9,3%) e U (7,2%). A frequência do haplótipo E, provavelmente originado a partir da recombinação dos haplótipos A e C, nos ameríndios brasileiros analisados é superior a encontrada em populações africanas e europeias. Além disso, a prevalência de cada haplótipo varia de acordo com a população ameríndia analisada: nos indivíduos das etnias Guarani-Kaiowá e Guarani-Ñandeva, o haplótipo A é o observado mais comumente, enquanto que nos indivíduos das etnias Kaingang e Xavante o haplótipo C é o mais prevalente. Através da aplicação do Teste de Neutralidade de Ewens-Watterson, foi concluído que os haplótipos observados não estão sob pressão de seleção (FRIEDRICH et al., 2012b).

Através de genotipagem dos polimorfismos *-13910 C>T* e *-22018 G>A* e sequenciamento da região do íntron 13 do gene *MCM6* em brasileiros de diferentes regiões do país, Friedrich e colaboradores (2012a) identificaram que o SNP *-13910 C>T*, de origem europeia, é a variante associada à LP observada mais frequentemente nos indivíduos analisados. A menor frequência foi constatada em afro-brasileiros de Porto Alegre (n = 182) e em populações miscigenadas de Belém (n = 200), no Pará, e de Recife (n = 262), em Pernambuco, nas frequências alélicas de 18,4%, 17,5% e 20,4%, respectivamente, enquanto a maior frequência foi reportada em euro-brasileiros de Porto Alegre (n = 337) (FRIEDRICH et al., 2012a). Além do SNP *-13910 C>T*, foram também identificadas em brasileiros as variantes associadas à LP *-14010*C* (presente em um indivíduo de ancestralidade africana de Porto Alegre) e *-14011*T* (presente em um indivíduo de Belém e em três indivíduos de Recife) (FRIEDRICH et al., 2012a). Nesse mesmo estudo, foram também observados os alelos *-13779*C* e *-13937*A*, – os quais não apresentam associação ao fenótipo comprovada – em indivíduos de Recife e Porto Alegre (FRIEDRICH et al., 2012a). Até o momento, a variante *-13779 G>C* foi reportada em um indivíduo LNP da etnia Somali, na Etiópia (INGRAM et al., 2009b), e em indivíduos do Oeste (n = 4, frequência alélica de 0,7%), Norte (n = 1, frequência alélica de 0,2%) e Sul (n = 50, frequência alélica de 4,3%) da Índia (GALLEGO ROMERO et al., 2012). Enquanto a variante *-13937 G>A* foi observada na África do Sul em dois indivíduos da população Xhosa (frequência alélica de 1,4%) (TORNIAINEN et al., 2009) e em um indivíduo da população Chagga (frequência alélica de 1,2%) (JONES et al., 2015); em apenas um

indivíduo no Chile (MORALES et al., 2011) e na população Mambila, de Camarões (frequência alélica de 0,8%) (JONES et al., 2015).

No estudo de Friedrich e colaboradores (2012a) também foram investigados quais haplótipos do gene *LCT*, de acordo com os marcadores estipulados por Hollox e colaboradores (2001), estavam presentes nos brasileiros analisados (n = 981). Ao todo, foram identificados 26 haplótipos, dos quais o haplótipo A foi o mais frequente em todos os grupos (34,3% nos afro-brasileiros de Porto Alegre; 46,0% nos euro-brasileiros de Porto Alegre; 38,0% na população miscigenada de Belém e 37,8% na população miscigenada de Recife) – assim como também já constatado para diferentes populações da África, Europa e Ásia (HOLLOX et al., 2001). Segundo os autores, a grande diversidade de haplótipos encontrada na população brasileira é compreensível considerando seu histórico de formação, contendo contribuições de ameríndios, africanos e europeus (FRIEDRICH et al., 2012a).

Outro estudo, incluindo a genotipagem do alelo *-13910*T* nas populações euro-brasileiras Menonita do Paraná (n = 151, grupo de ancestralidade alemã e holandesa e com histórico de acentuada endogamia) e de Curitiba (n = 292) – ambas da região Sul do Brasil – identificou a maior frequência desta variante de origem europeia já reportada no país, presente nas frequências alélicas de 65,0% e 33,0%, respectivamente (BOSCHMANN et al., 2016). Prevalência similar foi identificada por Oliveira e colaboradores (2020), também em populações euro-brasileiras de Curitiba (n = 122, frequência alélica de 34,0%).

2.8 COMPOSIÇÃO GENÉTICA DA POPULAÇÃO BRASILEIRA

As populações que compreendem o Brasil são notavelmente heterogêneas, sendo reflexo dos diversos povos que as originaram durante a formação do país e das diferentes dinâmicas de miscigenação existentes entre eles (ALVES-SILVA et al., 2000; RIBEIRO, 2014). De maneira geral, destacam-se três grandes populações na constituição genômica e cultural dos brasileiros contemporâneos: ameríndios, africanos e europeus (RIBEIRO, 2014).

As populações que povoaram inicialmente a América do Sul eram do Norte do continente e derivaram de povos asiáticos da região dos Montes Altai, no Sul da Sibéria, que se deslocaram para o continente americano no final do Pleistoceno (cerca de 16,5 mil anos atrás), através do estreito de Bering (GOEBEL; WATERS;

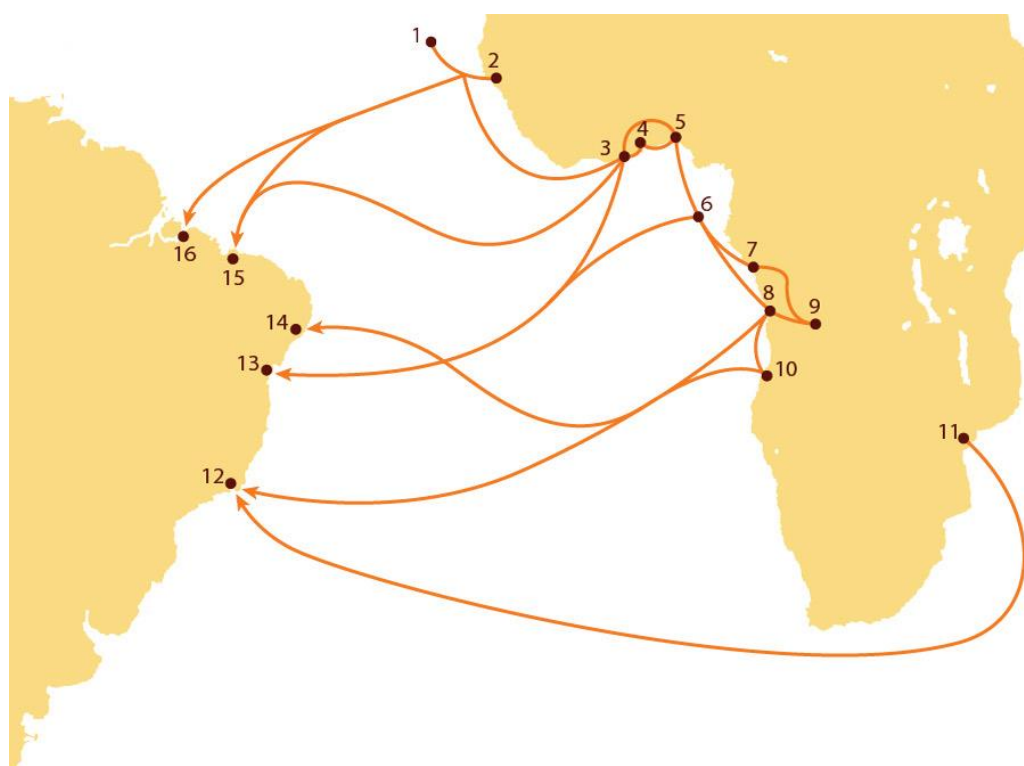
O'ROURKE, 2008; SALZANO, 2011). Esse evento de migração única promoveu um acentuado efeito gargalo de garrafa – caracterizado pela perda de diversidade genética da população decorrente da redução de seu tamanho efetivo (NEI; MARUYAMA; CHAKRABORTY, 1975) – que culminou na baixa variabilidade genética encontrada em ameríndios, além de proporcionar um gradiente na proporção de ancestralidade siberiana ao longo das Américas (diminuindo no sentido norte-sul) (WANG et al., 2007). Estima-se que o povoamento da América do Sul, e consequentemente do Brasil, tenha ocorrido a partir de 14 mil anos atrás, na região dos Andes (ROTHHAMMER; DILLEHAY, 2009). Embora evidências genômicas apontem para uma mesma população ancestral, não há consenso acerca das rotas migratórias utilizadas por esses povos para se dispersarem pelo restante do continente (REICH et al., 2012; ROTHHAMMER; DILLEHAY, 2009). É importante destacar, contudo, a bem estabelecida diferenciação constatada entre as populações dos Andes, Amazônia e Chacos, decorrente do isolamento promovido pelas barreiras geográficas existentes nestas localidades (ROTHHAMMER; DILLEHAY, 2009; SALZANO, 2011).

A partir do século XVI, com a chegada dos portugueses no território brasileiro, a dinâmica das populações nativas mudou drasticamente devido ao genocídio promovido pelos invasores, que ocasionou o extermínio de cerca de 90,0% da população original de 2 milhões de pessoas (de acordo com estimativas) (CHARNY, 2000; SALZANO; BORTOLINI, 2002). Nesse mesmo século foi iniciada a migração em massa de europeus, com aproximadamente 700 mil portugueses vindo para o Brasil entre 1500 e 1700 (IBGE, 2007).

Durante mais de 300 anos (entre os séculos XVI e XIX) ocorreu o tráfico negreiro transatlântico, promovendo a vinda de aproximadamente 4 milhões de africanos para serem escravizados em terras brasileiras (IBGE, 2007). Os africanos trazidos forçadamente vinham majoritariamente de duas regiões: Costa da Mina, que, atualmente, compreende os territórios de Gana, Togo, Benin e Nigéria, e Costa de Angola, que corresponde ao Norte da atual Angola (SOUZA, 2008) (FIGURA 11). Embora, a partir do século XVIII uma quantidade significativa de pessoas tenha vindo do golfo de Benin, a maioria partia dos portos de Angola, principalmente de Luanda. Quando desembarcavam no Brasil, os indivíduos podiam ser classificados de acordo com os portos de onde foram embarcados ou das feiras em que haviam sido negociados no interior da África – ignorando as etnias das quais eles eram

pertencentes – de modo que as designações brasileiras não refletiam a ancestralidade dessas populações (SOUZA, 2008). De maneira geral, na maior parte do território brasileiro, os escravizados pertenciam a povos bantos (de diferentes etnias) da África Central. Na Bahia, Maranhão, Minas Gerais e Rio de Janeiro, contudo, desembarcaram populações provenientes da África Ocidental, sendo formadas principalmente por grupos iorubás, pertencentes à família linguística nigero-congolesa (SOUZA, 2008).

FIGURA 11 - PRINCIPAIS ROTAS DO COMÉRCIO TRANSATLÂNTICO DE ESCRAVIZADOS PARA O BRASIL DURANTE OS SÉCULOS XVI AO XIX



FONTE: UFG (2015).

LEGENDA: 1. Cabo Verde; 2. Bissau (Guiné-Bissau); 3. São Jorge da Mina (Gana); 4. Ajudá (Benim); 5. Lagos (Nigéria); 6. São Tomé (São Tomé e Príncipe); 7. Cabinda (Angola); 8. Luanda (Angola); 9. Cassanje (Angola); 10. Benguela (Angola); 11. Moçambique; 12. Rio de Janeiro (Rio de Janeiro, Brasil); 13. Salvador (Bahia, Brasil); 14. Recife (Pernambuco, Brasil); 15. São Luís (Maranhão, Brasil); 16. Belém (Pará, Brasil)

O trabalho escravo, exercido pelos africanos e seus descendentes, perdurou oficialmente até 1888, com o decreto da abolição da escravatura (IBGE, 2007). Tal ato foi motivado por diversos processos, incluindo atos de resistência tomados pelos trabalhadores escravizados, como processos, fugas e organizações (CHALHOUB, 2011; FLORENTINO, 2003; MARQUESE, 2006), e a proibição do tráfico negreiro

transatlântico imposta pela Inglaterra a partir de 1850 (FREITAS, 2009). A necessidade de mão de obra para substituir o escravagismo aliada ao desejo de diminuir a proporção de negros – que correspondiam a mais da metade da população brasileira – impulsionou o processo de “branqueamento”, caracterizado pelo financiamento da política imigratória de europeus para o país desempenhado pelo governo brasileiro, culminando na vinda de mais de 3,5 milhões de imigrantes até 1929, abrangendo mais de 1 milhão de italianos e portugueses, e quantidades inferiores de espanhóis, alemães, poloneses e japoneses (IBGE, 2007; SANTOS, 2002). A lógica racista por trás desta política baseava-se no Determinismo Biológico e em teorias eugenistas, nos quais o aumento da população branca no Brasil promoveria a miscigenação dos imigrantes com os brasileiros, de modo que o país obtivesse uma “raça purificada” e “menos negra” (FREITAS, 2009).

Atualmente, de acordo com a classificação utilizada pelo IBGE (Instituto Brasileiro de Geografia e Estatística), a população brasileira é composta por 45,2% de brancos, 45,1% de pardos e 8,9% de pretos (IBGE, 2016). A distribuição dessas proporções, contudo, varia de acordo com a região do país, havendo predominância de brancos nas regiões Sul e Sudeste (76,7% e 52,9%, respectivamente) e de pardos nas porções Norte, Nordeste e Centro-Oeste (70,2%, 62,0% e 52,0%, respectivamente) (IBGE, 2016). É importante ressaltar que a metodologia pela qual tais dados foram obtidos baseia-se na autodeclaração dos participantes o que, segundo Piza e Rosemberg (1998) e Bastos e colaboradores (2009), pode suscitar muitos erros devido à deturpação da própria identidade ocasionada pelo racismo – tendendo a “embranquecer” sua classificação. Outros autores apontam que o mesmo ocorre quando a classificação é feita por heterodeterminação, sendo a classe social do entrevistado um importante componente utilizado para classificar sua cor, enfatizando a complexidade da classificação racial no Brasil (BASTOS et al., 2009; TELLES, 2002).

Apesar de características físicas como a cor da pele serem comumente utilizadas para prever a ancestralidade dos indivíduos, esse tipo de classificação induz a erros significativos – especialmente em países como o Brasil, em que a miscigenação ocorre amplamente desde sua colonização (SALZANO, 2004). Através de análises em populações do Rio Grande do Sul classificadas como “brancas”, foi identificada acentuada contribuição de marcadores de ancestralidade africana e ameríndia no DNA mitocondrial (mtDNA) desses indivíduos, enquanto que os

resultados de polimorfismos do cromossomo Y indicaram predominância de ancestralidade europeia (MARRERO et al., 2005). As diferenças no padrão de ancestralidade inferidos por marcadores uniparentais em brasileiros já haviam sido constatadas previamente (ALVES-SILVA et al., 2000; CARVALHO-SILVA et al., 2001; RIBEIRO-DOS-SANTOS et al., 2007) e são características também de outras populações latino-americanas miscigenadas e submetidas a processos de colonização (SANS, 2000). Tal similaridade identificada entre as populações é consequência de cruzamentos direcionais – majoritariamente não consentidos, devido ao estabelecimento de relações de poder desiguais entre homens invasores europeus e mulheres nativo-americanas e africanas, assim como seus respectivos descendentes – ocorridos a partir do final do século XV e início do século XVI (LACERDA, 2010; MERRIWETHER et al., 1997).

Parra e colaboradores (2003), utilizando dez alelos autossômicos população-específicos (PSAs, do inglês *population-specific alleles*), estimaram a proporção de ancestralidade africana dos indivíduos da comunidade rural de Queixadinha, em Minas Gerais, evidenciando a ausência de correlação existente entre a cor da pele e ancestralidade em brasileiros. Resultados similares foram obtidos em outros estudos com populações de todo Brasil (FERREIRA et al., 2006; LEITE et al., 2011; LINS et al., 2011; MOURA et al., 2015; MUNIZ et al., 2008; PENA et al., 2009, 2011; SOUZA et al., 2019), demonstrando que a metodologia mais fidedigna e, portanto, mais adequada para inferência de ancestralidade e estrutura populacional seriam os marcadores autossômicos (LEITE et al., 2011; LINS et al., 2011). A partir dessas análises, foi constatada uma composição genética significativamente heterogênea da população brasileira (previamente deduzida, considerando a história do país), havendo predominância de ancestralidade europeia em todas as regiões, com maiores porcentagens na região Sul (variando de 77,0% a 79,5%); seguida da ancestralidade africana, com porcentagens mais elevadas na região Nordeste (variando de 27,0% a 29,3%); e com a menor contribuição de ancestralidade ameríndia, se sobressaindo na região Norte (variando de 18,5% a 32,0%) (MOURA et al., 2015; PENA et al., 2011).

É importante salientar que, no Brasil, o racismo estabelecido baseia-se no fenótipo dos indivíduos, desconsiderando sua ancestralidade, sendo motivado, assim, por razões sociais e não biológicas, de modo que os estudos mencionados não podem

ser utilizados para corroborar o mito da democracia racial no país (KENT; WADE, 2015).

2.8.1 População afro-brasileira do Paraná

O povoamento da região que atualmente corresponde o estado do Paraná iniciou há aproximadamente 10 mil anos (PROUS, 1992), com proto-populações que deram origem aos povos que se estabeleceram nesta localidade – pertencentes aos troncos linguísticos Tupi e Gê (MARTINS, 1995). Os Tupis-Guarani, do tronco Tupi, se instituíram nas porções Norte, Noroeste e litorânea, enquanto as populações Kaingang e Xoklêng (Botocudos), do tronco Jê, se distribuíram de Norte a Sul, no centro e no litoral do estado (WACHOWICZ, 1967).

Com a chegada dos europeus, no final do século XV, o território paranaense tornou-se domínio espanhol devido à demarcação imposta pelo Tratado de Tordesilhas, com apenas uma pequena porção ao Leste do estado sob domínio português (MARTINS, 1995; SCHMIDT, 1996). Somente no século XVII o Paraná passou a pertencer à Coroa Portuguesa, com estabelecimento de populações luso-brasileiras no litoral do estado e, posteriormente, na região de Curitiba (ALEGRO et al., 2013; SWAIN, 1988). Esses povos gradualmente se expandiram pelo restante do estado, com consequente aniquilação de territórios e comunidades ameríndias (MOTA, 1992; SWAIN, 1988). A partir da metade do século XIX, houve estabelecimento dessas populações no Norte do Paraná. As porções Noroeste e Sudoeste do estado foram as últimas a serem ocupadas, já no século XX (SILVA, 2016).

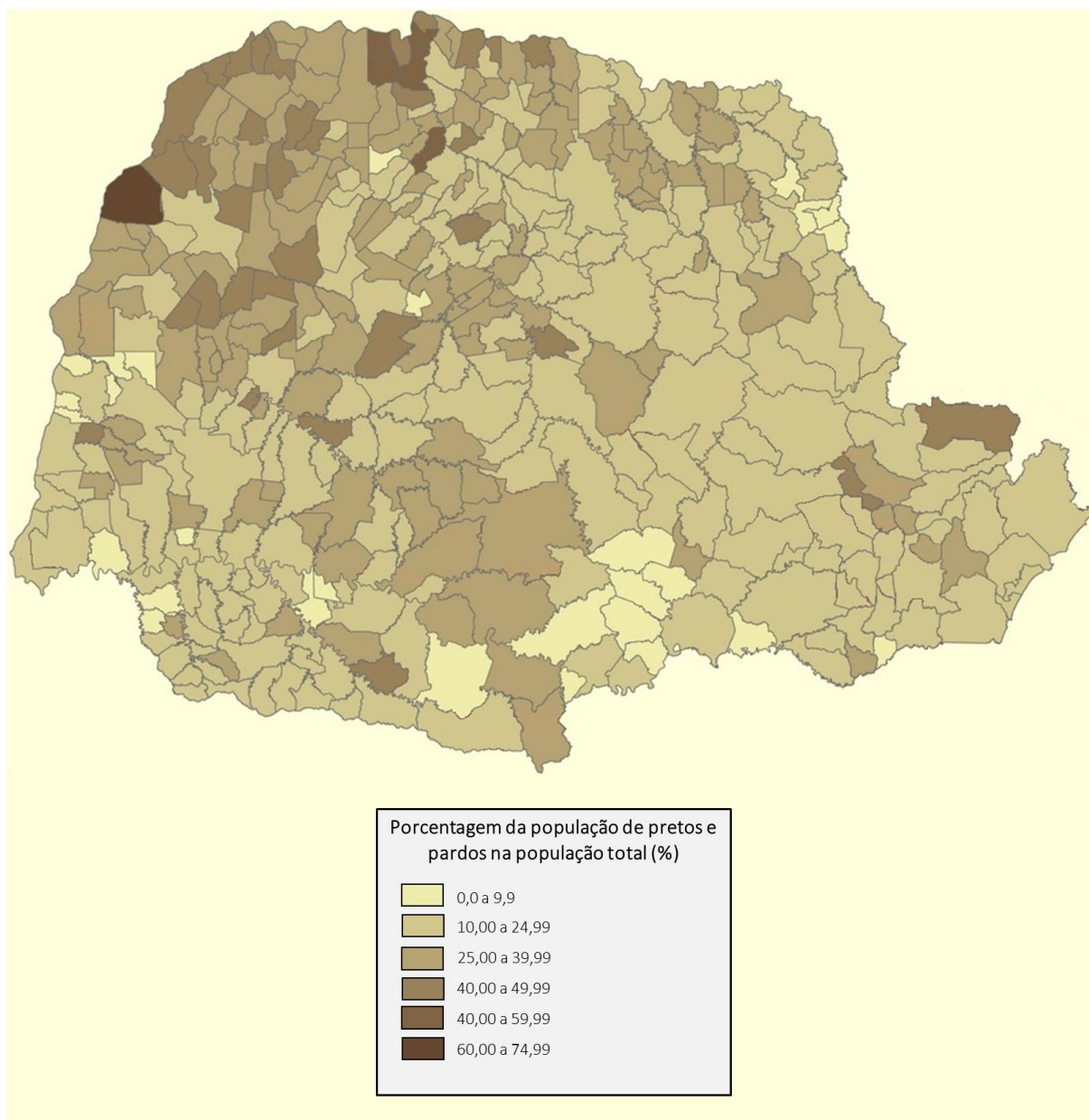
Além dos espanhóis e portugueses mencionados, outras populações europeias foram importantes para a formação da população paranaense. Em 1829, teve início a formação das primeiras colônias do estado formadas por imigrantes alemães, italianos, ucranianos e poloneses – dos quais grande parte tinha sua vinda financiada por instituições privadas e públicas (ALEGRO et al., 2013). Tais colônias, normalmente formadas por mais de uma etnia, tinham como objetivo instituir a agricultura de subsistência para suprir as necessidades do Paraná (PRIORI et al., 2012). De acordo com Martins (1995), estima-se que entre 1829 e 1929 o estado tenha recebido cerca de 130 mil imigrantes europeus.

Segundo Raggio, Bley e Trauczynski (2018), na segunda metade do século XVI já haviam populações negras no território paranaense, escravizadas por colonizadores portugueses que buscavam ouro no litoral. Em 1780, no primeiro levantamento realizado no estado, foi identificado que indivíduos negros e pardos equivaliam a mais de 40% dos paranaenses (MARTINS, 1995). Entre 1640 e 1700 estima-se que 25% da população do Paraná era composta por negros escravizados, e entre 1798 e 1830 a proporção era de cerca de 20% (FELIPE, 2018; GUTIÉRREZ, 1988). Apesar de menos expressiva do que no restante do país, a mão-de-obra escravizada era essencial para a economia colonial e amplamente empregada no estado (RAGGIO; BLEY; TRAUCZYNSKI, 2018). De acordo com o recenseamento de 1804 nas cidades de Antonina e Castro, os escravizados eram majoritariamente brasileiros, ou seja, descendentes de africanos nascidos no Brasil e, dentre os poucos africanos identificados, havia predominância de indivíduos pertencentes a grupos bantos, além de menores proporções de sudaneses (GUTIÉRREZ, 1988).

Na época dos processos abolicionistas, a quantidade de negros escravizados – oficialmente – diminuiu de maneira acentuada no Paraná: de 10500 em 1872 para 3600 em 1887. Contudo, esta redução populacional está associada à morte e transferência dos indivíduos, e não à alforria concedida pelos escravocratas (LAU FILHO, 2018). No século XX, a população preta e parda do estado do Paraná voltou a aumentar, com elevadas taxas de crescimento nos anos 1950 e 1960 (DELGADO, 2018).

De acordo com as estimativas mais recentes (IBGE, 2019), o Paraná possui 11,4 milhões de habitantes, dos quais 3,2% e 25,1% são pretos e pardos, respectivamente (IBGE, 2011) (FIGURA 12). Com isso, é o estado do Sul do país com a maior população negra, equivalente a 28,3% (DELGADO, 2018).

FIGURA 12 - DISTRIBUIÇÃO DA POPULAÇÃO NEGRA NOS MUNICÍPIOS DO ESTADO DO PARANÁ



FONTE: Adaptado de Instituto de Terras, Cartografia e Geologia do Paraná (2010).

Através de polimorfismos dos genes do sistema *HLA*, Probst e colaboradoras (2000) estimaram a origem de populações euro-brasileiras do Paraná. Em euro-brasileiros, a ancestralidade demonstrou ser majoritariamente europeia (80,6%), enquanto que nos afro-brasileiros havia predominância de ancestralidade africana (49,5%) e europeia (41,8%). Similarmente, Braun-Prado e colaboradores (2000) identificaram que os alelos do sistema *HLA* característicos de populações europeias e subsaarianas eram encontrados em todos os grupos populacionais do estado. Alelos de origem asiática e ameríndia apresentavam baixas frequências nessas populações.

Manta e colaboradores (2013), utilizando polimorfismos de inserção e deleção (indels) como marcadores autossômicos informativos de ancestralidade (AIM-Indels), estimaram as proporções de ancestralidade da população paranaense – desconsiderando seu fenótipo. Os resultados obtidos assemelharam-se aos encontrados na população brasileira total, mencionados previamente, havendo predominância europeia (71,0%), seguida de africana (17,5%) e ameríndia (11,5%). Os autores demonstraram também que os paranaenses amostrados tinham uma das mais baixas distâncias genéticas em relação à europeus ($F_{ST} = 0,0125$) – diferentemente do que foi encontrado em demais populações brasileiras (F_{ST} variando de 0,011 a 0,052 em populações da Região Sudeste; de 0,0451 a 0,248 em populações da Região Centro-Oeste; de 0,048 a 0,050 em populações da Região Nordeste e de 0,081 a 0,229 em populações da Região Norte). A distância genética foi calculada a partir do F_{ST} , que compara a variação genética existente entre duas populações com a variação encontrada na população total (HOLSINGER; WEIR, 2009).

2.8.2 Comunidades Quilombolas

De acordo com Moura (1992), diversos tipos de manifestações de resistência foram organizadas pelos escravizados ao longo dos mais de 300 anos de duração, oficial, da política escravocrata brasileira. O autor denomina tais manifestações de “quilombagem”, que incluíam guerrilhas, insurreições e os quilombos.

A definição clássica de quilombo constitui-se da habitação conjunta de escravizados fugitivos, ocorrendo em todo o território brasileiro – embora em diferentes magnitudes e estruturas (CHAGAS, 2001; MOURA, 1986). As várias comunidades quilombolas espalhadas por todo o Brasil – apesar de localizadas em áreas afastadas – não eram organizações isoladas, de modo que outras classes marginalizadas da população brasileira também as integravam, como indígenas e negros livres perseguidos, mulheres sem profissão e prostitutas, desertores do serviço militar, dentre outros (MOURA, 1986, 1992). Apesar disso, Carneiro (2001) elenca os quilombos como locais de reafirmação da cultura africana, cuja expressão e manifestação na sociedade vigente era reprimida. O próprio termo “quilombo” deriva da palavra de origem banta *ki-lombo*, que era a designação dada pelos povos

Imbangalas aos acampamentos em que viviam, na região que corresponde a atual Angola (MUNANGA, 2001; SOUZA, 2008).

Com a abolição do regime escravocrata, os quilombos deixaram de ser um local de fuga a passam a constituir comunidades camponesas, das quais apenas algumas conseguiram se manter no mesmo território devido, principalmente, à Lei das Terras de 1850, que assegurava o direito à terra somente mediante sua compra, expulsando demais ocupantes (FIABANI, 2007). Conseqüentemente, muitas dessas populações – juntamente com povos indígenas e outras comunidades tradicionais – foram marginalizadas socialmente desde então (LEITE, 2008).

Através de esforços dos movimentos sociais negros, a Constituição Brasileira de 1988 passou a reconhecer legalmente as Comunidades Quilombolas, conferindo-lhes o direito à terra (LEITE, 2008) – uma vez que esta é inerente à própria identidade desses grupos (GOMES, 2010; SARMENTO, 2007). Desta forma, alguns autores sugerem uma atualização na definição de quilombo, passando a ser entendido como “direito à terra, enquanto suporte de residência e sustentabilidade [...] nas diversas unidades de agregação das famílias e núcleos populacionais compostos majoritariamente, mas não exclusivamente, de afrodescendentes” (LEITE, 2008). Porém, apesar de terem seu direito garantido pela Constituição Brasileira, tais comunidades continuam lutando pela sua efetivação e são constantemente ameaçadas de expropriação, a partir tanto da criação de projetos de lei visando revogar tal direito, quanto da invasão dessas propriedades desempenhada por agentes expropriadores – muitas vezes assegurados por instituições legais (GOMES, 2010; LEITE, 2010).

É importante salientar que as Comunidades Quilombolas não ficaram isoladas das demais populações brasileiras, embora o grau de interação com os povos vizinhos seja variável (SCHMITT; TURATTI; CARVALHO, 2002). Bortolini e colaboradores (1995) evidenciaram tal variação ao analisarem, através da análise de polimorfismos de 20 genes, as proporções de ancestralidade de quatro Comunidades Quilombolas de diferentes regiões (Curiepe, Trombetas, Cametá e Paredão, localizadas na Venezuela, Pará, Tocantins e Rio Grande do Sul, respectivamente). Os valores obtidos oscilaram entre 46,7% a 80,2% para ancestralidade africana (46,7% em Curiepe; 56,4% em Trombetas; 48,0% em Cametá e 80,2% em Paredão), 17,9% a 36,2% para ancestralidade europeia (36,2% em Curiepe; 23,8% em Trombetas; 17,9% em Cametá e 19,8% em Paredão) e 0,0 a 34,1% para ancestralidade ameríndia

(17,1% em Curiepe; 19,8% em Trombetas; 34,1% em Cameté e 0,0% em Paredão). Um outro estudo, envolvendo a análise de 48 polimorfismos indels em dez comunidades do Vale do Ribeira, no estado de São Paulo, – com ocorrência de fluxo gênico entre elas – obteve variações similares: 7,3 a 69,5% para ancestralidade africana, 12,9 a 68,3% para ancestralidade europeia e 7,3 a 58,5% para ancestralidade ameríndia (KIMURA et al., 2013).

Ainda, através de análises de sete AIMs (Marcadores Informativos de Ancestralidade), Luizon (2007) investigou as proporções de ancestralidade genômica em três Comunidades Quilombolas de diferentes regiões do Brasil: Comunidade Rural de São Gonçalo (BA), Comunidade da Barra (BA) e Comunidade de Sertão do Valongo (SC). Assim como detectado nos trabalhos descritos anteriormente, o autor identificou predominância de ancestralidade africana em todos os quilombos (68,1% em São Gonçalo; 95,4% em Barra e 68,1% em Sertão do Valongo), porém em apenas duas delas foram constatadas contribuições de ancestralidade europeia (21,6% em São Gonçalo e 31,9% em Sertão do Valongo) e ameríndia (10,3% em São Gonçalo e 4,6% em Barra).

2.8.2.1 Sertão do Valongo

A Comunidade Quilombola de Sertão do Valongo localiza-se na zona rural do município litorâneo de Porto Belo, no Sul do estado de Santa Catarina (27°12'12"S 48°44'30"W). Estima-se que sua fundação tenha ocorrido por volta dos anos 1880, através de quatro casais formados por sete negros ex-escravizados e um homem branco (LEMES et al., 2015; SOUZA; CULPI, 1992). Teixeira (1990) sugere que, num período anterior a este, já haviam famílias de origem africana e não-africana estabelecidas na região em questão e que ela seria, na verdade, um ponto de convergência para ex-escravizados provenientes de localidades próximas.

De acordo com Souza e Culpi (2005), em 1995 a comunidade de Valongo era composta por 74 pessoas. Dados do Iphan (Instituto do Patrimônio Histórico e Artístico Nacional) apontam para a existência de 34 famílias compondo a comunidade, entre 2005 e 2008 (BORBA, 2016). Nestas famílias foi identificado um elevado grau de consanguinidade, com o coeficiente médio de endocruzamentos, F , equivalente a 0,081 (SOUZA; CULPI, 2005) – mais elevado que o estimado, na mesma época, para o estado de Santa Catarina ($F = 0,00032$) e para o Brasil ($F = 0,00088$) (FREIRE-

MAIA, 1990). A endogamia seria impulsionada tanto pela forte manifestação religiosa presente na comunidade (convertida à Igreja Adventista no início do século XX), que estimula somente casamentos entre seus membros, quanto da segregação racial por parte das populações vizinhas, compostas tradicionalmente por brancos católicos de origem alemã, italiana e portuguesa (BORBA, 2016; SOUZA; CULPI, 1992, 2005). Castells (2006) sugere ainda que os casamentos endogâmicos, praticados desde a fundação da comunidade, seriam também uma maneira de manter as terras dentro da mesma família, aspecto de grande importância considerando os dois propósitos da propriedade naquele contexto: trabalho (através da agricultura) e moradia.

O diminuto tamanho da população (total e efetiva, $n = 74$ e $n = 21$, respectivamente) de Sertão do Valongo e a pequena taxa de migração (estimada em 4%), constatados por Souza Culpi (2005), são indicativos da deriva genética nesta comunidade. Além disso, a ausência de vários alelos e haplótipos considerados polimórficos, corroboram a ocorrência de efeito fundador, caracterizado pela perda de variabilidade devido aos poucos indivíduos (que continham apenas uma fração de alelos) que originaram a população (MAYR, 1954).

Souza e Culpi (2005) estimaram a ancestralidade da população da comunidade de Valongo, através de marcadores autossômicos clássicos. Os indivíduos analisados possuíam alelos predominantemente de origem africana subsaariana (97,3%) com uma pequena parcela de alelos de origem europeia (2,7%). Luizon (2007) também estimou a ancestralidade desta população, através de AIMs, e, apesar de encontrar o mesmo padrão, os valores obtidos foram distintos dos encontrados no estudo descrito anteriormente (68,1% para ancestralidade africana e 31,9% para ancestralidade europeia).

3 OBJETIVOS

3.1 OBJETIVO GERAL

Caracterizar as populações afro-brasileiras de Curitiba e Região Metropolitana (Paraná) e da Comunidade Quilombola de Sertão de Valongo (Santa Catarina) quanto aos polimorfismos relacionados ao fenótipo da persistência da lactase.

3.2 OBJETIVOS ESPECÍFICOS

- a) Inferir os fenótipos dos indivíduos através da presença das variantes associadas à persistência da lactase, localizadas no íntron 13 do gene *MCM6*;
- b) Descrever as frequências alélicas, genotípicas, fenotípicas e haplotípicas das populações amostradas;
- c) Verificar se as frequências genotípicas das populações se encontram em equilíbrio de Hardy-Weinberg;
- d) Avaliar, através da estimativa dos valores de p , D' e r^2 , se os polimorfismos analisados se encontram em desequilíbrio de ligação;
- e) Comparar as frequências alélicas, genotípicas e haplotípicas encontradas entre as populações de Curitiba e Região Metropolitana e de Sertão do Valongo e verificar se há diferenças entre ambas;
- f) Comparar as frequências das variantes encontradas entre as populações afro-brasileiras analisadas e demais populações brasileiras;
- g) Comparar as frequências das variantes encontradas entre as populações afro-brasileiras analisadas e outras populações latino-americanas;
- h) Comparar as frequências das variantes encontradas entre as populações afro-brasileiras analisadas e populações africanas.

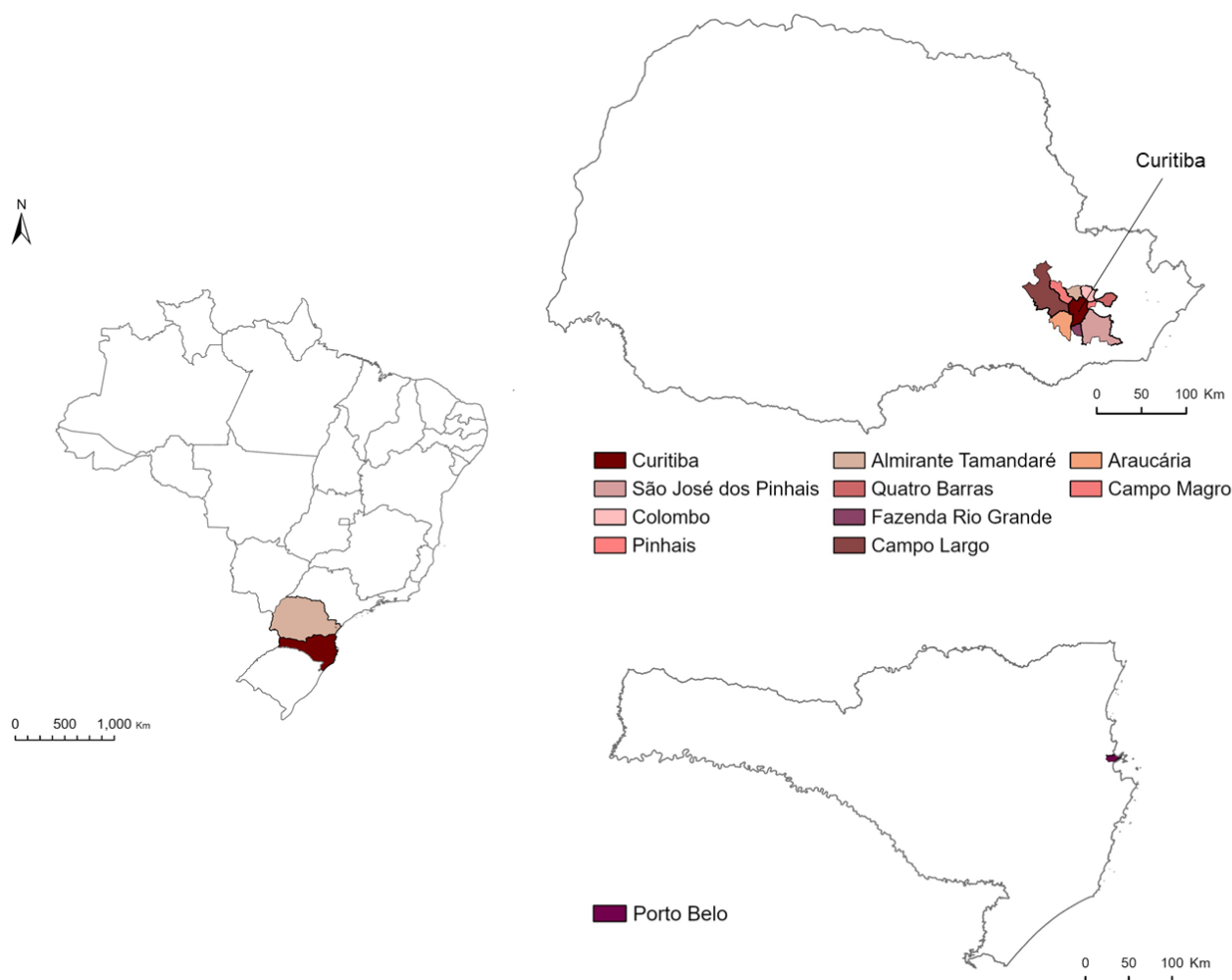
4 MATERIAIS E MÉTODOS

4.1 AMOSTRAS

Foram incluídas 338 amostras de indivíduos afro-brasileiros decorrentes de coletas realizadas previamente, sendo 319 provenientes de Curitiba e Região Metropolitana (PR) e 19 provenientes da Comunidade Quilombola de Sertão do Valongo, na região rural de Porto Belo (SC) (FIGURA 13). Dentre as 319 amostras de Curitiba e Região Metropolitana, 59 foram cedidas pela Prof^a. Dr^a. Maria Luiza Petzl-Erlar e são oriundas do Painel do Laboratório de Genética Molecular Humana (LGMH) da Universidade Federal do Paraná (UFPR), tendo sua coleta e uso aprovados pela Comissão Nacional de Ética em Pesquisa (Conep), número de parecer 2.045.739/2017; 260 foram cedidas pelo Prof. Dr. Ricardo Lehtonen Rodrigues de Souza e pertencem ao Laboratório de Polimorfismos e Ligação (LPL) da mesma instituição, com seu uso aprovado pela Conep, número de parecer 180/2001. Amostras provenientes de 19 indivíduos da Comunidade Quilombola de Sertão do Valongo foram cedidas pela Profa. Dra. Ilíada Rainha de Souza do Laboratório de Polimorfismos Genéticos (LAPOGE) da Universidade Federal de Santa Catarina (UFSC) e possuem aprovação pelo Comitê de Ética dessa mesma instituição.

A classificação dos indivíduos amostrados como afro-brasileiros foi realizada através de diferentes critérios. Na população de Curitiba e Região Metropolitana, baseou-se na heteroclassificação dos indivíduos, desempenhada pelos pesquisadores encarregados de coletar o material biológico em questão. Parte das amostras dessa população foram analisadas quanto à ancestralidade genômica, por meio dos polimorfismos de genes *HLA*, apresentando predominância de componentes genéticos africano (49,5%), seguido de europeu (41,8%) e ameríndio (8,7%) (BRAUN-PRADO et al., 2000; PROBST et al., 2000). A classificação da Comunidade Quilombola de Sertão do Valongo como afro-brasileira, por sua vez, foi realizada levando em conta o histórico de fundação da comunidade por ex-escravizados de ancestralidade predominantemente africana (LEMES et al., 2015; SOUZA; CULPI, 1992).

FIGURA 13 - LOCALIDADES DAS POPULAÇÕES AMOSTRADAS



FONTE: A autora (2020).

4.2 PCR GENE *MCM6*

A partir do DNA extraído foram realizadas PCR (do inglês *polymerase chain reaction*) do íntron 13 do gene *MCM6*, com amplificação de um fragmento de 594 pares de base através de dois iniciadores (RANCIARO et al., 2014) (TABELA 2). Devido a modificações nos reagentes utilizados no laboratório, foram utilizados dois protocolos de PCR: um contemplando as amostras do Painel do Laboratório de Genética Molecular Humana (UFPR) (TABELA 3), enquanto o outro foi utilizado para as amostras provenientes do Laboratório de Polimorfismos e Ligação (UFPR) e do Laboratório de Polimorfismos Genéticos (UFSC) (TABELA 4). Em cada reação de

PCR foram empregados controles positivo e negativo, que respectivamente, correspondiam a amostras de DNA de genótipo conhecido e soluções completas contendo água ao invés de DNA. As soluções foram submetidas ao termociclador sob as condições explicitadas na TABELA 5.

TABELA 2 - INICIADORES ESPECÍFICOS UTILIZADOS NA REAÇÃO DE PCR DO GENE *MCM6*

Iniciador direto (<i>forward</i>)	Iniciador reverso (<i>reverse</i>)
5' GGCAGGGGTTTGGAACTTTC 3' ¹	5' CTGTTGAATGCTCATACGACCA 3' ¹

Fonte: A autora (2020).

¹ RANCIARO et al. (2014).

TABELA 3 - PROTOCOLO DE PCR UTILIZADO PARA AS AMOSTRAS PROVENIENTES DO LABORATÓRIO DE GENÉTICA MOLECULAR HUMANA (UFPR)

Reagentes	[Inicial]	[Final]	Volume (µL)
Tampão (x)	10,00	1,50	1,50
MgCl ₂ (mM)	50,00	1,50	0,30
dNTP (mM)	2,00	0,20	1,00
Iniciador direto (10pmol/µL)	10,00	0,25	0,25
Iniciador reverso (10pmol/µL)	10,00	0,25	0,25
Taq Polimerase Convencional (U/µL)	5,00	0,02	0,04
DNA	-	-	1,50
Água (miliQ)	-	-	5,16
TOTAL	-	-	10,00

FONTE: A autora (2020).

TABELA 4 - PROTOCOLO DE PCR UTILIZADO PARA AS AMOSTRAS PROVENIENTES DO LABORATÓRIO DE LIGAÇÃO E POLIMORFISMOS (UFPR) E DO LABORATÓRIO DE POLIMORFISMOS GENÉTICOS (UFSC)

(Reagentes	[Inicial]	[Final]	Volume (µL)
Master Mix®	-	-	3,33
Iniciador direto (10pmol/µL)	10,00	0,20	0,20
Iniciador reverso (10pmol/µL)	10,00	0,20	0,20
DNA	-	-	1,50
Água (miliQ)	-	-	4,77
TOTAL	-	-	10,00

FONTE: A autora (2020).

NOTA: O Master Mix® foi recebido no laboratório através de doações. O cálculo para seu uso neste protocolo baseou-se na porcentagem de mix utilizado na reação previamente otimizada (TABELA 3).

TABELA 5 - PROGRAMA DO TERMOCICLADOR UTILIZADO NA PCR DO GENE *MCM6*

Etapa	Nº de ciclos	Temperatura (°C)	Tempo
Desnaturação inicial	1	94	2 min
Desnaturação		94	15 s
Anelamento	30	55	30 s
Extensão		72	1 min
Extensão final	1	72	1 min
Espera	1	12	5 min

FONTE: A autora (2020).

As amplificações resultantes foram observadas em corrida eletroforética a 80 volts (V) durante 45 minutos, em gel de agarose 1% corado com *Unisafe Dye®* (Uniscience) – fluorescente sob luz ultravioleta. As amostras que possuíam fragmentos amplificados, verificados através dos padrões eletroforéticos, foram selecionadas para sequenciamento (n = 276, dos quais 258 são provenientes dos afro-brasileiros de Curitiba e Região Metropolitana e 18 são provenientes dos afro-brasileiros da Comunidade Quilombola de Sertão do Valongo).

4.3 SEQUENCIAMENTO

Os produtos de PCR das amostras selecionadas foram submetidos a purificação com a finalidade de eliminar reagentes excedentes da reação (TABELA 6). A reação foi incubada no termociclador por 60 minutos à temperatura de 37°C seguido de 15 minutos à temperatura de 80°C.

TABELA 6 - CONCENTRAÇÃO DOS REAGENTES PARA A REAÇÃO DE PURIFICAÇÃO

Reagentes	Volume (µL)
Exonuclease I	0,40
FastAP	1,00
Água (miliQ)	0,60
TOTAL	2,00

FONTE: A autora (2020).

Para a reação de sequenciamento (TABELA 7) foram utilizadas duas soluções, uma com o iniciador direto e a outra com o iniciador reverso (TABELA 2). As reações

ocorreram em placas de 96 poços e passaram por um programa no termociclador com duração total de 150 minutos (TABELA 8).

TABELA 7 - CONCENTRAÇÃO DOS REAGENTES PARA A REAÇÃO DE SEQUENCIAMENTO

Reagentes	Volume (μL)
Água (miliQ)	5,60
Iniciador (10pmol/ μL)	0,20
Tampão 5x ¹	1,90
Big Dye®	0,30
TOTAL	8,00

FONTE: A autora (2020).

¹ Tampão preparado *in house*.

TABELA 8 - PROGRAMA DO TERMOCICLADOR PARA A REAÇÃO DE SEQUENCIAMENTO

Etapa	Nº de ciclos	Temperatura ($^{\circ}\text{C}$)	Tempo
Desnaturação inicial	1	95	1 min
Desnaturação		95	10 s
Anelamento	25	50	5 s
Extensão		60	4 min
Espera	1	4	1 min

FONTE: A autora (2020).

Posteriormente as soluções passaram por uma nova purificação por precipitação: a placa contendo as reações de sequenciamento foi centrifugada. Em seguida, se adicionou 2,5 μL de EDTA (ácido etilenodiamino tetra-acético) com a concentração de 125 milimolar (mM) e 30 μL de etanol absoluto. Logo após, estando fechada com borracha específica, passou por agitação manual 30 vezes; permaneceu em temperatura ambiente por 15 minutos e foi centrifugada a 18 $^{\circ}\text{C}$ na velocidade de 3220 *rcf* (força centrífuga relativa) por 45 minutos. Com a borracha removida, foi virada abruptamente e centrifugada invertida por 6 segundos, a fim de eliminar o sobrenadante. Então, acrescentou-se 30 μL de etanol 70,0% e a placa foi fechada com a mesma borracha e centrifugada a 25 $^{\circ}\text{C}$ na velocidade de 3220 *rcf* por 5 minutos. Novamente, com a borracha removida, foi virada manualmente de forma rápida; centrifugada de maneira invertida por 6 segundos e permaneceu coberta, em temperatura ambiente, durante 20 minutos com o propósito de finalizar a evaporação do etanol adicionado previamente. Por fim, foi acrescentado 10 μL de Hi-Di

Formamida e as reações de sequenciamento foram injetadas no sequenciador *3500 Series Genetic Analyzers (The Applied Biosystems™)*.

4.4 ANÁLISES BIOINFORMÁTICAS

As sequências obtidas foram alinhadas à sequência consenso de referência GRCh38.p12, disponibilizada no *NCBI (National Center for Biotechnology Information)*, no software *Mutation Surveyor®*, versão 3.30 (*SoftGenetics*). Através do alinhamento, foi possível identificar as variantes nas posições de interesse. No total, foram analisados dez *loci* cuja associação ao fenótipo da persistência da lactase já foi demonstrada ou sugerida, são eles: -13730, -13779, -13907, -13910, -13913, -13915, -13937, -14009, -14010 e -14011.

As frequências alélicas e genóticas, a verificação do equilíbrio de Hardy-Weinberg e os parâmetros estimados para o desequilíbrio de ligação entre *loci* (valor de p , D' e r^2) foram examinados utilizando o pacote "*genetics*", v. 1.3.8.1.2. (WARNES, 2019), implementado na linguagem de programação R (R DEVELOPMENT CORE TEAM, 2020). A estimativa de diversidade genética obtida através dos valores de heterozigosidade foi examinada utilizando o pacote "*pegas*", v. 0.13 (PARADIS, 2010), também implementado na linguagem de programação R (R DEVELOPMENT CORE TEAM, 2020).

A inferência dos fenótipos foi realizada através de contagem direta, concebendo os alelos das posições -13907, -13910, -13915, -14009, -14010 e -14011 de cada genótipo, considerando as correlações descritas na literatura mencionadas previamente (TABELA 9). As posições -13730, -13913, -13937 e -13779 foram desconsideradas para a atribuição do fenótipo LP devido às inconsistências, até o momento, de evidências que corroborem suas associações.

TABELA 9 - ALELOS UTILIZADOS PARA ATRIBUIÇÃO DO FENÓTIPO LP

Posição	Alelo
-13907	G
-13910	T
-13915	G
-14009	G
-14010	C
-14011	T

FONTE: A autora (2020).

Através do algoritmo PHASE (STEPHENS; SMITH; DONNELLY, 2001), implementado no *software* DNASP, versão 6.12.03 (ROZAS et al., 2017), foi realizada a reconstrução dos haplótipos (considerando as posições -13730, -13779, -13907, -13910, -13913, -13915, -13937, -14009, -14010 e -14011). Neste mesmo *software*, foram obtidas as frequências haplotípicas.

Para determinação da diferenciação genética entre as populações afro-brasileiras amostradas foi utilizado o coeficiente de endogamia F_{ST} , que baseia-se na diferença existente entre as frequências alélicas das populações (HOLSINGER; WEIR, 2009). Tal coeficiente foi obtido pelo *software* Arlequin, versão 3.5.2.2 (EXCOFFIER; LISCHER, 2010).

Para determinar a diferenciação entre as frequências fenotípicas, haplotípicas, alélicas e genotípicas entre as populações afro-brasileiras de Curitiba e Região Metropolitana e da Comunidade Quilombola de Sertão do Valongo, foi realizado o teste χ^2 para a igualdade de proporções através do pacote “*stats*”, v. 4.0.2, implementado na linguagem de programação R (R DEVELOPMENT CORE TEAM, 2020). Este mesmo teste foi utilizado para a comparação das frequências alélicas dos afro-brasileiros mencionados com as frequências alélicas de populações brasileiras, latino-americanas e africanas.

5 RESULTADOS

Dentre as 276 amostras sequenciadas, 259 foram selecionadas para análises – excluindo 17 amostras pertencentes à população de Curitiba e Região Metropolitana por serem provenientes de parentes consanguíneos de até 1º grau. Tal critério de exclusão não foi aplicado às amostras da Comunidade Quilombola de Sertão do Valongo pois a elevada consanguinidade é característica da população.

5.1 FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS

Na população afro-brasileira de Curitiba e Região Metropolitana, foram identificados alelos monomórficos, não associados ao fenótipo da persistência da lactase, nas posições -13779, -13907, -13913, -13937, -14009 e -14010. Apesar dos alelos -13915*T e -14011*C serem monomórficos pois suas frequências são superiores a 99,0%, foram encontradas também nesses *loci* as variantes marcadoras para a LP -13915*G e -14011*T, presentes, cada uma, em um único indivíduo. Foram observados alelos polimórficos somente em dois dos dez *loci* analisados, nas posições -13730 e -13910, incluindo a presença da variante associada à LP de origem europeia, -13910*T, em 81 indivíduos. Nesta população foram identificadas ainda as substituições não associadas ao fenótipo -13732 C>A (rs182609517) e -13752 C>T (rs893578023), em um indivíduo cada.

Na população de Sertão do Valongo, alelos polimórficos foram identificados somente na posição -13730, apresentando as variantes T e G. Nas demais posições analisadas, observou-se apenas alelos monomórficos não associados à persistência da lactase (TABELA 10).

TABELA 10 - FREQUÊNCIAS ALÉLICAS DAS VARIANTES DO ÍNTRON 13 DO GENE *MCM6*

<i>Locus</i>	Alelo	REGIÃO DE CURITIBA		SERTÃO DO VALONGO	
		Número de indivíduos (n = 241)	Frequência	Número de indivíduos (n = 18)	Frequência
-13730	T	240	0,975	18	0,833
	G	11	0,025	6	0,167
-13732	C	241	0,998	18	1
	A	1	0,002	0	0
-13752	C	241	0,998	18	1
	T	1	0,002	0	0
-13779	G	241	1	18	1
	C	0	0	0	0
-13907	C	241	1	18	1
	G	0	0	0	0
-13910	C	232	0,813	18	1
	T	81	0,187	0	0
-13913	T	241	1	18	1
	C	0	0	0	0
-13915	T	241	0,998	18	1
	G	1	0,002	0	0
-13937	G	241	1	18	1
	A	0	0	0	0
-14009	T	241	1	18	1
	G	0	0	0	0
-14010	G	241	1	18	1
	C	0	0	0	0
-14011	C	241	0,998	18	1
	T	1	0,002	0	0

FONTE: A autora (2020).

NOTA: Em negrito estão destacados os alelos associados à persistência da lactase.

Nos indivíduos da população de Curitiba e Região Metropolitana, foram observados quatro genótipos associados ao fenótipo da persistência da lactase: -13910 C/T, -13910 T/T, -13915 G/T e -14011 C/T (sendo os dois primeiros os mais abundantes). As frequências genotípicas dos *loci* que apresentam mais de um alelo – -13730, -13910, -13915 e -14011 – encontram-se em equilíbrio de Hardy-Weinberg ($p = 0,13$, $p = 0,83$, $p = 1$ e $p = 1$, respectivamente) (TABELA 11).

Nenhum genótipo associado ao traço LP foi observado na população de Sertão do Valongo. As frequências genotípicas da posição -13730, único *locus* polimórfico observado nesta população, também estão em equilíbrio de Hardy-Weinberg ($p = 1$) (TABELA 11).

TABELA 11 - FREQUÊNCIAS GENOTÍPICAS DOS *LOCI* DO ÍNTRON 13 DO GENE *MCM6* ANALISADOS

<i>Locus</i>	Genótipo	REGIÃO DE	SERTÃO DO
		CURITIBA	VALONGO
		Frequência (n)	Frequência (n)
	<i>TT</i>	0,954 (230)	0,667 (12)
-13730	<i>GT</i>	0,041 (10)	0,333 (6)
	<i>GG</i>	0,004 (1)	
-13732	<i>CC</i>	0,996 (240)	1 (18)
	<i>CA</i>	0,004 (1)	
-13752	<i>CC</i>	0,996 (240)	1 (18)
	<i>CT</i>	0,004 (1)	
-13779	<i>GG</i>	1 (241)	1 (18)
-13907	<i>CC</i>	1 (241)	1 (18)
	<i>CC</i>	0,664 (160)	1 (18)
-13910	<i>CT</i>	0,299 (72)	
	<i>TT</i>	0,037 (9)	
-13913	<i>TT</i>	1 (241)	1 (18)
-13915	<i>TT</i>	0,996 (240)	1 (18)
	<i>GT</i>	0,004 (1)	
-13937	<i>GG</i>	1 (241)	1 (18)
-14009	<i>TT</i>	1 (241)	1 (18)
-14010	<i>GG</i>	1 (241)	1 (18)
-14011	<i>CC</i>	0,996 (240)	1 (18)
	<i>CT</i>	0,004 (1)	

FONTE: A autora (2020).

NOTA: Em negrito estão destacados os genótipos associados à persistência da lactase e suas frequências.

5.2 FREQUÊNCIAS FENOTÍPICAS E HAPLÓTIPOS LP

Através de inferência fenotípica, foi identificado que ambas as populações afro-brasileiras analisadas apresentam majoritariamente o fenótipo de não persistência da lactase. Na população de Curitiba e Região Metropolitana, contudo, o fenótipo da persistência da lactase está presente em 34,0% ($n = 82$) dos indivíduos analisados. Na população de Sertão do Valongo, por sua vez, o fenótipo da não persistência da lactase foi o único encontrado (TABELA 12). Tais diferenças nas frequências fenotípicas são significativas ($p = 0,006$ para o fenótipo LP e $p = 0,006$ para o fenótipo LNP, após correção de continuidade).

TABELA 12 - FREQUÊNCIAS FENOTÍPICAS DAS POPULAÇÕES AFRO-BRASILEIRAS DE CURITIBA E REGIÃO METROPOLITANA E DE SERTÃO DO VALONGO

Fenótipo	REGIÃO DE CURITIBA		SERTÃO DO VALONGO	
	Número de indivíduos (n = 241)	Frequência	Número de indivíduos (n = 18)	Frequência
LP	82	0,340	0	0
LNP	159	0,660	18	1
TOTAL	241	1	18	1

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

As variantes alélicas que conferem o fenótipo LP na população de Curitiba e Região Metropolitana estão presentes em três haplótipos distintos, dos quais um é de origem europeia e um de origem africana, classificados assim devido à presença dos alelos -13910*T e -13915*G, respectivamente. O alelo -14011*T – cuja determinação da população em que surgiu não é determinada – também associado *in vitro* à persistência da lactase, está inserido num haplótipo que contém a variante europeia -13910*T – impedindo a determinação precisa de sua origem (LIEBERT et al., 2016, 2017). Vale ressaltar que o haplótipo contendo o alelo de origem europeia é responsável pela ocorrência do fenótipo em 97,8% dos indivíduos LP (TABELA 13).

TABELA 13 - HAPLÓTIPOS CONTENDO AS VARIANTES ASSOCIADAS AO FENÓTIPO LP NA POPULAÇÃO DA REGIÃO DE CURITIBA

Haplótipos										Origem	Frequência ¹ (n cromossomos)
-13730	-13779	-13907	-13910	-13913	-13915	-13937	-14009	-14010	-14011		
T	G	C	<u>I</u>	T	T	G	T	G	C	EUR	0,978 (88)
T	G	C	<u>I</u>	T	T	G	T	G	<u>I</u>	-	0,011 (1)
T	G	C	C	T	<u>G</u>	G	T	G	C	AFR	0,011 (1)

FONTE: A autora (2020).

LEGENDA: AFR = origem africana; EUR = origem europeia.

NOTA: Em negrito e sublinhado estão destacados os alelos marcadores para o fenótipo da persistência da lactase.

¹ A frequência foi calculada considerando apenas os indivíduos com o fenótipo da persistência da lactase (n = 82), havendo 90 cromossomos contendo os alelos LP.

5.3 HAPLÓTIPOS DA REGIÃO INTENSIFICADORA DO GENE *LCT* E DESEQUILÍBRIO DE LIGAÇÃO

Além dos haplótipos associados ao fenótipo da persistência da lactase descritos anteriormente, foram identificados outros quatro haplótipos na população de Curitiba e Região Metropolitana, diferindo nas posições -13730, -13732 e -13752. Dois destes haplótipos também foram identificados nos indivíduos da Comunidade Quilombola de Sertão do Valongo, sendo os únicos encontrados nesta população. Sendo assim, foram identificados sete haplótipos na população de Curitiba e Região Metropolitana e apenas dois na população da Sertão do Valongo. Além disso, foi observada predominância do mesmo haplótipo, contendo apenas alelos ancestrais, nas duas populações, com frequências de 78,8% na população da Região de Curitiba e de 83,3% em Sertão do Valongo. Tais frequências, de acordo com o teste χ^2 para a igualdade de proporções, são equivalentes ($p = 0,665$, com correção de continuidade aplicada) – diferentemente do outro haplótipo comum a ambas as populações, cujas frequências são significativamente diferentes ($p < 0,001$) (TABELA 14).

TABELA 14 - HAPLÓTIPOS PRESENTES NAS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS

Haplótipos												Fenótipo inferido	Origem	REGIÃO DE CURITIBA	SERTÃO DO VALONGO
-13730	-13732	-13752	-13779	-13807	-13910	-13913	-13915	-13937	-14009	-14010	-14011			Frequência (n cromossomos)	Frequência (n cromossomos)
T	C	C	G	C	C	T	T	G	T	G	C	LNP	-	0,788 (379)	0,833 (30)
T	C	C	G	C	<u>I</u>	T	T	G	T	G	C	LP	EUR	0,183 (88)	0
G	C	C	G	C	C	T	T	G	T	G	C	LNP	AFR ¹	0,023 (11)	0,167 (6)
T	C	C	G	C	<u>I</u>	T	T	G	T	G	<u>I</u>	LP	-	0,002 (1)	0
T	C	C	G	C	C	T	<u>G</u>	G	T	G	C	LP	AFR	0,002 (1)	0
T	A	C	G	C	C	T	T	G	T	G	C	LNP	AFR ²	0,002 (1)	0
T	C	T	G	C	C	T	T	G	T	G	C	LNP	AFR ³	0,002 (1)	0

FONTE: A autora (2020).

LEGENDA: AFR = origem africana; EUR = origem europeia; LP = Lactase persistente, LNP = Lactase não persistente.

NOTA: Em negrito e sublinhado estão destacados os alelos marcadores para o fenótipo da persistência da lactase.

¹ Haplótipo considerado de origem africana considerando a ocorrência do alelo -13730*G exclusivamente em populações africanas.

² Haplótipo considerado de origem africana considerando a ocorrência do alelo -13732*A exclusivamente em populações de origem africana.

³ Haplótipo considerado de origem africana considerando a ocorrência do alelo -13752*T exclusivamente em populações de origem africana.

As análises de desequilíbrio de ligação foram realizadas somente para a população de Curitiba e Região Metropolitana pois, como já mencionado, foi identificado apenas uma variante na população da Comunidade Quilombola de Sertão do Valongo, impedindo as comparações par a par. As combinações envolvendo as

variantes -13732^*A , -13752^*T , 13915^*G e -14011^*T foram excluídas do cálculo pois, por estarem presentes em heterozigose num único indivíduo cada, seus resultados não seriam informativos. Desta forma, o desequilíbrio de ligação foi calculado entre o par de alelos -13730^*G e -13910^*T , os quais não estão em desequilíbrio ($D' = 0,98$, $p = 0,098$, $r^2 = 0,075$).

5.4 DIVERSIDADE E DISTÂNCIA GENÉTICA

A diversidade genética, medida a partir da heterozigosidade esperada média (H_{eM}) para todos os *loci* analisados, encontrada nas populações afro-brasileiras de Curitiba e Região Metropolitana ($H_{eM} = 0,037$) e Sertão do Valongo ($H_{eM} = 0,024$) não difere estatisticamente ($p = 1$, com aplicação de correção de continuidade).

Os maiores coeficientes de heterozigosidade identificados são provenientes da posição -13730 nos indivíduos de Sertão do Valongo e da posição -13910 nos indivíduos da Região de Curitiba. Para essas populações, os valores de heterozigosidade observada e esperada avaliados para cada um dos *loci* não são estatisticamente diferentes (TABELA 15).

TABELA 15 - HETEROZIGOSIDADE ESPERADA E OBSERVADA NAS POPULAÇÕES DA REGIÃO DE CURITIBA E DE SERTÃO DO VALONGO

População	Locus	H_e	H_o	p
REGIÃO DE CURITIBA	-13730	0,049	0,041	0,592
	-13732	0,004	0,004	1
	-13752	0,004	0,004	1
	-13910	0,304	0,299	0,850
	-13915	0,004	0,004	1
	-14011	0,004	0,004	1
SERTÃO DO VALONGO	-13730	0,286	0,333	0,661

FONTE: A autora (2020).

LEGENDA: H_e = Heterozigosidade esperada; H_o = Heterozigosidade observada.

NOTA: Valor de p calculado a partir do teste de hipóteses para igualdade de proporções, considerando H_o como a igualdade de proporções.

Apesar da similaridade existente entre a heterozigosidade esperada média das populações de Curitiba e Região Metropolitana e de Sertão do Valongo, a distância genética entre ambas – obtida pelo valor de F_{ST} – indica uma diferenciação moderada entre elas ($F_{ST} = 0,119$).

5.5 COMPARAÇÕES POPULACIONAIS

5.5.1 Afro-brasileiros de Curitiba e Região e de Sertão do Valongo

Foi observado, através do teste χ^2 para igualdade de proporções, que as frequências dos alelos -13730*G e -13910*T são estatisticamente diferentes ($p < 0,001$ e $p = 0,004$, respectivamente) entre as populações da Região de Curitiba e de Sertão do Valongo. Em contrapartida, as frequências dos alelos -13732*A, -13752*T, -13915*G e -14011*T são estatisticamente semelhantes para ambas as populações ($p = 1$, em todos os casos, com correção de continuidade aplicada).

Quanto às frequências genótípicas, foram identificadas diferenças estatísticas significativas nas proporções dos genótipos -13730 G/T; -13730 T/T e -13910 C/C entre as populações afro-brasileiras paranaense e catarinense analisadas, sendo $p < 0,02$ em todos os casos (TABELA 16).

TABELA 16 - COMPARAÇÃO DAS FREQUÊNCIAS GENOTÍPICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS DE CURITIBA E REGIÃO E DE SERTÃO DO VALONGO

<i>Locus</i>	<i>Genótipo</i>	<i>Valor de p¹</i>
	GG	1
-13730	GT	< 0,001
	TT	< 0,001
-13732	CC	1
	CA	1
-13752	CC	1
	CT	1
	TT	0,867
-13910	CT	0,014
	CC	< 0,001
-13915	GT	1
	TT	1
-14011	CT	1
	CC	1

FONTE: A autora (2020).

NOTA: Estão destacados em negrito valores de p significativos ($< 0,05$), representando frequências estatísticas diferentes.

¹ Valores de p após correção de continuidade, devido ao tamanho reduzido da amostra de Sertão do Valongo.

5.5.2 Populações brasileiras

Considerando o alelo *-13910*T*, as frequências identificadas na população de Curitiba e Região Metropolitana (18,7%) são estatisticamente similares de frequências já reportadas em populações de Belém (PA), Recife (PE), Salvador (BA), afro-brasileiros de Porto Alegre (RS) e Pelotas (RS). Além disso, também não foram identificadas diferenças estatísticas significativas com relação às frequências de indivíduos pardos e pretos, provenientes do estudo de Mattar e colaboradores (2009), que não indicava a origem dos brasileiros em questão. No que se refere às frequências do alelo *-13910*T* na população de Sertão do Valongo (0,0%), foram identificadas proporções similares somente em populações de Fortaleza, em povos ameríndios das etnias Guarani-Kaiowá, Guarani-Ñandeva Kaingang e Xavante das regiões Sul e Central do país e em afro-brasileiros oriundos do já mencionado estudo de Mattar e colaboradores (2009) (TABELA 18).

Para os demais alelos já identificados em populações brasileiras (*-13779*C*, *-13937*A*, *-14010*C* e *-14011*T*), as frequências observadas nos afro-brasileiros analisados são estatisticamente similares das frequências reportadas para outras populações (TABELA 18).

TABELA 17 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE *MCM6* EM POPULAÇÕES BRASILEIRAS

Localidade	População	N	Frequência (%)					Referência
			<i>-13779 G>C</i>	<i>-13910 C>T</i>	<i>-13937 G>A</i>	<i>-14010 G>C</i>	<i>-14011 C>T</i>	
Belém (PA)	População geral	200	0	17,5	0	0	0,5	Friedrich et al. (2012b)
Curitiba (PR)	Euro-brasileiros	292	-	33,0	-	-	-	Boschmann et al. (2016)
Fortaleza (CE)	População geral	119	-	23,9	-	-	-	Ponte et al. (2016)
	Ameríndios e Nipo-brasileiros	98	-	29,1	-	-	-	Hartwig et al. (2016)
Pelotas (RS)	Afro-brasileiros	426	-	16,2	-	-	-	Hartwig et al. (2016)
	Euro-brasileiros	2158	-	38,7	-	-	-	Hartwig et al. (2016)
	Indivíduos miscigenados ¹	161	-	25,5	-	-	-	Hartwig et al. (2016)
Porto Alegre (RS)	Afro-brasileiros	182	0	18,4	0	0,3	0	Friedrich et al. (2012b)
	Euro-brasileiros	337	0	29,5	0	0	0	Friedrich et al. (2012b)
Recife (PE)	População geral	262	0,4	20,4	0,4	0	0,8	Friedrich et al. (2012b)
	Guarani-Kaiowá	84	0	0,6	0	0	0	Friedrich et al. (2012a)
Região Central	Guarani-Ñandeva	59	0	7,6	0	0	0	Friedrich et al. (2012a)
	Xavante	101	0	0,5	0	0	0	Friedrich et al. (2012a)
	Euro-brasileiros	122	-	34,4	-	-	-	Oliveira et al. (2020)
Região Sul	Kaingang	72	0	4,9	0	0	0	Friedrich et al. (2012a)
	Menonitas	151	-	65,0	-	-	-	Boschmann et al. (2016)
Salvador (BA)	Pacientes com Sínd. Metabólica	134	-	18,7	-	-	-	Araujo et al. (2019)
	Branco	399	-	24,7	-	-	-	Mattar et al. (2009)
Não identificada	Nipo-brasileiros	53	-	0	-	-	-	Mattar et al. (2009)
	Pardos	65	-	23,1	-	-	-	Mattar et al. (2009)
	Pretos	50	-	12,0	-	-	-	Mattar et al. (2009)

FONTE: A autora (2020).

TABELA 18 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E DEMAIS POPULAÇÕES BRASILEIRAS

Localidade	População	Valor de <i>p</i>		Referência
		REGIÃO DE CURITIBA SERTÃO DO VALONGO ¹		
-13779 G>C				
Recife (PE)	População geral	0,337	1	Friedrich et al. (2012b)
-13910 C>T				
Belém (PA)	População geral	0,653	0,012	Friedrich et al. (2012b)
Curitiba (PR)	Euro-brasileiros	< 0,001	< 0,001	Boschmann et al. (2016)
Fortaleza (CE)	População geral	0,023	0,057	Ponte et al. (2016)
	Ameríndios e Nipo-brasileiros	0,003	< 0,001	Hartwig et al. (2016)
Pelotas (RS)	Afro-brasileiros	0,249	0,017	Hartwig et al. (2016)
	Euro-brasileiros	< 0,001	< 0,001	Hartwig et al. (2016)
Porto Alegre (RS)	Indivíduos miscigenados ²	0,021	0,001	Hartwig et al. (2016)
	Afro-brasileiros	0,922	0,010	Friedrich et al. (2012b)
	Euro-brasileiros	< 0,001	< 0,001	Friedrich et al. (2012b)
Recife (PE)	População geral	0,485	0,005	Friedrich et al. (2012b)
	Guarani-Kaiowá	< 0,001	1	Friedrich et al. (2012a)
Região Central do Brasil	Guarani-Nandeva	0,004	0,193	Friedrich et al. (2012a)
	Xavante	< 0,001	1	Friedrich et al. (2012a)
	Euro-brasileiros	< 0,001	< 0,001	Oliveira et al. (2020)
Região Sul do Brasil	Kaingang	< 0,001	0,386	Friedrich et al. (2012a)
	Menonitas	< 0,001	< 0,001	Boschmann et al. (2016)
Salvador (BA)	Pacientes com Sínd. Metabólica	0,996	0,009	Araujo et al. (2019)
	Branços	0,015	0,001	Mattar et al. (2009)
Não identificada	Nipo-brasileiros	< 0,001	*	Mattar et al. (2009)
	Pardos	0,262	0,003	Mattar et al. (2009)
	Pretos	0,110	0,067	Mattar et al. (2009)
-13937 G>A				
Recife (PE)	População geral	0,337	1	Friedrich et al. (2012b)
-14010 G>C				
Porto Alegre (RS)	Afro-brasileiros	0,250	1	Friedrich et al. (2012b)
-14011 C>T				
Belém (PA)	População geral	0,458	1	Friedrich et al. (2012b)
Porto Alegre (RS)	Afro-brasileiros	0,385	*	Friedrich et al. (2012b)
	Euro-brasileiros	0,237	*	Friedrich et al. (2012b)
Recife (PE)	População geral	0,210	1	Friedrich et al. (2012b)
	Guarani-Kaiowá	0,555	*	Friedrich et al. (2012a)
Região Central do Brasil	Guarani-Nandeva	0,621	*	Friedrich et al. (2012a)
	Xavante	0,584	*	Friedrich et al. (2012a)
Região Sul do Brasil	Kaingang	0,512	*	Friedrich et al. (2012a)

FONTE: A autora (2020).

NOTA: Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Devido ao tamanho reduzido da amostra, foi aplicada correção de continuidade para os valores de *p* obtidos com as comparações das frequências alélicas da população da Comunidade Quilombola de Sertão do Valongo.

5.5.3 Populações latino-americanas

As frequências do alelo **-13910*T** na maioria das populações miscigenadas do Chile (denominadas *Mestízos*) analisadas são estatisticamente similares das

encontradas na população de Curitiba e Região Metropolitana (18,7%) – com exceção de apenas uma população miscigenada da região de Araucanía, descrita por Fernández e Flores (2014). Além disso, foi identificada similaridade estatística também com as frequências deste alelo na população Rapanui, nativa da Ilha de Páscoa. Com relação à população da Comunidade de Sertão do Valongo (0,0%), não foram encontradas diferenças estatísticas para as frequências alélicas descritas em populações chilenas ameríndias e na população miscigenada da província de Elquí, na região de Coquimbo (TABELA 19, TABELA 20).

Foi identificada diferença estatística significativa em relação às frequências do alelo *-13910*T* na população afro-brasileira de Curitiba e Região Metropolitana quando comparada com a população colombiana caribenha, não havendo diferenças quando a comparação foi realizada com a frequência desse alelo na população de Sertão do Valongo (TABELA 19, TABELA 20).

No Equador, não foram constatadas diferenças estatísticas significativas entre a frequência do alelo *-13910*T* observada na população de Sertão do Valongo e as frequências de afro-equatorianos, de populações ameríndias e de populações miscigenadas da região amazônica do país. Para a proporção da variante na população de Curitiba e Região Metropolitana, entretanto, não foi identificada diferença estatística significativa somente quando comparada com a frequência de populações miscigenadas da região de *Sierra* (TABELA 19, TABELA 20).

Para as populações mexicanas analisadas, não foi observada diferença estatística significativa entre a frequência do alelo *-13910*T* na população de Curitiba e Região Metropolitana e as frequências em populações miscigenadas das cidades de Tepic e Guadalajara. Para a frequência de Sertão do Valongo, não foi constatada diferença em relação às frequências das populações ameríndias Nahuas e Maya (TABELA 19, TABELA 20).

TABELA 19 - FREQUÊNCIAS ALÉLICAS DO POLIMORFISMO -13910 C>T EM POPULAÇÕES LATINO-AMERICANAS

País	Localidade	População	N	Frequência (%) -13910 C>T	Referência
Chile	Araucanía	Mapuche ¹	29	5,0	Fernández; Flores (2014)
Chile	Araucanía	<i>Mestizos</i>	115	12,5	Fernández; Flores (2014)
Chile	Araucanía (Temuco)	<i>Mestizos</i>	115	17,5	Fernández et al. (2016)
Chile	Coquimbo	<i>Mestizos</i>	437	22,0	Montalva et al. (2019)
Chile	Coquimbo (Choapa)	<i>Mestizos</i>	201	2,0	Fernández et al. (2016)
Chile	Coquimbo (Elquí)	<i>Mestizos</i>	13	15,0	Fernández et al. (2016)
Chile	Coquimbo (Limarí)	<i>Mestizos</i>	223	23,0	Fernández et al. (2016)
Chile	Ilha de Páscoa	Rapanui ²	86	13,0	Fernández et al. (2016)
Chile	Não identificada	Ameríndios	43	5,8	Morales et al. (2011)
Chile	Não identificada	Hispânicos	216	22,3	Morales et al. (2011)
Chile	Santiago	<i>Mestizos</i>	116	22,8	Fernández et al. (2016)
Colômbia	Região caribenha	População geral	128	10,0	Mendonza Torres et al. (2012)
Equador	Amazônia	<i>Mestizos</i>	48	7,3	Paz-Y-Miño et al. (2016)
Equador	Costa	Afro-equatorianos	49	6,1	Paz-Y-Miño et al. (2016)
Equador	Costa	<i>Mestizos</i>	124	12,5	Paz-Y-Miño et al. (2016)
Equador	Sierra	Afro-equatorianos	76	10,5	Paz-Y-Miño et al. (2016)
Equador	Sierra	Ameríndios	128	7,4	Paz-Y-Miño et al. (2016)
Equador	Sierra	<i>Mestizos</i>	316	15,2	Paz-Y-Miño et al. (2016)
México	Jalisco	Nahuas ³	86	1,2	Ojeda-Granados et al. (2016)
México	Jalisco (Cuquío)	<i>Mestizos</i>	99	26,8	Ojeda-Granados et al. (2016)
México	Jalisco (Guadalajara)	<i>Mestizos</i>	699	20,2	Ojeda-Granados et al. (2016)
México	Jalisco (Villa Purificación)	<i>Mestizos</i>	32	32,8	Ojeda-Granados et al. (2016)
México	Não identificada	Maya	25	2,0	Bersaglieri et al. (2004)
México	Não identificada	Pima	25	0	Bersaglieri et al. (2004)
México	Nayarit	Huicholes ³	95	0	Ojeda-Granados et al. (2016)
México	Nayarit (Tepic)	<i>Mestizos</i>	185	15,7	Ojeda-Granados et al. (2016)

FONTE: A autora (2020).

¹ População ameríndia do Chile.² População tradicional da Ilha de Páscoa, de ancestralidade australasiana.³ Populações ameríndias do México.

TABELA 20 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E OUTRAS POPULAÇÕES LATINO-AMERICANAS

País	Localidade	População	Valor de <i>p</i>		Referência
			REGIÃO DE CURITIBA	SERTÃO DO VALONGO ¹	
-13910 C>T					
Chile	Araucanía	Mapuche ²	0,010	0,433	Fernández; Flores (2014)
Chile	Araucanía	<i>Mestízos</i>	0,043	0,049	Fernández; Flores (2014)
Chile	Araucanía (Temuco)	<i>Mestízos</i>	0,679	0,014	Fernández et al. (2016)
Chile	Coquimbo	<i>Mestízos</i>	0,152	0,003	Montalva et al. (2019)
Chile	Coquimbo (Choapa)	<i>Mestízos</i>	0,645	0,006	Fernández et al. (2016)
Chile	Coquimbo (Elquí)	<i>Mestízos</i>	0,674	0,056	Fernández et al. (2016)
Chile	Coquimbo (Limari)	<i>Mestízos</i>	0,097	0,002	Fernández et al. (2016)
Chile	Ilha de Páscoa	Rapanui ³	0,079	0,049	Fernández et al. (2016)
Chile	Não identificada	Ameríndios	0,003	0,329	Morales et al. (2011)
Chile	Não identificada	Hispânicos	0,183	0,003	Morales et al. (2011)
Chile	Santiago	<i>Mestízos</i>	0,192	0,003	Fernández et al. (2016)
Colômbia	Região caribenha	População geral	0,002	0,091	Mendonza Torres et al. (2012)
Equador	Amazônia	<i>Mestízos</i>	0,006	0,219	Paz-Y-Miño et al. (2016)
Equador	Costa	Afro-equatorianos	0,002	0,219	Paz-Y-Miño et al. (2016)
Equador	Costa	<i>Mestízos</i>	0,034	0,050	Paz-Y-Miño et al. (2016)
Equador	Sierra	Afro-equatorianos	0,019	0,089	Paz-Y-Miño et al. (2016)
Equador	Sierra	Ameríndios	< 0,001	0,184	Paz-Y-Miño et al. (2016)
Equador	Sierra	<i>Mestízos</i>	0,123	0,022	Paz-Y-Miño et al. (2016)
México	Jalisco	Nahuas ⁴	< 0,001	1	Ojeda-Granados et al. (2016)
México	Jalisco (Cuquio)	<i>Mestízos</i>	0,019	0,001	Ojeda-Granados et al. (2016)
México	Jalisco (Guadalajara)	<i>Mestízos</i>	0,476	0,005	Ojeda-Granados et al. (2016)
México	Jalisco (Villa Purificación)	<i>Mestízos</i>	0,008	< 0,001	Ojeda-Granados et al. (2016)
México	Não identificada	Maya	0,003	1	Bersaglieri et al. (2004)
México	Não identificada	Pima	0,001	*	Bersaglieri et al. (2004)
México	Nayarit	Huicholes ⁴	< 0,001	*	Ojeda-Granados et al. (2016)
México	Nayarit (Tepic)	<i>Mestízos</i>	0,253	0,021	Ojeda-Granados et al. (2016)

FONTE: A autora (2020).

NOTA: Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Devido ao tamanho reduzido da amostra, foi aplicada correção de continuidade para os valores de *p* obtidos com as comparações das frequências alélicas da população da Comunidade Quilombola de Sertão do Valongo.

² População ameríndia do Chile.

³ População tradicional da Ilha de Páscoa, de ancestralidade australasiana.

⁴ Populações ameríndias do México.

Através de dados disponibilizados pelo 1000 Genomes Project – Fase 3 (THE 1000 GENOMES PROJECT CONSORTIUM, 2015) e pelo *gnomAD* (*The Genome Aggregation Database*) - versão 3 (KARCZEWSKI et al., 2020), foi possível realizar comparações com outras frequências alélicas além da variante -13910*T. Mediante as comparações feitas com as frequências dos alelos -13730*G (2,5%), -13910*T (18,7%), -13913*C (0,0%), -13915*G (0,2%) e -14011*T (0,2%) na população afro-brasileira de Curitiba e Região Metropolitana, não foram observadas diferenças estatísticas com as frequências de nenhum dos alelos na população de ancestralidade mexicana de Los Angeles (EUA). Para as demais populações comparadas, foi identificado ao menos um alelo com frequência estatisticamente diferente. De acordo

com as comparações realizadas com a população de Sertão do Valongo (frequências alélicas de 16,7% para o alelo -13730*G e de 0,0% para os alelos -13910*T, -13913*C, -13915*G e -14011*T), constatou-se que as frequências das variantes -13730*G e -13910*T são estatisticamente diferentes das frequências nos demais grupos analisados. Os alelos -13913*C, -13915*G e -14011*T, por sua vez, não apresentaram diferenças estatísticas com relação às suas frequências nas populações comparadas (TABELA 21, TABELA 22).

TABELA 21 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE *MCM6* EM POPULAÇÕES LATINO-AMERICANAS DO 1000 GENOMES PROJECT E *gnomAD*

País / Localidade	População	N	Frequência (%)								Referência
			-13730 T>G	-13779 G>C	-13910 C>T	-13913 T>C	-13915 T>G	13937 G>A	-14010 G>C	-14011 C>T	
Colômbia	CLM	94	0,5	0	30,9	0	0	0	0	0	1000 Genomes Project - Fase 3
Estados Unidos	MXL	64	0,8	0	24,2	0	0	0	0	0	1000 Genomes Project - Fase 3
Peru	PEL	85	0	0	10,6	0	0	0	0	0	1000 Genomes Project - Fase 3
Porto Rico	PUR	104	0,5	0	20,7	0	1,4	0	0	0	1000 Genomes Project - Fase 3
Latinos / Americanos miscigenados		13650	0,8	0	23,1	0,1	0,2	0	0	0,1	<i>gnomAD</i> - versão 3

FONTE: A autora (2020).

LEGENDA: CLM = Colombianos em Medellín (Colômbia); MXL = Ancestralidade Mexicana em Los Angeles (Estados Unidos); PEL = Peruanos em Lima (Peru); PUR = Porto-riquenhos em Porto Rico.

TABELA 22 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E POPULAÇÕES LATINO-AMERICANAS DO 1000 GENOMES PROJECT E *gnomAD*

População	-13730 T>G		-13910 C>T		-13913 T>C		-13915 T>G		-14011 C>T	
	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹
CLM ²	0,099	< 0,001	0,001	< 0,001	-	-	0,532	*	0,532	*
MXL ²	0,234	< 0,001	0,162	0,002	-	-	0,606	*	0,606	*
PEL ²	0,038	< 0,001	0,015	0,086	-	-	0,552	*	0,552	*
PUR ²	0,075	< 0,001	0,541	0,006	-	-	0,050	1	0,511	*
Latinos / Americanos miscigenados ³	< 0,001	< 0,001	0,022	0,002	0,515	1	0,879	1	0,203	1

FONTE: A autora (2020).

LEGENDA: CLM = Colombianos em Medellín (Colômbia); MXL = Ancestralidade Mexicana em Los Angeles (Estados Unidos); PEL = Peruanos em Lima (Peru); PUR = Porto-riquenhos em Porto Rico.

NOTA: Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Devido ao tamanho reduzido da amostra, foi aplicada correção de continuidade para os valores de *p* obtidos com as comparações das frequências alélicas da população da Comunidade Quilombola de Sertão do Valongo.

² Dados obtidos do 1000 Genomes Project - Fase 3 (THE 1000 GENOMES PROJECT CONSORTIUM, 2015).

³ Dados obtidos do *gnomAD* - versão 3 (KARCZEWSKI et al., 2020).

5.5.4 Populações africanas

A frequência do alelo -13730*G observada na população afro-brasileira de Curitiba e Região Metropolitana (2,5%) foi estatisticamente similar às frequências

desta variante na maioria das populações africanas analisadas, divergindo apenas das descritas para a população Jaali do Sudão e para as populações gerais da África Central, Ocidental e Setentrional (TABELA 23, TABELA 24). Quanto às comparações realizadas com as mesmas populações, porém com as frequências do alelo -13730*G na Comunidade Quilombola de Sertão do Valongo (16,7%), foram identificadas diferenças estatísticas significativas em populações da Angola, Etiópia, Moçambique e Sudão bem como na população geral do Sul do continente africano (TABELA 23, TABELA 25).

Para a população de Curitiba e Região, as frequências dos alelos -13907*G (0,0%) e -13910*T (18,7%) foram estatisticamente diferentes de todas as frequências comparadas de populações africanas (TABELA 23, TABELA 24). Para a população de Sertão do Valongo, por sua vez, a frequência da variante -13907*G (0,0%) demonstrou diferenças estatisticamente significativas das relatadas para os povos Afar e Tigre da Etiópia, e para a população Beja Beni Amer do Sudão. Para o alelo -13910*T (0,0%), não foram observadas diferenças com relação à sua frequência nos grupos comparados (TABELA 23, TABELA 25).

Similar ao relatado para o alelo -13910*T, foram observadas diferenças estatísticas significativas na maioria das populações africanas analisadas com relação à frequência da variante -13913*C na população de Curitiba e Região Metropolitana (0,0%) (TABELA 23, TABELA 24). Na população de Sertão do Valongo – em que a frequência da variante -13913*C também é de 0,0% –, nenhuma das comparações realizadas com as frequências do alelo em questão em populações africanas demonstrou diferença estatística (TABELA 23, TABELA 25).

A frequência do alelo -13915*G na população afro-brasileira de Curitiba e Região (0,2%) não difere estatisticamente das frequências comparadas de populações das porções Leste, Sul e Central do continente africano, divergindo, contudo, das reportadas em populações do Norte da África, e da Etiópia e Moçambique, na África Oriental (TABELA 23, TABELA 24). De maneira semelhante, a frequência da variante -13915*G na população de Sertão do Valongo (0,0%) difere estatisticamente apenas da frequência de povos da África Oriental, mais precisamente da Etiópia e Sudão (TABELA 23, TABELA 25).

Quanto a frequência do alelo -13937*A na população de Curitiba e Região Metropolitana (0,0%), foram observadas diferenças estatísticas significativas em relação à sua frequência em duas das quatro populações comparadas, sendo estas

provenientes das regiões Sul e Leste da África (TABELA 23, TABELA 24). Em contrapartida, nas comparações da frequência do alelo -13937*A na população de Sertão do Valongo (0,0%) com as mesmas quatro populações, não foram constatadas diferenças estatísticas (TABELA 23, TABELA 25).

As frequências dos alelos -14009*G e -14010*C na população afro-brasileira paranaense estudada (0,0% para ambas as variantes) apresentam diferenças estatísticas significativas das frequências identificadas na maioria das populações africanas comparadas, provenientes de todas as regiões do continente (TABELA 23, TABELA 24). A comparação desses alelos, porém com as frequências observadas em Sertão do Valongo (também 0,0% para ambas as variantes), indicam que há diferenças estatísticas significativas somente com a frequência do alelo -14009*G no povo Beja Beni Amer, do Sudão, e com a frequência do alelo -14010*C em populações da Angola e da África Oriental (TABELA 23, TABELA 25).

As frequências do alelo -14011*T descritas nas populações africanas analisadas (TABELA 23) não diferem estatisticamente das frequências desta variante observadas nas populações afro-brasileiras de Curitiba e Região Metropolitana (0,2%) (TABELA 24) e da Comunidade Quilombola de Sertão do Valongo (0,0%) (TABELA 25).

As frequências dos alelos -13730*G, -13907*G, 13910*T, -13913*C, -13915*G, 13937*A, -14009*G, -14010*C e -14011*T em populações de origem ou, de fato, africanas, disponibilizadas no *1000 Genomes Project* – Fase 3 (TABELA 26) (THE 1000 GENOMES PROJECT CONSORTIUM, 2015), quando comparadas com as frequências alélicas observadas nos afro-brasileiros de Curitiba e Região (2,5%, 0,0%, 18,7%, 0,0%, 0,2%, 0,0%, 0,0%, 0,0% e 0,2%, respectivamente), demonstraram diferenças significativas quanto à frequência das variantes -13730*G e -13910*T em todos os grupos comparados, com exceção da população Mende, de Serra Leoa, e dos afro-americanos do Sudoeste dos Estados Unidos, para o primeiro e segundo alelo citados, respectivamente. Além disso, foi constatada diferença estatística também quanto à frequência do alelo -13937*A na população estadunidense mencionada (TABELA 26, TABELA 27). As comparações das frequências alélicas das mesmas populações africanas com as observadas na Comunidade de Sertão do Valongo (16,7% para o alelo -13730*G e 0,0% para os demais), resultaram na identificação de apenas duas populações com diferenças estatísticas significativas. Uma delas foi a população Mende, de Serra Leoa, em relação a frequência do alelo -

13730*G, e a segunda, a população afro-americana do Sudoeste dos Estados Unidos, em relação às frequências dos alelos -13910*T e -13937*A (TABELA 26, TABELA 27).

Similarmente, para comparações realizadas com dados de indivíduos africanos / afro-americanos do *gnomAD*, versão 3 (KARCZEWSKI et al., 2020), foram identificadas diferenças estatísticas significativas quanto às frequências dos alelos -13730*G e -13910*T entre os afro-brasileiros analisados e o grupo comparado. Foi identificada também diferença com a frequência do alelo -13915*G na população de Curitiba e Região. Para as demais variantes comparadas, não foram observadas diferenças estatísticas significativas (TABELA 26, TABELA 27).

TABELA 23 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE MCM6 EM POPULAÇÕES AFRICANAS

País / Região	População	N	Frequência (%)										Referência
			-13730 T>G	-13779 G>C	-13907 C>G	-13910 C>T	-13913 T>C	-13915 T>G	-13937 G>A	-14009 T>G	-14010 G>C	-14011 C>T	
África Austral	População geral	63	0	0	0	0	1,6	0	2,4	0	5,6	0	Ranciaro et al. (2014)
África Central	População geral	229	11,8	0	0	3,7	0,2	0,4	0	0,2	0	0	Ranciaro et al. (2014)
África Ocidental	População geral	26	9,6	0	0	0	0	0	0	0	0	1,9	Ranciaro et al. (2014)
África Oriental	População geral	444	3,5	0,1	1,8	0	1,0	5,5	0,2	0	19,3	0	Ranciaro et al. (2014)
África Setentrional	População geral	57	7,0	0	8,8	10,5	0,9	6,1	0	7,0	0	0	Ranciaro et al. (2014)
Angola	Himba	69	3,6	-	-	7,2	-	0	0,7	-	8,7	-	Pinto et al. (2016)
Angola	Kuvale	81	5,6	-	-	0,6	-	0	0	-	3,7	-	Pinto et al. (2016)
Angola	Kwepe	34	0	-	-	0	-	0	0	-	4,4	-	Pinto et al. (2016)
Angola	Kwisi	20	0	-	-	0	-	0	0	-	17,5	-	Pinto et al. (2016)
Angola	Tjimba	15	0	-	-	0	-	0	0	-	23,3	-	Pinto et al. (2016)
Angola	Twa	18	0	-	-	0	-	0	0	-	19,4	-	Pinto et al. (2016)
Angola	!Xun	20	-	-	-	0	5,0	-	-	-	5,0	-	Breton et al. (2014)
Botsuana	Tuu ¹	42	-	-	-	0	3,6	-	-	-	2,4	0	Macholdt et al. (2014)
Botsuana e Namíbia	Khoe ¹	226	-	-	-	0,2	2,4	-	-	-	11,3	0	Macholdt et al. (2014)
Botsuana e Namíbia	Kx'a ¹	73	-	-	-	0	0,7	-	-	-	4,1	0	Macholdt et al. (2014)
Botsuana, Namíbia e Zâmbia	Bantu ¹	116	-	-	-	0	0	-	-	-	3,9	0,4	Macholdt et al. (2014)
Camarões	Mambila	20	7,5	-	0	0	-	0	-	0	-	-	Liebert et al. (2017)
Etiópia	Afar	62	3,2	-	24,2	0	-	18,5	-	0,8	-	-	Liebert et al. (2017)
Etiópia	Amhara	54	5,6	-	1,9	0	1,9	5,6	-	4,6	0	0	Jones et al. (2013)
Etiópia	Oromo Borana	104	2,4	-	7,7	0	0	22,1	-	1,0	4,8	-	Jones et al. (2015)
Etiópia	Oromo Harar	88	2,8	-	10,2	1,1	0,6	12,5	-	0,6	1,7	-	Jones et al. (2015)
Etiópia	Oromo Salale	88	2,3	-	2,8	0	3,4	9,1	-	3,4	0,6	-	Jones et al. (2015)
Etiópia	Oromo Wallaga-Begi	86	2,3	-	4,7	0	1,2	10,5	-	1,2	0	-	Jones et al. (2015)
Etiópia	Shabo	21	2,4	-	0	0	-	0	-	0	-	-	Liebert et al. (2017)
Etiópia	Tigre	44	2,3	-	19,3	0	0	11,4	-	3,4	0	0	Jones et al. (2013)
Etiópia	Wolayita	26	1,9	-	7,8	0	0	11,5	-	3,9	1,9	0	Jones et al. (2013)
Gana	Asante	20	7,5	-	0	0	-	0	-	0	-	-	Liebert et al. (2017)
Malawi	Chewa	20	2,5	-	0	0	-	0	-	0	-	-	Liebert et al. (2017)
Moçambique	Makua	14	3,6	-	-	0	-	0	0	-	0	-	Pinto et al. (2016)
Moçambique	Nyanja	11	9,1	-	-	0	-	0	0	-	0	-	Pinto et al. (2016)
Moçambique	Shangaan	23	0	-	-	0	-	0	2,2	-	2,2	-	Pinto et al. (2016)
Moçambique	Tswa	18	0	-	-	0	-	0	0	-	0	-	Pinto et al. (2016)
Moçambique	Yao	15	3,3	-	-	0	-	0	0	-	0	-	Pinto et al. (2016)
Sudão	Beja Beni Amer	64	3,9	-	17,2	0	-	26,6	-	13,3	-	-	Liebert et al. (2017)
Sudão	Jaali	38	7,9	-	1,3	1,3	-	19,7	-	9,2	-	-	Liebert et al. (2017)
Tanzânia	Chagga	41	4,9	-	0	0	-	0	-	0	-	-	Liebert et al. (2017)

FONTE: A autora (2020).

¹ Referente a populações que falam línguas pertencentes às famílias linguísticas Tuu, Khoe, Kx'a e Bantu.

TABELA 24 - VALORES DE *P* COM RELAÇÃO À SIMILARIDADE DAS FREQUÊNCIAS ALÉLICAS ENTRE A POPULAÇÃO AFRO-BRASILEIRA DA REGIÃO DE CURITIBA E POPULAÇÕES AFRICANAS

	País	População	-13730 T>G	-13907 C>G	-13910 C>T	-13913 T>C	-13915 T>G	-13937 G>A	-14009 T>G	-14010 G>C	-14011 C>T	Referência	
	-	População geral	0,074	*	< 0,001	0,006	0,609	0,001	*	< 0,001	0,609	Ranciaro et al. (2014)	
África Austral	Botsuana	Tuu ¹	-	-	< 0,001	< 0,001	-	-	-	0,001	0,676	Macholdt et al. (2014)	
	Botsuana e Namíbia	Khoe ¹	-	-	< 0,001	0,001	-	-	-	< 0,001	0,332	Macholdt et al. (2014)	
	Botsuana e Namíbia	Kx'a ¹	-	-	< 0,001	0,069	-	-	-	< 0,001	0,582	Macholdt et al. (2014)	
	Botsuana, Namíbia e Zâmbia	Bantu ¹	-	-	< 0,001	*	-	-	-	< 0,001	0,597	Macholdt et al. (2014)	
	-	População geral	< 0,001	*	< 0,001	0,305	0,533	*	0,305	*	0,329	Ranciaro et al. (2014)	
África Central	-	Himba	0,472	-	< 0,001	-	0,592	0,061	*	< 0,001	-	Pinto et al. (2016)	
	-	Kuvale	0,057	-	< 0,001	-	0,562	*	*	< 0,001	-	Pinto et al. (2016)	
	-	Kwepe	0,188	-	< 0,001	-	0,707	*	*	< 0,001	-	Pinto et al. (2016)	
	Angola	Kwisi	0,313	-	0,003	-	0,773	*	*	< 0,001	-	Pinto et al. (2016)	
	-	Tjimba	0,382	-	0,009	-	0,800	*	*	< 0,001	-	Pinto et al. (2016)	
	-	Twa	0,338	-	0,004	-	0,784	*	*	< 0,001	-	Pinto et al. (2016)	
	-	IXun	-	-	0,003	< 0,001	-	-	-	< 0,001	-	Breton et al. (2014)	
	Camargões	Mambila	0,068	*	0,003	-	0,773	-	*	-	-	Liebert et al. (2017)	
	-	População geral	0,005	*	0,001	*	0,742	*	*	*	0,054	Ranciaro et al. (2014)	
	Gana	Asante	0,068	*	0,003	-	0,773	-	*	-	-	Liebert et al. (2017)	
África Ocidental	-	População geral	0,310	0,003	< 0,001	0,027	< 0,001	0,297	*	< 0,001	0,175	Ranciaro et al. (2014)	
	-	Afar	0,648	< 0,001	< 0,001	*	< 0,001	-	0,048	-	-	Liebert et al. (2017)	
	-	Amhara	0,094	0,003	< 0,001	0,003	< 0,001	-	< 0,001	*	0,636	Jones et al. (2013)	
	-	Oromo Borana	0,947	< 0,001	< 0,001	*	< 0,001	-	0,031	< 0,001	-	Jones et al. (2015)	
	-	Oromo Harar	0,802	< 0,001	< 0,001	0,098	< 0,001	-	0,098	0,004	-	Jones et al. (2015)	
	Etiópia	Oromo Salale	0,873	< 0,001	< 0,001	< 0,001	< 0,001	-	< 0,001	0,098	-	Jones et al. (2015)	
	-	Oromo Wallaga-Begi	0,905	< 0,001	< 0,001	0,018	< 0,001	-	0,018	*	-	Jones et al. (2015)	
	-	Shabo	0,965	*	0,002	-	0,768	-	*	-	-	Liebert et al. (2017)	
	-	Tigre	0,904	< 0,001	< 0,001	*	< 0,001	-	< 0,001	*	0,669	Jones et al. (2013)	
	África Oriental	Wolayita	0,801	< 0,001	0,001	*	< 0,001	-	< 0,001	0,002	0,742	Jones et al. (2013)	
	Malawi	Chewa	0,997	*	0,003	-	0,773	-	*	-	-	Liebert et al. (2017)	
	-	Makua	0,724	-	0,012	-	0,809	*	*	*	-	Pinto et al. (2016)	
	Moçambique	Nyanja	0,065	-	0,025	-	0,831	*	*	*	-	-	Pinto et al. (2016)
Shangaan		0,279	-	0,001	-	0,757	0,001	*	0,001	-	-	Pinto et al. (2016)	
Tswa		0,338	-	0,004	-	0,784	*	*	*	-	-	Pinto et al. (2016)	
-		Yao	0,776	-	0,009	-	0,803	*	*	*	-	-	Pinto et al. (2016)
Sudão		Beja Beni Amer	0,387	< 0,001	< 0,001	-	< 0,001	-	< 0,001	-	-	-	Liebert et al. (2017)
-		Jaali	0,013	0,012	< 0,001	-	< 0,001	-	< 0,001	-	-	-	Liebert et al. (2017)
Tanzânia		Chagga	0,229	*	< 0,001	-	0,680	-	*	-	-	-	Liebert et al. (2017)
África Setentrional	-	População geral	0,016	< 0,001	0,038	0,040	< 0,001	*	< 0,001	*	0,626	Ranciaro et al. (2014)	

FONTE: A autora (2020).

NOTA: Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Referente a populações que falam línguas pertencentes às famílias linguísticas Tuu, Khoe, Kx'a e Bantu.

TABELA 25 - VALORES DE *P* COM RELAÇÃO À SIMILARIDADE DAS FREQUÊNCIAS ALÉLICAS ENTRE A POPULAÇÃO AFRO-BRASILEIRA DA COMUNIDADE QUILOMBOLA DE SERTÃO DO VALONGO E POPULAÇÕES AFRICANAS

	País	População	-13730 T>G	-13907 C>G	-13910 C>T	-13913 T>C	-13915 T>G	-13937 G>A	-14009 T>G	-14010 G>C	-14011 C>T	Referência
	-	População geral	< 0,001	*	*	1	*	0,815	*	0,327	*	Ranciaro et al. (2014)
África Austral	Botsuana	Tuu ¹	-	-	*	0,610	-	-	-	0,876	*	Macholdt et al. (2014)
	Bostuana e Namíbia	Khoe ¹	-	-	1	0,716	-	-	-	0,065	*	Macholdt et al. (2014)
	Bostuana e Namíbia	Kx'a ¹	-	-	*	1	-	-	-	0,474	*	Macholdt et al. (2014)
	Botsuana, Namíbia e Zâmbia	Bantu ¹	-	-	*	*	-	-	-	0,481	1	Macholdt et al. (2014)
	-	População geral	0,550	*	0,483	1	1	*	1	*	*	Ranciaro et al. (2014)
África Central		Himba	0,013	-	1	-	*	1	*	0,143	-	Pinto et al. (2016)
		Kuvale	0,054	-	1	-	*	*	*	0,525	-	Pinto et al. (2016)
		Kwepe	0,002	-	*	-	*	*	*	0,507	-	Pinto et al. (2016)
	Angola	Kwisi	0,024	-	*	-	*	*	*	0,025	-	Pinto et al. (2016)
		Tjimba	0,055	-	*	-	*	*	*	0,008	-	Pinto et al. (2016)
		Twa	0,033	-	*	-	*	*	*	0,017	-	Pinto et al. (2016)
		!Xun	-	-	*	0,521	-	-	-	0,521	-	Breton et al. (2014)
África Ocidental	Camarões	Mambila	0,379	*	*	-	*	-	*	-	-	Liebert et al. (2017)
	-	População geral	0,512	*	*	*	*	*	*	*	1	Ranciaro et al. (2014)
	Gana	Asante	0,379	*	*	-	*	-	*	-	-	Liebert et al. (2017)
	-	População geral	< 0,001	0,872	*	1	0,285	1	*	0,007	*	Ranciaro et al. (2014)
		Afar	0,011	0,002	*	-	0,012	-	1	-	-	Liebert et al. (2017)
		Amhara	0,082	1	*	1	0,336	-	0,431	*	*	Jones et al. (2013)
		Oromo Borana	0,001	0,175	*	*	0,004	-	1	0,375	-	Jones et al. (2015)
		Oromo Harar	0,003	0,093	1	1	0,052	-	1	0,988	-	Jones et al. (2015)
	Etiópia	Oromo Salale	0,001	0,674	*	0,567	0,125	-	0,567	1	-	Jones et al. (2015)
		Oromo Wallaga-Begi	0,001	0,399	*	1	0,088	-	1	*	-	Jones et al. (2015)
África Oriental		Shabo	0,071	*	*	-	*	-	*	-	-	Liebert et al. (2017)
		Tigre	0,011	0,011	*	*	0,081	-	0,633	*	*	Jones et al. (2013)
		Wolayita	0,035	0,237	*	*	0,093	-	0,643	1	*	Jones et al. (2013)
	Malawi	Chewa	0,083	*	*	-	*	-	*	-	-	Liebert et al. (2017)
		Makua	0,207	-	*	-	*	*	*	*	-	Pinto et al. (2016)
		Nyanja	0,675	-	*	-	*	*	*	*	-	Pinto et al. (2016)
	Moçambique	Shangaan	0,014	-	*	-	*	1	*	1	-	Pinto et al. (2016)
		Tswa	0,033	-	*	-	*	*	*	*	-	Pinto et al. (2016)
		Yao	0,177	-	*	-	*	*	*	*	-	Pinto et al. (2016)
		Beja Beni Amer	0,020	0,017	*	-	0,001	-	0,045	-	-	Liebert et al. (2017)
	Jaali	0,283	1	1	-	0,010	-	0,144	-	-	Liebert et al. (2017)	
	Chagga	0,079	*	*	-	*	-	*	-	-	Liebert et al. (2017)	
África Setentrional	-	População geral	0,160	0,145	0,094	1	0,285	*	0,227	*	*	Ranciaro et al. (2014)

FONTE: A autora (2020).

NOTA: Devido ao tamanho reduzido da amostra, foi aplicada correção de continuidade para os valores de *p* obtidos com as comparações das frequências alélicas da população da Comunidade Quilombola de Sertão do Valongo. Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Referente a populações que falam línguas pertencentes às famílias linguísticas Tuu, Khoe, Kx'a e Bantu.

TABELA 26 - FREQUÊNCIAS ALÉLICAS DOS ALELOS DERIVADOS DOS POLIMORFISMOS DO GENE *MCM6* EM POPULAÇÕES DE ORIGEM AFRICANA DO *1000 GENOMES PROJECT* E *gnomAD*

País / Localidade	População	N	Frequência (%)								Referência	
			-13730 T>G	-13779 G>C	-13910 C>T	-13913 T>C	-13915 T>G	13937 G>A	-14010 G>C	-14011 C>T		
Barbados	ACB	96	11,5	0	6,8	0	0	0	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Estados Unidos	ASW	61	9,8	0	17,2	0	0	0,8	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Gâmbia	GWD	113	6,2	0	0,9	0	0	0	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Nigéria	ESN	99	9,6	0	0	0	0	0	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Nigéria	YRI	108	6,5	0	0	0	0	0	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Quênia	LWK	99	8,1	0	0	0,5	0	0	8,6	1,0	0	<i>1000 Genomes Project</i> - Fase 3
Serra Leoa	MSL	85	2,9	0	0	0	0	0	0	0	0	<i>1000 Genomes Project</i> - Fase 3
Africanos / Afro-americanos		42044	6,3	0	12,5	0,01	0,01	0,1	0,1	0,5	0,5	<i>gnomAD</i> - versão 3

FONTE: A autora (2020).

LEGENDA: ACB = Afro-caribenhos em Barbados; ASW = Indivíduos de ancestralidade africana do Sudoeste dos Estados Unidos; ESN = Esan na Nigéria; GWD = Gambianos da Divisão Oeste de Gâmbia; LWK = Luhya em Webuye, no Quênia; MSL = Mende em Serra Leoa; YRI = Yoruba em Ibadan, na Nigéria.

TABELA 27 - DIFERENCIAÇÃO DAS FREQUÊNCIAS ALÉLICAS ENTRE AS POPULAÇÕES AFRO-BRASILEIRAS ANALISADAS E POPULAÇÕES DE ORIGEM AFRICANA DO 1000 GENOMES PROJECT E *gnomAD*

População	-13730 T>G		-13907 C>G		-13910 C>T		-13913 T>C		-13915 T>G		-13937 G>A		-14009 T>G		-14010 G>C		-14011 C>T	
	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹	<i>P</i> _{CURITIBA}	<i>P</i> _{VALONGO} ¹
ACB ²	< 0,001	0,551	-	-	< 0,001	0,224	*	*	0,528	*	*	*	-	-	*	*	0,528	*
ASW ²	< 0,001	0,404	-	-	0,710	0,017	*	*	0,615	*	0,047	1	-	-	*	*	0,615	*
ESN ²	< 0,001	0,332	-	-	< 0,001	*	*	*	0,521	*	*	*	-	-	*	*	0,521	*
GWD ²	0,015	0,063	-	-	< 0,001	1	*	*	0,493	*	*	*	-	-	*	*	0,493	*
LWK ²	0,001	0,189	-	-	< 0,001	*	0,118	1	0,521	*	*	*	-	-	0,118	1	0,151	1
MSL ²	0,751	0,004	-	-	< 0,001	*	*	*	0,552	*	*	*	-	-	*	*	0,552	*
YRI ²	0,010	0,078	-	-	< 0,001	*	*	*	0,503	*	*	*	-	-	*	*	0,503	*
Africanos / Afro-americanos ³	0,001	0,027	0,811	1	< 0,001	0,044	0,811	1	0,001	1	0,440	1	0,830	1	0,793	1	0,116	1

FONTE: A autora (2020).

LEGENDA: ACB = Afro-caribenhos em Barbados; ASW = Indivíduos de ancestralidade africana do Sudoeste dos Estados Unidos; ESN = Esan na Nigéria; GWD = Gambianos da Divisão Oeste de Gâmbia; LWK = Luhya em Webuye, no Quênia; MSL = Mende em Serra Leoa; YRI = Yoruba em Ibadan, na Nigéria. NOTA: Estão destacados em negrito valores de *p* significativos (< 0,05), representando frequências estatísticas diferentes. Com * estão sinalizadas as comparações de frequências que não puderam ser realizadas por serem idênticas entre os componentes analisados.

¹ Devido ao tamanho reduzido da amostra, foi aplicada correção de continuidade para os valores de *p* obtidos com as comparações das frequências alélicas da população da Comunidade Quilombola de Sertão do Valongo.

² Dados obtidos do 1000 Genomes Project - Fase 3 (THE 1000 GENOMES PROJECT CONSORTIUM, 2015).

³ Dados obtidos do *gnomAD* - versão 3 (KARCZEWSKI et al., 2020).

6 DISCUSSÃO

6.1 POLIMORFISMOS ASSOCIADOS AO FENÓTIPO LP EM POPULAÇÕES AFRO-BRASILEIRAS DO SUL DO BRASIL

Foram identificados polimorfismos associados ao fenótipo da persistência da lactase somente na população afro-brasileira de Curitiba e Região Metropolitana, sendo eles os alelos -13910^*T , -13915^*G e -14011^*T de origem europeia, africana e não determinada, respectivamente. É conspícua a prevalência da variante de origem europeia nesta população, responsável pela persistência da lactase em 97,8% dos casos. Consequentemente, o haplótipo preditor para o fenótipo LP mais abundante nestes indivíduos também é de origem europeia que, por apresentar apenas este alelo associado ao fenótipo LP, indica ausência de recombinação nesta população (ao menos para esta região genômica). Tais prevalências decorrem da elevada proporção de ancestralidade genômica europeia observada em populações dos estados do Sul do Brasil (MOURA et al., 2015; PENA et al., 2011; SOUZA et al., 2019), nos quais este componente de ancestralidade varia de 71,0%, 72,9% e 79,7% nas populações gerais dos estados do Paraná, Rio Grande do Sul e Santa Catarina, respectivamente (MANTA et al., 2013).

A prevalência do polimorfismo $-13910 C>T$ em populações da região Sul do Brasil já foi relatada em outros trabalhos e é a maior encontrada no país, sendo identificada a frequência alélica de 65,0% na população Menonita do Paraná, e de 38,7%, 33,0% e 29,5% em euro-brasileiros de Pelotas (RS), Curitiba e Porto Alegre (RS), nesta ordem (BOSCHMANN et al., 2016; FRIEDRICH et al., 2012a; HARTWIG et al., 2016). Deve-se destacar que a frequência do alelo nos afro-brasileiros de Curitiba e Região Metropolitana analisados (18,7%) é estatisticamente diferente, sendo mais baixa do que a encontrada nos euro-brasileiros em questão ($p < 0,001$ em todos os casos). A frequência do polimorfismo de origem europeia nos afro-brasileiros de Curitiba, contudo, não difere estatisticamente das frequências observadas em outras populações afro-brasileiras, incluindo as de Porto Alegre (frequência alélica de 18,4%, $p = 0,922$) (FRIEDRICH et al., 2012a), de Pelotas (frequência alélica de 16,2%, $p = 0,249$) (HARTWIG et al., 2016) e dos indivíduos pretos de localidade não identificada do estudo de Mattar e colaboradores (2008) (frequência alélica de 12,0%, $p = 0,110$). Foram também observadas frequências estatisticamente similares nas

populações gerais de Belém (frequência alélica de 17,5%, $p = 0,653$) (FRIEDRICH et al., 2012a) e nas populações da região Nordeste do Brasil, Recife (frequência alélica de 20,4%, $p = 0,485$) (FRIEDRICH et al., 2012a) e Salvador (frequência alélica de 18,7%, $p = 0,996$) (ARAUJO et al., 2019) – região cujo componente de ancestralidade africana é o mais elevado do país, com média de 27,0% da ancestralidade total (MOURA et al., 2015). Desta forma, foi identificada similaridade entre a frequência do alelo -13910^*T na população afro-brasileira de Curitiba e Região e nas demais populações afro-brasileiras em que a variante já foi descrita – com exceção apenas da população da Comunidade Quilombola de Sertão do Valongo, descrita no presente trabalho (frequência alélica de 0,0%, $p = 0,004$).

A ausência da substituição nucleotídica de origem europeia $-13910 T>C$ nos afro-brasileiros de Sertão do Valongo ratifica a já descrita pouca contribuição de europeus e/ou de indivíduos de ascendência europeia na formação e desenvolvimento desta população (SOUZA; CULPI, 2005; TEIXEIRA, 1990).

A presença da mutação $-13915 T>G$ em um único indivíduo da população de Curitiba e Região Metropolitana (frequência de 0,2%) é inédita em populações brasileiras. Apesar deste alelo estar presente em populações do Norte, Centro e Leste africano, sua maior prevalência é observada em populações da Península Arábica e Oriente Médio – sugerindo sua origem nesta região, com posterior inserção em populações da África Oriental através de fluxo gênico (RANCIARO et al., 2014). Sendo assim, a ocorrência do alelo na população brasileira analisada – apesar da baixa frequência – evidencia sua ancestralidade africana, podendo estar associado a contribuições de populações do Centro e Leste africano – regiões das quais africanos escravizados foram trazidos forçadamente para o Brasil entre os séculos XVI e XIX, durante o tráfico negreiro transatlântico (KLEIN, 1987; SOUZA, 2008). Além disso, o haplótipo no qual a variante referida está inserida não apresenta o polimorfismo de origem europeia -13910^*T , indicando também ausência de recombinação.

O alelo -14011^*T , por sua vez, é observado em baixas frequências em todas as populações nas quais já foi identificado, sendo 2,5% a maior, observada na população Kgalagadi de Botsuana (MACHOLDT et al., 2014). No Brasil, foi identificado em três indivíduos de Recife (frequência alélica de 0,8%) e em um indivíduo de Belém (frequência alélica de 0,5%) (FRIEDRICH et al., 2012a), em frequências que não diferiram estatisticamente das relatadas para a população de Curitiba e Região Metropolitana ($p = 0,210$ e $0,458$, respectivamente). Da mesma

forma, não foram observadas diferenças estatísticas significativas com relação à frequência do alelo nas demais populações latino-americanas (TABELA 20, TABELA 22) e africanas comparadas (TABELA 24, TABELA 25, TABELA 27). Portanto, as frequências observadas na população afro-brasileira de Curitiba estão de acordo com as frequências desse alelo em diferentes localidades.

Por não ter origem determinada e por já ter sido identificado em populações africanas, asiáticas, europeias e da Península Arábica (AL-ABRI et al., 2012; JONES et al., 2013; LIEBERT et al., 2017; MACHOLDT et al., 2014), não é possível especular qual teria sido a população de origem responsável por inserir a mutação *-14011 C>T* na população afro-brasileira de Curitiba e Região. Sua ocorrência num mesmo haplótipo contendo o alelo de origem europeia *-13910*T* já foi relatada por Liebert e colaboradores (2017) que, por também o identificarem em haplótipos provenientes de contextos evolutivos distintos, sugeriram que a mutação em questão teria ocorrido mais de uma vez em populações diferentes – não tratando-se, desse modo, de um haplótipo recombinante.

O polimorfismo *-13730 T>G* – apesar de não ter associação confirmada ao fenótipo da persistência da lactase – está também inserido na região intensificadora do gene *LCT*. Este foi o único polimorfismo identificado comum aos afro-brasileiros de Curitiba e Região Metropolitana (frequência alélica de 2,5%) e da Comunidade Quilombola de Sertão do Valongo (frequência alélica de 16,7%), embora em frequências estatisticamente diferentes ($p < 0,001$) – sendo a primeira vez que o alelo foi observado em brasileiros. O haplótipo no qual o alelo *-13730*G* está inserido é o mesmo em ambas as populações. Além disso, este alelo não está em desequilíbrio de ligação com o polimorfismo de origem europeia *-13910 C>T*, indicando ausência de recombinação nessa região genômica. Como o alelo *-13730*G* foi observado apenas em povos de origem africana pertencentes a diferentes partes do continente africano (com exceção da porção Sul) (INGRAM et al., 2009a; JONES et al., 2013; LIEBERT et al., 2017; RANCIARO et al., 2014), sua presença em ambas as populações afro-brasileiras analisadas corrobora a ancestralidade africana destes indivíduos. Ainda, há uma correspondência entre as localidades contendo populações que apresentam o polimorfismo *-13730 T>G* e as regiões das quais africanos escravizados eram trazidos para o Brasil, como Moçambique, Gana e Angola (KLEIN, 1987; SOUZA, 2008).

A frequência do alelo -13730*G na população de Sertão do Valongo (16,7%) é a mais alta já observada no planeta, sendo superior às outras maiores frequências identificadas provenientes de populações da África Central (11,8%) e de afro-caribenhos de Barbados (11,5%), reportados por Ranciaro e colaboradores (2014) e pelo *1000 Genomes Project* (2015), respectivamente. É importante destacar que, embora a frequência de Sertão do Valongo seja superior, ela não difere estatisticamente das outras duas populações mencionadas ($p = 0,550$ e $0,551$, respectivamente). A elevada frequência identificada na Comunidade de Sertão do Valongo, apesar de possivelmente ter sido influenciada pelo pequeno tamanho amostral ($n = 18$), é reflexo do seu pequeno tamanho populacional efetivo ($n = 21$), assegurado pelo isolamento da comunidade (SOUZA; CULPI, 1992, 2005). Desta forma, mesmo alelos em baixas frequências na população que a originou podem ter aumentado sua frequência após o isolamento, como consequência do efeito fundador (HATZIKOTOULAS; GILLY; ZEGGINI, 2014). Fenômeno equivalente já foi observado na população Menonita do Sul do Brasil, cujo histórico de formação é marcado por sucessivos eventos de gargalo de garrafa, elevada endogamia e isolamento. Tais fatores culminaram numa frequência do alelo -13910*T superior à encontrada na população alemã – um dos grupos que a formou originalmente no século XVI (BOSCHMANN, 2012).

A presença de demais mutações de origem africana associadas à persistência da lactase – além dos alelos -13730*G e -13915*G – nas duas populações afro-brasileiras analisadas era esperada, considerando que muitos dos africanos trazidos forçadamente para o Brasil durante os séculos XVI e XIX eram provenientes das mesmas regiões nas quais estima-se que os polimorfismos -13907 C>G, -13913 T>C, -14009 T>G e -14010 G>C tenham surgido, na África Oriental (KLEIN, 1987; TISHKOFF et al., 2007). Confirmando a ocorrência de componentes genéticos originários do leste africano em brasileiros, Gouveia e colaboradores (2020) estimaram, a partir de marcadores nucleares, a ancestralidade africana em populações de Salvador (BA), Bambuí (MG) e Pelotas (RS). Os autores identificaram que a ancestralidade africana predominante nos indivíduos analisados remonta a populações do Centro-Oeste da África, no caso da população nordestina, e do Leste-Sul da África, no caso das populações do Sudeste e Sul brasileiro (GOUVEIA et al., 2020). Resultados similares foram obtidos a partir de análises em sequências de DNA mitocondrial em populações afro-brasileiras do Rio de Janeiro (RJ) e Porto Alegre

(RS) (HÜNEMEIER et al., 2007). É válido destacar que, ao contrário do que ocorre na Europa com relação à variante -13910^*T , a distribuição dos alelos preditivos para o fenótipo LP no continente africano é mais influenciada pela origem da população do que pela a região geográfica na qual ela está estabelecida. Deste modo, populações que apresentam ancestralidade distintas, mas que habitam áreas próximas, podem apresentar frequências alélicas bastante discrepantes. Tal aspecto foi evidenciado por Macholdt e colaboradores (2014), que observaram o alelo -14010^*C na frequência de 36,0% na população Nama (língua Khoe) e de 2,5% na população Taa (língua Tuu) da região Oeste da Namíbia – mesmo ambas habitando regiões vizinhas do mesmo país. Neste caso, a ausência de polimorfismos de origem africana nos afro-brasileiros do Sul do Brasil avaliados indica que as populações africanas trazidas para o país não possuíam tais alelos ou os apresentavam em baixa frequência, sugerindo duas situações: a vinda de africanos exclusivamente não pertencentes a grupos pastoralistas ou a ação da deriva genética na vinda de grupos pastoralistas para o Brasil, em que as frequências das variantes associadas à LP em questão eram reduzidas nos grupos que migraram e, ao longo das gerações, foram perdidas ou mantidas em baixa frequência na população – não sendo detectadas pela amostragem realizada. A ocorrência da variante alélica -13915^*G em afro-brasileiros de Curitiba e Região e a presença do alelo -14010^*C em afro-brasileiros de Porto Alegre (FRIEDRICH et al., 2012a) corroboram a segunda situação. Além disso, aspectos como a presença de haplótipos de origem banta – grupo linguístico que inclui diversas populações com tradições pastoralistas – em brasileiros (AURICCHIO et al., 2007; VICARI; DE MELLO; FIGUEIREDO, 2005; ZAGO; FIGUEIREDO; OGO, 1992), bem como o relato em fontes históricas sobre a vinda desses povos para o país (SLENES, 1992) e sua conseqüente contribuição para a cultura brasileira (APONTES, 2010; MUKUNA, 1978; RODRIGUES, 2010) também apontam para a segunda hipótese.

As substituições nucleotídicas $-13732 C>A$ e $-13752 C>T$, observadas em afro-brasileiros de Curitiba e Região, em um indivíduo cada, já foram identificadas em populações africanas. O polimorfismo $-13732 C>A$ foi observado em afro-americanos do Sudoeste dos Estados Unidos, na população Luhya, do Quênia, e na população Mende, de Serra Leoa, nas frequências de 0,8% ($n = 1$), 0,5% ($n = 1$) e de 1,2% ($n = 2$), respectivamente (THE 1000 GENOMES PROJECT CONSORTIUM, 2015). Da mesma forma, o polimorfismo $-13752 C>T$ está presente na frequência de 0,3% em

africanos / afro-americanos descritos no *gnomAD* – versão 3 (KARCZEWSKI; MARTIN, 2020). Não há outras informações, tanto funcionais quanto populacionais, acerca dos alelos em questão até o momento.

6.2 DESEQUILÍBRIO DE LIGAÇÃO

O desequilíbrio de ligação é definido pela combinação não ao acaso de alelos localizados em diferentes *loci* (NORDBORG; TAVARÉ, 2001). Para estimá-lo nas populações afro-brasileiras de Curitiba e Região, foram utilizados parâmetros distintos: r^2 e D' que, conseqüentemente, representam aspectos diferentes do LD (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003). O r^2 representa o valor da correlação entre alelos num mesmo haplótipo e, portanto, indica a ocorrência (ou não) de eventos de recombinação e mutação que se sucederam no bloco haplotípico analisado (ANEXO 1) (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003; HILL; ROBERTSON, 1968). Desta forma, o valor de r^2 só será igual a um (isto é, indicando desequilíbrio de ligação) caso as frequências alélicas investigadas sejam idênticas (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003). Por outro lado, o parâmetro D' é dado pela razão da diferença entre as frequências haplotípicas observadas e esperadas pelas frequências alélicas observadas – sendo informativo sobre a ocorrência de eventos de recombinação (ANEXO 1) (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003; LEWONTIN, 1964).

O valor de r^2 , calculado a partir das frequências alélicas do par de variantes -13730*G e -13910*T ($r^2 = 0,075$), sugere que os alelos em questão estarão presentes em haplótipos distintos e que, além disso, seu surgimento ocorreu em contextos evolutivos diferentes (FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003). Tal constatação era esperada, considerando tanto a ausência de haplótipos recombinantes, contendo ambos os alelos nos afro-brasileiros de Curitiba e Região observados, quanto a origem do polimorfismo -13730 T>G em populações africanas e do polimorfismo -13910 C>T em populações europeias. É importante destacar que o baixo valor de r^2 e o alto valor de D' (0,98), analisados em conjunto, indicam diferenças nas frequências alélicas observadas (18,7% e 2,5%, para os alelos -13730*G e -13910*T, respectivamente) (WRAY; VISSCHER, 2008).

6.3 DIVERSIDADE E DISTÂNCIA GENÉTICA ENTRE OS AFRO-BRASILEIROS ANALISADOS

A ocorrência de mais alelos distribuídos numa maior quantidade de haplótipos na população afro-brasileira de Curitiba de Região Metropolitana ($n = 7$), quando comparada com a população afro-brasileira da Comunidade Quilombola de Sertão do Valongo ($n = 2$), são indicativos da maior variabilidade existente no *pool* genético da primeira. Através da heterozigosidade média ($H_{eM} = 0,037$ e $0,024$, para Curitiba e Região e Sertão do Valongo, respectivamente), foi observado que de fato há uma maior diversidade genética na população de Curitiba. Porém, tal diversidade não difere estatisticamente da encontrada em Sertão do Valongo ($p = 1$) – o que possivelmente decorre da desigualdade no tamanho amostral de ambas e dos poucos *loci* utilizados na determinação da diversidade ($n = 10$).

A menor diversidade genética encontrada na Comunidade de Sertão do Valongo concorda com o histórico de isolamento ao qual a população está submetida desde sua fundação, no final do século XIX, motivado tanto pela sua localização geográfica afastada, quanto pelas diferenças socioculturais com as comunidades vizinhas – de ascendência predominantemente europeia (CRUZ; HANAZAKI, 2008; SOUZA; CULPI, 2005). Tais aspectos, sobretudo os referentes às diferenças socioculturais estabelecidas, marcadas pela forte religiosidade presente na comunidade juntamente com a segregação racial exercida pelas populações que habitam regiões próximas a ela, resultaram no alto coeficiente de endocruzamento encontrado neste grupo ($F = 0,081$) – explicando sua baixa diversidade genética (CASTELLS, 2007; SOUZA; CULPI, 2005). Além disso, sua formação por apenas oito indivíduos evidencia a ocorrência de efeito fundador, como já sugerido por Souza e Culpí (2005), o que assegurou a baixa variabilidade genética nas gerações posteriores. Essa pouca variabilidade também já foi relatada para outras regiões genômicas nesta mesma população (SOUZA; CULPI, 2005).

Em contrapartida, a maior diversidade genética encontrada nos afro-brasileiros de Curitiba e Região Metropolitana já era esperada considerando o não isolamento e o caráter miscigenado da população, previamente relatado nos trabalhos de Probst e colaboradores (2000) e Braun-Prado e colaboradores (2000). Nestes trabalhos, foram avaliadas a ancestralidade genômica de parte das amostras incluídas no presente

estudo, relatando a contribuição de 49,5%, 41,8% e 8,7% de componentes genéticos africano, europeu e ameríndio, respectivamente.

A distância genética entre as populações de Curitiba e Região Metropolitana e de Sertão do Valongo, avaliada através do valor de F_{ST} , indica, de acordo com Wright (1949), uma diferenciação moderada entre ambas ($F_{ST} = 0,119$). Tal diferenciação é reflexo de estruturas populacionais distintas entre os grupos afro-brasileiros comparados, decorrente principalmente do contraste existente na formação e nas dinâmicas socioculturais de cada um. Ao contrário da população de Curitiba e Região Metropolitana, os afro-brasileiros da Comunidade Quilombola de Sertão do Valongo mantiveram um estilo de vida com pouco contato com pessoas de fora da comunidade desde sua fundação. Essa característica, aliada a outros fatores como baixa taxa de migração, elevado grau de endocruzamento e pequeno tamanho populacional, assegurou as diferenças existentes entre sua estrutura populacional e a dos afro-brasileiros do Paraná (SOUZA; CULPI, 2005). As diferenças entre as frequências dos alelos -13730*G e -13910*T e, conseqüentemente, dos genótipos -13730 G/T, -13730 T/T, -13910 C/T e -13910 C/C corroboram a diferenciação mencionada entre as populações em questão.

6.4 COMPARAÇÕES POPULACIONAIS

6.4.1 Entre populações latino-americanas

As frequências dos alelos localizados na região intensificadora do gene *LCT* -13779*C, -13937*A, -14010*C e -14011*T em populações brasileiras nas quais já foram identificados não diferiram estatisticamente das frequências observadas nos afro-brasileiros de Curitiba e Região Metropolitana e da Comunidade Quilombola de Sertão do Valongo (TABELA 18). Isso ocorre porque tais mutações estão presentes em menos de cinco indivíduos em todas as populações, o que indica a raridade da sua ocorrência em brasileiros – incluindo os de ascendência majoritariamente africana do sul do país analisados.

Como já descrito, a frequência do alelo -13910*T nos afro-brasileiros de Curitiba e Região (18,7%) não diferiu estatisticamente das observadas em outras populações afro-brasileiras. Essa similaridade é reflexo da semelhança entre tais populações, decorrente não somente da ancestralidade europeia compartilhada, mas

também da sua inserção em centros urbanos como Curitiba, Porto Alegre e Pelotas. A não constatação de diferenças estatísticas entre as frequências encontradas nas populações do Nordeste do Brasil, de Salvador e Recife – também grandes centros urbanos –, são resultantes dos mesmos fatores pois, embora as populações dessas cidades não tenham sido diferenciadas quanto à sua ancestralidade e/ou fenótipo nos trabalhos em que foram descritas, a contribuição dos componentes genéticos de origem africana nesta região são, em média, os mais elevados do país (MACHADO, 2008; MOURA et al., 2015). As diferenças estatísticas com relação a frequência do alelo -13910^*T em euro-brasileiros, ameríndios e nipo-brasileiros (TABELA 18) era esperada considerando a composição genética distinta de cada um desses grupos em relação aos afro-brasileiros de Curitiba e Região. Essa distinção ocorre apesar do caráter miscigenado atribuído historicamente à população brasileira pois, como evidenciado pelos dados do EPIGEN-Brasil (TARAZONA SANTOS et al., 2016), cruzamentos preferenciais em função da ancestralidade e aspectos relacionados à ela são típicos no país – assegurando tal diferenciação. Fatores como a discriminação racial são determinantes para o estabelecimento desta dinâmica de cruzamento (TELLES, 2004).

Por sua vez, a frequência do alelo -13910^*T na população de Sertão do Valongo (0,0%) não diferiu estatisticamente das frequências relatadas para a população geral de Fortaleza (frequência alélica de 23,9%, $p = 0,057$) (PONTE et al., 2016), para os ameríndios das etnias Guarani-Kaiowá (frequência alélica de 0,6%, $p = 1$), Guarani-Ñandeva (frequência alélica de 7,6%, $p = 0,193$), Kaingang (frequência alélica de 4,9%, $p = 0,386$) e Xavante (frequência alélica de 0,5%, $p = 1$) (FRIEDRICH et al., 2012b), bem como para os indivíduos pretos (frequência alélica de 12,0%, $p = 0,067$) do estudo de Mattar e colaboradores (2008), cuja localidade a qual pertencem não foi identificada. A baixa frequência da variante, observada nesses grupos, explica tal similaridade – impulsionada também pelo pequeno número de amostras da população de Sertão do Valongo ($n = 18$). A não similaridade com as frequências encontradas em outras populações afro-brasileiras, em contrapartida, provavelmente se deve ao isolamento da comunidade valonguense. A dinâmica de isolamento impediu a ocorrência de fluxo gênico recente e, conseqüentemente, resultou numa estrutura populacional distinta – assim como citado anteriormente, com relação às diferenças com a população afro-brasileira de Curitiba e Região.

A ausência de diferenças estatísticas entre as frequências do alelo *-13910*T* na população de Curitiba e Região Metropolitana (18,7%) e em onze das 15 populações latino-americanas miscigenadas comparadas (TABELA 20) – incluindo populações do Chile, Equador e México – indica que há uma estrutura populacional similar entre eles. Tal similaridade é reflexo dos processos de colonização europeia associado ao extermínio de populações nativas e de uso de mão de obra de africanos escravizados em todo o continente e ocorre apesar das proporções de ancestralidade genômica dos componentes ameríndios, africanos e europeus variarem nos diferentes países devido à história demográfica de cada um (BERTONI, 2011; MOYA, 2018; SALZANO; SANS, 2014). Além disso, o mesmo padrão de não similaridade com as frequências observado em ameríndios brasileiros é encontrado também com relação a grupos nativos de demais países da América Latina comparados (TABELA 20). É importante salientar que as diferenças nas proporções de ancestralidade entre a população afro-brasileira em questão e as populações miscigenadas e de origem africana do Equador são notáveis – sobretudo com relação às porções europeias –, sendo as contribuições nessas populações, respectivamente, equivalentes a 19,0% e 16,0% europeu, 73,0% e 28,0% ameríndio, 8,0% e 56,0% africano (GONZÁLEZ-ANDRADE et al., 2007). Esse contraste pode explicar as diferenças estatísticas referente a comparação da frequência do alelo *-13910*T* com essas populações (TABELA 20).

Quanto às comparações com populações latino-americanas do *1000 Genomes Project* (TABELA 21), foram observadas diferenças estatísticas significativas entre as frequências dos alelos *-13730*G* e *-13910*T* nos afro-brasileiros de Curitiba e Região (2,5% e 18,7%, respectivamente) e nos Peruanos de Lima, no Peru, decorrentes da ausência da mutação na última (TABELA 22). De acordo com Martin e colaboradores (2017), o conjunto de amostras referente à população peruana apresenta 77,0% de componentes genéticos de ancestralidade ameríndia, sendo a maior proporção da referida ancestralidade encontrada em todas as populações do *1000 Genomes Project*. Sendo assim, a elevada contribuição de ancestralidade europeia explica a ausência da variante de origem africana *-13730*G* e a baixa frequência do polimorfismo de origem europeia *-13910 C>T* nesta população, ocasionando as desigualdades entre as frequências. Foram também identificadas diferenças estatísticas com relação à frequência do alelo *-13910*T* na população de afro-brasileira de Curitiba e Região e nos Colombianos de Medellín, na Colômbia. Isso

decorre, possivelmente, devido à elevada contribuição de ancestralidade europeia também identificada nesses indivíduos, propiciando uma maior frequência alélica (GENOVESE et al., 2013; GRAVEL et al., 2013). Com relação às comparações realizadas referente à frequência das variantes -13730^*G e -13910^*T na população de Sertão do Valongo (16,7% e 0,0%, respectivamente) e em populações latino-americanas (TABELA 19, TABELA 21), foram identificadas diferenças estatísticas em todos os casos (TABELA 20, TABELA 22). Tais diferenças, como já exposto anteriormente, é consequência da baixa variabilidade genética e estrutura populacional isolada da Comunidade de Sertão do Valongo. Para os alelos -13915^*G e -14011^*T , não foram encontradas diferenças estatísticas entre nenhuma das populações comparadas (TABELA 22), o que já era esperado devido às suas baixas frequências em todas elas.

A constatação de diferenças estatísticas com relação às frequências dos alelos -13730^*G e -13910^*T quando comparados os dois grupos afro-brasileiros da região Sul e os dados da população latino-americana geral, do *gnomAD*, possivelmente decorre da discrepância entre os tamanhos amostrais. Isso não ocorre para as demais variantes porque suas respectivas frequências são inferiores a 0,1% nos demais latino-americanos, tornando-as similares às suas baixas frequências nos afro-brasileiros paranaenses e catarinenses comparados (TABELA 21).

6.4.2 Entre populações africanas

As frequências dos alelos -13730^*T , -13907^*G , -13910^*T , -13913^*C , -13915^*G e -14009^*G comparadas entre a população geral do Norte da África (TABELA 23) e os afro-brasileiros de Curitiba e Região Metropolitana (frequências alélicas de 2,5%, 0,0%, 18,7%, 0,0%, 0,2% e 0,0%, respectivamente) demonstraram ser estatisticamente diferentes ($p = 0,016$, $< 0,001$, $0,038$, $0,040$, $< 0,001$ e $< 0,001$, respectivamente) (TABELA 24). Tal aspecto decorre da pouca contribuição que os povos dessa região do continente africano exerceram na composição da população brasileira, já que não estavam inclusos entre os grupos que eram trazidos para o Brasil durante do comércio transatlântico de escravizados (SOUZA, 2008). A única variante que não diferiu estatisticamente, contudo, foi a -14011^*T (frequência alélica de 0,2% na população de Curitiba e Região, $p = 0,626$) e isto se deve às baixas frequências em que o polimorfismo está presente nas populações nas quais já foi identificado,

como já descrito anteriormente. Pelo mesmo motivo, este alelo não demonstrou diferenças estatísticas com relação à nenhuma das frequências em todas as populações comparadas (TABELA 24).

O polimorfismo de origem africana *-13730 T>G* está presente em frequências similares às da população de Curitiba e Região Metropolitana (2,5%) na maioria dos povos africanos comparados (TABELA 24). Tal similaridade foi constatada inclusive em populações cujas localidades coincidem com as que sabidamente englobam grupos trazidos para o Brasil durante o período do tráfico transatlântico, como Angola e Moçambique. Contudo, foram observadas diferenças estatísticas com relação às frequências reportadas para populações provenientes de regiões que também incluíam grupos trazidos para o país entre os séculos XVI e XIX, como África Central ($p < 0,001$) e Ocidental ($p = 0,005$) e a população de Jaali do Sudão, na África Oriental ($p = 0,013$) (TABELA 24). Como as diferenças estão restritas a apenas alguns dos grupos comparados nessas localidades, tal diferenciação não invalida as similaridades com as demais populações destas regiões relatadas anteriormente. Sendo assim, a frequência do polimorfismo *-13730 T>G* nos afro-brasileiros de Curitiba e Região Metropolitana pode ser explicada pela ancestralidade comum com populações africanas. Para a população de Sertão do Valongo, foram identificadas diferenças estatísticas significativas com relação a frequência do alelo *-13730*G* (16,7%) numa maior quantidade de populações, incluindo grupos da Angola, Etiópia, Moçambique e Sudão (TABELA 25). Porém, ao contrário do que foi descrito para os afro-brasileiros de Curitiba, as frequências de Sertão do Valongo diferiram por serem superiores às observadas nas populações mencionadas, sendo provavelmente consequência do efeito fundador na origem da comunidade.

A frequência do alelo *-13910*T* dos afro-brasileiros de Curitiba e Região Metropolitana (18,7%) difere estatisticamente das frequências de todas as populações africanas comparadas (TABELA 24), evidenciando que sua inserção nesta população seria decorrente da miscigenação com europeus. A ausência deste alelo na população da Comunidade de Sertão do Valongo, decorrente dos fatores mencionados anteriormente, promoveu a não constatação de diferenças estatísticas com relação a frequência desse alelo nas populações africanas em que foi identificado.

Nas populações africanas, as frequências dos alelos *-13907*G*, *-13913*C*, *-14009*G* e *-14010*C* (TABELA 23) são estatisticamente superiores às encontradas nos afro-brasileiros de Curitiba e Região Metropolitana (frequência de 0,0% para todos

os alelos em questão) (TABELA 24). Apesar das variantes referidas também estarem ausentes na população de Sertão do Valongo, seu pequeno tamanho amostral ($n = 18$) resultou em diferenças estatísticas com relação a uma menor quantidade de populações comparadas (TABELA 25).

A frequência do alelo -13915^*G nos afro-brasileiros de Curitiba e Região Metropolitana (0,2%) não diferiu estatisticamente da maior parte das populações africanas comparadas (TABELA 24), com exceção de grupos da Etiópia ($p = 0,012$ na população Afar e $p = 0,004$ na população Oromo, proveniente da região de Borana) e Sudão ($p = 0,001$ e $0,10$ nas populações Beja Beni Amer e Jaali, respectivamente) nos quais sua frequência é elevada (TABELA 23). O mesmo é válido para a frequência da Comunidade Quilombola de Sertão do Valongo (0,0%), embora também com um número menor de populações (TABELA 25). Como já descrito, provavelmente eventos de deriva genética teriam sido responsáveis por diminuir a frequência do alelo -13915^*G em afro-brasileiros, já que sua presença – mesmo em poucos indivíduos – assegura que a variante tenha chegado ao país através dos processos migratórios promovidos pelo comércio transatlântico de escravizados.

Similar ao encontrado nas populações latino-americanas, a maior parte das populações africanas do *1000 Genomes Project* e do *gnomAD* diferiram estatisticamente apenas para as frequências dos alelos -13730^*G e -13910^*T nas populações afro-brasileiras comparadas (TABELA 27). Além das variantes mencionadas, foram constatadas diferenças também quanto à frequência do alelo -13915^*G na população africana / afro-americana do *gnomAD*, e quanto à frequência do alelo -13937^*A nos afro-americanos do Sudoeste dos Estados Unidos do *1000 Genomes Project*. A restrição majoritária de diferenças estatísticas às frequências dos alelos -13730^*G e -13910^*T possivelmente decorre deles serem as variantes em maior frequência identificadas nos afro-brasileiros analisados.

7 CONCLUSÕES

Nos afro-brasileiros de Curitiba e Região Metropolitana, assim designados a partir de heterodeterminação, foram identificados os polimorfismos associados ao fenótipo da persistência da lactase -13910 C>T, -13915 T>G e -14011 C>T, bem como os polimorfismos -13730 T>G, -13732 C>A e -13752 C>T que, embora não tenham associação ao fenótipo em questão confirmada, localizam-se na região intensificadora a montante do gene *LCT*, assim como as demais variantes mencionadas. Em contrapartida, nos afro-brasileiros da Comunidade Quilombola de Sertão do Valongo, classificados assim devido ao histórico de fundação da comunidade por indivíduos de ancestralidade predominantemente africana, não foram identificados alelos associados a este fenótipo, sendo a variante -13730*G o único polimorfismo encontrado na população.

O alelo -13910*T, de origem europeia, é o mais prevalente dentre os identificados na população de Curitiba e Região Metropolitana. Sua maior frequência, quando comparada à dos demais alelos, é consequência da elevada proporção de ancestralidade europeia em populações brasileiras – sobretudo na região Sul do país. Esta contribuição proeminente de ancestralidade europeia decorre da imigração acentuada de populações deste continente para o Brasil desde o século XVI, acentuada no período pós-abolição do regime escravocrata através do estímulo para a vinda gerado pelo financiamento do governo brasileiro. Além disso, a ausência da variante -13910*T na população de Sertão do Valongo corrobora a pouca contribuição que populações europeias e/ou de origem europeia exerceram na formação e desenvolvimento da comunidade, assegurando sua ancestralidade predominantemente africana.

A ocorrência do polimorfismo -13730 T>G, presente exclusivamente em populações da África, nos afro-brasileiros analisados corrobora sua ancestralidade africana. Tal alelo, inclusive, foi observado nos indivíduos de Sertão do Valongo na maior frequência já identificada no planeta, sendo consequência do efeito fundador na constituição da comunidade em questão, formada por menos de dez indivíduos no final do século XIX e mantendo-se com tamanho populacional reduzido desde então.

A baixa frequência do alelo de origem africana -13915*G assim como a ausência de demais alelos de origem africana associados ao fenótipo LP nos afro-brasileiros de Curitiba e Região e de Sertão do Valongo analisados são decorrentes

de eventos de deriva genética na vinda forçada dos africanos durante o comércio transatlântico de escravizados, entre os séculos XVI e XIX. As frequências desses alelos nas populações que migraram para o Brasil provavelmente eram baixas, de modo que suas respectivas frequências nas gerações subsequentes foram reduzidas. Desta forma, o aparecimento de haplótipos recombinantes contendo alelos de origem africana e europeia torna-se mais raro na população – evidenciado pela sua ausência nos indivíduos amostrados. Contudo, foram analisadas apenas uma fração das populações afro-brasileiras selecionadas, de modo que o tamanho da amostra pode ter sido insuficiente para constatar os alelos mencionados anteriormente.

A presença dos alelos -13915^*G e -14011^*T indicam que, apesar da identificação da variante -13910^*T ser assertiva para o diagnóstico da hipolactasia primária em grande parte da população analisada, seu uso não permite a determinação da condição em todos os casos. Portanto, é recomendável incluir mais alelos da região intensificadora do gene *LCT*, além do -13910^*T , para que o diagnóstico apresente maior eficiência – especialmente nos afro-brasileiros de Curitiba e Região Metropolitana.

A menor diversidade genética encontrada na população afro-brasileira da Comunidade Quilombola de Sertão do Valongo, quando comparada com a população de Curitiba e Região, e a diferenciação genética moderada entre elas evidenciam a estrutura populacional distinta que cada uma apresenta. Enquanto a população de Sertão do Valongo manteve-se isolada, devido tanto à sua localização geográfica afastada, quanto às diferenças socioculturais com comunidades vizinhas, a população de Curitiba e Região esteve mais sujeita à fluxo gênico. Além disso, aspectos como a elevada taxa de endocruzamento encontrada em Sertão do Valongo favoreceram ainda mais a diminuição de sua diversidade genética.

Foram identificadas similaridades entre as frequências alélicas da população de Curitiba e Região e demais populações miscigenadas da América Latina, decorrente das semelhanças existentes nos processos que as constituíram nos últimos séculos. Tais processos incluem o extermínio de populações nativas a partir da chegada dos europeus e posterior inserção de grupos africanos pelo tráfico transatlântico de escravizados. O mesmo não foi constatado para a Comunidade Quilombola de Sertão do Valongo devido ao caráter isolado deste grupo. Foram observadas frequências similares também com relação às frequências alélicas dos

afro-brasileiros estudados e de populações africanas, provavelmente derivado de sua ancestralidade comum.

REFERÊNCIAS

- AHMAD, M.; FLATZ, G. Prevalence of Primary Adult Lactose Malabsorption in Pakistan. **Human Heredity**, v. 34, n. 2, p. 69–75, 1984.
- AL-ABRI, A. et al. Distribution of the Lactase Persistence-Associated Variant Alleles -13910*T and -13915*G among the People of Oman and Yemen. **Human Biology**, v. 84, n. 3, p. 271–286, 2012.
- ALEGRO, R. C. et al. **Temas e Questões para o ensino de História do Paraná**. Londrina: Editora da Universidade Estadual de Londrina, 2013.
- ALEXANDER, J. Islam, archaeology and slavery in africa. **World Archaeology**, v. 33, n. 1, p. 44–60, 2001.
- ALHARBI, O.; EL-SOHEMY, A. Lactose Intolerance (*LCT* -13910C>T) Genotype Is Associated with Plasma 25-Hydroxyvitamin D Concentrations in Caucasians: A Mendelian Randomization Study. **The Journal of Nutrition**, v. 147, n. 6, p. 1063–1069, 2017.
- ALLENTOFT, M. E. et al. Population genomics of Bronze Age Eurasia. **Nature**, v. 522, p. 167–172, 2015.
- ALVES-SILVA, J. et al. The ancestry of Brazilian mtDNA lineages. **American Journal of Human Genetics**, v. 67, n. 2, p. 444–461, 2000.
- ALZATE, H.; GONZÁLEZ, H.; GUZMÁN, J. Lactose Intolerance in South American Indians. **American Journal of Clinical Nutrition**, v. 22, n. 2, p. 122–123, 1969.
- AMBROSE, S. H. Chronology of the later stone age and food production in East Africa. **Journal of Archaeological Science**, v. 25, n. 4, p. 377–392, 1998.
- ANDERSON, B.; VULLO, C. Did malaria select for primary adult lactase deficiency? **Gut**, v. 35, n. 10, p. 1487–1489, 1994.
- ANGUITA-RUIZ, A.; AGUILERA, C. M.; GIL, Á. Genetics of Lactose Intolerance: An Updated Review and Online Interactive World Maps of Phenotype and Genotype Frequencies. **Nutrients**, v. 12, n. 2689, 2020.
- APONTES, S. A. Acomodação de palavras Bantu em Português: algumas consequências morfofonológicas. **Revista Philologus**, v. 16, n. 46, p. 42–58, 2010.
- ARAUJO, E. M. et al. Genetic and Oral Tests for the Diagnosis of Lactose Intolerance in Mixed-Ancestry Brazilians with Metabolic Syndrome. **Lifestyle Genomics**, v. 12, n. 1–6, p. 1–9, 2019.
- ARNHOLD, R. G.; PERMAN, J. A.; NURSE, G. T. Persistent high intestinal lactase activity in Papua New Guinea. The breath hydrogen test in a Sepik population. **Annals of Human Biology**, v. 8, n. 5, p. 481–484, 1981.
- ARNOLD, J. et al. L'intolerance au lactose chez l'adulte au Senegal. **Comptes**

Rendus des Seances de la Societe de Biologie, v. 174, n. 6, p. 983–992, 1980.

ASP, N. G. et al. Complete Deficiency of Brush-Border Lactase in Congenital Lactose Malabsorption. **The Lancet**, v. 302, n. 7824, p. 329–330, 1973.

AURICCHIO, M. T. B. DE M. et al. Frequency and Origins of Hemoglobin S Mutation in African-Derived Brazilian Populations. **Human Biology**, v. 79, n. 6, p. 667–677, 2007.

AURICCHIO, S. et al. Isolated Intestinal Lactase Deficiency in the Adult. **The Lancet**, v. 273, n. 7083, p. 1126–1128, 1963.

BABU, J. et al. Frequency of lactose malabsorption among healthy southern and northern Indian populations by genetic analysis and lactose hydrogen breath and tolerance tests. **American Journal of Clinical Nutrition**, v. 91, n. 1, p. 140–146, 2010.

BASTOS, J. L. et al. Does the way i see you affect the way I see myself? associations between interviewers' and interviewees' "color/race" in southern Brazil. **Cadernos de Saúde Pública**, v. 25, n. 10, p. 2111–2124, 2009.

BAYLESS, T. M.; BROWN, E.; PAIGE, D. M. Lactase Non-persistence and Lactose Intolerance. **Current Gastroenterology Reports**, v. 19, n. 5, 2017.

BAYLESS, T. M.; ROSENSWEIG, N. S. A Racial Difference in the Incidence of Lactase Deficiency. **JAMA**, v. 197, n. 12, p. 968–972, 1966.

BAYOUMI, R. et al. Beja And Nilotes: Nomadic pastoralist groups in the Sudan with opposite distributions of the adult lactase phenotypes. **American Journal of Physical Anthropology**, v. 58, n. 2, p. 173–178, 1982.

BAYOUMI, R. A. L. et al. Distribution of the lactase phenotypes in the population of the Democratic Republic of the Sudan. **Human Genetics**, v. 57, n. 3, p. 279–281, 1981.

BEJA-PEREIRA, A. et al. Gene-culture coevolution between cattle milk protein genes and human lactase genes. **Nature Genetics**, v. 35, n. 4, p. 311–313, 2003.

BERNARDES-SILVA, C. F. et al. Lactase persistence/non-persistence variants, *C/T_13910* and *G/A_22018*, as a diagnostic tool for lactose intolerance in IBS patients. **Clinica Chimica Acta**, v. 386, n. 1–2, p. 7–11, 2007.

BERSAGLIERI, T. et al. Genetic signatures of strong recent positive selection at the lactase gene. **American Journal of Human Genetics**, v. 74, n. 6, p. 1111–1120, 2004.

BERTONI, B. Admixture Mapping and Genetic Technologies: Perspectives from Latin America. In: GIBBON, S.; SANTOS, R. V.; SANS, M. (Eds.). . **Racial identities, genetic ancestry, and health in South America: Argentina, Brazil, Colombia and Uruguay**. Nova York: Palgrave Macmillan, 2011.

BEYERLEIN, L. et al. Correlation between symptoms developed after the oral ingestion of 50 g lactose and results of hydrogen breath testing for lactose intolerance. **Alimentary Pharmacology and Therapeutics**, v. 27, n. 8, p. 659–665, 2008.

- BLOOM, G.; SHERMAN, P. W. Dairying barriers affect the distribution of lactose malabsorption. **Evolution and Human Behavior**, v. 26, n. 4, p. 301–312, 2005.
- BODLAJ, G. et al. Genotyping of the Lactase-Phlorizin Hydrolase -13910 Polymorphism by LightCycler PCR and Implications for the Diagnosis of Lactose Intolerance. **Clinical Chemistry**, v. 52, n. 1, p. 145–148, 2006.
- BOLIN, T. D.; DAVIS, A. E. Primary Lactase Deficiency: Genetic or Acquired? **The American Journal of Digestive Diseases**, v. 15, n. 7, p. 679–692, 1970.
- BOLL, W.; WAGNER, P.; MANTEI, N. Structure of the chromosomal gene and cDNAs coding for lactase-phlorizin hydrolase in humans with adult-type hypolactasia or persistence of lactase. **American Journal of Human Genetics**, v. 48, n. 5, p. 889–902, 1991.
- BORBA, F. M. Usos do passado e apropriações patrimoniais em territórios quilombolas de Santa Catarina. **Revista Confluências Culturais**, v. 5, n. 2, p. 138, 2016.
- BORTOLINI, M. C. et al. Evolutionary relationships between black South American and African populations. **Human biology; an international record of research**, v. 67, n. 4, p. 547–559, 1995.
- BOSCHMANN, S. E. **Análise genética da persistência da lactase nas populações Menonita e euro-brasileira do Sul do Brasil**. 101f. Dissertação (Mestrado em Genética) – Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2012.
- BOSCHMANN, S. E. et al. The Frequency of the *LCT** -13910C>T Polymorphism Associated with Lactase Persistence Diverges among Euro-Descendant Groups from Brazil. **Medical Principles and Practice**, v. 25, n. 1, p. 18–20, 2016.
- BRAND, J. C. et al. Lactose malabsorption in Australian aborigines. **American Journal of Clinical Nutrition**, v. 37, n. 3, p. 449–452, 1983.
- BRAUN-PRADO, K. et al. HLA class I polymorphism, as characterised by PCR-SSOP, in a Brazilian exogamic population. **Tissue Antigens**, v. 56, n. 5, p. 417–427, 2000.
- BRETON, G. et al. Lactase persistence alleles reveal partial East African ancestry of Southern African Khoe pastoralists. **Current Biology**, v. 24, n. 8, p. 852–858, 2014.
- BULHÕES, A. C. et al. Correlation between lactose absorption and the C/T-13910 and G/A-22018 mutations of the lactase-phlorizin hydrolase (*LCT*) gene in adult-type hypolactasia. **Brazilian Journal of Medical and Biological Research**, v. 40, n. 11, p. 1441–1446, 2007.
- BÜNING, C. et al. The C/C-13910 and G/G-22018 genotypes for adult-type hypolactasia are not associated with inflammatory bowel disease. **Scandinavian Journal of Gastroenterology**, v. 38, n. 5, p. 538–542, 2003.
- BURGER, J. et al. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 10, p. 3736–3741, 2007.

BURGIO, G. R. et al. Prevalence of primary adult lactose malabsorption and awareness of milk intolerance in Italy. **The American Journal of Clinical Nutrition**, v. 39, p. 100–104, 1984.

BUTTRISS, J. Nutritional properties of fermented milk products. **International Journal of Dairy Technology**, v. 50, n. 1, p. 21–27, 1997.

CALÓGERAS, P. **Formação histórica do Brasil**. 3. ed. São Paulo: Companhia Editora Nacional, 1938.

CAMPBELL, A. K.; WAUD, J. P.; MATTHEWS, S. B. The molecular basis of lactose intolerance. **Science Progress**, v. 92, n. 3–4, p. 241–287, 2009.

CARNEIRO, E. Singularidade dos Quilombos. In: MOURA, C. (Ed.). **Os Quilombos na Dinâmica Social do Brasil**. Maceió: Editora da Universidade Federal de Alagoas, 2001. p. 11–20.

CARRERA, E.; NESHEIM, M. C.; CROMPTON, D. W. T. Lactose maldigestion preschool children Ascaris-infected preschool children. **The American Journal of Clinical Nutrition**, v. 39, p. 255–264, 1984.

CARVALHO-SILVA, D. R. et al. The phylogeography of Brazilian Y-chromosome lineages. **American Journal of Human Genetics**, v. 68, n. 1, p. 281–286, 2001.

CASKEY, D. A. et al. Effects of age on lactose malabsorption in Oklahoma Native Americans as determined by breath H₂ analysis. **The American Journal of Digestive Diseases**, v. 22, n. 2, p. 113–116, 1977.

CASTELLS, A. N. G. A cor do lugar: o Sertão do Valongo como patrimônio cultural. **Ilha: Revista de Antropologia**, v. 8, n. 1,2, p. 417–439, 2006.

CASTELLS, A. N. G. Políticas de patrimônio: entre exclusão e o direito à cidadania. **O público e o privado**, n. 10, p. 63–76, 2007.

CAVALLI-SFORZA, L. T. et al. Primary adult lactose malabsorption in Italy: Regional differences in prevalence and relationship to lactose intolerance and milk consumption. **American Journal of Clinical Nutrition**, v. 45, n. 4, p. 748–754, 1987.

CEDERLUND, A. et al. Lactose in Human Breast Milk an Inducer of Innate Immunity with Implications for a Role in Intestinal Homeostasis. **PLoS ONE**, v. 8, n. 1, p. e53876, 2013.

CHAGAS, M. DE F. A política do reconhecimento dos “remanescentes das comunidades dos quilombos”. **Horizontes Antropológicos**, v. 7, n. 15, p. 209–235, 2001.

CHALHOUB, S. **Visões da liberdade: uma história das últimas décadas da escravidão na Corte**. São Paulo: Companhia das Letras, 2011.

CHARNY, I. W. **Encyclopedia of Genocide - Volume 1 A-H**. 1. ed. Santa Barbara: ABC-CLIO, 2000.

CHIN, E. L. et al. Association of lactase persistence genotypes (rs4988235) and ethnicity with dairy intake in a healthy U.S. population. **Nutrients**, v. 11, n. 8, p. 1–23, 2019.

CHRISTIN, P. A.; WEINREICH, D. M.; BESNARD, G. Causes and evolutionary significance of genetic convergence. **Trends in Genetics**, v. 26, n. 9, p. 400–405, 2010.

CHUNG, M. H.; MCGILL, D. B. Lactase Deficiency in Orientals. **Gastroenterology**, v. 54, n. 2, p. 225–226, 1968.

COELHO, M. et al. On the edge of Bantu expansions: MtDNA, y chromosome and lactase persistence genetic variation in southwestern Angola. **BMC Evolutionary Biology**, v. 9, n. 1, p. 1–18, 2009.

COOK, G. C. Lactase Activity in Newborn and Infant Baganda. **British Medical Journal**, v. 1, p. 527–530, 1967.

COOK, G. C. Did Persistence of Intestinal Lactase Into Adult Life Originate on the Arabian Peninsula? **Royal Anthropological Institute of Great Britain and Ireland**, v. 13, n. 3, p. 418–427, 1978.

COOK, G. C.; AL-TORKI, M. T. High Intestinal Lactase Concentrations in Adult Arabs in Saudi Arabia. **British Medical Journal**, v. 3, n. 5976, p. 135–136, 1975.

COOK, G. C.; KAJUBI, S. K. Tribal incidence of lactase deficiency in Uganda. **Lancet**, v. 287, n. 7440, p. 725–729, 1966.

COPLEY, M. S. et al. Dairying in antiquity. III. Evidence from absorbed lipid residues dating to the British Neolithic. **Journal of Archaeological Science**, v. 32, n. 4, p. 523–546, 2005.

CORDAIN, L.; HICKEY, M. S.; KIM, K. Malaria and ricketts represent selective forces for the convergent evolution of adult lactase persistence. In: GEPTS, P. et al. (Eds.). . **Biodiversity in Agriculture: Domestication, Evolution, and Sustainability**. Cambridge: Cambridge University Press, 2012. p. 299–308.

COX, J. A.; ELLIOTT, F. G. Primary adult lactose intolerance in the Kivu Lake area: Rwanda and the Bushi. **The American Journal of Digestive Diseases**, v. 19, n. 8, p. 714–724, 1974.

CRAIG, O. E. et al. Did the first farmers of central and eastern Europe produce dairy foods? **Antiquity**, v. 79, n. 306, p. 882–894, 2005.

CRAIG, O. E. et al. Ancient lipids reveal continuity in culinary practices across the transition to agriculture in Northern Europe. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, n. 44, p. 17910–17915, 2011.

CRUZ, L. M.; HANAZAKI, N. The Ethnobotany of an Afro-Brazilian Community at Sertão do Valongo, Santa Catarina, Brazil. **Functional Ecosystems and Communities**, v. 2, n. 1, p. 59–65, 2008.

CUATRECASAS, P.; LOCKWOOD, D. H.; CALDWELL, J. R. Lactase Deficiency in the adult: A Common Occurrence. **The Lancet**, v. 285, n. 7375, p. 14–18, 1965.

DAHLQVIST, A. et al. Intestinal Lactase Deficiency and Lactose Intolerance in Adults. Preliminary Report. **Gastroenterology**, v. 45, p. 488–491, 1964.

DAVEY SMITH, G. et al. Lactase persistence-related genetic variant: Population substructure and health outcomes. **European Journal of Human Genetics**, v. 17, n. 3, p. 357–367, 2009.

DE BASSI, R. M. A. **Casamentos consangüíneos em populações brasileiras e alguns parâmetros migracionais e etários, associados a casamentos, em Curitiba**. 77f. Tese (Mestrado em Genética) – Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 1983.

DELGADO, P. R. Negros no Paraná na primeira década do século XXI: características demográficas e desigualdades raciais. In: RAGGIO, A. Z.; BLEY, R. B.; TRAUZYNSKI, S. C. (Eds.). **Abordagem histórica sobre a população negra no estado do Paraná**. Curitiba: SEJU, 2018. p. 46–74.

DIAMOND, J. **Armes, Germes e Aço**. 23. ed. Rio de Janeiro: Record, 2019.

DISSANAYAKE, A. S. et al. Prevalence of primary adult lactose malabsorption in the Eastern province of Saudi Arabia. **Annals of Saudi Medicine**, v. 10, n. 6, p. 598–601, 1990.

DONG, S. K.; LONG, R. J.; KANG, M. Y. Milking and milk processing: traditional technologies in the yak farming system of the Qinghai-Tibetan Plateau, China. **International Journal of Dairy Technology**, v. 56, n. 2, p. 86–93, 2003.

DUNCAN, I. W.; SCOTT, E. M. Lactose intolerance in Alaskan Indians and Eskimos. **American Journal of Clinical Nutrition**, v. 25, n. 9, p. 867–868, 1972.

DUNNE, J. et al. First dairying in green Saharan Africa in the fifth millennium BC. **Nature**, v. 486, n. 7403, p. 390–394, 2012.

DUNPHY, J. V. et al. Intestinal Lactase Deficit in Adults. **Gastroenterology**, v. 49, n. 1, p. 12–21, 1965.

DURHAM, W. H. **Coevolution: Genes, Culture and Human Diversity**. 1. ed. Stanford: Stanford University Press, 1991.

ENATTAH, N. S. et al. Identification of a variant associated with adult-type hypolactasia. **Nature Genetics**, v. 30, n. 2, p. 233–237, 2002.

ENATTAH, N. S. et al. Evidence of still-ongoing convergence evolution of the lactase persistence *T-13910* alleles in humans. **American Journal of Human Genetics**, v. 81, n. 3, p. 615–625, 2007.

ENATTAH, N. S. et al. Independent Introduction of Two Lactase-Persistence Alleles into Human Populations Reflects Different History of Adaptation to Milk Culture. **American Journal of Human Genetics**, v. 82, n. 1, p. 57–72, 2008.

EVERSHED, R. P. et al. Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. **Nature**, v. 455, n. 7212, p. 528–531, 2008.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564–567, 2010.

FAN, S. et al. Review of Recent Human Adaptation. **Science**, v. 354, n. 6308, 2016.

FELIPE, D. A. A presença negra na história do Paraná: pelo direito à memória. In: RAGGIO, A. Z.; BLEY, R. B.; TRAUZYNSKI, S. C. (Eds.). **Abordagem histórica sobre a população negra no estado do Paraná**. Curitiba: SEJU, 2018. p. 7–24.

FERGUSON, A.; MAXWELL, J. D. Genetic aetiology of lactose intolerance. **The Lancet**, v. 290, n. 7508, p. 188–190, 1967.

FERNÁNDEZ, C. I. et al. Lactase non-persistence and general patterns of dairy intake in indigenous and mestizo Chilean populations. **American Journal of Human Biology**, v. 28, n. 2, p. 213–219, 2016.

FERNÁNDEZ, C. I.; FLORES, S. V. Brief communication: Lactase persistence and dairy intake in Mapuche and Mestizo populations from Southern Chile. **American Journal of Physical Anthropology**, v. 155, n. 3, p. 482–487, 2014.

FERREIRA, L. B. et al. Genomic Ancestry of a Sample Population From the State of São Paulo, Brazil. **American Journal of Human Biology**, v. 18, p. 702–705, 2006.

FEWLASS, H. et al. Chemical evidence of dairying by hunter-gatherers in highland Lesotho in the late first millennium AD. **Nature Human Behaviour**, 2020.

FIABANI, A. O quilombo antigo e o quilombo contemporâneo: verdades e construções. **XXIV Simpósio Nacional de História**, p. 256–257, 2007.

FIGUEROA, R. B. et al. Intestinal lactase deficiency in an apparently normal Peruvian population. **The American Journal of Digestive Diseases**, v. 16, n. 10, p. 881–889, 1971.

FLATZ, G. et al. Distribution of physiological adult lactase phenotypes, lactose absorber and malabsorber, in Germany. **Human Genetics**, v. 62, n. 2, p. 152–157, 1982.

FLATZ, G. Genetics of lactose digestion in humans. **Advances in Human Genetics**, v. 16, p. 1–77, 1987.

FLATZ, G.; ROTTHAUWE, H. W. Lactose Nutrition and Natural Selection. **The Lancet**, v. 302, n. 7820, p. 76–77, 1973.

FLATZ, G.; SAENGUDOM, C. Lactose Tolerance in Asians: a Family Study. **Nature**, v. 224, p. 915–916, 1969.

FLATZ, G.; SCHILDGE, C.; SEKOU, H. Distribution of adult lactase phenotypes in the Tuareg of Niger. **American Journal of Human Genetics**, v. 38, n. 4, p. 515–520,

1986.

FLINT-GARCIA, S. A.; THORNSBERRY, J. M.; BUCKLER IV, E. S. Structure of Linkage Disequilibrium in Plants. **Annual Review of Plant Biology**, v. 54, p. 357–374, 2003.

FLORENTINO, M. De escravos, forros e fujões no Rio de Janeiro imperial. **Revista USP2**, n. 59, p. 104–115, 2003.

FREIRE-MAIA, N. Genetic effects in Brazilian populations due to consanguineous marriages. **American Journal of Medical Genetics**, v. 35, n. 1, p. 115–117, 1990.

FREITAS, F. O Determinismo Biológico do século XIX , e o Genético do Século XX. **Revista África e Africanidades**, v. 2, n. 5, 2009.

FRIEDRICH, D. C. et al. Several Different Lactase Persistence Associated Alleles and High Diversity of the Lactase Gene in the Admixed Brazilian Population. **PLoS ONE**, v. 7, n. 9, p. 1–8, 2012a.

FRIEDRICH, D. C. et al. Stability or variation? Patterns of lactase gene and its enhancer region distributions in Brazilian Amerindians. **American Journal of Physical Anthropology**, v. 147, n. 3, p. 427–432, 2012b.

GABRIEL, S. B. et al. The structure of haplotype blocks in the human genome. **Science**, v. 296, n. 5576, p. 2225–2229, 2002.

GADE, D. W. **Nature and culture in the Andes**. 1. ed. Madison: University of Wisconsin Press, 1999.

GALLEGO ROMERO, I. et al. Herders of indian and european cattle share their predominant allele for lactase persistence. **Molecular Biology and Evolution**, v. 29, n. 1, p. 249–260, 2012.

GAMBA, C. et al. Genome flux and stasis in a five millennium transect of European prehistory. **Nature Communications**, v. 5, p. 1–9, 2014.

GENOVESE, G. et al. Mapping the human reference genome's missing sequence by three-way admixture in Latino Genomes. **American Journal of Human Genetics**, v. 93, n. 3, p. 411–421, 2013.

GILAT, T. et al. Genetics of Primary Adult Lactase Deficiency. **Gastroenterology**, v. 64, n. 4, p. 562–568, 1973.

GOEBEL, T.; WATERS, M. R.; O'ROURKE, D. H. The Late Pleistocene dispersal of modern humans in the Americas. **Science**, v. 319, n. 5869, p. 1497–1502, 2008.

GOMES, L. Justiça seja feita: Direito Quilombola ao Território. In: ALMEIDA, A. W. B. et al. (Eds.). . **Caderno de debates Nova Cartografia Social: Territórios quilombolas e conflitos**. Manaus: Projeto Nova Cartografia Social da Amazônia / UEA Edições, 2010. p. 187–196.

GONÇALVES, V. F. et al. The phylogeography of African Brazilians. **Human Heredity**,

v. 65, n. 1, p. 23–32, 2008.

GONZÁLEZ-ANDRADE, F. et al. Sex-specific genetic admixture of Mestizos, Amerindian Kichwas, and Afro-Ecuadorans from Ecuador. **Human Biology**, v. 79, n. 1, p. 51–77, 2007.

GOUVEIA, M. H. et al. Origins, Admixture Dynamics, and Homogenization of the African Gene Pool in the Americas. **Molecular Biology and Evolution**, v. 37, n. 6, p. 1–10, 2020.

GRAVEL, S. et al. Reconstructing Native American Migrations from Whole-Genome and Whole-Exome Data. **PLoS Genetics**, v. 9, n. 12, 2013.

GREGORY, W. K. **The order of mammals**. 1. ed. Nova York: The Trustees, 1910.

GUTIÉRREZ, H. Crioulos e Africanos no Paraná, 1798-1830. **Revista Brasileira de História**, v. 8, n. 16, p. 161–188, 1988.

HAEMMERLI, U. P. et al. Acquired Milk Intolerance in the Adult Caused by Lactose Malabsorption Due to a Selective Deficiency of Intestinal Lactase Activity. **American Journal of Medicine**, v. 38, p. 7–30, 1965.

HARRISON, G. G. Primary Adult Lactase Deficiency: A Problem in Anthropological Genetics. **American Anthropologist**, v. 77, n. 4, p. 812–835, 1975.

HARTWIG, F. P. et al. Association of lactase persistence genotype with milk consumption, obesity and blood pressure: A Mendelian randomization study in the 1982 Pelotas (Brazil) Birth Cohort, with a systematic review and meta-analysis. **International Journal of Epidemiology**, v. 45, n. 5, p. 1573–1587, 2016.

HARVEY, C. B. et al. Regional localization of the lactase-phlorizin hydrolase gene, *LCT*, to chromosome 2q21. **Annals of Human Genetics**, v. 57, n. 3, p. 179–185, 1993.

HARVEY, C. B. et al. DNA polymorphisms in the lactase gene: Linkage Disequilibrium across the 70-kb region. **European Journal of Human Genetics**, v. 3, n. 1, p. 27–41, 1995.

HARVEY, C. B. et al. Characterisation of a human homologue of a yeast cell division cycle gene, *MCM6*, located adjacent to the 5' end of the lactase gene on chromosome 2q21. **FEBS Letters**, v. 398, n. 2–3, p. 135–140, 1996.

HARVEY, C. B. et al. Lactase haplotype frequencies in Caucasians: Association with the lactase persistence/non-persistence polymorphism. **Annals of Human Genetics**, v. 62, n. 3, p. 215–223, 1998.

HASSAN, H. Y. et al. Genetic diversity of lactase persistence in East African populations. **BMC Research Notes**, v. 9, n. 1, p. 1–7, 2016.

HATZIKOTOULAS, K.; GILLY, A.; ZEGGINI, E. Using population isolates in genetic association studies. **Briefings in Functional Genomics**, v. 13, n. 5, p. 371–377, 2014.

HE, T. et al. The role of colonic metabolism in lactose intolerance. **European Journal**

of **Clinical Investigation**, v. 38, n. 8, p. 541–547, 2008.

HERTZLER, S. R.; SAVAIANO, D. A. Colonic adaptation to daily lactose feeding in lactose maldigesters reduces lactose intolerance. **American Journal of Clinical Nutrition**, v. 64, n. 2, p. 232–236, 1996.

HEYER, E. et al. Lactase persistence in Central Asia: Phenotype, genotype, and evolution. **Human Biology**, v. 83, n. 3, p. 379–392, 2011.

HIJAZI, S. S. et al. Distribution of adult lactase phenotypes in Bedouins and in urban and agricultural populations of Jordan. **Tropical and Geographical Medicine**, v. 35, n. 2, p. 157–161, 1983.

HILL, S. C.; MOHAMMAD, T. R.; TOOMASKIVISILD. Brief communication: Effect of nomadic subsistence practices on lactase persistence associated genetic variation in Kuwait. **American Journal of Physical Anthropology**, v. 152, n. 1, p. 140–144, 2013.

HILL, W. G.; ROBERTSON, A. Linkage Disequilibrium in Finite Populations. **Theoretical and Applied Genetics**, v. 38, p. 226–231, 1968.

HO, M. W.; POVEY, S.; SWALLOW, D. M. Lactase polymorphism in adult British natives: Estimating allele frequencies by enzyme assays in autopsy samples. **American Journal of Human Genetics**, v. 34, n. 4, p. 650–657, 1982.

HOLDEN, C.; MACE, R. Phylogenetic Analysis of the Evolution of Lactose Digestion in Adults. **Human Biology**, v. 69, n. 5, p. 605–628, 1997.

HOLLOX, E. J. et al. Lactase haplotype diversity in the old world. **American Journal of Human Genetics**, v. 68, n. 1, p. 160–172, 2001.

HOLSINGER, K. E.; WEIR, B. S. Genetics in geographically structured populations: Defining, estimating and interpreting FST. **Nature Reviews Genetics**, v. 10, n. 9, p. 639–650, 2009.

HOLZEL, A.; SCHWARZ, V.; SUTCLIFFE, K. W. Defective lactose absorption causing malnutrition in infancy. **The Lancet**, p. 1126–1128, 1959.

HUBÁČEK, J. A. et al. Frequency of adult type-associated lactase persistence *LCI13910C/T* genotypes in the Czech/Slav and Czech Roma/Gypsy populations. **Genetics and Molecular Biology**, v. 40, n. 2, p. 450–452, 2017.

HÜNEMEIER, T. et al. Niger-Congo Speaking Populations and the Formation of the Brazilian Gene Pool: mtDNA and Y-Chromosome Data. **American Journal of Physical Anthropology**, v. 133, n. 2, p. 854–867, 2007.

HUSSEIN, L.; EZZILARAB, A. The frequency distribution of lactose malabsorption among adult populations from the Eastern and Western Egyptian Deserts. **Biochemical Genetics**, v. 32, n. 9–10, p. 331–342, 1994.

IBGE. **Brasil: 500 anos de povoamento**. Rio de Janeiro: IBGE, 2007.

IBGE. **Censo Demográfico 2010. Características da População e dos Domicílios**.

Rio de Janeiro: IBGE, 2011.

IBGE. **Pesquisa Nacional por Amostra de Domicílios (PNAD): Síntese de indicadores 2015**. Rio de Janeiro: IBGE, 2016.

IBGE. **Estimativas da população residente com data de referência 1º de julho de 2019**. Brasília: IBGE, 2019.

IMTIAZ, F. et al. The *T/G 13915* variant upstream of the lactase gene (*LCT*) is the founder allele of lactase persistence in an urban Saudi population. **Journal of Medical Genetics**, v. 44, n. 10, p. 2–5, 2007.

INGRAM, C. J. E. et al. A novel polymorphism associated with lactose tolerance in Africa: Multiple causes for lactase persistence? **Human Genetics**, v. 120, n. 6, p. 779–788, 2007.

INGRAM, C. J. E. et al. Lactose digestion and the evolutionary genetics of lactase persistence. **Human Genetics**, v. 124, n. 6, p. 579–591, 2009a.

INGRAM, C. J. E. et al. Multiple rare variants as a cause of a common phenotype: Several different lactase persistence associated alleles in a single ethnic group. **Journal of Molecular Evolution**, v. 69, n. 6, p. 579–588, 2009b.

INSTITUTO DE TERRAS, CARTOGRAFIA E GEOLOGIA DO PARANÁ (ITCG). **População negra e comunidades quilombolas no estado do Paraná**. Disponível em: <https://www.iat.pr.gov.br/sites/agua-terra/arquivos_restritos/files/documento/2020-04/populacao_negra.pdf>. Acesso em: 7 jul. 2020.

ITAN, Y. et al. The origins of lactase persistence in Europe. **PLoS Computational Biology**, v. 5, n. 8, p. 17–19, 2009.

ITAN, Y. et al. A worldwide correlation of lactase persistence phenotype and genotypes. **BMC Evolutionary Biology**, v. 10, n. 36, 2010.

JENSEN, R. G.; BLANC, B.; PATTON, S. B. Particulate Constituents in Human and Bovine Milks. In: JENSEN, R. G. (Ed.). **Handbook of Milk Composition**. 1. ed. Cambridge: Academic Press, 1995. p. 50–62.

JENSEN, T. G. K. et al. The -14010*C variant associated with lactase persistence is located between an Oct-1 and HNF1 α binding site and increases lactase promoter activity. **Human Genetics**, v. 130, n. 4, p. 483–493, 2011.

JEONG, C. et al. Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 48, p. E11248–E11255, 2018.

JESSE, F. et al. “I Hope Your Cattle Are Well”: Archaeological Evidence for Early Cattle-centred Behaviour in the Eastern Sahara of Sudan and Chad. In: BOLLING, M.; SCHNEGG, M.; WOTZKA, H.-P. (Eds.). **Pastoralism in Africa: Past, Present and Future**. 1. ed. Nova York: Library of Congress Cataloging-in-Publication Data, 2013. p. 63–103.

JONES, B. L. et al. Diversity of lactase persistence alleles in Ethiopia: Signature of a soft selective sweep. **American Journal of Human Genetics**, v. 93, n. 3, p. 538–544, 2013.

JONES, B. L. et al. Diversity of lactase persistence in African milk drinkers. **Human Genetics**, v. 134, n. 8, p. 917–925, 2015.

KADWELL, M. et al. Genetic analysis reveals the wild ancestors of the llama and the alpaca. **Proceedings of the Royal Society B: Biological Sciences**, v. 268, n. 1485, p. 2575–2584, 2001.

KARCZEWSKI, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. **Nature**, v. 581, n. 7809, p. 434–443, 2020.

KARCZEWSKI, K. J.; MARTIN, A. R. Analytic and Translational Genetics. **Annual Review of Biomedical Data Science**, v. 18, n. 4, p. 217–241, 2020.

KENT, M.; WADE, P. Genetics against race: Science, politics and affirmative action in Brazil. **Social Studies of Science**, v. 45, n. 6, p. 816–838, 2015.

KERLIN, P.; WONG, L. Lactose tolerance despite hypolactasia in adult coeliac disease. **Journal of Gastroenterology and Hepatology**, v. 2, n. 3, p. 233–237, 1987.

KEUSCH, G. T. et al. Lactase deficiency in Thailand: effect of prolonged lactose feeding. **The American Journal of Clinical Nutrition**, v. 22, n. 5, p. 638–641, 1969.

KIMURA, L. et al. Genomic ancestry of rural African-derived populations from Southeastern Brazil. **American Journal of Human Biology**, v. 25, n. 1, p. 35–41, 2013.

KLEIN, H. S. A Demografia do Tráfico Atlântico de Escravos para o Brasil. **Estudos Econômicos**, v. 17, n. 2, p. 129–149, 1987.

KOZLOV, A. I. Hypolactasia in the Indigenous Populations of Northern Russia. **International Journal of Circumpolar Health**, v. 57, n. 1, p. 18–21, 1998.

KRETCHMER, N. Lactose and lactase – A historical perspective. **Gastroenterology**, v. 61, n. 6, p. 805–813, 1971.

KRETCHMER, N. Lactose and Lactase. **Scientific American**, v. 227, n. 4, p. 70–79, 1972.

KRUSE, T. A. et al. The human lactase-phlorizin hydrolase gene is located on chromosome 2. **FEBS Letters**, v. 240, n. 1–2, p. 123–126, 1988.

KRÜTTLI, A. et al. Ancient DNA analysis reveals high frequency of european lactase persistence allele (*T-13910*) in medieval Central Europe. **PLoS ONE**, v. 9, n. 1, 2014.

KUHN, N. J.; CARRICK, D. T.; WILDE, C. J. Lactose Synthesis: The Possibilities of Regulation. **Journal of Dairy Science**, v. 63, n. 2, p. 328–336, 1980.

KUOKKANEN, M. et al. Transcriptional regulation of the lactase-phlorizin hydrolase

gene by polymorphisms associated with adult-type hypolactasia. **Gut**, v. 52, n. 5, p. 647–652, 2003.

LACERDA, M. B. **Colonização dos corpos: Ensaio sobre o público e o privado. Patriarcalismo, patrimonialismo, personalismo e a violência contra as mulheres na formação do Brasil**. 114f. Dissertação (Mestrado em Direito) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.

LAU FILHO, E. Aspectos históricos da presença do negro no Paraná. In: RAGGIO, A. Z.; BLEY, R. B.; TRAUZYNSKI, S. C. (Eds.). **Abordagem histórica sobre a população negra no estado do Paraná**. Curitiba: SEJU, 2018. p. 38–45.

LEBENTHAL, E.; ANTONOWICZ, I.; SHWACHMAN, H. Correlation of lactase activity, lactose tolerance and milk consumption in different age groups. **The American Journal of Clinical Nutrition**, v. 28, p. 595–600, 1975.

LEFÈVRE, C. M.; SHARP, J. A.; NICHOLAS, K. R. Evolution of Lactation: Ancient Origin and Extreme Adaptations of the Lactation System. **Annual Review of Genomics and Human Genetics**, v. 11, n. 1, p. 219–238, 2010.

LEICHTER, J. Lactose tolerance in a Slavic population. **The American Journal of Digestive Diseases**, v. 17, n. 1, p. 73–76, 1972.

LEITE, I. B. O projeto político quilombola: Desafios, conquistas e impasses atuais. **Revista Estudos Feministas**, v. 16, n. 3, p. 965–977, 2008.

LEITE, I. B. Humanidades Insurgentes: Conflitos e Criminalização dos Quilombos. In: ALMEIDA, A. W. B. et al. (Eds.). **Caderno de debates Nova Cartografia Social: Territórios quilombolas e conflitos**. Manaus: Projeto Nova Cartografia Social da Amazônia / UEA Edições, 2010. p. 18–41.

LEITE, T. K. M. et al. Genomic ancestry, self-reported “color” and quantitative measures of skin pigmentation in Brazilian admixed siblings. **PLoS ONE**, v. 6, n. 11, p. 1–9, 2011.

LEMBER, M. et al. Lactase non-persistence and milk consumption in Estonia. **World Journal of Gastroenterology**, v. 12, n. 45, p. 7329–7331, 2006.

LEMES, R. B. et al. Estimation of inbreeding and substructure levels in African-derived Brazilian quilombo populations. **Human Biology**, v. 86, n. 4, p. 276–288, 2015.

LEONARDI, M. et al. The evolution of lactase persistence in Europe. A synthesis of archaeological and genetic evidence. **International Dairy Journal**, v. 22, n. 2, p. 88–97, 2012.

LEWINSKY, R. H. et al. *T-13910* DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. **Human Molecular Genetics**, v. 14, n. 24, p. 3945–3953, 2005.

LEWONTIN, R. C. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. **Genetics**, v. 49, n. 1, p. 49–67, 1964.

LIEBERT, A. et al. In Vitro Functional Analyses of Infrequent Nucleotide Variants in the Lactase Enhancer Reveal Different Molecular Routes to Increased Lactase Promoter Activity and Lactase Persistence. **Annals of Human Genetics**, v. 80, n. 6, p. 307–318, 2016.

LIEBERT, A. et al. World-wide distributions of lactase persistence alleles and the complex effects of recombination and selection. **Human Genetics**, v. 136, n. 11–12, p. 1445–1453, 2017.

LIN, C. J.; BARBOSA, A. S. Técnicas de Análise da Regulação da Transcrição Gênica e suas Aplicações na Endocrinologia Molecular. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 46, n. 4, p. 330–340, 2002.

LINS, T. C. et al. Genetic heterogeneity of self-reported ancestry groups in an admixed Brazilian population. **Journal of Epidemiology**, v. 21, n. 4, p. 240–245, 2011.

LLOYD, M. et al. Regulation of intestinal lactase in adult hypolactasia. **The Journal of Clinical Investigation**, v. 89, n. 2, p. 524–529, 1992.

LOKKI, A. I. et al. Lactase persistence genotypes and malaria susceptibility in Fulani of Mali. **Malaria Journal**, v. 10, n. 9, 2011.

LOMER, M. C. E.; PARKES, G. C.; SANDERSON, J. D. Review article: Lactose intolerance in clinical practice - myths and realities. **Alimentary Pharmacology and Therapeutics**, v. 27, n. 2, p. 93–103, 2008.

LUIZON, M. R. **Dinâmica da mistura étnica em comunidades remanescentes de quilombo brasileiras**. 115f. Tese (Doutorado em Genética) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

MACFARLANE, G. T.; CUMMINGS, J. H. Probiotics and prebiotics: can regulating the activities of intestinal bacteria benefit health? **British Medical Journal**, v. 318, n. 7189, p. 999–1003, 1999.

MACHADO, T. M. B. **Ancestralidade em Salvador - BA**. 85f. Dissertação (Mestrado em Biotecnologia em Saúde e Medicina Investigativa) – Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, 2008.

MACHOLDT, E. et al. Tracing pastoralist migrations to southern Africa with lactase persistence alleles. **Current Biology**, v. 24, n. 8, p. 875–879, 2014.

MACHOLDT, E. et al. New insights into the history of the *C-14010* lactase persistence variant in Eastern and Southern Africa. **American Journal of Physical Anthropology**, v. 156, n. 4, p. 661–664, 2015.

MALMSTRÖM, H. et al. High frequency of lactose intolerance in a prehistoric hunter-gatherer population in northern Europe. **BMC Evolutionary Biology**, v. 10, n. 1, 2010.

MANTA, F. S. DE N. et al. Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. **PLoS ONE**, v. 8, n. 9, p. 1–11, 2013.

MARQUESE, R. DE B. A dinâmica da escravidão no Brasil: resistência, tráfico negreiro

e alforrias, séculos XVII a XIX. **Novos Estudos - CEBRAP**, n. 74, p. 107–123, 2006.

MARRERO, A. R. et al. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. **American Journal of Human Biology**, v. 17, n. 4, p. 496–506, 2005.

MARTIN, A. R. et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. **American Journal of Human Genetics**, v. 100, n. 4, p. 635–649, 2017.

MARTINS, R. **História do Paraná**. Curitiba: Travessa dos Editores, 1995.

MARTON, A.; XUE, X.; SZILAGYI, A. Meta-analysis: The diagnostic accuracy of lactose breath hydrogen or lactose tolerance tests for predicting the North European lactase polymorphism *C/T-13910*. **Alimentary Pharmacology and Therapeutics**, v. 35, n. 4, p. 429–440, 2012.

MATHIESON, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. **Nature**, v. 528, p. 499–503, 2015.

MATTAR, R. et al. Single nucleotide polymorphism *C/T-13910*, located upstream of the lactase gene, associated with adult-type hypolactasia: Validation for clinical practice. **Clinical Biochemistry**, v. 41, n. 7–8, p. 628–630, 2008.

MATTAR, R. et al. Frequency of *LCT -13910C>T* single nucleotide polymorphism associated with adult-type hypolactasia/lactase persistence among Brazilians of different ethnic groups. **Nutrition Journal**, v. 8, n. 1, p. 1–3, 2009.

MATTAR, R. et al. *LCT-22018G>A* single nucleotide polymorphism is a better predictor of adult-type hypolactasia/lactase persistence in Japanese-Brazilians than *LCT-13910C>T*. **Clinics**, v. 65, n. 12, p. 1399–1400, 2010.

MATTAR, R. et al. Comparison of Quick Lactose Intolerance Test in duodenal biopsies of dyspeptic patients with single nucleotide polymorphism *LCT-13910C>T* associated with primary hypolactasia/lactase-persistence. **Acta Cirurgica Brasileira**, v. 28, n. suppl 1, p. 77–82, 2013.

MATTAR, R.; MAZO, D. F. D. C. Intolerância à lactose: mudança de paradigmas com a Biologia Molecular. **Revista da Associação Médica Brasileira**, v. 56, n. 2, p. 230–236, 2010.

MATTAR, R.; MAZO, D. F. DE C.; CARRILHO, F. J. Lactose intolerance: Diagnosis, genetic, and clinical factors. **Clinical and Experimental Gastroenterology**, v. 5, n. 1, p. 113–121, 2012.

MATTHEWS, S. B. et al. Systemic lactose intolerance: A new perspective on an old problem. **Postgraduate Medical Journal**, v. 81, n. 953, p. 167–173, 2005.

MAYR, E. Change of genetic environment and evolution. In: HUXLEY, J. S.; HARDY, A. C.; FORD, E. B. (Eds.). **Evolution as a Process**. Londres: Allen & Unwin, 1954. p. 157–180.

MCBEAN, L. D.; MILLER, G. D. Allaying fears and fallacies about lactose intolerance. **Journal of the American Dietetic Association**, v. 98, n. 6, p. 671–676, 1998.

MCCRACKEN, R. D. Lactase Deficiency: An Example of Dietary Evolution. **Current Anthropology**, v. 12, n. 4/5, p. 479–517, 1971.

MCINTOSH, S. K.; SCHEINFELDT, L. B. It's Getting Better All the Time: Comparative Perspectives from Oceania and West Africa on Genetic Analysis and Archaeology. **African Archaeological Review**, v. 29, n. 2–3, p. 131–170, 2012.

MELONI, T. et al. Primary Lactase Deficiency and Past Malarial Endemicity in Sardinia. **Italian Journal of Gastroenterology Hepatologic**, v. 30, n. 5, p. 490–493, 1998.

MENDEL, L. B.; MITCHELL, P. H. Chemical Studies on Growth - I. The inverting enzymes of the alimentary tract, especially in the embryo. **American Journal of Physiology**, v. 20, p. 81–96, 1907.

MERRIWETHER, D. A. et al. Mitochondrial Versus Nuclear Admixture Estimates Demonstrate a Past History of Directional Mating. **American Journal of Physical Anthropology**, v. 102, p. 153–159, 1997.

MÉTNEKI, J. et al. A study of lactose absorption capacity in twins. **Human Genetics**, v. 67, n. 3, p. 296–300, 1984.

MIN-SHENG, P. et al. Lactase persistence may have an independent origin in Tibetan populations from Tibet, China. **Journal of human genetics**, v. 57, n. 6, p. 394–397, 2012.

MONTALTO, M. et al. Management and treatment of lactose malabsorption. **World Journal of Gastroenterology**, v. 12, n. 2, p. 187–191, 2006.

MONTALVA, N. et al. Adaptation to milking agropastoralism in Chilean goat herders and nutritional benefit of lactase persistence. **Annals of Human Genetics**, v. 83, n. 1, p. 11–22, 2019.

MONTGOMERY, R. K. et al. Lactose and lactase - Who is lactose intolerant and why? **Journal of Pediatric Gastroenterology and Nutrition**, v. 45, n. SUPPL. 2, p. S131–S137, 2007.

MORALES, E. et al. The European lactase persistence genotype determines the lactase persistence state and correlates with gastrointestinal symptoms in the Hispanic and Amerindian Chilean population: a case-control and population-based study. **BMJ Open**, v. 1, n. e000125, 2011.

MOREIRA, A. L. G. et al. Teor de lactose em leites UHT que declaram ser zero lactose. In: STEVENS, D. et al. (Org). **JC na Escola Ciência, Tecnologia e Sociedade: Ciência Alimentando o Brasil**. 2. Ed. São Paulo: Centro Paula Souza, 2018.

MOTA, L. T. **Presença e resistência Kaingang no Paraná**. 412f. Dissertação (Mestrado em Ciências Sociais) - Pontifícia Universidade Católica de São Paulo, São Paulo, 1992.

MOURA, C. **Os quilombos e a rebelião negra**. 5. ed. São Paulo: Editora Brasiliense, 1986. v. 12

MOURA, C. **História do negro brasileiro**. 2. ed. São Paulo: Editora Ática, 1992.

MOURA, R. R. et al. Meta-analysis of Brazilian genetic admixture and comparison with other Latin America countries. **American Journal of Human Biology**, v. 27, n. 5, p. 674–680, 2015.

MOYA, J. Migração e formação histórica da América Latina em perspectiva global. **Sociologias**, v. 20, n. 49, p. 24–68, 2018.

MUKUNA, K. W. O Contato Musical Transatlântico: Contribuição Bantu na Música Popular Brasileira. **África: Revista do Centro de Estudos Africanos da USP**, v. 1, n. 1, p. 97–101, 1978.

MULCARE, C. A. et al. The T allele of a single-nucleotide polymorphism 13.9 kb upstream of the lactase gene (*LCT*) (C - 13.9kbT) does not predict or cause the lactase-persistence phenotype in africans. **American Journal of Human Genetics**, v. 74, n. 6, p. 1102–1110, 2004.

MUNANGA, K. Origem e histórico do quilombo em África. In: MOURA, C. (Ed.). **Os Quilombos na Dinâmica Social do Brasil**. Maceió: Editora da Universidade Federal de Alagoas, 2001. p. 21–34.

MUNIZ, Y. C. N. et al. Genomic ancestry in urban Afro-Brazilians. **Annals of Human Biology**, v. 35, n. 1, p. 104–111, 2008.

MUSTAPHA, A.; HERTZLER, S. R.; SAVAIANO, D. A. Lactose: Nutritional Significance. In: FOX, P. F. (Ed.). **Advanced Dairy Chemistry - Volume 3**. 2. ed. Londres: Chapman & Hall, 1997. p. 127–154.

MYLES, S. et al. Genetic evidence in support of a shared Eurasian-North African dairying origin. **Human Genetics**, v. 117, n. 1, p. 34–42, 2005.

MYO-KHIN et al. Ineffectiveness of breath methane excretion as a diagnostic test for lactose malabsorption. **Journal of Pediatric Gastroenterology and Nutrition**, v. 28, n. 5, p. 474–479, 1999.

NAGY, D. et al. Prevalence of adult-type hypolactasia as diagnosed with genetic and lactose hydrogen breath tests in Hungarians. **European Journal of Clinical Nutrition**, v. 63, n. 7, p. 909–912, 2009.

NAGY, D. et al. Comparison of lactase persistence polymorphism in ancient and present-day Hungarian populations. **American Journal of Physical Anthropology**, v. 145, n. 2, p. 262–269, 2011.

NEI, M.; MARUYAMA, T.; CHAKRABORTY, R. The Bottleneck Effect and Genetic Variability in Populations. **Evolution**, v. 29, n. 1, p. 1–10, 1975.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. 6. ed. Porto Alegre: Artmed, 2014.

NEVANLINNA, H. R. The Finnish population structure A genetic and genealogical study. **Hereditas**, v. 71, n. 2, p. 195–235, 1972.

NICHOLS, B. L. et al. Congenital maltase-glucoamylase deficiency associated with lactase and sucrase deficiencies. **Journal of Pediatric Gastroenterology and Nutrition**, v. 35, n. 4, p. 573–579, 2002.

NOLAN, D. J. et al. Genetic adult lactase persistence is associated with risk of Crohn's Disease in a New Zealand population. **BMC Research Notes**, v. 3, n. 339, 2010.

NORDBORG, M.; TAVARÉ, S. Linkage disequilibrium: what history has to tell us. **TRENDS in Genetics**, v. 18, n. 2, p. 1–11, 2001.

NORIO, R.; NEVANLINNA, H. R.; PERHEENTUPA, J. Hereditary diseases in Finland; rare flora in rare soul. **Annals of Clinical Research**, v. 5, n. 3, p. 109–141, 1973.

NOSE, O. et al. Breath hydrogen test for detecting lactose malabsorption in infants and children. Prevalence of lactose malabsorption in Japanese children and adults. **Archives of Disease in Childhood**, v. 54, n. 6, p. 436–440, 1979.

OJEDA-GRANADOS, C. et al. Association of lactase persistence genotypes with high intake of dairy saturated fat and high prevalence of lactase non-persistence among the Mexican population. **Journal of Nutrigenetics and Nutrigenomics**, v. 9, n. 2–4, p. 83–94, 2016.

OLALDE, I. et al. The Beaker phenomenon and the genomic transformation of northwest Europe. **Nature**, v. 555, n. 7695, p. 190–196, 2018.

OLDS, L. C.; AHN, J. K.; SIBLEY, E. -13915*G DNA polymorphism associated with lactase persistence in Africa interacts with Oct-1. **Human Genetics**, v. 129, n. 1, p. 111–113, 2011.

OLIVEIRA, L. C. et al. Lack of Association between *LCT_rs140433552*CA>del* Indel Polymorphism and Lactose Intolerance in a Southern Brazilian Population. **Lifestyle Genomics**, v. 13, n. 4, 2020.

OUTRAM, A. K. et al. The Earliest Horse Harnessing and Milking. **Science**, v. 323, n. March, p. 1332–1335, 2009.

PARADIS, E. Pegas: An R package for population genetics with an integrated-modular approach. **Bioinformatics**, v. 26, n. 3, p. 419–420, 2010.

PARRA, F. C. et al. Color and genomic ancestry in Brazilians. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 1, p. 177–182, 2003.

PAZ-Y-MIÑO, C. et al. A study of the molecular variants associated with lactase persistence in different Ecuadorian ethnic groups. **American Journal of Human Biology**, v. 28, n. 6, p. 774–781, 2016.

PENA, S. D. J. et al. DNA tests probe the genomic ancestry of Brazilians. **Brazilian Journal of Medical and Biological Research**, v. 42, n. 10, p. 870–876, 2009.

PENA, S. D. J. et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. **PLoS ONE**, v. 6, n. 2, p. e17063, 2011.

PHILLIPSON, D. W. **African Archaeology**. 3. ed. Nova York: Cambridge University Press, 2005.

PICCIANO, M. F. Nutrient composition of human milk. **Pediatric Clinics of North America**, v. 48, n. 1, p. 53–67, 2001.

PINTO, J. C. et al. Food and pathogen adaptations in the Angolan Namib desert: Tracing the spread of lactase persistence and human African trypanosomiasis resistance into southwestern Africa. **American Journal of Physical Anthropology**, v. 161, n. 3, p. 436–447, 2016.

PIZA, E.; ROSEMBERG, F. Cor nos censos brasileiros. **Revista USP**, n. 40, p. 122–137, 1998.

PLANTINGA, T. S. et al. Low prevalence of lactase persistence in Neolithic South-West Europe. **European Journal of Human Genetics**, v. 20, n. 7, p. 778–782, 2012.

PONTE, P. R. L. et al. Clinical evaluation, biochemistry and genetic polymorphism analysis for the diagnosis of lactose intolerance in a population from northeastern Brazil. **Clinics**, v. 71, n. 2, p. 82–89, 2016.

PORTER, L.; DUNN, C. H. Some studies on sugar infant feeding. **American Journal of Diseases of Children**, v. 10, n. 2, p. 77–86, 1915.

POTTER, N. N.; HOTCHKISS, J. H. Milk and Milk Products. In: POTTER, N. N.; HOTCHKISS, J. H. (Eds.). **Food Science**. 5. ed. Boston: Springer, 1995.

POULTER, M. et al. The causal element for the lactase persistence/non-persistence polymorphism is located in a 1 Mb region of linkage disequilibrium in Europeans. **Annals of Human Genetics**, v. 67, n. 4, p. 298–311, 2003.

PRIEHODOVÁ, E. et al. Lactase persistence variants in Arabia and in the African Arabs. **Human Biology**, v. 86, n. 1, p. 7–18, 2014.

PRIORI, A. et al. **História do Paraná (séculos XIX e XX)**. Maringá: Editora da Universidade Estadual de Maringá, 2012.

PROBST, C. M. et al. HLA polymorphism and evaluation of European, African, and Amerindian contribution to the white and mulatto populations from Paraná, Brazil. **Human Biology**, v. 72, n. 4, p. 597–617, 2000.

PROUS, A. **Arqueologia Brasileira**. Brasília: Editora Universidade de Brasília, 1992.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, 2020. Disponível em: < <http://www.R-project.org> >.

RAB, S. M.; BASEER, A. High intestinal lactase concentration in adult Pakistanis.

British Medical Journal, v. 1, n. 6007, p. 436, 1976.

RAGGIO, A. Z.; BLEY, R. B.; TRAUZYNSKI, S. C. **Abordagem histórica sobre a população negra no estado do Paraná**. 1. ed. Curitiba: SEJU, 2018. v. 2.

RANCIARO, A. [Comunicação pessoal]. 17 mar. 2020.

RANCIARO, A. et al. Genetic origins of lactase persistence and the spread of pastoralism in Africa. **American Journal of Human Genetics**, v. 94, n. 4, p. 496–510, 2014.

RAO, D. R. et al. Prevalence of lactose maldigestion - Influence and interaction of age, race, and sex. **Digestive Diseases and Sciences**, v. 39, n. 7, p. 1519–1524, 1994.

RAZ, M. et al. Frequency of *LCT-13910C/T* and *LCT-22018G/A* single nucleotide polymorphisms associated with adult-type hypolactasia/lactase persistence among Israelis of different ethnic groups. **Gene**, v. 519, n. 1, p. 67–70, 2013.

REICH, C. M.; ARNOULD, J. P. Y. Evolution of Pinnipedia lactation strategies: A potential role for α -lactalbumin? **Biology Letters**, v. 3, n. 5, p. 546–549, 2007.

REICH, D. et al. Reconstructing Native American population history. **Nature**, v. 488, n. 7411, p. 370–374, 2012.

RHODE, D. et al. Yaks, yak dung, and prehistoric human habitation of the Tibetan Plateau. **Developments in Quaternary Science**, v. 9, n. C, p. 205–224, 2007.

RIBEIRO-DOS-SANTOS, Â. K. C. et al. Nucleotide variability of HV-I in Afro-descendants populations of the Brazilian Amazon Region. **Forensic Science International**, v. 167, n. 1, p. 77–80, 2007.

RIBEIRO, D. **O povo brasileiro: A formação e o sentido do Brasil**. 1. ed. São Paulo: Global Editora, 2014.

ROBAYO-TORRES, C. C.; NICHOLS, B. L. Molecular differentiation of congenital lactase deficiency from adult-type hypolactasia. **Nutrition Reviews**, v. 65, n. 2, p. 95–98, 2007.

RODRIGUES, R. N. **Os africanos no Brasil**. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2010.

ROMAGNUOLO, J.; SCHILLER, D.; BAILEY, R. J. Using breath tests wisely in a gastroenterology practice: An evidence-based review of indications and pitfalls in interpretation. **American Journal of Gastroenterology**, v. 97, n. 5, p. 1113–1126, 2002.

ROSENKRANZ, W. et al. Distribution of human adult lactase phenotypes in the population of Austria. **Human Genetics**, v. 62, n. 2, p. 158–161, 1982.

ROTHHAMMER, F.; DILLEHAY, T. D. The late pleistocene colonization of South America: An interdisciplinary perspective. **Annals of Human Genetics**, v. 73, n. 5, p. 540–549, 2009.

ROZAS, J. et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. **Molecular Biology and Evolution**, v. 34, n. 12, p. 3299–3302, 2017.

SABETI, P. C. et al. Detecting recent positive selection in the human genome from haplotype structure. **Nature**, v. 419, p. 832–837, 2002.

SAHI, T. et al. Recessive Inheritance of Adult-Type Lactose Malabsorption. **The Lancet**, v. 302, n. 7833, p. 823–826, 1973.

SAHI, T. Dietary lactose and the aetiology of human small-intestinal hypolactasia. **Gut**, v. 19, n. 11, p. 1074–1086, 1978.

SAHI, T. Hypolactasia and Lactase Persistence. **Scandinavian Journal of Gastroenterology**, v. 29, n. 202, p. 1–6, 1994a.

SAHI, T. Genetics and epidemiology of adult-type hypolactasia. **Scandinavian Journal of Gastroenterology**, v. 29, n. S202, p. 7–20, 1994b.

SAHI, T.; LAUNIALA, K.; LAITINEN, H. Hypolactasia in a fixed cohort of Young finnish adults: A follow-up study. **Scandinavian Journal of Gastroenterology**, v. 18, n. 7, p. 865–870, 1983.

SALZANO, F. M. Interethnic variability and admixture in Latin America - Social implications. **Revista de Biologia Tropical**, v. 52, n. 3, p. 405–415, 2004.

SALZANO, F. M. The Prehistoric Colonization of the Americas: Evidence and Models. **Evolution: Education and Outreach**, v. 4, n. 2, p. 199–204, 2011.

SALZANO, F. M.; BORTOLINI, M. C. **The Evolution and Genetics of Latin American Populations**. 1. ed. Nova York: Cambridge University Press, 2002.

SALZANO, F. M.; SANS, M. Interethnic admixture and the evolution of Latin American populations. **Genetics and Molecular Biology**, v. 37, n. 1, p. 151–170, 2014.

SANS, M. Admixture Studies in Latin America: From the 20th to the 21st Century. **Human Biology**, v. 72, n. 1, p. 155–177, 2000.

SANTOS, S. A. Historical roots of the “whitening” of Brazil. **Latin American Perspectives**, v. 29, n. 1, p. 61–82, 2002.

SARMENTO, D. A garantia do Direito à posse dos remanescentes de quilombos antes da desapropriação. In: DUPRAT, D. (Ed.). **Pareceres Jurídicos - Direito dos Povos e das Comunidades Tradicionais**. Manaus: UEA, 2007. p. 77–103.

SAVILAHTI, E.; LAUNIALA, K.; KUITUNEN, P. Congenital lactase deficiency. **Archives of Disease in Childhood**, v. 58, p. 246–252, 1983.

SCHMIDT, M. A. M. **Histórias do cotidiano paranaense**. Curitiba: Letraviva, 1996.

SCHMITT, A.; TURATTI, M. C. M.; CARVALHO, M. C. P. DE. A atualização do conceito de quilombo: identidade e território nas definições teóricas. **Ambiente & Sociedade**, v. 5, n. 10, p. 129–136, 2002.

SÉGUREL, L. et al. Why and when was lactase persistence selected for? Insights from Central Asian herders and ancient DNA. **PLoS Biology**, v. 18, n. 6, p. e3000742, 2020.

SÉGUREL, L.; BON, C. On the Evolution of Lactase Persistence in Humans. **Annual Review of Genomics and Human Genetics**, v. 18, n. 1, p. 297–319, 2017.

SENEWIRATNE, B.; THAMBIPILLAI, S.; PERERA, H. Intestinal Lactase Deficiency in Ceylon (Sri Lanka). **Gastroenterology**, v. 72, n. 6, p. 1257–1259, 1977.

SILVA, A. C. A colonização do território paranaense e o dinamismo dos municípios da frente Norte. **Revista GEOMAE**, v. 7, n. 1, p. 9–20, 2016.

SIMOONS, F. J. The Non-Milking Area of Africa. **Anthropos**, v. 49, n. 1,2, p. 58–66, 1954.

SIMOONS, F. J. Primary adult lactose intolerance and the milking habit: A problem in biological and cultural interrelations - I. Review of the medical research. **The American Journal of Digestive Diseases**, v. 14, n. 12, p. 819–836, 1969.

SIMOONS, F. J. Primary adult lactose intolerance and the milking habit: A problem in biologic and cultural interrelations - II. A culture historical hypothesis. **The American Journal of Digestive Diseases**, v. 15, n. 8, p. 695–710, 1970.

SIMOONS, F. J. The geographic hypothesis and lactose malabsorption. **The American Journal of Digestive Diseases**, v. 23, n. 11, p. 963–980, 1978.

SLENES, R. W. “Malungu, ngoma vem!”: África coberta e descoberta do Brasil. **Revista USP**, v. 12, n. 1, p. 48–67, 1992.

SMITH, A. B. Origins and Spread of African Pastoralism in Africa. **Annual Review of Anthropology**, v. 21, n. 1, p. 125–141, 1992.

SOUZA, A. M. et al. A systematic scoping review of the genetic ancestry of the Brazilian population. **Genetics and Molecular Biology**, v. 42, n. 3, p. 495–508, 2019.

SOUZA, I. R.; CULPI, L. Valongo, an isolated Brazilian Black community. I. Structure of the population. **Brazilian Journal of Genetics**, v. 15, p. 439–447, 1992.

SOUZA, I. R.; CULPI, L. Valongo, genetic studies on an isolated Afro-Brazilian community. **Genetics and Molecular Biology**, v. 28, n. 3, p. 402–406, 2005.

SOUZA, M. M. **África e Brasil Africano**. 2. ed. São Paulo: Ática, 2008.

STEPHENS, M.; SMITH, N. J.; DONNELLY, P. A new statistical method for haplotype reconstruction from population data. **American Journal of Human Genetics**, v. 68, n. 4, p. 978–989, 2001.

STORHAUG, C. L.; FOSSE, S. K.; FADNES, L. T. Country, regional, and global estimates for lactose malabsorption in adults: a systematic review and meta-analysis. **The Lancet Gastroenterology and Hepatology**, v. 2, n. 10, p. 738–746, 2017.

SVERRISDÓTTIR, O. Ó. et al. Direct estimates of natural selection in Iberia indicate

calcium absorption was not the only driver of lactase persistence in Europe. **Molecular Biology and Evolution**, v. 31, n. 4, p. 975–983, 2014.

SWAIN, T. N. Fronteiras do Paraná: da colonização à migração. In: AUBERTIN, C. (Ed.). **Fronteiras**. Brasília: Editora Universidade de Brasília, 1988. p. 19–37.

SWALLOW, D. M. Genetics of Lactase Persistence and Lactose Intolerance. **Annual Review of Genetics**, v. 37, p. 197–219, 2003.

TANDON, K.; SINGH, D. S. Lactose intolerance in North and South Indians. **The American Journal of Clinical Nutrition**, v. 34, p. 943–946, 1981.

TARAZONA SANTOS, E. et al. Brasil e a Idiosincrasia da Miscigenação. **Revista da Universidade Federal de Minas Gerais**, v. 22, n. 1.2, p. 232–249, 2016.

TEIXEIRA, V. R. **De negros a adventistas, em busca da salvação: Estudo de um grupo rural de Santa Catarina**. 97f. Dissertação (Mestrado em Antropologia Social) – Universidade Federal de Santa Catarina, Blumenau, 1990.

TELLES, E. E. Racial ambiguity among the Brazilian population. **Ethnic and Racial Studies**, v. 25, n. 3, p. 415–441, 2002.

TELLES, E. E. **Race in another America - the significance of skin color in Brazil**. Princeton: Princeton University Press, 2004.

THE 1000 GENOMES PROJECT CONSORTIUM. A global reference for human genetic variation. **Nature**, v. 526, n. 7571, p. 68–74, 2015.

THOMAS, S. et al. Age dependency of the lactase persistence and lactase restriction phenotypes among children in Sri Lanka and Britain. **Journal of Tropical Pediatrics**, v. 36, n. 2, p. 80–85, 1990.

TISHKOFF, S. A. et al. Convergent adaptation of human lactase persistence in Africa and Europe. **Nature Genetics**, v. 39, n. 1, p. 31–40, 2007.

TISHKOFF, S. A. et al. The Genetic Structure and History of Africans and African Americans. **Science**, v. 324, n. 5930, p. 1035–1044, 2009.

TORNIAINEN, S. et al. Screening of variants for lactase persistence/non-persistence in populations from South Africa and Ghana. **BMC Genetics**, v. 10, p. 1–5, 2009.

TROELSEN, J. T. et al. An Upstream Polymorphism Associated with Lactase Persistence has Increased Enhancer Activity. **Gastroenterology**, v. 125, n. 6, p. 1686–1694, 2003.

TROELSEN, J. T. Adult-type hypolactasia and regulation of lactase expression. **Biochimica et Biophysica Acta - General Subjects**, v. 1723, n. 1–3, p. 19–32, 2005.

UNIVERSIDADE FEDERAL DE GOIÁS (UFG). CENTRO INTEGRADO DE APRENDIZAGEM EM REDE (CIAR). **África em Arte-Educação [e-book]**. Goiânia: Gráfica da UFG, 2015.

UPTON, J.; GEORGE, P. The prevalence of lactose intolerance (adult hypolactasia) in a randomly selected New Zealand population. **New Zealand Medical Journal**, v. 123, n. 1308, p. 117–118, 2010.

VICARI, P.; DE MELLO, A. B.; FIGUEIREDO, M. S. Effects of hydroxyurea in a population of Brazilian patients with sickle cell anemia. **American Journal of Hematology**, v. 78, n. 3, p. 243–244, 2005.

VIGNE, J.-D. The origins of animal domestication and husbandry: A major change in the history of humanity and the biosphere. **Comptes Rendus Biologies**, v. 334, n. 3, p. 171–181, 2011.

VIGNE, J.-D.; HELMER, D. Was milk a “secondary product” in the Old World Neolithisation process? Its role in the domestication of cattle, sheep and goats. **Anthropozoologica**, v. 42, n. 2, p. 9–40, 2007.

VILLAKO, K.; MAAROOS, H. Clinical picture of hypolactasia and lactose intolerance. **Scandinavian Journal of Gastroenterology**, v. 29, n. S202, p. 36–54, 1994.

WACHOWICZ, R. C. **História do Paraná**. Curitiba: Editora dos Professores, 1967.

WALKER, C.; THOMAS, M. G. The evolution of lactose digestion. In: PAQUES, M.; LINDNER, C. (Eds.). **Lactose: Evolutionary Role, Health Effects and Applications**. [s.l.] Academic Press, 2019. p. 1–48.

WANG, S. et al. Genetic variation and population structure in Native Americans. **PLoS Genetics**, v. 3, n. 11, p. 2049–2067, 2007.

WANG, Y. et al. The lactase persistence/non-persistence polymorphism is controlled by a cis-acting element. **Human Molecular Genetics**, v. 4, n. 4, p. 657–662, 1995.

WANG, Y. et al. The genetically programmed down-regulation of lactase in children. **Gastroenterology**, v. 114, n. 6, p. 1230–1236, 1998.

WARNES, G. et al. **genetics: Population Genetics**. R package version 1.3.8.1.2. 2019. Disponível em: <<https://CRAN.R-project.org/package=genetics>>.

WHEELER, J. C. Evolution and present situation of South American Camelidae. **Biological Journal of the Linnean Society**, v. 54, p. 271–295, 1995.

WILKIN, S. et al. Dairy pastoralism sustained eastern Eurasian steppe populations for 5,000 years. **Nature Ecology and Evolution**, v. 4, n. 3, p. 346–355, 2020.

WITAS, H. W. et al. Hunting for the *LCT-13910*T* allele between the middle neolithic and the middle ages suggests its absence in dairying LBK people entering the Kuyavia region in the 8th millennium BP. **PLoS ONE**, v. 10, n. 4, p. 1–24, 2015.

WITTENBERG, D. F.; MOOSA, A. Lactose maldigestion - age-specific prevalence in black and Indian children. **South African Medical Journal**, v. 78, n. 8, p. 470–472, 1990.

WRAY, N. R.; VISSCHER, P. M. Population genetics and its relevance to gene

mapping. In: NEALE, B. M. et al. (Eds.). . **Statistical Genetics: Gene Mapping through Linkage and Association**. [s.l.] Taylor & Francis Group, 2008. p. 87–112.

WRIGHT, S. The Genetical Structure of Populations. **Annals of Eugenics**, v. 15, n. 1, p. 323–354, 1949.

YANG, Y. et al. Proteomics evidence for kefir dairy in Early Bronze Age China. **Journal of Archaeological Science**, v. 45, n. 1, p. 178–186, 2014.

YOSHIDA, Y. et al. Studies on the etiology of milk intolerance in Japanese adults. **Gastroenterologia Japonica**, v. 10, n. 1, p. 29–34, 1975.

ZAGO, M. A.; FIGUEIREDO, M. S.; OGO, S. H. Bantu β s cluster haplotype predominates among Brazilian Blacks. **American Journal of Physical Anthropology**, v. 88, n. 3, p. 295–298, 1992.

APÊNDICE 1 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA EUROPA

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Alemanha	Bayern / Alemães	433	85,2	14,8	Flatz et al. (1982)
Alemanha	Berlim / Alemães	15	93,3	6,7	Flatz et al. (1982)
Alemanha	Hessen / Alemães	70	87,1	12,9	Flatz et al. (1982)
Alemanha	Niedersachsen e Bremen / Alemães	328	85,7	14,3	Flatz et al. (1982)
Alemanha	Nordrhein-Westfalen / Alemães	346	87,3	12,7	Flatz et al. (1982)
Alemanha	Rheinland-Pfalz / Alemães	56	91,1	8,9	Flatz et al. (1982)
Alemanha	Schleswig e Hamburg / Alemães	255	87,5	12,5	Flatz et al. (1982)
Áustria	Região Leste / Austríacos	181	77,3	22,7	Itan et al. (2010); Rosenkranz et al. (1982)
Áustria	Região Oeste / Austríacos	166	85,5	14,5	Itan et al. (2010); Rosenkranz et al. (1982)
Dinamarca	Hellerup / Dinamarqueses	761	97,0	3,0	Busk et al. (1975); Gudmand-Høyer et al. (1969)
Finlândia	Helsinki / Falantes de finlandês	449	83,0	17,0	Jussila (1969); Jussila et al. (1970); Sahi (1974)
Finlândia	Helsinki / Falantes de suíço	91	92,0	8,0	Sahi (1974)
Itália	Região Central / Italianos	65	82,0	18,0	Cavalli-Sforza et al. (1987); Itan et al. (2010)
Itália	Região Norte / Italianos	208	49,0	51,0	Burgio et al. (1984); Itan et al. (2010)
Itália	Região Sul / Italianos	99	46,0	54,0	Itan et al. (2010); Rinaldi et al. (1984)
Itália	Sardenha	47	15,0	85,0	Itan et al. (2010); Meloni et al. (1998)
Rep. Tcheca	/ Tchechos	17	82,0	18,0	Leichter (1972)
Suécia	/ Suecos	400	96,0	4,0	Dahlqvist; Lindquist (1971)
Suíça	Zurique / Suíços	746	77,0	23,0	Beyerlein et al. (2008)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

**APÊNDICE 2 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM
POPULAÇÕES DA ÁFRICA**

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
África do Sul	Askham / População miscigenada	20	45,0	55,0	Breton et al. (2014)
África do Sul	Askham / ≠ Khomani	20	20,0	80,0	Breton et al. (2014)
África do Sul	Colesberg / Karretjie	20	20,0	80,0	Breton et al. (2014)
África do Sul	Colesberg / População miscigenada	20	15,0	85,0	Breton et al. (2014)
África do Sul	Região Sudeste / Falantes Bantu	41	17,0	83,0	Breton et al. (2014)
África do Sul	Wellington / População miscigenada	20	45,0	55,0	Breton et al. (2014)
Angola	Menogue / !Xun	20	10,0	90,0	Breton et al. (2014)
Congo	Luozi / Manyanga	11	0	100,0	Breton et al. (2014)
Namíbia	Tsumkwe / Ju/'hoansi	20	5,0	95,0	Breton et al. (2014)
Namíbia	Windhoek / Herero	14	0	100,0	Breton et al. (2014)
Namíbia	Windhoek / Nama	22	50,0	50,0	Breton et al. (2014)
Níger	/ Tuareg	118	87,3	12,7	Flatz; Schildge; Sekou (1986)
Senegal	/ População miscigenada	-	71,0	29,0	Arnold et al. (1980)
Quênia	/ Borana	8	60,0	40,0	Tishkoff et al. (2007)
Quênia	/ Burji	8	37,5	62,5	Tishkoff et al. (2007)
Quênia	/ El Molo	9	44,4	55,6	Tishkoff et al. (2007)
Quênia	/ Gabra	9	88,9	11,1	Tishkoff et al. (2007)
Quênia	/ Kikuyu	2	50,0	50,0	Tishkoff et al. (2007)
Quênia	/ Konso	6	33,3	66,7	Tishkoff et al. (2007)
Quênia	/ Maasai	32	71,9	28,1	Tishkoff et al. (2007)
Quênia	/ Marakwet	6	50,0	50,0	Tishkoff et al. (2007)
Quênia	/ Nandi	4	0	100,0	Tishkoff et al. (2007)
Quênia	/ Ogiek	11	54,5	45,5	Tishkoff et al. (2007)
Quênia	/ Pokot	14	42,9	57,1	Tishkoff et al. (2007)
Quênia	/ Rendille	8	62,5	37,5	Tishkoff et al. (2007)
Quênia	/ Sabaot	6	50,0	50,0	Tishkoff et al. (2007)
Quênia	/ Samburu	9	88,9	11,1	Tishkoff et al. (2007)
Quênia	/ Sengwer	16	12,5	87,5	Tishkoff et al. (2007)
Quênia	/ Somali	1	100,0	0	Tishkoff et al. (2007)
Quênia	/ Tugen	16	50,0	50,0	Tishkoff et al. (2007)
Quênia	/ Turkana	12	33,3	66,7	Tishkoff et al. (2007)
Quênia	/ Wata	1	100,0	0	Tishkoff et al. (2007)
Quênia	/ Yaaku	14	57,0	43,0	Tishkoff et al. (2007)
Ruanda	/ Hutu	36	42,0	58,0	Cox; Elliot (1974)
Ruanda	/ Shi	28	4,0	96,0	Cox; Elliot (1974)
Ruanda	/ Tutsi	27	92,2	7,8	Cox; Elliot (1974)
Ruanda	/ Twa	22	23,0	77,0	Cox; Elliot (1974)
Sudão	/ Ama	2	50,0	50,0	Tishkoff et al. (2007)
Sudão	/ Beja Beni Amer	6	100,0	0	Tishkoff et al. (2007)
Sudão	/ Beja Hadandawa	11	81,8	18,2	Tishkoff et al. (2007)
Sudão	/ Dinka	9	66,7	33,3	Tishkoff et al. (2007)
Sudão	/ Koalib	1	100,0	0	Tishkoff et al. (2007)
Sudão	/ Liguri-Logorik	1	0	100,0	Tishkoff et al. (2007)
Sudão	/ Masalit	1	100,0	0	Tishkoff et al. (2007)

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Sudão	/ Nuer	4	50,0	50,0	Tishkoff et al. (2007)
Sudão	/ Shilook	8	37,5	62,5	Tishkoff et al. (2007)
Sudão	Norte do Nilo / Dongolawi	16	18,7	81,3	Bayoumi et al. (1981)
Sudão	Norte do Nilo / Núbia	21	33,3	66,7	Bayoumi et al. (1981)
Sudão	Norte do Nilo / Shaygi	42	38,1	61,9	Bayoumi et al. (1981)
Sudão	Região Central / Jaali	113	53,1	46,9	Bayoumi et al. (1981)
Sudão	Região Central / Nuba	58	20,7	79,3	Bayoumi et al. (1981)
Sudão	Região Norte / Kahli	21	61,9	38,1	Bayoumi et al. (1981)
Sudão	Região Norte / Beja	303	83,2	16,8	Bayoumi et al. (1982)
Sudão	Região Norte / Gomoeria	31	67,7	32,3	Bayoumi et al. (1981)
Sudão	Região Sudoeste / Habbani	19	47,4	52,6	Bayoumi et al. (1981)
Sudão	Região Sudoeste / Misseri	20	40,0	60,0	Bayoumi et al. (1981)
Sudão	Região Sul / Nilóticos	282	25,5	74,5	Bayoumi et al. (1982)
Tanzânia	/ Akie	14	42,9	57,1	Tishkoff et al. (2007)
Tanzânia	/ Burunge	18	33,3	66,7	Tishkoff et al. (2007)
Tanzânia	/ Datog	4	0	100,0	Tishkoff et al. (2007)
Tanzânia	/ Dorobo	10	40,0	60,0	Tishkoff et al. (2007)
Tanzânia	/ Fiome	12	8,3	91,7	Tishkoff et al. (2007)
Tanzânia	/ Hadza	18	50,0	50,0	Tishkoff et al. (2007)
Tanzânia	/ Iraqw	21	85,7	14,3	Tishkoff et al. (2007)
Tanzânia	/ Maasai	17	58,8	41,2	Tishkoff et al. (2007)
Tanzânia	/ Mbugu	30	33,3	66,7	Tishkoff et al. (2007)
Tanzânia	/ Mbugwe	13	30,8	69,2	Tishkoff et al. (2007)
Tanzânia	/ Pare	10	60,0	40,0	Tishkoff et al. (2007)
Tanzânia	/ Rangi	35	48,6	51,4	Tishkoff et al. (2007)
Tanzânia	/ Samba'a	3	0	100,0	Tishkoff et al. (2007)
Tanzânia	/ Sandawe	30	26,7	73,3	Tishkoff et al. (2007)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

**APÊNDICE 3 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM
POPULAÇÕES DO ORIENTE MÉDIO**

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Arábia Saudita	Província Leste / Árabes urbanos	109	43,0	57,0	Dissanayake et al. (1990)
Arábia Saudita	Província Leste / Beduínos	21	81,0	19,0	Dissanayake et al. (1990)
Arábia Saudita	Província Leste / Iemenitas	17	53,0	47,0	Dissanayake et al. (1990)
Egito	Região de Sinai / Egípcios	172	65,7	34,3	Hussein; Ezzilarab (1994)
Jordânia	/ Beduínos	162	76,0	24,0	Hijazi et al. (1983)
Jordânia	Região Oeste / Jordanianos e Palestinos	148	25,0	75,0	Hijazi et al. (1983)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

**APÊNDICE 4 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM
POPULAÇÕES DA ÁSIA**

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Índia	Região Norte / Indianos	124	72,6	27,4	Tandon; Singh (1981)
Índia	Região Sul / Indianos	60	33,4	66,6	Tandon; Singh (1981)
Japão	Hirosaki / Japoneses	102	1,0	99,0	Yoshida et al. (1975)
Paquistão	Região Norte / Baloochi	32	37,5	62,5	Ahmad; Flatz (1984)
Paquistão	Região Norte / Kashmir	27	29,6	70,4	Ahmad; Flatz (1984)
Paquistão	Região Norte / Punjabi	322	41,0	59,0	Ahmad; Flatz (1984)
Paquistão	Região Norte / Sindhi	33	42,4	57,6	Ahmad; Flatz (1984)
Paquistão	Região Sul / Baloochi	4	100,0	0	Rab; Baseer (1976)
Paquistão	Região Sul / Mohajir	15	98,0	0,2	Rab; Baseer (1976)
Paquistão	Região Sul / Pathan	15	93,3	2	Rab; Baseer (1976)
Paquistão	Região Sul / Punjabi	9	100,0	0	Rab; Baseer (1976)
Paquistão	Região Sul / Sindhi	12	100,0	0	Rab; Baseer (1976)
Rússia	Oeste da Sibéria / Mansi	81	28,0	72,0	Itan et al. (2010); Kozlov (1998)
Rússia	Oeste da Sibéria / Nenets	9	22,0	78,0	Itan et al. (2010); Kozlov (1998)
Rússia	Região Noroeste / Komi-Izhems	56	38,0	62,0	Itan et al. (2010); Kozlov (1998)
Rússia	Região Norte / Khanty	115	29,0	71,0	Itan et al. (2010); Kozlov (1998)
Rússia	Região Oeste / Udmurtia	30	60,0	40,0	Itan et al. (2010); Kozlov (1998)
Uzbequistão	Bukhara / Tajiko-Uzbek	100	11,0	89,0	Heyer et al. (2011)
Uzbequistão	Gazli / Kazakh	83	25,3	74,7	Heyer et al. (2011)
Sri Lanka	/ Cingaleses	200	27,5	72,5	Senewiratne; Thambipillai; Perera (1977)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

**APÊNDICE 5 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM
POPULAÇÕES DA AMÉRICA**

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Chile	Araucanía / Mapuche	29	10,0	90,0	Fernández et al. (2016)
Chile	Araucanía / Temuco	115	30,0	70,0	Fernández et al. (2016)
Chile	Choapa / <i>Mestízos</i>	201	36,0	64,0	Fernández et al. (2016)
Chile	Elqui / <i>Mestízos</i>	13	23,0	77,0	Fernández et al. (2016)
Chile	Ilha de Páscoa / Rapanui	86	25,0	75,0	Fernández et al. (2016)
Chile	Limarí / <i>Mestízos</i>	223	39,0	61,0	Fernández et al. (2016)
Chile	Santiago / <i>Mestízos</i>	116	40,0	60,0	Fernández et al. (2016)
Equador	Amazônia / <i>Mestízos</i>	48	12,5	87,5	Paz-Y-Miño et al. (2016)
Equador	Costa / Afro-equadorianos	49	12,2	87,8	Paz-Y-Miño et al. (2016)
Equador	Costa / <i>Mestízos</i>	124	22,6	77,4	Paz-Y-Miño et al. (2016)
Equador	Serra / Afro-equadorianos	76	18,4	81,6	Paz-Y-Miño et al. (2016)
Equador	Serra / Ameríndios	128	12,5	87,5	Paz-Y-Miño et al. (2016)
Equador	Serra / <i>Mestízos</i>	316	26,9	73,1	Paz-Y-Miño et al. (2016)
Estados Unidos	Alaska / Esquimós	36	16,6	83,3	Ducan; Scott (1972)
Estados Unidos	Oklahoma / Ameríndios	60	17,5	82,5	Caskey et al. (1977)
México	Jalisco / Cuquio	99	45,5	54,5	Ojeda-Granados et al. (2016)
México	Jalisco / Guadalajara	699	35,9	64,1	Ojeda-Granados et al. (2016)
México	Jalisco / Nahuas	86	2,3	97,7	Ojeda-Granados et al. (2016)
México	Jalisco / Villa Purificación	32	56,3	43,7	Ojeda-Granados et al. (2016)
México	Nayarit / Huicholes	95	0	100,0	Ojeda-Granados et al. (2016)
México	Nayarit / Tepic	182	31,4	68,6	Ojeda-Granados et al. (2016)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

APÊNDICE 6 – FREQUÊNCIAS FENOTÍPICAS OBSERVADAS EM POPULAÇÕES DA OCEANIA

País	Localidade / População	N	Frequência (%)		Referência
			LP	LNP	
Austrália	Região Oeste / Aborígenes	40	16,0	84,0	Brand et al. (1983)
Ilhas Salomão		44	1,0	99,0	Bersaglieri et al. (2004); Itan et al. (2010); Storhaug; Fosse; Fadnes (2017)
Papua-Nova Guiné	Região Noroeste / Sepik	35	22,9	77,1	Arnhold; Perman; Nurse (1981)

FONTE: A autora (2020).

LEGENDA: LP = Lactase persistente, LNP = Lactase não persistente.

APÊNDICE 7 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA EUROPA E DE ANCESTRALIDADE EUROPEIA

País	Localidade / População	N	Frequência alélica (%)	Referência
Alemanha	/ Alemães	417	52,0	Büning et al. (2003)
Áustria	/ Austríacos	220	58,0	Bodlaj et al. (2006)
Canadá	Toronto / Euro-canadenses	720	48,1	Alharbi; El-Sohemy (2017)
Escócia	/ Mulheres britânicas	516	81,9	Davey Smith et al. (2009)
Estados Unidos	/ Euro-americanos	24	77,2	Bersaglieri et al. (2004)
Estados Unidos	/ Euro-americanos	128	59,8	Chin et al. (2019)
Estônia	Väike-Maarja / Estônios	355	51,4	Lember et al. (2006)
França	Região Sudoeste / Bascos	24	66,7	Bersaglieri et al. (2004)
Finlândia	/ Finlandeses	1876	58,0	Enattah et al. (2008); Itan et al. (2010)
Finlândia	Região Leste / Finlandeses	77	55,0	Enattah et al. (2008); Itan et al. (2010)
Finlândia	Região Oeste / Finlandeses	308	62,0	Enattah et al. (2008); Itan et al. (2010)
Finlândia e Suécia	/ Finlandeses e Suecos	180	81,5	Bersaglieri et al. (2004)
Hungria	/ Húngaros	296	62,2	Itan et al. (2010); Nagy et al. (2009)
Inglaterra	Região Central / Britânicas	684	75,9	Davey Smith et al. (2009)
Inglaterra	Região Norte / Britânicas	1168	75,1	Davey Smith et al. (2009)
Inglaterra	Região Sudeste / Britânicas	947	69,8	Davey Smith et al. (2009)
Nova Zelândia	/ Neozelandeses	945	69,1	Nolan et al. (2010)
Nova Zelândia	/ Neozelandeses	1064	72,0	Upton; George (2010)
Reino Unido	Ilhas Orkney / Britânicos	16	68,8	Bersaglieri et al. (2004)
Rep. Tcheca	/ Ciganos	300	27,0	Hubáček et al. (2017)
Rep. Tcheca	/ Tchecos	288	76,0	Hubáček et al. (2017)

FONTE: A autora (2020).

APÊNDICE 8 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA ÁFRICA

País	Localidade / População	N	Frequência alélica (%)	Referência
Argélia	Mzab / Mozabite	30	21,7	Bersaglieri et al. (2004)
Argélia	Mzab / Mozabite	22	22,7	Ranciaro et al. (2014)
Camarões	/ Fulani	91	39,0	Ingram et al. (2007)
Camarões	/ Fulbe	49	11,2	Mulcare et al. (2004)
Camarões	/ Hausa	18	13,9	Mulcare et al. (2004)
Camarões	/ Kwanja	70	0	Mulcare et al. (2004)
Camarões	/ Mambila	122	0,4	Mulcare et al. (2004)
Camarões	/ Nos (Nsaw)	126	0	Mulcare et al. (2004)
Camarões	/ Yamba	21	0	Mulcare et al. (2004)
Camarões	/ Outras populações	128	4,3	Mulcare et al. (2004)
Etiópia	/ Anuak (Anywak)	108	0	Mulcare et al. (2004)
Etiópia	/ Nuer	119	0	Mulcare et al. (2004)
Etiópia	/ Outras populações	1	0	Mulcare et al. (2004)
Malawi	/ Chewa	84	0	Mulcare et al. (2004)
Malawi	/ Ngoni	14	0	Mulcare et al. (2004)
Malawi	/ Tumbuka	58	0	Mulcare et al. (2004)
Malawi	/ Yao	49	0	Mulcare et al. (2004)
Malawi	/ Outras populações	58	0	Mulcare et al. (2004)
Nigéria	/ Ibibio	110	0	Mulcare et al. (2004)
Nigéria	/ Oron	44	0	Mulcare et al. (2004)
Nigéria	/ Outras populações	22	0	Mulcare et al. (2004)
Saara Ocidental	/ Saharawi	57	26,0	Enattah et al. (2007)
Saara Ocidental	/ Saharawi	22	23,0	Enattah et al. (2008)
Senegal	/ Wolof	69	0	Mulcare et al. (2004)
Senegal	/ Manjak	93	0	Mulcare et al. (2004)
Senegal	/ Outras populações	19	2,6	Mulcare et al. (2004)
Sudão	/ Fulani	44	48,0	Enattah et al. (2008)
Sudão	Região Norte / Ga'ali	30	0	Mulcare et al. (2004)
Sudão	Região Norte / Shaigi	11	0	Mulcare et al. (2004)
Sudão	Região Norte / Outras populações	88	0	Mulcare et al. (2004)
Sudão	Região Sul / Dinka	34	0	Mulcare et al. (2004)
Sudão	Região Sul / Nuer	13	0	Mulcare et al. (2004)
Sudão	Região Sul / Outras populações	73	0	Mulcare et al. (2004)
Uganda	/ Musesse	22	0	Mulcare et al. (2004)
Uganda	/ Outras populações	18	0	Mulcare et al. (2004)

FONTE: A autora (2020).

APÊNDICE 9 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA ÁSIA E ORIENTE MÉDIO

País	Localidade / População	N	Frequência alélica (%)	Referência
Arábia Saudita	/ Sauditas	248	0,4	Enattah et al. (2008)
Índia	/ Falantes de línguas austro-asiáticas	331	1,1	Gallego Romero et al. (2012)
Índia	/ Falantes de línguas dravídicas	853	7,6	Gallego Romero et al. (2012)
Índia	/ Falantes de línguas indo-europeias	896	18,4	Gallego Romero et al. (2012)
Índia	/ Falantes de línguas tibeto-birmanesas	120	0,8	Gallego Romero et al. (2012)
Índia	/ Populações isoladas	60	2,5	Gallego Romero et al. (2012)
Índia	Ilhas de Andamã / Andamaneses	24	1,5	Gallego Romero et al. (2012)
Índia	Região Central / Indianos	179	5,6	Gallego Romero et al. (2012)
Índia	Região Leste / Indianos	310	1,8	Gallego Romero et al. (2012)
Índia	Região Nordeste / Indianos	139	0,7	Gallego Romero et al. (2012)
Índia	Região Norte / Indianos	77	19,5	Babu et al. (2010)
Índia	Região Norte / Indianos	290	15,9	Gallego Romero et al. (2012)
Índia	Região Oeste / Indianos	468	21,3	Gallego Romero et al. (2012)
Índia	Região Sul / Indianos	76	6,6	Babu et al. (2010)
Índia	Região Sul / Indianos	864	8,5	Gallego Romero et al. (2012)
Irã	/ Iranianos	21	10,0	Enattah et al. (2007)
Irã	/ Qashqai	10	5,0	Enattah et al. (2007)
Jordânia	/ Jordanianos	112	5,4	Enattah et al. (2008)
Paquistão	/ Balooch	25	36,0	Bersaglieri et al. (2004)
Paquistão	/ Balooch	19	34,0	Enattah et al. (2007)
Paquistão	/ Balti	23	0	Enattah et al. (2007)
Paquistão	/ Brahui	25	34,0	Bersaglieri et al. (2004)
Paquistão	/ Brahui	30	27,0	Enattah et al. (2007)
Paquistão	/ Burusho	25	10,0	Bersaglieri et al. (2004)
Paquistão	/ Burusho	30	2,0	Enattah et al. (2007)
Paquistão	/ Hazara	25	8,0	Bersaglieri et al. (2004)
Paquistão	/ Hazara	14	4,0	Enattah et al. (2007)
Paquistão	/ Kalash	25	0	Bersaglieri et al. (2004)
Paquistão	/ Kalash	30	0	Enattah et al. (2007)
Paquistão	/ Kashmiri	20	12,0	Enattah et al. (2007)
Paquistão	/ Makrani	25	34,0	Bersaglieri et al. (2004)
Paquistão	/ Makrani	29	17,0	Enattah et al. (2007)
Paquistão	/ Mohannes	29	28,0	Enattah et al. (2007)
Paquistão	/ Parsi	29	14,0	Enattah et al. (2007)
Paquistão	/ Pathan	25	30,0	Bersaglieri et al. (2004)
Paquistão	/ Pathan	28	30,0	Enattah et al. (2007)
Paquistão	/ Sindhi	25	32,0	Bersaglieri et al. (2004)
Paquistão	/ Sindhi	28	41,0	Enattah et al. (2007)
Rússia	República de Komi / Komis	10	15,0	Enattah et al. (2007)
Rússia	República de Mordovia / Erzas	30	27,0	Enattah et al. (2007)
Rússia	República de Mordovia / Mokshas	30	28,0	Enattah et al. (2007)
Rússia	República de Udmúrtia / Udmurtes	30	33,0	Enattah et al. (2007)

FONTE: A autora (2020).

APÊNDICE 10 – FREQUÊNCIAS ALÉLICAS DO SNP -13910 C>T EM POPULAÇÕES DA AMÉRICA E OCEANIA

País	Localidade / População	N	Frequência alélica (%)	Referência
Chile	Araucanía / Mapuche	29	5,0	Fernández; Flores (2014)
Chile	Araucanía / <i>Mestizos</i>	115	12,5	Fernández; Flores (2014)
Equador	Amazônia / <i>Mestizos</i>	48	7,3	Paz-Y-Miño et al. (2016)
Equador	Costa / Afro-equadorianos	49	6,1	Paz-Y-Miño et al. (2016)
Equador	Costa / <i>Mestizos</i>	124	12,5	Paz-Y-Miño et al. (2016)
Equador	Serra / Afro-equadorianos	76	10,5	Paz-Y-Miño et al. (2016)
Equador	Serra / Ameríndios	128	7,4	Paz-Y-Miño et al. (2016)
Equador	Serra / <i>Mestizos</i>	316	15,2	Paz-Y-Miño et al. (2016)
México	Jalisco / Cuquio	99	26,8	Ojeda-Granados et al. (2016)
México	Jalisco / Guadalajara	699	20,0	Ojeda-Granados et al. (2016)
México	Jalisco / Nahuas	86	1,2	Ojeda-Granados et al. (2016)
México	Jalisco / Villa Purificación	32	32,8	Ojeda-Granados et al. (2016)
México	Nayarit / Huicholes	95	0	Ojeda-Granados et al. (2016)
México	Nayarit / Tepic	182	15,7	Ojeda-Granados et al. (2016)
Papua-Nova Guiné	Bougainville / Melanésios	44	0	Bersaglieri et al. (2004)
Papua-Nova Guiné	/ Papuas	34	0	Bersaglieri et al. (2004)

FONTE: A autora (2020).

APÊNDICE 11 – FREQUÊNCIAS ALÉLICAS DO SNP -13907 C>G

País	Localidade / População	N	Frequência alélica (%)	Referência
Arábia Saudita	/ Sauditas	248	0,8	Enattah et al. (2008)
Arábia Saudita	Região do Nilo / Sauditas	7	2,2	Priehodová et al. (2014)
Etiópia	/ Afar	10	20,0	Ingram et al. (2007)
Etiópia	/ Afar	72	25,0	Jones et al. (2015)
Etiópia	/ Afar	37	29,7	Ingram et al. (2009b)
Etiópia	/ Amhara	19	5,3	Ingram et al. (2007)
Etiópia	/ Amhara	71	5,6	Jones et al. (2015)
Etiópia	/ Beta Israel	18	8,3	Ranciaro et al. (2014)
Etiópia	/ Maale	61	1,6	Jones et al. (2015)
Etiópia	/ Manjo	30	1,8	Jones et al. (2015)
Etiópia	/ Oromo	74	6,8	Jones et al. (2015)
Etiópia	/ Somali	9	5,6	Ingram et al. (2007)
Etiópia	/ Tigre	44	19,3	Jones et al. (2013)
Etiópia	/ Wolayita	26	7,8	Jones et al. (2013)
Iêmen	Al Hudaydah / Iemenitas	78	1,5	Priehodová et al. (2014)
Iêmen	Hadramaut / Iemenitas	83	1,5	Liebert et al. (2017)
Quênia	/ Borana	17	11,8	Ranciaro et al. (2014)
Quênia	/ Gabra	41	8,5	Ranciaro et al. (2014)
Quênia	/ Garreh	19	7,9	Ranciaro et al. (2014)
Quênia	/ Ilchamus Maasai	10	5,0	Ranciaro et al. (2014)
Quênia	/ Mumonyot Maasai	7	7,1	Ranciaro et al. (2014)
Quênia	/ Orma	24	2,1	Ranciaro et al. (2014)
Quênia	/ Rendille	12	4,2	Ranciaro et al. (2014)
Quênia	/ Samburu	9	6,3	Ranciaro et al. (2014)
Quênia	/ Wata	22	2,3	Ranciaro et al. (2014)
Quênia	/ Yaaku	14	3,6	Ranciaro et al. (2014)
Sudão	/ Beja Beni Amer	82	6,1	Ingram et al. (2007)
Sudão	/ Beja Beni Amer	77	13,0	Jones et al. (2015)
Sudão	/ Dounglawi	6	8,3	Ingram et al. (2007)
Sudão	/ Beja Beni Amer	6	25,0	Ranciaro et al. (2014)
Sudão	/ Beja Hadandawa	11	18,2	Ranciaro et al. (2014)
Sudão	/ Fulani	39	1,3	Hassan et al. (2016); Liebert et al. (2017)
Sudão	/ Jaali	88	0	Ingram et al. (2007)
Sudão	/ Jaali	64	8,0	Jones et al. (2015)
Sudão	/ Shaigi	9	0	Ingram et al. (2007)
Sudão	Karima / Shwaiga	38	2,6	Hassan et al. (2016)
Sudão	Kerma / Mahas	39	1,3	Hassan et al. (2016); Liebert et al. (2017)
Sudão	New Halfa / Shokrya	37	2,5	Hassan et al. (2016)
Sudão	Região Norte / Jaali	20	5,0	Enattah et al. (2008)
Sudão	Sinkat / Beja Beni Amer	37	14,9	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Wadi Halfa / Halfawien	39	1,3	Hassan et al. (2016); Liebert et al. (2017)

FONTE: A autora (2020).

APÊNDICE 12 – FREQUÊNCIAS ALÉLICAS DO SNP -13915 T>G

País	Localidade / População	N	Frequência alélica (%)	Referência
Arábia Saudita	/ Beduínos	46	48,9	Ingram et al. (2007)
Arábia Saudita	Região Central / Sauditas	90	60,6	Imtiaz et al. (2007)
Arábia Saudita	Região Leste / Sauditas	82	61,6	Imtiaz et al. (2007)
Arábia Saudita	Região Norte / Sauditas	86	64,5	Imtiaz et al. (2007)
Arábia Saudita	Região Oeste / Sauditas	92	57,6	Imtiaz et al. (2007)
Arábia Saudita	Região Sul / Sauditas	82	52,4	Imtiaz et al. (2007)
Camarões	/ Baggara	31	3,2	Ranciaro et al. (2014)
Camarões	/ Shuwa	16	6,3	Ingram et al. (2007)
Camarões	/ Shuwa	51	5,9	Jones et al. (2015)
Chade	Região de Mao / Baggara	27	46,3	Priehodová et al. (2014)
Egito	El Hayz / Árabes	34	1,5	Priehodová et al. (2014)
Equador	/ Mestízos	488	0,5	Paz-Y-Miño et al. (2016)
Etiópia	/ Afar	10	15,0	Ingram et al. (2007)
Etiópia	/ Afar	72	18,8	Jones et al. (2015)
Etiópia	/ Amhara	19	13,2	Ingram et al. (2007)
Etiópia	/ Amhara	71	4,2	Jones et al. (2015)
Etiópia	/ Beta Israel	18	2,8	Ranciaro et al. (2014)
Etiópia	/ Maale	61	4,1	Jones et al. (2015)
Etiópia	/ Oromo	74	7,4	Jones et al. (2015)
Etiópia	/ Somali	9	5,6	Ingram et al. (2007)
Iêmen	/ Iemenitas Judaicos	23	15,2	Ranciaro et al. (2014)
Iêmen	Al Hudaydah / Iemenitas	43	65,1	Priehodová et al. (2014)
Iêmen	Hadramaut / Iemenitas	40	26,3	Priehodová et al. (2014)
Iêmen	Ilha de Soqotra / Iemenitas	65	58,5	Priehodová et al. (2014)
Iêmen	Região de Hajja / Iemenitas	34	61,8	Priehodová et al. (2014)
Iêmen	Região de Hudeida / Iemenitas	66	59,1	Priehodová et al. (2014)
Índia	Região Sul / Indianos	864	0,8	Gallego Romero et al. (2012)
Jordânia	/ Beduínos	23	39,1	Ingram et al. (2007)
Kuwait	Al Ahmadi / Ajman	37	57,0	Hill; Mohammad; Kivisild (2013)
Kuwait	Al Farwaniyah / Mutran	29	55,0	Hill; Mohammad; Kivisild (2013)
Mali	Koro / Fulani	162	0,6	Lokki et al. (2011)
Nigéria	Região de Ngala / Shuwa	53	6,6	Priehodová et al. (2014)
Omã	/ Omanis de origem asiática	96	0	Al-Abri et al. (2012)
Omã	/ Iemenitas	239	54,8	Al-Abri et al. (2012)
Omã	Região Norte / Árabes	342	13,6	Al-Abri et al. (2012)
Omã	Região Sul / Árabes Dhofari	210	27,6	Al-Abri et al. (2012)
Palestina	/ Beduínos	19	13,2	Ingram et al. (2007)
Palestina	/ Drusos	14	10,7	Ingram et al. (2007)
Palestina	/ Palestinos	22	4,5	Ranciaro et al. (2014)
Palestina	Região Urbana / Árabes	81	4,9	Ingram et al. (2007)
Quênia	/ Borana	17	8,8	Ranciaro et al. (2014)
Quênia	/ Burji	8	6,3	Ranciaro et al. (2014)
Quênia	/ El Molo	16	3,1	Ranciaro et al. (2014)
Quênia	/ Gabra	41	19,5	Ranciaro et al. (2014)
Quênia	/ Garreh	19	15,8	Ranciaro et al. (2014)

País	Localidade / População	N	Frequência alélica (%)	Referência
Quênia	/ Konso	6	8,3	Ranciaro et al. (2014)
Quênia	/ Marakwet	7	7,1	Ranciaro et al. (2014)
Quênia	/ Orma	24	12,5	Ranciaro et al. (2014)
Quênia	/ Pokot	14	3,6	Ranciaro et al. (2014)
Quênia	/ Rendille	12	8,3	Ranciaro et al. (2014)
Quênia	/ Saboot	6	8,3	Ranciaro et al. (2014)
Quênia	/ Samburu	8	5,6	Ranciaro et al. (2014)
Quênia	/ Sanye	14	7,1	Ranciaro et al. (2014)
Quênia	/ Wata	22	15,9	Ranciaro et al. (2014)
Sudão	/ Beja Beni Amer	82	24,4	Ingram et al. (2007)
Sudão	/ Beja Beni Amer	77	24,0	Jones et al. (2015)
Sudão	/ Beja Beni Amer	6	16,7	Ranciaro et al. (2014)
Sudão	/ Beja Hadandawa	11	18,2	Ranciaro et al. (2014)
Sudão	/ Jaali	64	15,6	Jones et al. (2015)
Sudão	/ Shaigi	9	5,6	Ingram et al. (2007)
Sudão	El-Fashir / Darfurians	49	1,0	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Kadugli / Nuba	40	3,5	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Karima / Shwaiga	38	13,2	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Kerma / Mahas	39	3,8	Hassan et al. (2016); Liebert et al. (2017)
Sudão	New Halfa / Shokrya	40	30,0	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Região do Nilo / Árabes	46	7,6	Priehodová et al. (2014)
Sudão	Região Leste / Rashaayda	52	76,9	Priehodová et al. (2014)
Sudão	Shendi / Gaali	40	11,5	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Sinkat / Beja Beni Amer	37	33,7	Hassan et al. (2016); Liebert et al. (2017)
Sudão	Wadi Halfa / Halfawien	39	11,5	Hassan et al. (2016); Liebert et al. (2017)

FONTE: A autora (2020).

APÊNDICE 13 – FREQUÊNCIAS ALÉLICAS DO SNP -14009 T>G

País	Localidade / População	N	Frequência alélica (%)	Referência
Camarões	/ Baggara	31	1,6	Liebert et al. (2017); Ranciaro et al. (2014)
Chade	/ Abbala	27	3,7	Liebert et al. (2017); Ranciaro et al. (2014)
Chade	Região de Mao / Baggara	27	3,7	Priehodová et al. (2014)
Egito	El Hayz / Árabes	34	5,9	Priehodová et al. (2014)
Etiópia	/ Afar	37	1,3	Ingram et al. (2009b)
Etiópia	/ Afar	72	0,7	Jones et al. (2015)
Etiópia	/ Amhara	71	2,6	Jones et al. (2015)
Etiópia	/ Maale	61	5,7	Jones et al. (2015)
Etiópia	/ Manjo	30	3,3	Jones et al. (2015)
Etiópia	/ Oromo	74	3,4	Jones et al. (2015)
Etiópia	/ Somali	37	1,3	Ingram et al. (2009b)
Iêmen	Al-Akhkum / Iemenitas	43	1,2	Priehodová et al. (2014)
Iêmen	/ Iemenitas Judaicos	23	1,0	Liebert et al. (2017); Ranciaro et al. (2014)
Israel	/ Árabes não-Beduínos	80	0,6	Ingram et al. (2009b)
Kuwait	/ Kuwaitianos	33	1,5	Liebert et al. (2017)
Líbano	/ Libaneses	24	1,0	Liebert et al. (2017); Ranciaro et al. (2014)
Nigéria	Região de Ngala / Shuwa	53	0,9	Priehodová et al. (2014)
Omã	/ Omanis de origem asiática	96	1,5	Al-Abri et al. (2012)
Palestinos	/ Palestinos	22	1,0	Liebert et al. (2017); Ranciaro et al. (2014)
Paquistão	/ Hazara	12	1,0	Liebert et al. (2017); Ranciaro et al. (2014)
Paquistão	/ Kalash	16	1,0	Liebert et al. (2017); Ranciaro et al. (2014)
Sudão	/ Beja Beni Amer	81	10,5	Ingram et al. (2009b)
Sudão	/ Beja Beni Amer	77	11,7	Jones et al. (2015)
Sudão	/ Beja Beni Amer	6	23,5	Liebert et al. (2017); Ranciaro et al. (2014)
Sudão	/ Beja Hadandawa	11	23,5	Liebert et al. (2017); Ranciaro et al. (2014)
Sudão	/ Jaali	86	6,4	Ingram et al. (2009b)
Sudão	/ Jaali	64	7,0	Jones et al. (2015)
Sudão	/ Shaigi	9	16,7	Ingram et al. (2009b); Liebert et al. (2017)
Sudão	Região do Nilo / Árabes	46	13,0	Priehodová et al. (2014)

FONTE: A autora (2020).

APÊNDICE 14 – FREQUÊNCIAS ALÉLICAS DO SNP -14010 G>C

País	Localidade / População	N	Frequência alélica (%)	Referência
África do Sul	/ !Xhosa	21	14,3	Ranciaro et al. (2014)
África do Sul	/ !Xun	11	4,5	Ranciaro et al. (2014)
África do Sul	Askham / ≠ Khomani	20	5,0	Breton et al. (2014)
África do Sul	Askham / População miscigenada	20	22,5	Breton et al. (2014)
África do Sul	Colesberg / Karretjie	20	10,0	Breton et al. (2014)
África do Sul	Colesberg / População miscigenada	20	2,5	Breton et al. (2014)
África do Sul	Região Sudeste / Falantes Bantu	41	7,3	Breton et al. (2014)
África do Sul	Wellington / População miscigenada	20	10,0	Breton et al. (2014)
Angola	/ !Xun	36	1,4	Pinto et al. (2016)
Angola	/ Himba	69	8,7	Pinto et al. (2016)
Angola	/ Kuvale	54	6,0	Coelho et al. (2009); Liebert et al. (2017)
Angola	/ Kuvale	81	3,7	Pinto et al. (2016)
Angola	/ Kwepe	34	4,4	Pinto et al. (2016)
Angola	/ Kwisi	20	17,5	Pinto et al. (2016)
Angola	/ Nyaneka-Nkhumbi	153	3,0	Coelho et al. (2009); Liebert et al. (2017)
Angola	/ Ovimbundu	96	1,0	Coelho et al. (2009); Liebert et al. (2017)
Angola	/ Tjimba	15	23,3	Pinto et al. (2016)
Angola	/ Twa	18	19,4	Pinto et al. (2016)
Angola	Menogue / !Xun	20	5,0	Breton et al. (2014)
Botsuana	/ !lAni	11	9,1	Macholdt et al. (2014)
Botsuana	/ G!ui	17	8,8	Macholdt et al. (2014)
Botsuana	/ Gllana	10	20,0	Macholdt et al. (2014)
Botsuana	/ Naro	19	5,3	Macholdt et al. (2014)
Botsuana	/ Shua	27	7,4	Macholdt et al. (2014)
Botsuana	/ Tshwa	15	16,7	Macholdt et al. (2014)
Botsuana	/ Tswana	18	2,8	Macholdt et al. (2014)
Botsuana	Região Leste / Taa	11	4,5	Macholdt et al. (2014)
Botsuana	Região Norte / Ju!'hoan	21	2,4	Macholdt et al. (2014)
Botsuana	Região Sul / Ju!'hoan	26	5,8	Macholdt et al. (2014)
Botsuana	Região Oeste / Taa	20	2,5	Macholdt et al. (2014)
Iêmen	Al-Akhkum / Iemenitas	43	1,2	Priehodová et al. (2014)
Iêmen	Hadramaut / Iemenitas	83	0,6	Liebert et al. (2017)
Iêmen	Hodeida / Iemenitas	66	0,8	Priehodová et al. (2014)
Moçambique	/ Shangaan	23	2,2	Pinto et al. (2016)
Namíbia	/ !Xun	19	5,3	Macholdt et al. (2014)
Namíbia	/ !lXo	19	7,9	Macholdt et al. (2014)
Namíbia	/ Damara	34	4,4	Macholdt et al. (2014)
Namíbia	/ Haillom	40	8,8	Macholdt et al. (2014)
Namíbia	/ Herero	21	7,1	Macholdt et al. (2014)
Namíbia	/ Himba	16	12,5	Macholdt et al. (2014)
Namíbia	/ Nama	25	36,0	Macholdt et al. (2014)
Namíbia	Tsumkwe / Ju!'hoansi	20	2,5	Breton et al. (2014)
Namíbia	Windhoek / Nama	22	20,5	Breton et al. (2014)
Quênia	/ Boni	24	20,8	Ranciaro et al. (2014)
Quênia	/ Borana	17	8,8	Ranciaro et al. (2014)
Quênia	/ Burji	8	6,3	Ranciaro et al. (2014)

País	Localidade / População	N	Frequência alélica (%)	Referência
Quênia	/ El Molo	16	21,9	Ranciaro et al. (2014)
Quênia	/ Gabra	41	4,9	Ranciaro et al. (2014)
Quênia	/ Il'gvesi Maasai	11	68,2	Ranciaro et al. (2014)
Quênia	/ Ilchamus Maasai	10	55,0	Ranciaro et al. (2014)
Quênia	/ Keyo	2	50,0	Ranciaro et al. (2014)
Quênia	/ Kikuyu	7	50,0	Ranciaro et al. (2014)
Quênia	/ Konso	6	8,3	Ranciaro et al. (2014)
Quênia	/ Luo	22	4,5	Ranciaro et al. (2014)
Quênia	/ Marakwet	7	35,7	Ranciaro et al. (2014)
Quênia	/ Mumonyot Maasai	7	50,0	Ranciaro et al. (2014)
Quênia	/ Nandi	4	25,0	Ranciaro et al. (2014)
Quênia	/ Ogiek	11	36,4	Ranciaro et al. (2014)
Quênia	/ Orma	24	22,9	Ranciaro et al. (2014)
Quênia	/ Pare	18	19,4	Ranciaro et al. (2014)
Quênia	/ Pokot	14	28,6	Ranciaro et al. (2014)
Quênia	/ Purko Maasai	1	100,0	Ranciaro et al. (2014)
Quênia	/ Rendille	12	8,3	Ranciaro et al. (2014)
Quênia	/ Sabaot	6	16,7	Ranciaro et al. (2014)
Quênia	/ Samburu	9	27,8	Ranciaro et al. (2014)
Quênia	/ Sanye	14	3,6	Ranciaro et al. (2014)
Quênia	/ Sengwer	16	6,3	Ranciaro et al. (2014)
Quênia	/ Taita	10	10,0	Ranciaro et al. (2014)
Quênia	/ Tugen	15	23,3	Ranciaro et al. (2014)
Quênia	/ Turkana	13	19,1	Ranciaro et al. (2014)
Quênia	/ Wata	22	15,9	Ranciaro et al. (2014)
Quênia	/ Yaaku	14	53,6	Ranciaro et al. (2014)
Tanzânia	/ Akie	14	21,4	Ranciaro et al. (2014)
Tanzânia	/ Burunge	17	38,2	Ranciaro et al. (2014)
Tanzânia	/ Datoga	4	62,5	Ranciaro et al. (2014)
Tanzânia	/ Dorobo	10	40,0	Ranciaro et al. (2014)
Tanzânia	/ Iraqw	26	51,9	Ranciaro et al. (2014)
Tanzânia	/ Fiome	10	55,0	Ranciaro et al. (2014)
Tanzânia	/ Maasai	19	44,7	Ranciaro et al. (2014)
Tanzânia	/ Mbugu	29	31,0	Ranciaro et al. (2014)
Tanzânia	/ Mbugwe	12	29,2	Ranciaro et al. (2014)
Tanzânia	/ Pare	10	10,0	Ranciaro et al. (2014)
Tanzânia	/ Rangi	35	30,0	Ranciaro et al. (2014)
Tanzânia	/ Sandawe	31	14,5	Ranciaro et al. (2014)
Zâmbia	/ Nkoya	16	3,1	Macholdt et al. (2014)

FONTE: A autora (2020).

APÊNDICE 15 – FREQUÊNCIAS ALÉLICAS DO SNP -14011 C>T

País	Localidade / População	N	Frequência alélica (%)	Referência
-	África Ocidental	26	1,9	Ranciaro et al. (2014)
Botsuana	/ Kgalagadi	20	2,5	Macholdt et al. (2014)
Etiópia	/ Etiópes de origem não identificada	89	0,6	Jones et al. (2013)
Etiópia	/ Falantes de línguas cushíticas	47	1,1	Jones et al. (2013)
Índia	Região Norte / Indianos	463	0,8	Gallego Romero et al. (2012)
Índia	Região Oeste / Indianos	291	2,1	Gallego Romero et al. (2012)
Omã	Região Norte / Árabes	342	1,0	Al-Abri et al. (2012)

FONTE: A autora (2020).

ANEXO 1 – PARÂMETROS DE DESEQUILÍBRIO DE LIGAÇÃO

$$D_{ab} = (\pi_{AB} - \pi_A\pi_B).$$

$$r^2 = \frac{(D_{ab})^2}{\pi_A\pi_a\pi_B\pi_b}.$$

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A\pi_b, \pi_a\pi_B)} \text{ for } D_{ab} < 0;$$

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A\pi_B, \pi_a\pi_b)} \text{ for } D_{ab} > 0.$$

FONTE: FLINT-GARCIA; THORNSBERRY; BUCKLER IV, 2003.

LEGENDA: A e a = Alelos localizados no *locus* 1; B e b = Alelos localizados no *locus* 2; π_{AB} , π_{Ab} , π_{aB} e π_{ab} = Frequências haplotípicas; π_A , π_a , π_B , e π_b = Frequências alélicas.