



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
DEPARTAMENTO DE ARQUEOLOGIA
CURSO DE BACHARELADO EM ARQUEOLOGIA

Ana Cristina de Lima Pedrozo Cardouzo

**IDENTIFICAÇÃO DE MARCADORES GENÔMICOS ASSOCIADOS À
TUBERCULOSE EM INDIVÍDUOS ANTIGOS DAS AMÉRICAS**

Recife
2024

Ana Cristina de Lima Pedrozo Cardouzo

**IDENTIFICAÇÃO DE MARCADORES GENÔMICOS ASSOCIADOS À
TUBERCULOSE EM INDIVÍDUOS ANTIGOS DAS AMÉRICAS**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Arqueologia da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Arqueologia.

Orientador (a): Prof. Dr. Henry Socrates Lavallo Sullasi

Coorientador (a): Dr. André Luiz Campelo dos Santos

Recife

2024

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Cardouzo, Ana Cristina de Lima Pedrozo.

Identificação de Marcadores Genômicos Associados à Tuberculose em
Indivíduos Antigos das Américas / Ana Cristina de Lima Pedrozo Cardouzo. -
Recife, 2024.

73 : il., tab.

Orientador(a): Henry Socrates Lavallo Sullasi

Coorientador(a): André Luiz Campelo dos Santos

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Filosofia e Ciências Humanas, Arqueologia -
Bacharelado, 2024.

Inclui referências.

1. Arqueologia. 2. Arqueogenética. 3. Arqueologia da Doença . 4.
Tuberculose. 5. Marcadores Genéticos. 6. Bioinformática. I. Sullasi, Henry
Socrates Lavallo . (Orientação). II. Santos , André Luiz Campelo dos .
(Coorientação). IV. Título.

930 CDD (22.ed.)

Ana Cristina de Lima Pedrozo Cardouzo

**IDENTIFICAÇÃO DE MARCADORES GENÔMICOS ASSOCIADOS À
TUBERCULOSE EM INDIVÍDUOS ANTIGOS DAS AMÉRICAS**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Arqueologia da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Arqueologia.

Aprovado em: 28/11/2024

BANCA EXAMINADORA

Prof. Dr. Henry Socrates Lavalle Sullasi (Orientador)
Universidade Federal de Pernambuco

Dr. André Luiz Campelo dos Santos (Coorientador)
Universidade Atlântica da Flórida

Me. Jouldes Matos Duarte (Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Allysson Allan de Farias (Examinador Externo)
Universidade Federal do Ceará

AGRADECIMENTOS

Primeiramente, agradeço a Deus por ter me sustentado durante toda essa jornada. Junto a ele, agradeço a minha família, a base de tudo, em especial, minha mãe, minha irmã e meu pai por sempre me apoiarem e me ensinarem o que é ser resiliente. Junto a eles, agradeço às minhas tias, tios e avós, por todo o suporte durante os momentos difíceis e os sacrifícios que fizeram para que eu seguisse um sonho.

Agradeço também aos meus amigos, aos antigos e aos novos, por estarem comigo durante esse percurso, especialmente, a Júlia Cardouzo por sempre me lembrar quem eu sou, e a Yasmin Koury por todas as conversas e risos compartilhados.

Obrigada a João Rhannyé, Thaysa Gisella, Marcela Vitória, Giovanna Lima, Izabela Ingrid, Kimberly Pimentel e Isabela Maia por terem trilhado esse caminho comigo e por terem tornado o Recife um lar.

Agradeço também o corpo docente da graduação em Arqueologia da Universidade Federal de Pernambuco por todos os ensinamentos compartilhados. Em especial, agradeço ao meu orientador, professor Henry, por todas as orientações, oportunidades e por ter apresentado André, meu coorientador, a qual dedico eternos agradecimentos.

Dentre os docentes, agradeço muitíssimo ao professor Bruno Tavares por todas as conversas e conselhos. O senhor se tornou um grande amigo durante o percurso da graduação. Junto a isso, estendo os agradecimentos aos amigos que fiz no GEAGT – Ingrid, Rafael, João Ricardo e Carol – obrigada por terem tirado de mim as risadas mais sinceras.

Agradeço também ao LEARQ, Laboratório de Estudos Arqueométricos, por todo o suporte durante a elaboração da presente pesquisa, especialmente, a Rafael Leal e Joulde por todos os surtos com inglês e bioinformática. Foi maravilhoso partilhar esses momentos com vocês.

Por fim, estendo eternos agradecimentos aos meus companheiros de tatame que, indiretamente, foram fundamentais na minha construção como pessoa e pesquisadora. Em especial, agradeço ao meu sensei, Guto Mesquita, as minhas meninas: Adriana, Isadora, Kamilly e Malu por não me deixarem desistir, e aos meus parceiros Carlos Arthur, Carlo Farias, Radson, Paulo, Cauê, Enzo, Tobias e Túlio por todas as risadas, ensinamentos e lutas. OSS!

“Todo interesse pela doença e pela morte é apenas mais uma expressão de interesse pela vida” - Thomas Mann

RESUMO

A Arqueogenética é área da Arqueologia que nos conduz ao limiar entre biologia e humanidades. Sendo um campo em rápida e franca expansão, é fascinante que uma das fontes mais importantes sobre o passado humano esteja dentro de nós. Outro pequeno fator, ou melhor, ser, que também se esconde dentro de nós são os patógenos – bactérias, vírus, protozoários e outros seres que mudaram, mais que qualquer outro fator, a história humana. Dentre eles está o causador da temida Tuberculose, uma das doenças que mais matou pessoas ao longo da história, sendo por isso conhecida como “a maior assassina de todos os tempos”. Todavia, apesar de ser uma das doenças infecciosas mais estudadas ao longo dos anos, pouco ainda se compreende sobre essa infecção, a relação patógeno-hospedeiro e outras nuances. Nesse contexto, considerando a hipótese de que a Tuberculose existia nas Américas antes do contato europeu, o presente trabalho objetiva apontar marcadores genéticos que estão associados à suscetibilidade ou à resistência à doença no continente Americano antes da chegada dos colonizadores europeus, e com isso tecer hipóteses e questionamentos sobre a sua distribuição e prevalência no continente. Para tal, realizou-se um extenso levantamento bibliográfico acerca da Tísica Pulmonar e das alterações genéticas associadas à mesma. Com uma abordagem da Arqueogenética, tendo um DNA antigo como fonte, junto a um olhar da Arqueologia da Doença, visto que estudar as patologias de um grupo auxilia na reconstrução e compreensão de aspectos do cotidiano desse grupo, aplicaram-se técnicas da Bioinformática para buscar no conjunto amostral analisado – sendo este referente a 224 indivíduos advindos de distintos sítios arqueológicos das Américas, trabalhados por Santos (2020) em sua tese e por Lindo e colegas (2018) em um artigo – as alterações genéticas outrora levantadas. Em razão disso, através da elaboração de gráficos para uma melhor visualização dos resultados, foi possível observar uma maior predominância de alterações que causam suscetibilidade à Tuberculose pelos povos originários da América, o que pode ter sido um fator de causa para o grande número de mortes que ocorreu com a chegada dos colonizadores e da variante europeia da doença. Destarte, a pesquisa também levanta alguns questionamentos acerca da origem da Tuberculose no continente Americano e reforça a importância de se debruçar sobre a doença e seu patógeno, bem como a relação de patógeno-hospedeiro que o mesmo, possivelmente, possui com a espécie humana, acreditando que este seja um campo frutífero para futuros estudos.

Palavras-chave: Arqueogenética, Arqueologia da Doença, Bioinformática, Tuberculose

ABSTRACT

Archaeogenomics is an area of Archeology that takes us to the threshold between biology and the humanities. As a rapidly expanding field, it is fascinating that one of the most important sources about the human past is within us. Another small factor, or rather being, that also hides within us are pathogens – bacteria, viruses, protozoa and other microorganisms that have changed, more than any other factor, human history. Among them is the cause of the dreaded Tuberculosis, one of the diseases that killed the most people throughout history and is therefore known as “the biggest killer of all time”. However, despite being one of the most studied infectious diseases, little is still understood about this infection, the pathogen-host relationship and other nuances. In this context, considering the hypothesis that Tuberculosis existed in the Americas before European contact, the present work aims to point out genetic markers that are associated with susceptibility or resistance to the disease on the American continent before the arrival of European colonizers and with this we weave hypotheses and questions about its distribution and prevalence on the continent. To this end, an extensive bibliographic survey was carried out on Pulmonary Phthisis and the genetic changes associated with it. With an Archeogenomics approach, having ancient DNA as a source, along with a look at the Archeology of Disease, since studying the pathologies of a group helps in the reconstruction and understanding of aspects of the daily life of that group, Bioinformatics techniques were applied, through specific environments, software and tools, to search the analyzed sample set – this referring to 224 individuals coming from different archaeological sites in the Americas and worked on by Santos (2020) in his thesis and by Lindo and colleagues (2018) in an article – the genetic changes previously raised. As a result, through the creation of graphs for better visualization of the results, we find a predominance of variants that caused susceptibility to Tuberculosis among the original peoples of the Americas, which may have been a factor in the large number of deaths that occurred with the arrival of colonizers and the European variant of Tuberculosis. Therefore, the research also raises some questions about the origin of Tuberculosis on the American continents and reinforces the importance of focusing on the disease and its pathogen, as well as the pathogen-host relationship that it possibly has with the human species, believing that this is a fruitful field for future studies.

Keywords: Archaeogenomics, Archeology of Disease, Bioinformatics, Tuberculosis

LISTA DE ILUSTRAÇÕES

Figura 1 -	Representação de um SNP (Polimorfismo de Nucleotídio Único) tipo mais comum de polimorfismo.	15
Figura 2 -	Micrografia eletrônica de Varredura do <i>Mycobacterium tuberculosis</i>	20
Figura 3 -	Representação Cartográfica da saída da África das linhagens de Tuberculose Humana	22
Figura 4 -	Exemplo de Dendograma (árvore filogenética) de uma planta do gênero <i>Protium</i>	28
Figura 5 -	Distribuição das Amostras	38
Figura 6 -	Arquivo VCF ao ser aberto na plataforma da planilha do google/excel	40
Figura 7 -	Distribuição dos SNPs no continente americano	43
Gráfico 1 -	Proporção de variantes associadas à Suscetibilidade e Resistência a Tuberculose no continente americano antes da colonização	46
Gráfico 2 -	Proporção de variantes associadas à Suscetibilidade e Resistência na América do Norte antes do contato	46
Gráfico 3 -	Proporção de variantes associadas à Suscetibilidade e Resistência na América do Sul antes do contato	47
Gráfico 4 -	Dispersão dos SNPs de Suscetibilidade e Resistência nos países analisados	48
Gráfico 5 -	Proporção dos genes analisados predominantes nas Américas pré-contato	48
Gráfico 6 -	Proporção dos genes predominantes na América do Sul antes do contato	49

Gráfico 7 -	Proporção dos genes predominantes na América do Norte antes do contato	50
Gráfico 8 -	Proporção das Variantes (SNP) associados à TB presentes na América pré-contato	51
Gráfico 9 -	Proporção das variantes (SNP) associados à TB presentes na América do Norte	52
Gráfico 10 -	Proporção das variantes (SNP) associados à TB presentes na América do Sul	52
Gráfico 11 -	Proporção de variantes associadas à Suscetibilidade versus Resistência nas Américas no Intervalo 1	53
Gráfico 12 -	Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Norte no Intervalo 1	54
Gráfico 13 -	Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Sul no Intervalo 1	54
Gráfico 14 -	Proporção de variantes associadas à Suscetibilidade versus Resistência nas Américas no Intervalo 2	55
Gráfico 15 -	Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Norte no Intervalo 2	55
Gráfico 16 -	Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Sul no Intervalo 2	56
Gráfico 17 -	Proporção de genes associados à TB nas Américas no Intervalo 1	56
Gráfico 18 -	Proporção de genes associados à TB na América do Sul no Intervalo 1	57
Gráfico 19 -	Proporção de genes associados à TB na América do Norte no Intervalo 1	57

Gráfico 20 -	Proporção de genes associados à TB nas Américas no	58
	Intervalo 2	
Gráfico 21 -	Proporção de genes associados à TB na América do Sul no	58
	Intervalo 2	
Gráfico 22 -	Proporção de genes associados à TB na América do Norte	59
	Intervalo 2	
Gráfico 23 -	Proporção dos SNP associados à TB na América no	59
	Intervalo 1	
Gráfico 24 -	Proporção dos SNP associados à TB na América no	60
	Intervalo 2	
Gráfico 25 -	Proporção dos SNP associados à TB na América do Sul no	61
	Intervalo 1	
Gráfico 26 -	Proporção dos SNP associados à TB na América do Sul no	61
	Intervalo 2	
Gráfico 27 -	Proporção dos SNP associados à TB na América do Norte	62
	no Intervalo 1	
Gráfico 28 -	Proporção dos SNP associados à TB na América do Norte no	62
	Intervalo 2	

LISTA DE TABELAS

Tabela 1 -	Linhagens das cepas de Tuberculose em Humanos	21
Tabela 2 -	Polimorfismos associados à suscetibilidade à Tuberculose	36
Tabela 3 -	Polimorfismos associados à resistência à Tuberculose	36
Tabela 4 -	Polimorfismos encontrados no processo da “Chamada de Variante”	42
Tabela 5 -	Datações e Sítios dos Indivíduos Analisados	43

LISTA DE ABREVIATURAS E SIGLAS

AP	Antes do Presente
MTBC	Complexo <i>Mycobacterium tuberculosis</i>
M.tb/MTB	<i>Mycobacterium tuberculosis</i> (em humanos)
SNP	Polimorfismo de Nucleotídeo Único
TB	Tuberculose
TLR	Receptores do tipo Toll/ Toll Like Receptors

SUMÁRIO

1 INTRODUÇÃO	14
1.1 JUSTIFICATIVA	15
1.2 PROBLEMÁTICA	17
1.3 OBJETIVOS	18
2 REFERENCIAL TEÓRICO	19
2.1 Tuberculose: “A Maior Assassina de Todos os Tempos”	19
2.2 Arqueologia da Doença: “Existe Doença Desde o Momento em que se Existe Vida”	24
2.3 Um olhar sobre o Passado: DNA antigo como Fonte	27
2.4 O Casamento entre duas Ciências: Conhecendo a Bioinformática	30
3 MATERIAIS E MÉTODOS	34
3.1 Conhecendo os Caminhos: Apresentando os Polimorfismos	35
3.3 Aplicando a Bioinformática: Chamando uma Variante	38
4 RESULTADOS E DISCUSSÃO	42
5 CONSIDERAÇÕES FINAIS	66
6 REFERÊNCIAS BIBLIOGRÁFICAS	69

1 INTRODUÇÃO

As doenças infecciosas foram, e ainda são, uma das principais causas de mortalidade humana e, portanto, representam uma forte pressão seletiva na evolução humana (Murci, 2019). Dentre essas doenças, uma das mais danosas à espécie humana é a Tuberculose (TB). Considerada a mais mortal da era comum, com mais de um bilhão de mortes nos últimos 2.000 anos (Kerner *et al.*, 2021), recebendo assim a alcunha de ‘Maior Assassina de Todos os Tempos’ (Zimmerman, Barry; Zimmerman, David, 2022), a TB em humanos é uma doença específica causada pela infecção por um membro do complexo *Mycobacterium tuberculosis* – MTBC – (Baker *et al.*, 2015).

Estando presente em toda história evolutiva da espécie humana moderna, *Homo sapiens*, a TB vem sendo estudada, e discutida, há muitos anos. Um grande marco de seus estudos foi a descoberta da doença, e seu patógeno, em africanos e nativos americanos antes da colonização europeia surpreendendo os pesquisadores que acreditavam que esses grupos só teriam desenvolvido a TB após o contato com os colonos, como havia acontecido com a Varíola (Baker *et al.*, 2015; Luna *et al.*, 2020).

Os novos estudos trazem à luz evidências genéticas (relacionadas ao DNA e suas estruturas) e bioarqueológicas (associadas a remanescentes ósseos humanos) que corroboram a hipótese da existência da TB no continente Americano antes da colonização europeia, bem como sugerem uma possível presença de outras cepas da TB nas Américas muito antes do período de contato (Joseph *et al.*, 2023). Entretanto, até o presente momento, a sua prevalência no continente, principalmente na América do Sul, não é bem compreendida.

De uma perspectiva bioarqueológica, a identificação da lesão óssea da TB em remanescentes ósseos antigos pode fornecer inúmeras informações, desde rotas de disseminação até associação com outros indicadores esqueléticos de estresse (Luna *et al.*, 2020). Entretanto, a maioria das doenças infecciosas afeta os tecidos moles do corpo. No caso da TB, apenas uma porção – 3% a 5% dos pacientes – desenvolve lesões ósseas (Pálfi *et al.*, 2015).

Nesse cenário, um avanço memorável para os estudos de Paleopatologia, foram as notáveis contribuições dos métodos genéticos aplicados à caracterização e detecção de doenças infecciosas em espécimes arqueológicos (Bos *et al.*, 2014 *apud* Nelson *et al.*, 2020). Isso porque, as abordagens genéticas desenvolvidas na última década revelaram, em parte,

que diferenças na composição genética do hospedeiro humano podem resultar em alterações na susceptibilidade do indivíduo a fenótipos complexos relacionados ao sistema imunológico (Casanova e Abel, 2018 *apud* Murci, 2019). Nesse contexto, muitos estudos apontaram para o papel da genética do hospedeiro na suscetibilidade, ou resistência, à TB (Berrington e Hawn, 2007 *apud* Pálfi *et al.*, 2015; Kerner *et al.*, 2021).

Destarte, tendo em vista que as lesões esqueléticas consistentes com a TB podem passar despercebidas no registro arqueológico (Berrington e Hawn, 2007 *apud* Kerner *et al.*, 2021), o que limita compreender a prevalência da doença no continente Americano, os avanços nos métodos genéticos aplicados a espécimes arqueológicos, por sua vez, tornam novas abordagens possíveis. Dentre elas, sendo o interesse do presente trabalho, a análise de Polimorfismos de Nucleotídeo Único – SNP ¹(Figura 1), associados à TB no continente Americano antes da colonização, através da perspectiva da Arqueogenética – estudo do passado humano utilizando-se de técnicas da genética molecular (Renfrew e Bahn, 2005).

Figura 1 - Representação de um SNP (Polimorfismo de nucleotídeo único) tipo mais comum de polimorfismo.

...GATCTGCATCGTCGGCAATTGTCATGCGCATC... (Indivíduo A)

...GATCTGCATCGTCGGCGTTGTCATGCGCATC... (Indivíduo B)

Fonte: Santos (2020, p. 69).

1.1 JUSTIFICATIVA

A arqueologia é uma ciência que se debruça sobre a materialidade da cultura humana, seja ela no passado ou no presente, bem como a imaterialidade atrelada a essa materialidade (Reinhard, 2020). Nesse contexto, os campos e possibilidades de pesquisas são infinitas, sendo um deles o da Arqueogenética. Termo cunhado por Colin Renfrew, a Arqueogenética pode ser entendida como o estudo do passado humano utilizando-se de técnicas da genética molecular (Renfrew e Bahn, 2005); sendo notável, segundo os mesmo autores, que uma das fontes mais importantes sobre o ser humano esteja dentro dele mesmo.

Para além de seu próprio DNA, que é uma fonte inestimável de conhecimento sobre o passado, o ser humano também carrega consigo pequenas criaturas – patógenos – que

¹ Forma muito comum de alteração genética em sequências de DNA. Ocorre quando uma única base nucleotídica é substituída ao longo de uma sequência de DNA (Reis *et al.*, 2016).

exerceram, ao longo da história, forte pressão seletiva² sobre a evolução da espécie humana. Dentre esses patógenos, encontra-se o causador da temida Tísica Pulmonar, Tuberculose, em humanos, ou seja, o *Mycobacterium tuberculosis* (M.tb). Apesar de ser causada por uma bactéria, muitos estudos são voltados para as alterações genéticas que ocorrem no DNA humano tornando-o mais suscetível ou resistente à infecção.

Nesse sentido, a análise de DNA antigo, seja do patógeno ou do próprio ser humano, oferece informações importantes sobre uma das doenças mais mortais da história humana. Dentre essas pesquisas, destacam-se, para o prezado trabalho, aquelas que apontam para a existência da Tuberculose no continente americano antes da colonização. Isso porque, por muito tempo acreditou-se que a TB havia sido introduzida nas Américas por meio da colonização europeia, assim como aconteceu com a gripe e com a varíola. Todavia, os dados genéticos e bioarqueológicos contrapõem esse cenário apresentando novas possibilidades.

Dentre esses trabalhos, destacam-se a pesquisa de Luna *et al.* (2020), que por meio da análise bioarqueológica de um remanescente humano argentino conseguiu encontrar marcas da TB na Argentina há – pelo menos – 1300 anos antes do presente. Existe também o cenário das múmias peruanas, apresentadas por Krause e Trappe (2019) em seu livro ‘A jornada dos nossos genes: Uma história da Humanidade’, que por meio de análises arqueogenéticas encontraram nas múmias o DNA de uma cepa desconhecida de M.tb, chamada por eles de Cepa pré-colombiana. Por fim, cita-se o trabalho de Joseph *et al.* (2023) que, ao analisar DNA antigo de populações andinas do Equador, encontrou uma forte seleção positiva em uma alteração genética associada a uma maior suscetibilidade à Tuberculose. Nesse contexto, percebeu-se que a infecção poderia existir na região há, pelo menos, 3000 anos antes do presente, exercendo uma forte pressão sobre as populações agrícolas que ali viviam.

À vista disso, a pesquisa justifica-se como uma contribuição aos trabalhos na abordagem da Arqueogenética com um olhar voltado para a Arqueologia da Doença. Isso porque, no que tange os estudos da TB no continente Americano antes da colonização, principalmente no Brasil, existe uma carência de dados sobre o assunto que coloca limites na compreensão da prevalência e distribuição da doença no continente. Dessa forma, pretende-se aumentar o número de trabalhos que se apropriem da abordagem do uso do DNA antigo para investigações de grupos humanos pretéritos, já que a mesma não é de comum aplicação na ciência arqueológica nacional, o que se reflete na falta de trabalhos voltados para o tema no Brasil. Junto a isso, de forma continental, justifica-se como uma tentativa de acrescentar

² A pressão seletiva pode ser entendida como a influência do ambiente sobre a sobrevivência e a reprodução dos organismos, atuando sobre as variações fenotípicas nos indivíduos de uma população, favorecendo aquelas que possuem características adaptativas que aumentam sua aptidão reprodutiva (Quinelato, 2024)

novas informações ao estudo da Tuberculose e sua prevalência nas Américas antes da colonização.

Isso porque, caso se indentifique as alterações associadas a infecção e essas se apresentem em distintos indivíduos de um mesmo sítio arqueológico em um mesmo intervalo de tempo, seria possível inferir novos aspectos da Tuberculose nas populações originárias antes da colonização, bem como inferir novos questionamentos sobre a sua origem nessa parte do mundo.

Destarte, outra contribuição comum a trabalhos desse gênero, é o auxílio à comunidade médica, para além da científica, a manter o ritmo constante no chamado ‘corrida armamentista biológica’ (Krause e Trappe, 2019). Isso porque, o estudo das doenças e dos patógenos através do DNA antigo, de acordo com os mesmos autores, nos permite iniciar a compreensão de como eles evoluíram ao longo do tempo e como o DNA humano se adaptou, ou os neutralizou, ao longo dos séculos (Krause e Trappe, 2019).

1.2 PROBLEMÁTICA

Tendo em vista que as pesquisas atuais – através de análises de lesões ósseas – apontam para a presença da Tuberculose nas Américas em um período anterior ao contato, questionou-se se esses mesmo grupos já possuíam os polimorfismos genéticos que na atualidade são apontados, e discutidos, como associados à suscetibilidade, ou resistência, dos hospedeiros à infecção da TB.

Isso porque, considerando a porcentagem de pacientes que desenvolvem a lesão óssea associada à TB e como essas marcas podem passar despercebidas no registro arqueológico, os dados genéticos podem sugerir a presença da doença mesmo na ausência de manifestações ósseas; bem como pode trazer novas possibilidades de interpretação, e debate, quanto à prevalência, e distribuição, da infecção no continente Americano.

Logo, a pergunta guia da presente pesquisa é: os antigos grupos humanos, que povoaram a América, já apresentavam os marcadores genéticos que, na atualidade, são apontados como causadores de suscetibilidade ou resistência à Tuberculose? Se o fizeram:

- Quais seriam esses marcadores?
- Qual sua frequência?
- Qual sua dispersão no continente?

1.3 OBJETIVOS

Com o intuito de compreender a prevalência, e distribuição, dos dados genéticos associados à TB nas Américas, o trabalho objetiva, de forma geral, apontar marcadores genéticos que estão associados à suscetibilidade, ou à resistência, à TB no continente Americano antes da chegada dos colonizadores europeus. Para atingir tal propósito, alguns objetivos específicos foram delineados. São eles:

- Levantar e identificar os polimorfismos genéticos que na atualidade são apontados como associados a uma maior suscetibilidade ou resistência à TB;
- Buscar nas amostras analisadas os polimorfismos levantados;
- Apresentar visualmente como esses polimorfismos se distribuem no continente Americano.

2 REFERENCIAL TEÓRICO

No presente capítulo, dividido em 4 seções, apresentaremos os conceitos norteadores da presente pesquisa, suas aplicações, definições e discussões atuais. Tais informações foram advindas de uma pesquisa bibliográfica em cima de livros e artigos de distintos bancos de dados, tais como: National Institutes of Health (<https://www.nih.gov/>), National Center for Biotechnology (<https://www.ncbi.nlm.nih.gov/>), PubMed (<https://pubmed.ncbi.nlm.nih.gov/>), ScienceDirect (<https://www.sciencedirect.com/>), Scielo (<https://scielo.org/en/>) e Scientific reports (<https://www.nature.com/srep/>).

2.1 Tuberculose: “A Maior Assassina de Todos os Tempos”

No filme O Senhor dos Anéis e a Sociedade do Anel (2001), dirigido por Peter Jackson, a elfa Galadriel pontua, em uma de suas célebres frases, que: “até mesmo a menor das criaturas pode mudar o rumo do mundo”. Se pararmos para pensar, de forma semelhante são os microorganismos – bactérias, vírus, fungos e protozoários – que, apesar do tamanho diminuto de alguns, sempre estiveram conosco acompanhando e mudando a história humana, talvez, mais do que qualquer outro fator (Zimmerman, Barry; Zimmerman, David, 2022).

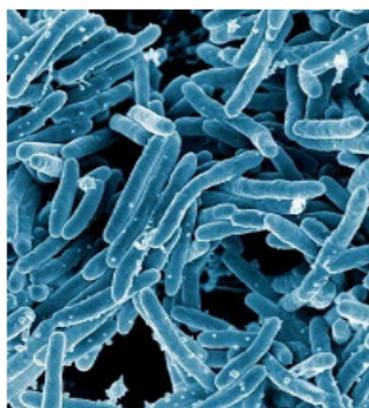
As doenças, como pontuam os irmãos Zimmerman (2022), existem, praticamente, desde o momento em que se existe vida e, ao longo da história, foram as responsáveis por dizimar grandes populações e escrever muitos percursos históricos. Dentre essas doenças se encontra a Tuberculose (TB), a patologia com a alcunha, de acordo com o livro Medicina Macabra 3 (Zimmerman, Barry; Zimmerman, David, 2022), de “a maior assassina de todos os tempos”.

Sendo considerada, antes da pandemia de COVID-19, a *causa mortis* mais comum entre todos os agentes infecciosos isolados (Chen *et al.*, 2023), a TB é uma das doenças infecciosas que mais mata pessoas portadoras do Vírus da Imunodeficiência Humana (HIV) no mundo (Who, 2016, *apud* Souza *et al.*, 2022), sendo também a responsável por cerca de 1,5 milhões de mortes anualmente (Who, 2016, *apud* Kerner *et al.*, 2021). Inserida no complexo de doenças infectocontagiosas e crônicas (Pedro, 2015, *apud* Souza *et al.*, 2022), a TB é causada, principalmente, pela bactéria *Mycobacterium tuberculosis* (M.tb) e atua afetando, sobretudo, os pulmões.

Todavia, apenas a infecção pelo patógeno não é suficiente para causar um processo patológico. Cerca de 90% dos indivíduos que apresentam a infecção latente – infectada com o patógeno, mas sem desenvolver sintomas – por TB nunca desenvolvem a mesma (OMS, 2019, *apud* Souza *et al.*, 2022). Isso sugere que outros fatores, tais quais: estado imunológico

do indivíduo, interações genéticas e fatores ambientais interferem na evolução, suscetibilidade e resistência à infecção (Souza *et al.*, 2022), caracterizando assim uma doença multifatorial e silenciosa. Isto se dá porque, no que tange os fatores genéticos, eles também podem flutuar em uma população ou em um único indivíduo, ou seja, em virtude de outros fatores – incluindo outros patógenos – um ser humano, ou um grupo, que antes era resistente à infecção torna-se mais suscetível ou resistente a mesma (Furin *et al.*, 2019). Tal situação pode ser exemplificada em pessoas com o vírus do HIV, que coincidentemente são um dos mais afetados pela TB (Zimmerman, Barry; Zimmerman, David, 2022), o que reflete o poder de um vírus e de uma bactéria, nesse caso, de um pequeno bacilo que coexiste com a espécie humana à muito tempo.

Figura 2 - Micrografia eletrônica de Varredura do *Mycobacterium tuberculosis*



Fonte: National Institute of Allergy and Infectious Diseases (NIAID, 2024).

Essa pequena bactéria, conhecida também como bacilo de Koch em virtude do descobridor do mesmo – Robert Koch em 1882, é um dos muitos membros do complexo *Mycobacterium tuberculosis* (MTBC)³, que possui uma gama de hospedeiros, incluindo, os humanos e outros mamíferos (Brosch *et al.*, 2002). Em consequência disso, a hipótese amplamente aceita para o ‘surgimento’ da TB em seres humanos, é que a mesma teria surgido no neolítico – graças à domesticação dos animais, tratando-se de uma transmissão zoonótica. Nesse contexto, o patógeno humano teria derivado do patógeno do gado – *Mycobacterium bovis* – por contaminação durante o período já citado (Baker *et al.*, 2015).

³ A Tuberculose pode ser causada por qualquer um dos sete membros, conhecidos, que compõem o complexo *Mycobacterium tuberculosis* (MTBC). Dentro desse complexo, existe a espécie *Mycobacterium tuberculosis* (M.tb) que é a principal causadora da TB em humanos. Essa espécie, em função de mutações, divide-se em distintas linhagens específicas de determinadas regiões.

Entretanto, novos cenários surgem. Isso porque, apesar do complexo MTBC apresentar inúmeras espécies, a TB humana é específica, ou seja, a bactéria associada ao ser humano não possui um reservatório animal, ou ambiental, conhecido até o presente momento (Comas *et al.*, 2014). Nesse cenário, com o auxílio de novas abordagens e análises de DNA antigo e biomarcadores, as novas hipóteses apontam para o surgimento da TB Humana há, pelo menos, 70.000 anos no continente Africano, tendo acompanhado assim as migrações dos humanos anatomicamente modernos para fora do continente (Comas *et al.*, 2014).

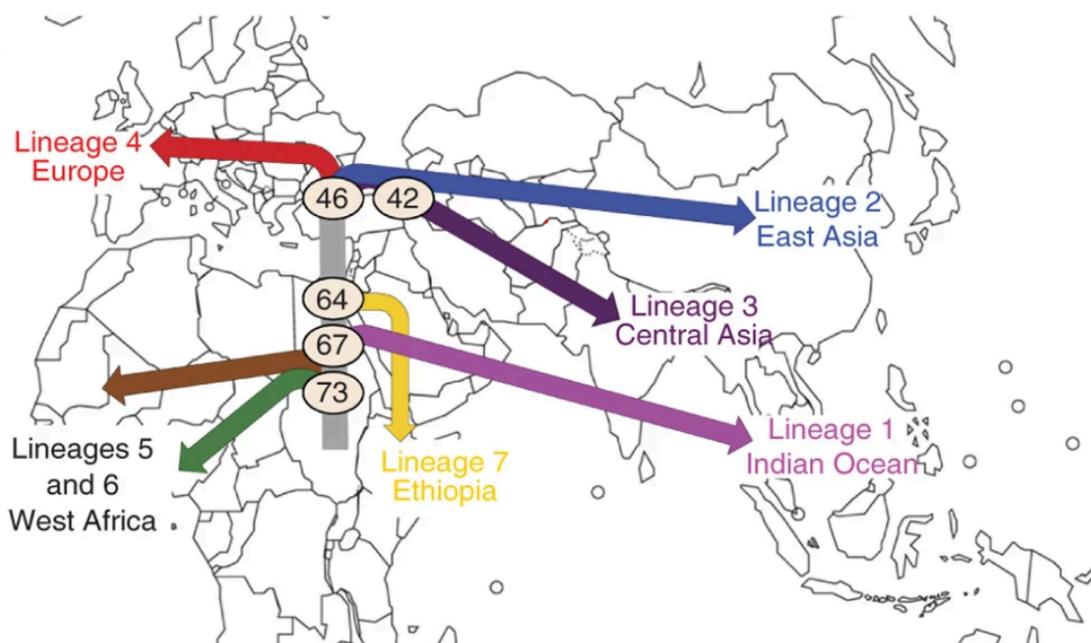
Nessa perspectiva, vem sendo discutida, e apresentada, a possibilidade do patógeno associado ao humano (*Mycobacterium tuberculosis*, M.tb) ser na verdade um descendente direto de uma cepa mais antiga que existiu antes da divisão das linhagens do *Mycobacterium africanum* (Baker *et al.*, 2015). Isso porque, até o presente momento, sabe-se que existem sete distintas linhagens de cepas de TB Humana distribuídas (Tabela 1 e Figura 3), e específicas, em diferentes regiões do globo (Comas *et al.*, 2014).

Tabela 1 - Linhagens das cepas de Tuberculose em humanos

LINHAGENS	LOCAIS ENDÊMICOS	DATAÇÃO APROXIMADA
L1	Regiões do Oceano Índico e parte da África	Aproximadamente 67.000 anos
L2	Leste da Ásia	Aproximadamente 46.000 anos
L3	Ásia Central	Aproximadamente 42.000 anos
L4	Europa e América	Aproximadamente 46.000 anos
L5	Oeste da África	Aproximadamente 73.000 anos
L6	Oeste da África	Aproximadamente 73.000 anos
L7	Etiópia	Aproximadamente 64.000 anos

Fonte: Autoria própria (2024) segundo os dados de Comas (2014).

Figura 3 - Representação cartográfica da saída da África das linhagens da Tuberculose Humana.



Fonte: Comas (2014, p. 5).

Isso sugere que o bacilo da TB já poderia ter sido um patógeno humano antes da domesticação no neolítico (Baker *et al.*, 2015), tendo ocorrido uma provável expansão do complexo MTBC, e da infecção, em consequência do aumento da densidade populacional que se seguiu ao estabelecimento dos primeiros assentamentos humanos durante o período (Comas *et al.*, 2014; Krause e Trappe, 2019). Nesse contexto, com as fortes indicações da antiguidade da infecção e sua relação com o ser humano, existem indícios que apontam para uma possível co-evolução entre o ser humano e o patógeno (Comas *et al.*, 2014), ou, pelo menos, uma intrínseca relação entre hospedeiro e patógeno (Alves, Prado e Takenami, 2022).

Nesse cenário, ao voltarmos nossos olhos para o continente Americano, muitas lacunas ainda existem com relação à TB – mesmo que o continente esteja representando 3% da carga mundial da infecção (Ministério da Saúde, 2017), com o Brasil sendo considerado o país com o maior número de casos no continente (Alves, Prado e Takenami, 2022). Na atualidade, análises filogenéticas sugerem que a linhagem filogenética L4, predominante na Europa, deixou o continente europeu durante o período de colonização, continuando a ser linhagem dominante nas Américas nos dias de hoje (Joseph *et al.*, 2023). Entretanto, existe um debate considerável sobre a presença da doença no continente pré-contato.

Isso em razão de, embora a prevalência e a distribuição geográfica da infecção não seja totalmente compreendida no continente, as evidências genéticas e bioarqueológicas sugerem a presença de TB em humanos há pelo menos 1.400 anos antes do presente nas

antigas Américas (Luna *et al.*, 2020; Joseph *et al.*, 2023). Junto a isso, também apontam para uma possível adaptação à infecção em partes dos Andes – Equador – há pelo menos 3.000 anos antes do presente (Joseph *et al.*, 2023). Outro forte indício dessa presença foi as múmias encontradas no Peru, de 2.000 mil anos, que apresentavam evidências genéticas e ósseas de infecção por Tuberculose (Krause e Trappe, 2019). Em razão disso, os trabalhos sugerem a possibilidade de uma linhagem zoonótica, anterior ao contato, ter existido nas Américas (Joseph *et al.*, 2023), sendo essa cepa, provavelmente oriunda da África, tendo chegado no continente pelo mar (Krause e Trappe, 2019).

Ademais, apesar de serem necessários novos estudos e abordagens para compreender a antiguidade, e surgimento, da TB nas Américas e no restante do globo, estudos apontam para uma coevolução entre patógeno e hospedeiro, sendo este um dado já conhecido. Nesse sentido, tendo em vista que para cenários de coevolução muito tempo é necessário, pode-se sugerir uma longa coexistência do *H.sapiens* com o *M.tuberculosis*. As hipóteses para se compreender essa relação são inúmeras – desde adaptação a grandes altitudes (Wu *et al.*, 2020; Joseph *et al.*, 2023) até um baixo impacto da seleção natural no MTBC (Comas *et al.*, 2014). Logo, pode-se pontuar que as pressões seletivas impostas pelos patógenos – nesse caso o da TB – foram fundamentais durante a evolução humana (Kerner *et al.*, 2021), mesmo que ainda não se compreenda como e porquê.

Hodiernamente, o que se tem são cada vez mais estudos voltados para a relação entre a TB e os genes. De acordo com Nascimento (2013 *apud* Souza *et al.*, 2022), há no mínimo 100 anos se discute e se estuda a influência da genética do indivíduo na suscetibilidade ou resistência à TB. A maior parte desses trabalhos busca pontuar alguns marcadores genéticos/polimorfismos – variações na sequência de DNA – que desempenham um papel na resposta imune dada ao processo de infecção. Grande parte desses estudos é realizada por profissionais da área da saúde e envolve a análise de material genético proveniente de populações humanas do presente.

Todavia, existem exceções, ainda dentro da área da saúde, que estendem essas pesquisas para análises de DNA antigo, como os de Kerner *et al.* (2021), Sabin *et al.* (2020) e Nogueira e Zorzetto (2013). No continente europeu, onde a TB é considerada uma das doenças infecciosas mais mortais já documentadas (Kerner *et al.*, 2021), destacam-se os trabalhos realizados em amostras de DNA antigo de Sabin *et al.* (2020) e Kerner *et al.* (2021) que debruçam-se sobre a evolução da doença junto à evolução da humanidade na Europa. O primeiro detalha sobre o DNA do patógeno encontrado em um remanescente do século XVII e o segundo argumenta sobre a alta taxa da doença na Europa nos últimos 2.000 mil anos.

No continente asiático destacam-se trabalhos sobre a população Han (Chinesa) e Tibetana. Nesses trabalhos, as análises são feitas em amostras de grupos atuais e comparados entre si, verificando a existência de polimorfismos mais comuns em um grupo do que em outro, mais recorrente em um sexo do que outro e daqueles que são comuns a todos os grupos, incluindo nos trabalhos europeus apresentados acima. Nesse cenário, destacam-se na Ásia as produções de Wu *et al.* (2020), que apresenta os polimorfismos mais recorrentes da Tuberculose em dois estudos de caso independentes na população Han e Tibetana, e os de Liu *et al.* (2020) que versam sobre essas mutações apenas na população Han (chinesa) atual.

Nas Américas, por sua vez, muitos estudos são feitos para se entender a coevolução da doença juntamente com a evolução do *H. sapiens*. Dentre esses trabalhos destaca-se Joseph *et al.* (2023), que se debruça sobre a adaptação aos Andes, encontrando também possíveis indícios de adaptação à TB, e o estudo de Luna *et al.* (2020) que, com um viés bioarqueológico, debruça-se sobre um remanescente ósseo argentino com marcas de TB. Além disso, para o Brasil, pode-se citar o de Souza *et al.* (2022) e o de Nogueira e Zorzetto (2013), que se debruçam sobre a TB – dentre outras doenças – no Brasil do Segundo Reinado sob o governo de Dom Pedro II (1840-1889).

Esses e outros estudos corroboram o que foi pesquisado por Berrington e Hawn (2007 *apud* Kerner *et al.*, 2021), que sugere que os fatores genéticos do hospedeiro influenciam a sua suscetibilidade ou resistência à TB. Logo, em um cenário em que a TB (re)emerge como uma ameaça mundial – com inúmeros focos de epidemia (Furin *et al.*, 2019), a compreensão sobre a mesma e seus inúmeros fatores, ou pelo menos a tentativa de compreender o universo da infecção, torna-se essencial para compreender uma parte do passado humano, seus reflexos no presente e, quem sabe, servir de auxílio no futuro.

2.2 Arqueologia da Doença: “Existe Doença Desde o Momento em que se Existe Vida”

As doenças são uma parte inevitável da vida. Tão inevitáveis quanto a própria morte (Brown *et al.*, 1996 *apud* Roberts e Manchester, 2005). De acordo com Teixeira (2020), o termo doença está associado ao conceito biológico de adaptação que pode ser entendido como a capacidade de ser sensível às variações do meio e produzir respostas capazes de adaptar-se. Logo, a doença seria a ausência de adaptação ao meio – físico, social e psíquico – no qual o indivíduo se sente mal e apresenta alterações orgânicas visíveis (Teixeira, 2020).

Seu ‘surgimento’, bem como o da saúde, ocorre desde o momento em que existe vida (Zimmerman, Barry; Zirmmerman, David, 2022), estando ambos intimamente ligados ao meio em que o indivíduo, ou a população, está inserida. Dentro dessa dinâmica com o ambiente, está o princípio de homeostasia – condição de estabilidade que o organismo precisa para a realização de suas funções de forma adequada para manter o equilíbrio do corpo (Humberto Maturana *apud* Teixeira, 2020). Destarte, como pontuado por Teixeira (2020), toda atividade contrária ao equilíbrio do corpo, ou seja, ao princípio de homeostasia, é considerado uma doença, ou melhor, um processo patológico.

Nessa perspectiva, Patologia é o estudo de toda e qualquer alteração no organismo humano, com o objetivo de explicar a origem e os sintomas dessas alterações que coloquialmente chamamos de doença, ou seja, é o estudo das doenças (Teixeira, 2020). Nesse contexto, uma área em crescente desenvolvimento é o da Arqueologia da Doença, esta que tem íntima relação com os estudos no campo da Paleopatologia.

Apesar de distintas, alguns pesquisadores tendem a tratá-las como sinônimos. Entretanto, para a presente pesquisa, elas serão tratadas como dois campos de estudos diferentes, mas ligados, uma vez que os estudos no campo da Arqueologia da Doença surgem dos avanços no campo da Paleopatologia. A Paleopatologia, nesse cenário, pode ser compreendida como o estudo da origem, evolução e progresso das doenças ao longo do tempo (Roberts e Manchester, 2005), sejam essas patologias em humanos ou não.

A Arqueologia da Doença, por sua vez, pode ser entendida como o estudo da evolução das doenças através do tempo analisando o modo como se manifestaram nos seres humanos e como os seres humanos se adaptaram às mudanças de seu ambiente (Roberts e Manchester, 2005; Japiassu, 1976; Larsen, 1997; Lima, 1994; Ortner, 2003; Roberts; Manchester, 2007; Souza, 2009 *apud* Silva, 2015). Logo, é o ramo da ciência que se debruça sobre as evidências iniciais – dados biológicos e culturais/bioarqueológicos – do estado de saúde de nossos ancestrais (Roberts e Manchester, 2005), sendo as suas principais fontes os remanescentes ósseos humanos e as iconografias; e os seus primeiros métodos de estudo a análise macroscópica ou a observação visual.

A título de exemplo, caso a distinção não tenha ficado clara, peguemos o cenário da Peste Negra na Idade Média na Europa. A Paleopatologia irá debruçar-se em como a Peste Negra teria afetado os indivíduos fisicamente, biologicamente, e as sequelas deixadas para as gerações futuras. A Arqueologia da Doença, por outro lado, ficaria responsável por estudar como a doença afetou culturalmente/socialmente as populações europeias e como elas responderam a isso. Dessa forma, a Paleopatologia teria como fonte de estudo os ossos,

textos escritos e o DNA (se conseguido), e a Arqueologia da Doença, para além dos vestígios bioarqueológicos de ossos e DNA, também iria se debruçar sobre a materialidade usada para lidar com a doença, tal como as máscaras e as vestimentas dos médicos da época.

Entretanto, apesar desta distinção, a Paleopatologia, como fonte primária de evidências para a investigação da saúde no passado (Silva, 2015), tem fundamentado a Arqueologia da Doença. De acordo com Buikstra e Roberts (2005), o desenvolvimento da Paleopatologia contou com 4 fases, sendo a quarta fase a que marca o início da distinção entre os dois campos, pois é nela que começam-se as análises populacionais, culturais/bioarqueológicas e a introdução de métodos biomoleculares – de DNA antigo e biomarcadores – como caminhos de análise (Roberts e Manchester, 2005). Sendo esse último considerado um avanço notável nos estudos de Paleopatologia, e Arqueologia da Doença, em razão das consideráveis contribuições dos métodos genéticos quando aplicados à detecção e caracterização de doenças infecciosas em espécimes arqueológicos; que agora se estendem até o período Neolítico (Bos *et al.*, 2014 *apud* Nelson, 2020).

Nessa perspectiva, tendo em vista que o estudo das doenças de um grupo auxiliam a reconstruir e compreender aspectos do cotidiano desse grupo (Bruikstra; Cook, 1980 *apud* Silva, 2015), a inserção de novas técnicas e metodologias conduziu a pesquisas que buscam compreender cada vez mais as populações, e não apenas o indivíduo, na interpretação das alterações dos padrões das doenças (Silva, 2015). Logo, na atualidade, a Paleopatologia vem trabalhando sob um viés interpretativo que é aplicado nos mais diversos campos da Arqueologia (Silva, 2015), sendo um deles, o da Arqueologia da Doença.

Contudo, apesar do amplo desenvolvimento e da gama de disciplinas que se debruçam sobre essas ciências, alguns pontos precisam ser recordados ao tratar de Arqueologia da Doença e Paleopatologia. Inicialmente, é recomendado evitar realizar inferências baseando-se em uma única fonte – principalmente as iconográficas – pois as artes e textos retratam, em sua maioria, patologias excepcionais não refletindo assim aquelas de caráter cotidiano (Roberts e Manchester, 2005).

Junto a isso, ao se debruçar sobre os remanescentes ósseos em busca de lesões, é preciso saber que doenças distintas podem deixar marcas similares nos ossos, que algumas doenças que existem hoje podem não ter existido no passado – da mesma forma que no passado podem existir doenças que hoje não existem mais – e que, principalmente no registro arqueológico, nem sempre o remanescente estará com as lesões que são características do registro clínico atual sobre a patologia (Roberts e Manchester, 2005).

Além disso, cabe pontuar a importância de se estudar – e interpretar – os remanescentes ósseos humanos dentro do contexto de sua cultura, ou seja, a sua cultura material, o período de tempo, a área geográfica e outros fatores culturais que auxiliam enormemente – para além dos ossos – na compreensão da história da doença (Roberts e Manchester, 2005). Isso porque, como dito anteriormente, o conceito de doença – assim como o de saúde – está intimamente ligado à concepção de adaptação (Teixeira, 2020).

Logo, torna-se fundamental a compreensão do ambiente, ou melhor, do meio – seja ele social, físico, ambiental ou psíquico- em que o indivíduo, ou a população, está inserida, bem como a interação entre essas partes, para se entender o desenvolvimento da doença (Roberts e Manchester, 2005; Teixeira, 2020), afinal a dieta, o clima e o ambiente eram diferentes no passado e podem ter predisposto grupos a distintas patologias (Roberts e Manchester, 2005).

Destarte, torna-se nítida a importância dos fatores culturais e ambientais na compreensão da evolução de patologias, bem como as notáveis portas que a introdução da análise de DNA antigo – e biomarcadores- abriu para os avanços da Arqueologia da Doença e da Paleopatologia; com ambos os campos também contribuindo para os conhecimentos na medicina moderna (Roberts e Manchester, 2005; Nelson *et al.*, 2020; Teixeira, 2020). Na próxima seção será apresentado como o estudo de DNA antigo pode contribuir para a Ciência Arqueológica e outras ciências.

2.3 Um olhar sobre o Passado: DNA antigo como Fonte

A Arqueologia é uma ciência que se debruça sobre a materialidade da cultura humana, seja ela no passado ou no presente, bem como a imaterialidade atrelada a essa materialidade (Reinhard, 2020). Nesse contexto, inúmeros conceitos e abordagens foram incorporados pela ciência arqueológica, sendo um deles o da Arqueogenética. Termo cunhado por Colin Renfrew, a Arqueogenética pode ser entendida como o estudo do passado humano utilizando-se de técnicas da genética molecular (Renfrew e Bahn; 2005), sendo notável – segundo os mesmo autores – que uma das fontes mais importantes sobre o passado humano esteja dentro de nós.

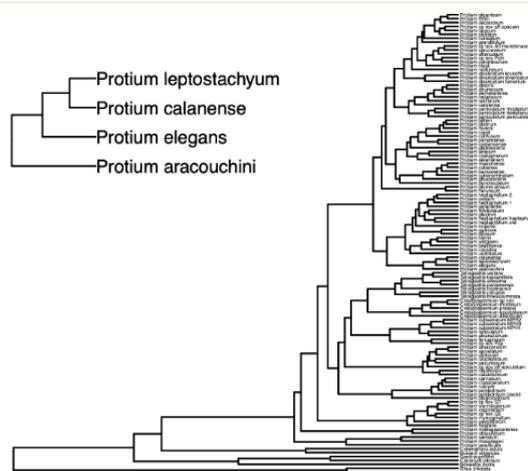
A priori, é um campo em franca e rápida expansão, com alguns de seus conceitos ainda sendo trabalhados. Seu desenvolvimento está associado com os avanços genéticos – como os estudos dos grupos sanguíneos humanos, e os estudos de bioquímica – sobre enzimas, biomoléculas, proteínas e outras, que juntos formam o campo conhecido como

“Genética Clássica” (Renfrew e Bahn; 2005). Todavia, apesar dos avanços nesse campo, foi apenas em 1953, com a decifração do código genético do DNA, que se abriram as possibilidades de estudos e métodos baseados no sequenciamento do DNA. Isso é, a análise e determinação das bases nitrogenadas – Adenina (A), Citosina (C), Guanina (G) e Timina (T) – que constituem a dupla hélice do material genético e, com isso, a informação genética básica dos indivíduos (Renfrew e Bahn; 2005).

Em razão disso, os primeiros estudos na área começaram a surgir na década de 1990, sendo em 1994 publicada a primeira síntese sobre o assunto: *The History and Geography of Humans Genes*, por Luca Cavalli-Sforza e colegas (Renfrew e Bahn; 2005), com o objetivo de apresentar – e compreender – a distribuição geográfica dos genes das populações europeias e as possíveis rotas migratórias desses grupos. Além disso, surgem também os métodos de linhagem, voltados para a busca de um ancestral comum nos estudos de evolução humana, e as rotas de dispersão dos primeiros grupos humanos. Nessa perspectiva, as análises Arqueogenéticas se davam em função do estudo do DNA mitocondrial – obtido apenas pela linhagem materna, ou seja, herdado pela mãe – e do cromossomo Y – que não sofre recombinação e é transmitido pela linhagem paterna, ou seja, transmitido de pai para filho (Renfrew e Bahn; 2005).

Ademais, os avanços Arqueogenéticos, permitiram também a elaboração de árvores filogenéticas – a partir de dendrogramas – que transmitisse de forma visual as redes evolutivas e, também, permitissem uma datação para as linhagens através do “relógio molecular” – evolução de um gene em uma taxa, aproximadamente, constante em distintas espécies (Renfrew e Bahn; 2005).

Figura 4 - Exemplo de um dendrograma (árvore filogenética) de uma planta do gênero *Protium*.



Fonte: Perdiz (2022).

Logo, torna-se nítido que essa abordagem permite estabelecer um claro esboço para a origem da nossa espécie – seja em âmbito espacial ou cronológico – e suas dispersões por todo o globo (Renfrew e Bahn; 2005). Tais trabalhos, na atualidade, continuam sendo reproduzidos e melhorados, associando-se às melhorias tecnológicas e metodológicas.

“Para a arqueogenética, as partes ‘inúteis’ do genoma são valiosas, porque elas nos permitem estabelecer o que chamamos de relógio molecular. {...}. Quanto mais tempo tiver passado, mais novas variantes terão se acumulado no DNA ou alterado sua frequência”.
(Krause e Trappe, 2019, p. 25).

Um grande marco para a Arqueogenética, foi o recebimento do Prêmio Nobel em Fisiologia ou Medicina pelo sueco Svante Paabo, no ano de 2022, por sua pesquisa voltada para o desenvolvimento, e sequenciamento, de técnicas que permitiram analisar, e sequenciar, o DNA antigo do Homem Neandertal (Strauss, 2022). Desde 2010 o Instituto Max Planck de Antropologia Evolutiva (Leipzig, Alemanha) faz história na ciência ao se debruçar sobre análise de DNA antigo, expandindo-a para para os mais diversos contextos cronológicos e geográficos (Strauss, 2022). Todavia, um dos primeiros trabalhos, que abriu caminho para estudos frutíferos no campo de antigos genomas humanos, foi lançado em 2010 com a publicação de um genoma quase completo advindo de um tufo de cabelo preservado, e escavado, no permafrost em Qeqertasussuk, Groenlândia (Rasmussen *et al.*, 2010). No Brasil, por sua vez, os primeiros trabalhos voltados para essa nova área do conhecimento são datados do ano de 2018 (Strauss, 2022).

Como exemplos, no cenário Americano, têm-se a tese de doutorado em Arqueologia de André Luiz Campelo dos Santos (Santos, 2020), que aborda a paleomigração no continente americano por uma perspectiva Arqueogenética. Em sua tese, foi realizada a análise de dados genéticos de mais de 200 indivíduos antigos cujos remanescentes foram evidenciados em sítios arqueológicos localizados em diversos países das Américas. Estes genomas antigos foram então comparados com um genoma de referência humano (Santos, 2020). Para além de seu trabalho, há também trabalhos de conclusão de curso (TCC) na graduação em Arqueologia da Universidade Federal de Pernambuco (UFPE), voltadas para uma abordagem da Arqueogenética; um deles voltado para uma análise comparativa das linhagens mitocondriais na América do Sul (Leal, 2021) e outra voltada para a identificação de ancestralidade Neandertal em indivíduos antigos do Brasil (Fernandes, 2021).

Outro exemplo, já citado nessa pesquisa, é o trabalho de Sophie K. Joseph *et al.* (Joseph *et al.*, 2023) que se debruça sobre a adaptação de povos andinos no Equador a elevadas altitudes, acabando por identificar regiões no genoma – com forte pressão seletiva positiva – envolvidas com a resposta imunitária associada a Tuberculose em povos indígenas do Equador há, pelo menos, 3000 anos (Joseph *et al.*, 2023). Esse e outros trabalhos como os de Kerner *et al.* (2021) e Wu *et al.* (2020) são exemplos de como a abordagem da Arqueogenética pode ser aplicada na identificação de marcadores genéticos de patologias nos estudos de Paleopatologia e Arqueologia, sendo os marcadores genéticos, ou polimorfismos, variações na sequência do DNA (Figura 1).

Nesse contexto, cria-se uma interdisciplinaridade com a área da Arqueologia da Doença – que, como apresentado anteriormente, é uma ciência que conta com uma gama de disciplinas (Roberts e Manchester, 2005) – e com a área da Saúde, que já utiliza de ferramentas como a bioinformática – união da biologia com as áreas da ciência da computação – para a análise, interpretação e processamento de dados biológicos. Entretanto, apesar dos crescentes avanços e melhorias nas análises de DNA antigo, cabe salientar que a extração da molécula de DNA antigo é algo complicado. Isso porque, o efeito do tempo no DNA é severo, muitas vezes, reduzindo a molécula a poucos pares de base, pequenos fragmentos de fita simples e universo de DNA contaminante (Strauss, 2022).

Logo, em virtude do severo dano que o tempo pode causar ao DNA, trabalhar em cima de amostras de DNA antigo requer uma série de processos e etapas que envolvem, muitas vezes, algoritmos complexos que necessitam de softwares especializados (Danecek *et al.*, 2011). Em razão disso, a Arqueogenética e outras ciências que se debruçam sobre o DNA, utilizam em suas análises o fruto do casamento perfeito entre a Biologia e a Informática: a Bioinformática, esta que será apresentada e discutida, brevemente, a seguir.

2.4 O Casamento entre duas Ciências: Conhecendo a Bioinformática

Nos filmes de Jurassic Park tem-se a representação sutil do fruto do casamento entre duas grandes ciências: Biologia e Informática, que ficou conhecido como Bioinformática. Podendo ser conceituada como a aplicação de ferramentas de computação e análise à captura e interpretação de dados biológicos (Bayat, 2002), a Bioinformática é um campo interdisciplinar em rápido crescimento, essencial para o gerenciamento de dados na biologia e medicina moderna (Bayat, 2002).

O termo, e a utilização, da Bioinformática remonta à década de 1950, com os primeiros sequenciamentos – e análises – de aminoácidos e proteínas (Minardi, 2023). Todavia, o termo só se torna popular com o projeto de sequenciamento do genoma humano – Projeto Genoma Humano 2000 – (Minardi, 2023). Em razão disso, na atualidade, esse ramo da ciência tornou-se bastante conhecido quando associado aos estudos genômicos (DNA). Uma vez que, o Projeto Genoma Humano, particularmente, foi uma das principais conquistas no campo da Bioinformática (Bayat, 2002).

Entretanto, apesar de, desde a sua origem, estar mais fortemente associada à área molecular, a Bioinformática também se debruça sobre estudos funcionais, biomoleculares, evolutivos, estruturais, produções de genes, interações gênicas, clínicos, dentre outros (Cassanego, 2020; Minardi, 2023). Dessa forma, de acordo com Bayat (2002), as aplicações dessa ciência são:

- Apresentação e análise de vias moleculares para compreender as interações gene-doença;
- Análise e previsão da estrutura e função de genes e proteínas;
- Análise de variações e expressão genética;
- Previsão e detecção de redes de regulação genética;
- Modelagem complexa de dinâmicas e redes reguladoras de genes

Por causa disso, essa ciência acaba por interagir com, e auxiliar, áreas da biologia, da medicina, da genética, da genômica, proteômica, transcriptômica, da engenharia, da física, da computação, da biomedicina, da química e muitas outras. De acordo com Cassanego (2020), algumas das áreas de atuação da Bioinformática são:

- Medicina Molecular;
- Análise Forense;
- Estudos sobre mudanças climáticas;
- Estudos evolutivos e filogenéticos;
- Ciências Veterinárias;
- Terapias de Genes

Logo, trata-se de um campo interdisciplinar em rápido crescimento e adaptação, pois, uma ferramenta de bioinformática – para ser considerada bem sucedida – precisa acompanhar os avanços nas tecnologias de sequenciamento (Danecek *et al.*, 2021).

Todavia, como teria surgido essa interação – que geraria o casamento perfeito – entre a Biologia e a Informática? Bem, do mesmo modo que outras ciências reuniram inúmeros dados ao longo da história, a biologia também reuniu uma grande quantidade. Áreas como proteômica, genômica, transcriptômica e outras, reuniram uma gama de dados que precisavam ser analisados. Isso porque “os dados são como diamantes brutos que precisam ser lapidados para produzir informações que gerem conhecimento” (Cassanego, 2020).

Dessa forma, a Bioinformática ‘nasce’ como a aplicação da computação para uma melhor compreensão e utilização efetiva dos dados biológicos e biomédicos (Minardi, 2023). Em razão disso, houve um crescimento na demanda de ferramentas computacionais para a análise, gerenciamento, integração, modelação e visualização dos dados biológicos, que está sempre em busca de ferramentas mais aprimoradas (Cassanego, 2020).

Essa busca tem sido um empreendimento global que vem resultando em redes computacionais que permitiram cada vez mais fácil e rápido acesso aos dados biológicos, e que consentiram para o desenvolvimento de programas e software para a realização da análise rápida e sem ‘esforço’ (Bayat, 2002). Nesse contexto, se a Bioinformática é uma ciência de aplicação de ferramentas para a análise de dados biológicos e biomédicos, vale salientar de forma breve que a Biologia Computacional, que pode ser vista como área irmã da Bioinformática, é a responsável pelo desenvolvimento de modelos e algoritmos que solucionem os problemas que surgem da Bioinformática (Minardi, 2023). Isso porque, as ricas informações contidas no DNA – sejam os dados de alinhamento ou as variantes que carregam – sofrem com uma série de etapas de processamento que, muitas vezes, requerem softwares especializados (Danecek et al, 2021).

Nessa perspectiva, as principais ferramentas ligadas à Bioinformática incluem programas de computador e recursos disponibilizados pela internet, estes que são cada vez mais aprimorados para o desenvolvimento de softwares complexos com funções de recuperar, classificar, analisar e armazenar dados de sequenciamento (Bayat, 2002). Com os avanços das tecnologias de sequenciamento do genoma, e os projetos de sequenciamento em larga escala, novos formatos de dados tornaram-se necessários para a interoperabilidade, análise eficiente e armazenamento compacto dos dados (Danecek et al, 2021).

Portanto, para a realização de tais trabalhos, geralmente, os profissionais se utilizam de programas computacionais como Anaconda – sua versão conhecida como Miniconda, ou a Bioconda – com pacotes de softwares e ferramentas específicas para dados biológicos, sendo todos baseados no gerenciador de ambiente CONDA e comandando pela linguagem de programação Python (Alves, 2006).

Nos ambientes criados em qualquer um desses programas, o pesquisador pode instalar ferramentas e pacotes de softwares específicos para o seu tipo de trabalho. Entre os formatos de arquivo mais utilizados no campo da Bioinformática estão o SAM (*Sequence Alignment/Map*), especializado em armazenar alinhamento de leitura, e o VCF (*Variant Call Format*), que é utilizado para registrar informações sobre variantes genéticas, como os SNP – Polimorfismos de Nucleotídeo Único (*Single Nucleotide Polymorphisms*) – (Danecek *et al.*, 2021).

Apesar da enorme contribuição que ambos os formatos tiveram, eles podem se tornar lentos para analisar ou ineficientes para armazenar os dados dependendo do seu tamanho (Danecek *et al.*, 2021). Nesse cenário, para melhorar a interoperabilidade e as análises, as partes binárias desses formatos são amplamente utilizadas pela Bioinformática: o formato BAM (*Binary Alignment/Map*) para os dados de alinhamento e o formato BCF (*Binary Call Format*), versão binária do VCF, projetada para manipular grandes volumes de dados de variantes genéticas de forma mais rápida e eficiente (Danecek *et al.*, 2021).

A primeira versão online de ambos os formatos, encontrados nos pacotes de ferramentas SAMtools (arquivo BAM) e BCFtools (arquivo VCF), apareceu em 1998; sendo mantida e melhorada desde então. Os pacotes incluem ferramentas de conversão, classificação, consulta, estatística, chamada de variante e outros (Danecek *et al.*, 2021).

Por fim, cabe salientar que apesar das melhorias e avanços tecnológicos, arquivos extremamente grandes têm processamentos que necessitam de um alto grau de paralelização em Clusters/agrupamento computacionais (Danecek *et al.*, 2021). No próximo capítulo, serão apresentados em detalhes os dados alvos deste processamento, e como eles podem ser informativos em estudos de Arqueogenética e Arqueologia da Doença.

3 MATERIAIS E MÉTODOS

Para a realização dos objetivos apresentados, a pesquisa passou por um processo de levantamento bibliográfico, tanto para a identificação dos polimorfismos associados à infecção quanto para a compreensão da Arqueogenética, da Bioinformática e da Arqueologia da Doença, bem como suas aplicações. No que concerne à compreensão dos conceitos e suas aplicações, a busca voltou-se para livros e produções científicas – em endereços eletrônicos de artigos científicos como PMC (<https://pmc.ncbi.nlm.nih.gov>), Nature (<https://www.nature.com>), PubMed (<https://pubmed.ncbi.nlm.nih.gov>) – sobre os temas, estes que foram triados, discutidos e apresentados no capítulo 2 do presente estudo.

Para o levantamento dos polimorfismos associados à infecção, a busca se deu de forma exploratória através de plataformas como: BMC (<https://www.biomedcentral.com/>), National Center for Biotechnology (<https://www.ncbi.nlm.nih.gov/>), National Library of Medicine (<https://www.nlm.nih.gov/>), ScienceDirect (<https://www.sciencedirect.com/>), Scientific Electronic Library Online (<https://scielo.org/en/>) e Scientific reports (<https://www.nature.com/srep/>), através de: palavras-chaves, resumos, introduções e leituras das produções científicas e artigos, tendo sido feito uma triagem dos polimorfismos debatidos nos textos para aqueles mais recorrentes em todas as obras trabalhadas. Sua apresentação, bem como as explicações por trás de cada escolha, será apresentada a posteriori na próxima seção deste capítulo.

Em segunda instância, com o levantamento dos polimorfismos mais recorrentes e a escolha das amostras a serem analisadas, sendo elas referentes aos 221 indivíduos trabalhados por Santos em sua tese de doutorado (Santos, 2020) e três indivíduos trabalhados por Lindo *et al.* (Lindo *et al.*, 2018) totalizando 224 indivíduos – foram utilizadas ferramentas da Bioinformática na busca dos polimorfismos.

Tal busca foi realizada fazendo uso do ambiente computacional de alta performance, chamado SDumont, para a análise genômica, bem como a utilização do programa Anaconda, em sua forma de Miniconda, para a instalação de pacotes e ferramentas computacionais que realizaram desde a leitura e preparação das amostras até a busca pelos polimorfismos estudados.

Dessa forma, equipando-se com a ferramenta computacional conhecida como BCFtools (Danecek *et al.*, 2021), utilizou-se da função de filtragem e busca pelos polimorfismos listados, especificamente associados à TB, no conjunto amostral aqui trabalhado.

O resultado final da busca dos polimorfismos na amostra leva à produção de uma tabela com todos os indivíduos que apresentam algum tipo de polimorfismo associado à Tuberculose. Dessa forma, é realizada a organização dos dados e a elaboração de diferentes tipos de gráficos, como dispersão e frequência, utilizando-se de ferramentas do programa Microsoft Excel.

Por fim, munindo-se das concepções teóricas adquiridas pelas leituras escolhidas, os resultados obtidos serão discutidos, dando-se foco na apresentação das implicações e possibilidades dos mesmos para a presente pesquisa, bem como pesquisas futuras. Nas próximas seções deste capítulo, apresentaremos os polimorfismos levantados, as amostras a serem analisadas e as principais ferramentas, e processos, da busca pelos polimorfismos nas amostras através da Bioinformática.

3.1 Conhecendo os Caminhos: Apresentando os Polimorfismos

Nos processos que atuam no dogma central da biologia molecular, tais como a duplicação, transcrição e tradução do código genético e sua transformação em proteínas, podem ocorrer mudanças – aleatórias – nas bases que constituem o DNA. Essas mudanças são conhecidas como mutações. Quando a frequência dessa mutação é maior que 2% em uma única população, essa variação passa a ser chamada de polimorfismo genético (Souza *et al.*, 2019), sendo o mais comum o Polimorfismo de Nucleotídeo Único (SNP).

Isso porque, a diferença genômica entre dois indivíduos, quando comparados, é de 0,01%, ou seja, 1 em cada 1000 apresenta alguma distinção com relação à posição das bases nitrogenadas – adenina, citosina, guanina e timina (Santos, 2020). Logo, os estudos genômicos consideram apenas esse único polimorfismo, pois analisar 99,9% de dados similares seria inviável (questões de tempo e monetárias), principalmente, quando se leva em conta o tamanho do genoma humano: mais de 3 bilhões de bases nitrogenadas (Santos, 2020).

Nessa perspectiva, a primeira parte da pesquisa foi dedicada a um levantamento acerca dos Polimorfismos de Nucleotídeo Único (SNP) associados à Tuberculose (TB). Para tal, buscaram-se textos que trabalhassem a análise de DNA – antigo e moderno – nas pesquisas sobre TB e sua relação com fatores genéticos. Através dos bancos de dados, tais como: BMC (<https://www.biomedcentral.com/>), National Center for Biotechnology (<https://www.ncbi.nlm.nih.gov/>), National Library of Medicine (<https://www.nlm.nih.gov/>), ScienceDirect (<https://www.sciencedirect.com/>), Scientific Electronic Library Online (<https://scielo.org/en/>) e Scientific reports (<https://www.nature.com/srep/>) – foi possível

chegar aos polimorfismos mais discutidos até o presente momento (março de 2024), em grupos modernos, associadas à suscetibilidade ou resistência à TB. São eles:

Tabela 2 - Polimorfismos associados à suscetibilidade à Tuberculose.

SNP	GENE	CROMOSSOMO	POSIÇÃO	Autores
rs34536443	TYK2 P1104A	19	10463118	Kerner et al, 2021
rs3804099	TLR2	4	154624656	Wu et al,2020;Varshney,2022
rs5743708	TLR2	4	154626317	Varshney,2022
rs4833095	TLR1	4	38799710	Varshney,2022
rs4986790	TLR4	9	120475302	Varshney,2022
rs4986791	TLR4	9	120475602	Varshney,2022
rs352139	TLR9	3	52258372	Varshney,2022
rs5743867	TOLLIP	11	1328351	Wu et al,2020

Fonte: Aatoria Própria (2024)

Tabela 3 - Polimorfismos associados à resistência à Tuberculose.

SNP	GENE	CROMOSSOMO	POSIÇÃO	Autores
rs11536889	TLR4	9	120478131	Wu et al, 2020
rs3750920	TOLLIP	11	1309956	Wu et al, 2020
rs5743618	TLR1	4	38798648	Varshney et al,2022
rs7873784	TLR4	9	120478936	Varshney et al,2022
rs5743810	TLR4	4	38830350	Varshney et al,2022
rs5743836	TLR9	3	52260782	Varshney et al,2022

Fonte: Aatoria Própria (2024).

A priori, fechou-se o número de polimorfismos a serem analisados na presente pesquisa em 14, sendo 8 relacionados à suscetibilidade e 6 à resistência. Nesse cenário, o levantamento – e a escolha – de polimorfismos corrobora os estudos que relataram que diferentes genes da via TLR – receptores do tipo Toll/Toll Like Receptors que compõe o sistema imunológico inato – (como os apresentados na tabela) ativam a resposta imune celular e podem afetar a susceptibilidade individual à TB (Wu *et al.*, 2020). Isso porque, os TLR são parceiros importantes do sistema imunológico que protegem o hospedeiro contra infecções microbianas (Varshney *et al.*, 2022), bem como são mediadores essenciais da resposta individual a diversos patógenos e desempenham um papel central na etiologia de doenças humanas comuns, incluindo à TB (Wu *et al.*, 2020).

Entretanto, cabe salientar que alguns polimorfismos no gene TLR ficaram de fora. Os SNP rs1898830 e rs3764880 (gene TLR2 e TLR8 respectivamente) – todos apontados como causadores de suscetibilidade à infecção da TB – foram descartados pelos seguintes motivos: O primeiro deles, referente ao SNP (rs1898830), deu-se em virtude desse polimorfismo ter sido discutido no trabalho de Wu *et al.* (2020), estando associado às populações Han e Tibetana. No entanto, análises mais recentes não confirmaram a sua relação direta com a suscetibilidade à TB, mesmo que tenham sido analisados (Varshney *et al.*, 2022; Chen *et al.*, 2023). O segundo motivo, referente ao rs3764880, se deu em virtude dessa alteração acontecer no cromossomo X, não sendo foco do presente trabalho abordar os cromossomos sexuais.

Ademais, cabe pontuar também, que dois dos polimorfismos considerados na análise, o primeiro relacionado à suscetibilidade e o segundo a resistência contra à TB, o SNP rs5743867 (gene TOLLIP) e o rs11536889 (gene TLR4), são associados a populações específicas. O do gene TOLLIP a populações Africanas e o do TLR4 populações tibetanas (Wu *et al.*, 2020). Além disso, cabe mencionar também, que para Chen *et al.* (2023) o polimorfismo TLR2 – rs5743708 – é o que apresenta uma significativa associação com o maior risco de Tuberculose. Por fim, cabe pontuar que nem todos os polimorfismos associados à TB estão relacionados com os TLR. É o caso do da variante P1104A (gene TYK2) discutida no trabalho de Kerner *et al.* (2021), no cenário da europa neolítica.

Outrossim, cabe ressaltar que, a busca pelos polimorfismos conduziu a cromossomos específicos (tabelas 2 e 3) que foram o alvo da presente pesquisa, ou seja, nas amostras antigas buscou-se nesses cromossomos os polimorfismos listados. Destarte, após um breve apanhado sobre o levantamento acerca dos polimorfismos associados à TB, na próxima seção iremos nos debruçar sobre as amostras que serão utilizadas no desenvolvimento da presente pesquisa.

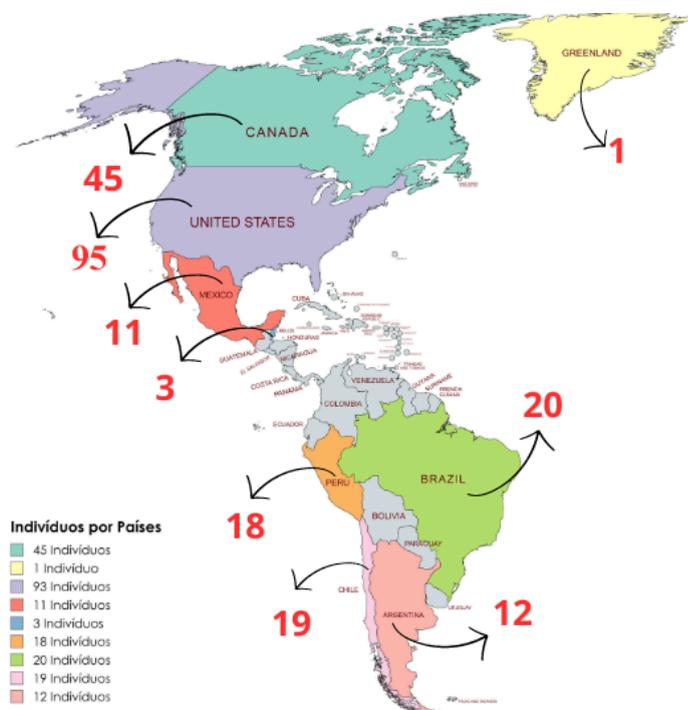
3.2 Um olhar sobre as Américas: Apresentando as Amostras

Como mencionado anteriormente, para a realização da presente pesquisa serão analisados indivíduos nativos-americanos já sequenciados geneticamente, em que os dados já se encontram publicados em trabalhos sobre genômica (Lindo *et al.*, 2018; Santos, 2020).

Nesse contexto, os remanescentes a serem trabalhados são provenientes de distintos sítios arqueológicos de 8 países do continente Americano, são eles: 12 da Argentina, 3 do Belize, 20 do Brasil, 45 do Canadá, 19 do Chile, 95 dos Estados Unidos da América (EUA), 1 da Groenlândia, 11 do México e 18 do Peru. Dessas amostras, 221 foram trabalhadas por

Santos (Santos, 2020) e as outras 3, localizadas especificamente no Peru, foram trabalhadas por Lindo *et al.* (Lindo *et al.*, 2018).

Figura 5 - Distribuição das amostras.



Fonte: Autoria Própria (2024)

Cronologicamente, de forma estimada, variam do Pleistoceno (>12.000 antes do presente) até os períodos de contato e pós-contato com os europeus, os últimos 500 anos aproximadamente (Santos *et al.*, 2022). Para mais detalhes sobre as amostras consultar os autores já citados (Lindo *et al.*, 2018; Santos, 2020). Destarte, na próxima e última seção sobre a metodologia, será apresentado, de forma sucinta, a aplicação da Bioinformática em um processo conhecido como “Chamada de Variante”, este que objetiva buscar nas amostras aqui apresentadas os polimorfismos discutidos anteriormente.

3.3 Aplicando a Bioinformática: Chamando uma Variante

A extração e análise de DNA antigo (*ancient DNA*), realizada através de ossos arqueológicos, tem se tornando comum na Europa e na América do Norte para a resolução de inúmeros problemas, incluindo a identificação de doenças (Hermann; Hummel, 1994; Hummel, 2007 *apud* Santos, 2020).

Para a obtenção dos dados genômicos, as amostras extraídas são submetidas ao processo de sequenciamento da Próxima Geração (Santos, 2020) que permite a obtenção de frações do genoma (Santos, 2020). Nessa perspectiva, partindo de uma análise comparativa,

os dados sequenciados das amostras são comparados com o genoma de referência humano – GRCh37 – pois essa foi a versão utilizada nas publicações originais desses genomas, disponível na internet (Santos, 2023), em busca de polimorfismos. Nesse processo, se faz necessário a utilização de técnicas e métodos – específicos e ideais – da Bioinformática.

Como ambiente ideal, no desenvolvimento da presente pesquisa, para análise do genoma completo, foi utilizado o supercomputador de alta performance para análise genômica, conhecido como SDumont, sendo que sua utilização só foi possível graças à aprovação do projeto do LEARQ na chamada interna 08/2023 (PROPESQI/UFPE). Nesse sentido, seguindo os tutoriais de instalação para Windows 10, disponibilizado pela equipe que cuida do supercomputador, foram instalados os aplicativos Winfix e Cisco VPN client.

Uma vez instalados, através do login e da senha específica para cada usuário, bem como códigos de linguagem computacional Python, torna-se possível acessar as páginas para a realização das análises. Cabe ressaltar que, cada usuário – do Projeto LEARQ que foi submetido na chamada interna de utilização do Supercomputador – possui uma cota individual de armazenamento. Entretanto, apesar de ser um valor considerável, não é o suficiente para análises de genoma maiores e mais complexos. Nessa perspectiva, para a realização da pesquisa, será utilizado a pasta compartilhada Scratch que possui um armazenamento maior do que a cota individual oferecida a cada usuário.

Contudo, cabe salientar que, apesar de possuir uma maior quantidade de armazenamento, todos os arquivos dessa pasta são excluídos após o tempo de 60 dias. Dessa forma, ficou acordado que a pesquisa seria realizada na pasta Scratch e os resultados, bem como outros dados necessários, seriam enviados para a pasta Home, que como dito anteriormente, cada usuário possui a sua.

Outrossim, no ambiente ideal do SDumont, e utilizando-se de ferramentas apropriadas como Anaconda, ou melhor dizendo, Miniconda, as amostras seriam preparadas – iniciando pelo genoma de referência – seguindo os moldes do que é exemplificado no tutorial AAAGs_2018⁴.

Nessa perspectiva, com a finalização da preparação do genoma de referência, inicia-se a preparação das amostras seguindo o mesmo tutorial utilizado para o de referência, ou seja, arquivos FASTA, FASTQ, SAM, BAM e VCF (Danecek *et al.*,2021). Nessa etapa, as amostras são organizadas, lidas e analisadas estatisticamente, dando ao pesquisador uma visão do estado de suas amostras e como isso pode influenciar no andamento da pesquisa.

⁴ https://github.com/Sagui-omics/AAAGs_2018/tree/master/Session2_ReadAlignment_VariantCalling

Além disso, é nesse momento que os arquivos das amostras são convertidos em arquivos mais leves – chamados de BAM – para facilitar o processamento dos mesmos.

Ademais, com as preparações feitas, realiza-se a comparação entre o genoma de referência e as amostras analisadas em busca de polimorfismos; no processo conhecido como “Chamada de Variante”. Tal mapeamento acontece, de forma resumida, da seguinte forma:

- Os arquivos Bam, após as análises métricas, são convertidos em arquivos GVCF e nesses arquivos serão listados todas as variantes encontradas nas amostras (AAAGs_2018);
- Utilizando-se de uma ferramenta específica – *Bcftools* – é realizado, nos arquivos GVCF, uma contagem de indivíduos, nome desses indivíduos e uma contagem de variantes/polimorfismos; gerando um arquivo VCF;
- Ao fim desse processo, o arquivo VCF gerado poderá ser aberto na plataforma do Excel, ou Planilhas do Google, local em que os dados serão trabalhados;
- Nessas plataformas, Excel ou Planilha do Google, os dados se apresentam – de forma geral – em uma linguagem binária de 0 e 1; sendo 0/0 referente a nenhum polimorfismo, ou seja, a amostra analisada está igual ao genoma de referência na parte em que foi analisada, e o 1/1 ou o 0/1 apontando para a presença do polimorfismo buscado.

Figura 6 - Arquivo VCF ao ser aberto na plataforma da planilha do google/excel.

A	B	C	D	E	F	G	H	I	J	K	L	M	N
#CHROM	POS	ID	REF	ALT	Kennewick	I0038	I0719	I1128	USR1	125	158	167	168
4	154608453	rs1898830	A	G	0/1
4	154624656	rs3804099	T	C	.	.	0/0	1/1	0/1	0/0	0/0	0/0	0/0
4	154626317	rs5743708	G	A
1	159005649	rs1772408	A	G	.	1/1	1/1	.	0/1
9	120478131	rs11536889	G	C
11	1309956	rs3750920	C	T	0/0	0/0

Fonte: Autoria Própria (2024)

Outrossim, com os dados inseridos na plataforma das Planilhas do Google, ou Excel, torna-se possível a construção de informações estatísticas – gráficos – de distribuição desses genes pelo continente Americano antes da colonização, sendo possível apontar a área de maior ou menor incidência. Além disso, também é passível de análise a frequência desses genes, bem como pontuar se há uma maior predominância de genes ligados à susceptibilidade ou resistência à infecção no continente Americano antes do contato.

Logo, tal análise permitirá responder ao questionamento realizado – pois seria possível ver se há ou não os polimorfismos ligados à TB – bem como, dependendo dos

marcadores conseguidos e suas distribuições, permitirá levantar questionamentos sobre a ancestralidade desses indivíduos e a evolução da doença no território estudado. Destarte, a próxima seção será dedicada à apresentação dos resultados encontrados após a análise das amostras assim como breves discussões e considerações sobre os mesmos.

4 RESULTADOS E DISCUSSÃO

A primeira etapa das análises refere-se a levantar quais SNP (Polimorfismo de Nucleotídeo Único) foram encontrados após a “Chamada de Variante”. Dentre os 14 polimorfismos levantados, sendo 8 de suscetibilidade e 6 de resistência, 11 foram evidenciados no processo descrito acima (Tabela 4), sendo tal achado um indicativo da presença dessas alterações genômicas nos indivíduos analisados.

Tabela 4 - Polimorfismos encontrados no processo da “Chamada de Variante”. Os vermelhos causam suscetibilidade à infecção, enquanto os azuis, resistência. A coluna dos cromossomos indica a localização cromossômica de cada SNP, estando os polimorfismos listados distribuídos nos cromossomos 3,4,9 e 11. A posição exata de cada SNP no cromossomo, descrita em coordenadas genômicas, é dada na coluna de posição. A coluna do gene faz referência a parte do que está sofrendo alteração. O ID, por sua vez, é como esse SNPs é catalogado nos bancos de dados da área molecular.

CROMOSSOMO	POSIÇÃO	GENE	ID
3	52258372	TLR9	rs352139
3	52260782	TLR9	rs5743836
4	38798648	TLR1	rs5743618
4	38799710	TLR4	rs4833095
4	38830350	TLR4	rs5743810
4	1,55E+08	TLR2	rs3804099
4	1,55E+08	TLR2	rs5743708
9	1,20E+08	TLR4	rs11536889
9	1,20E+08	TLR4	rs7873784
11	1309956	TOLLIP	rs3750920
11	132835	TOLLIP	rs5743867

Fonte: Autoria própria (2024)

Dentro das amostras analisadas, que totalizam 224 indivíduos, 77 apresentaram pelo menos 1 dos SNPs apontados, estes que se distribuem pela América da seguinte forma (Figura 7): 4 na Argentina, 2 em Belize, 9 no Brasil, 14 no Canadá, 8 no Chile, 25 nos Estados Unidos, 1 na Groenlândia e 14 no Peru. Nos 11 indivíduos analisados do México não foram encontradas nenhuma das variantes listadas. Esse resultado pode ser em função da degradação das amostras analisadas (seja em função do tempo, do ambiente na qual estiveram

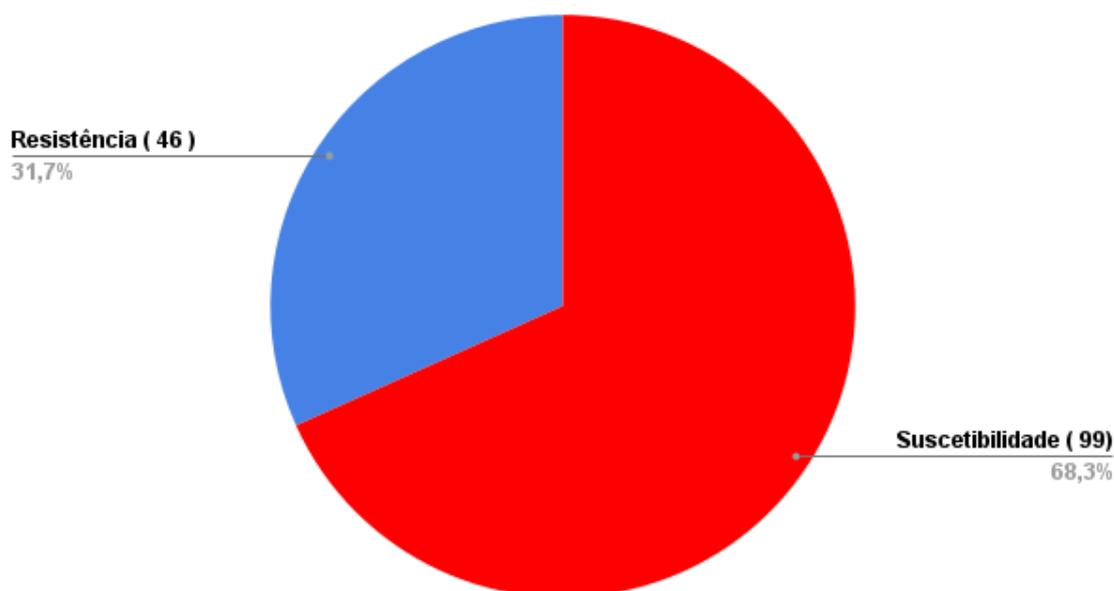
CP23	Brasil	Lapa do Santo	9.550
CP25	Brasil	Lapa do Santo	9.680-9.530
LAR001	Brasil	Laranjal	6.660-6.450
SUMIDOURO6	Brasil	Gruta do Sumidouro	>10.000
SUMIDOURO5	Brasil	Gruta do Sumidouro	>10.000
516	Canadá	GbTo-36	N/A
158	Canadá	GbTo-23	2.740-2.350
168	Canadá	GbTo-18	2.330-1.840
181	Canadá	GbTo-23	2.250-1.770
311	Canadá	GbTo-31	1.670-1.280
357	Canadá	GbTo-31	N/A
365	Canadá	GbTo-31	1.890-1.440
413	Canadá	GbTo-23	1.500-1.140
468	Canadá	GbTo-33	1.500-1.130
470	Canadá	GbTo-33	1.180-800
507	Canadá	GbTo-36	1.890-1.410
532	Canadá	GbTo-36	N/A
CK-13	Canadá	Lucier	4.853-4.837
19651	Canadá	Big Bar Lake	6.000
I11974	Chile	Los Rieles	11.140-10.730
I1752	Chile	Conchalí	650-540
I1753	Chile	Los Rieles	5.310-4.870
I1754	Chile	Conchalí	910-740
I2537	Chile	Pica Ocho	720-570
895	Chile	N/A	N/A
5832	Chile	Punta Santa Ana-1	6.500
A460	Chile	Cueva Ayayema	4.500
Kennewick	Estados Unidos	Columbia River, WA	8.400-8.690
I0719	Estados Unidos	Chaluka Midden, AK	700-310
I1118	Estados Unidos	Warm Cave, AK	500-130
I1125	Estados Unidos	Ship Rock Island, AK	710-440
I1128	Estados Unidos	Warm Cave, AK	560-250
I5320	Estados Unidos	Tochak McGrath, AK	N/A
I5321	Estados Unidos	Tochak McGrath, AK	N/A
I5319	Estados Unidos	Tochak McGrath, AK	900-550
USR1	Estados Unidos	Upward Sun River, AK	11.600-11.270
Anzick-1	Estados Unidos	Anzick, MT	12.722-12.590

Ala1	Estados Unidos	Sii Túupentak, CA	N/A
CR-01	Estados Unidos	Santa Cruz Island, CA	764-561
PS-09	Estados Unidos	Point Sal, CA	N/A
SC-03	Estados Unidos	San Clemente Island, CA	665-485
SC-05	Estados Unidos	San Clemente Island, CA	760-561
SM-02	Estados Unidos	San Miguel Island, CA	465-150
SN-11	Estados Unidos	San Nicolas Island, CA	714-517
SN-13	Estados Unidos	San Nicolas Island, CA	508-276
SN-17	Estados Unidos	San Nicolas Island, CA	4.774-4.342
SN-44	Estados Unidos	San Nicolas Island, CA	4.819-4.450
AHUR_2064	Estados Unidos	Spirit Cave, NV	10.000
Lovelock1	Estados Unidos	Lovelock Cave, NV	2.000
Lovelock2	Estados Unidos	Lovelock Cave, NV	1.900
Saqqaq	Groelândia	Qeqertasussuk	4.000
CP29	Peru	Cuncaicha	9.240-8.770
CP8	Peru	Cuncaicha	4.290-4.080
CUN008	Peru	Cuncaicha	3.370-3.180
I0038	Peru	Lauricocha	8.730-8.520
I0039	Peru	Lauricocha	8.580-8.420
I0041	Peru	Lauricocha	8.580-8.420
I1357	Peru	Botigiriyocc	880
I1484	Peru	Botigiriyocc	900-730
I1485	Peru	Cueva Yacotogia	1.160-960
I2261	Peru	La Galgada	4.230-3.980
I2551	Peru	Huayuncalla	910-740
IL2	Peru	Rio Uncallane	N/A
IL3	Peru	Rio Uncallane	1830
IL4	Peru	Rio Uncallane	1630

Fonte: Autoria própria (2024). Para dados das amostras, ver Lindo e colegas (2018) e Santos (2020).

Nesse contexto, após retirar todas as amostras em que não foi possível fazer nenhuma leitura, em virtude da degradação do próprio material genético analisado, elaboraram-se gráficos para uma melhor visualização dos resultados conseguidos. O primeiro gráfico realizado foi um de frequência, entre os SNP de suscetibilidade versus os de resistência, para se visualizar qual lado seria mais predominante nas Américas pré-contato.

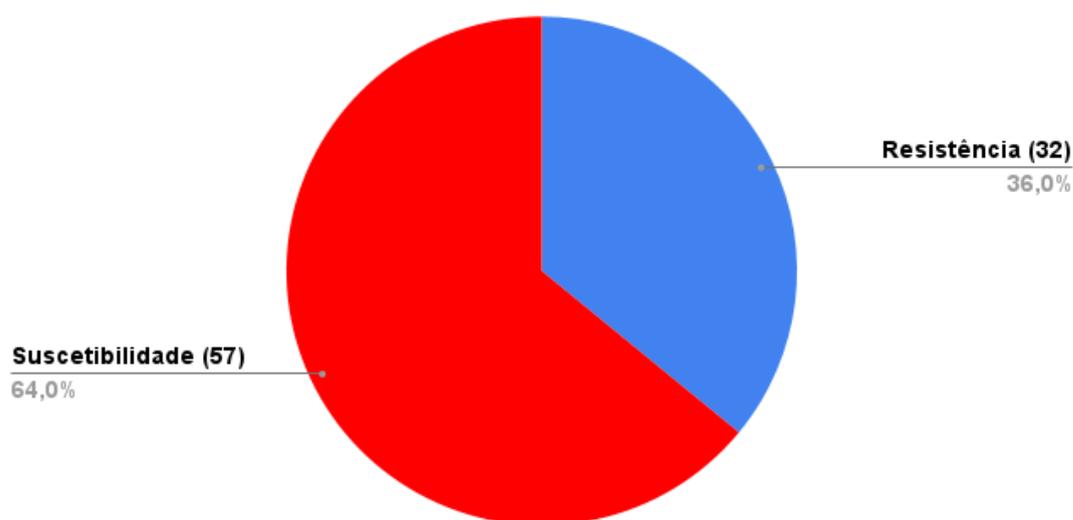
Gráfico 1 - Proporção de variantes associadas à Suscetibilidade e Resistência a Tuberculose no Continente Americano antes da Colonização.



Fonte: Autoria própria (2024)

Como demonstrado pelo gráfico acima, no território Americano antes da colonização havia uma maior suscetibilidade à infecção da Tuberculose, sendo este cenário visualizado nas Américas do Norte (Gráfico 2) e a do Sul (Gráfico 3).

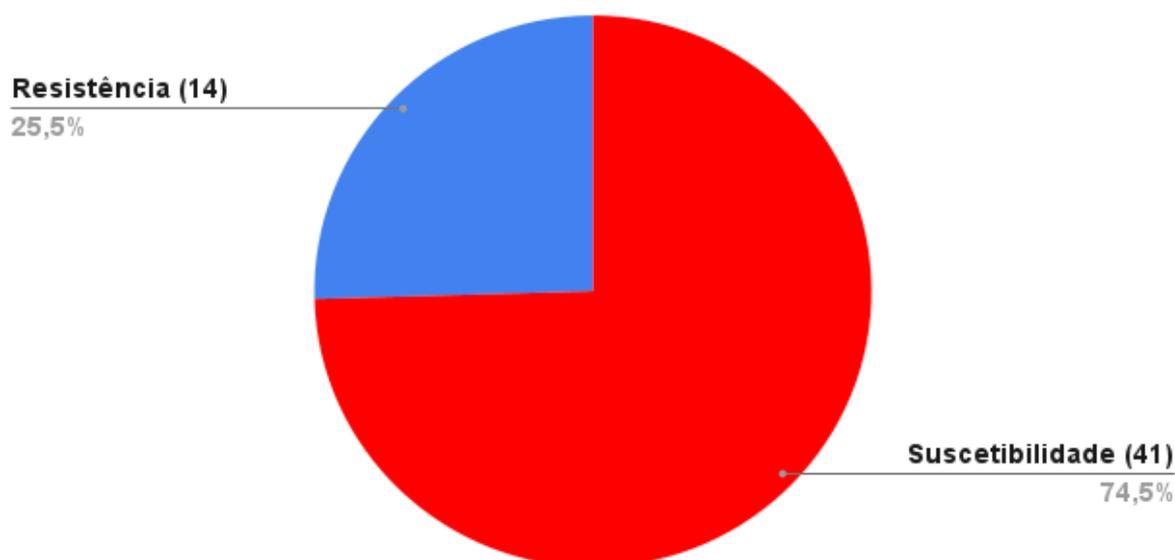
Gráfico 2 - Proporção de variantes associadas à Suscetibilidade e Resistência na América do Norte antes do contato.



Fonte: Autoria própria (2024)

Para a Mesoamérica, infelizmente, em virtude da não leitura de dados nos indivíduos do México, ficou-se apenas com as duas amostras de Belize. Se levarmos esses dados em voga, apesar do pequeno número de indivíduos sequenciados, para a parte central da América pode-se supor uma maior suscetibilidade à Tuberculose.

Gráfico 3 - Proporção de variantes associadas à Suscetibilidade e Resistência na América do Sul antes do contato.

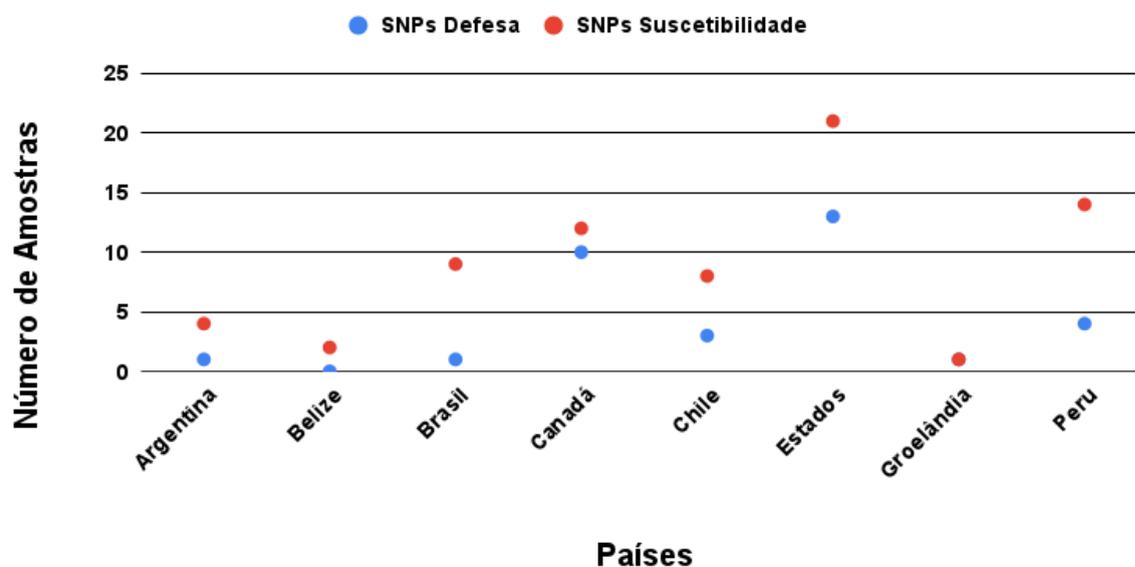


Fonte: Autoria própria (2024)

Em um quadro geral, em busca de se ter uma visualização de como esses SNP se distribuem pelos países, elaborou-se um gráfico de dispersão dessas alterações (Gráfico 4) em que é possível ver que os Estados Unidos, por possuir o maior número de indivíduos sequenciados, é o que apresenta o maior número de indivíduos com alterações – seja ela de suscetibilidade, resistência ou ambas.

Entretanto, uma maior quantidade absoluta de sequenciamentos não expressa uma maior ou menor suscetibilidade ou resistência à TB. Em razão disso, o Brasil, a Argentina e o Peru apresentam uma das maiores porcentagens de suscetibilidade à infecção (90%, 80%, 70%) respectivamente, e o Canadá, ao contrário dos países Sul americanos, apresenta a maior porcentagem (quase 50%) de alterações de resistência contra à infecção. Todavia, para inferências mais profundas sobre tal distribuição e dispersão, serão necessários novos sequenciamentos nos países em questão.

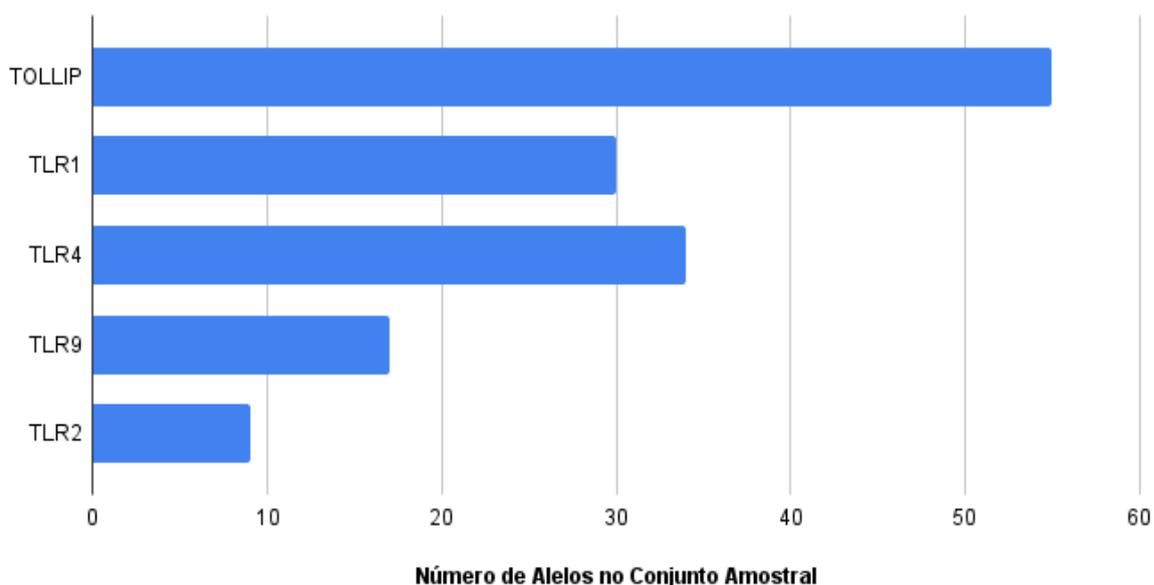
Gráfico 4 - Dispersão dos SNP de Suscetibilidade e Resistência nos países analisados.



Fonte: A autoria própria (2024)

Ademais, se focarmos nos genes que mais sofrem mutação, não fazendo distinção se essa alteração oferece uma maior suscetibilidade ou resistência a doença, o gene em foco acaba por ser o TOLLIP, uma proteína ‘Toll-interacting’ que atua como regulador das respostas imunológicas inatas que são mediadas pelo receptor Toll-Like – TLR – (Capelluto, 2012).

Gráfico 5 - Proporção dos genes analisados predominantes nas Américas pré-contato.



Fonte: A autoria própria (2024)

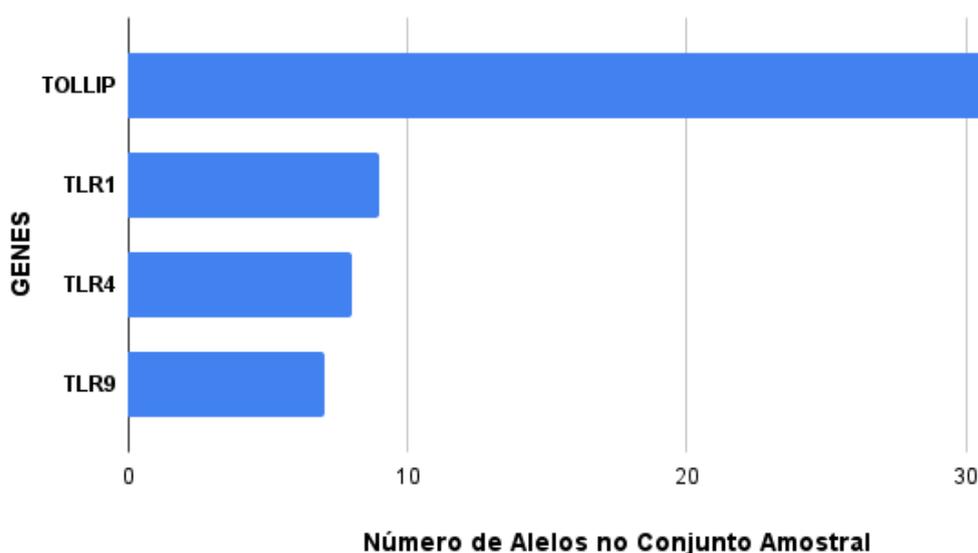
Para além deste, destacam-se também os genes TLR4 e TLR1, que são genes já estudados e associados a uma maior suscetibilidade à infecção (Shah *et al.*, 2013; Wu *et al.*,

2020). Tal contexto apresenta uma relação entre esses dois genes que é corroborada por outros trabalhos (Capelluto, 2012; Shah *et al.*, 2013; Wu *et al.*, 2020), uma vez que o TOLLIP atua como um crítico regulador da via TLR em humanos e essa regulação desempenha um papel na suscetibilidade à Tuberculose.

Isso porque, o patógeno causador da TB em humanos (*M.tb*) é inicialmente discernido pelo TLR4 e TLR2, posteriormente interagindo com a proteína de alteração Toll (TOLLIP) para ativação dos macrófagos (Wu *et al.*, 2020). Destarte, alterações nessa cadeia de sinalização, afeta a suscetibilidade do indivíduo à infecção, visto que a modificação nessa cadeia interrompe a produção de citocinas e células importantes para a defesa do corpo.

Por sua vez, ao olharmos separadamente cada América, observa-se na América do Sul a prevalência do gene TOLLIP (Gráfico 6), enquanto na América do Norte prevalece o TLR4 (Gráfico 7). A Mesoamérica, apesar de possuir apenas dois indivíduos sequenciados no Belize que permitiram a leitura, tende para uma possível predominância do Gene TOLLIP, sendo esse o único SNP encontrado nesses indivíduos. Tal contexto, salienta novamente que a quantidade absoluta de amostras não significa uma maior ou menor suscetibilidade ou resistência à infecção. Outro fator, que diferencia as Américas, é a América do Norte apresentar o maior índice de alterações no gene do TLR2, que é o mais associado a uma maior suscetibilidade a infecção por TB (Tientcheu, 2017; Wu *et al.*, 2020), enquanto a América do Sul – dentro da presente análise – não apresenta nenhum indivíduo com essas características.

Gráfico 6 - Proporção dos genes predominantes na América do Sul antes do contato.

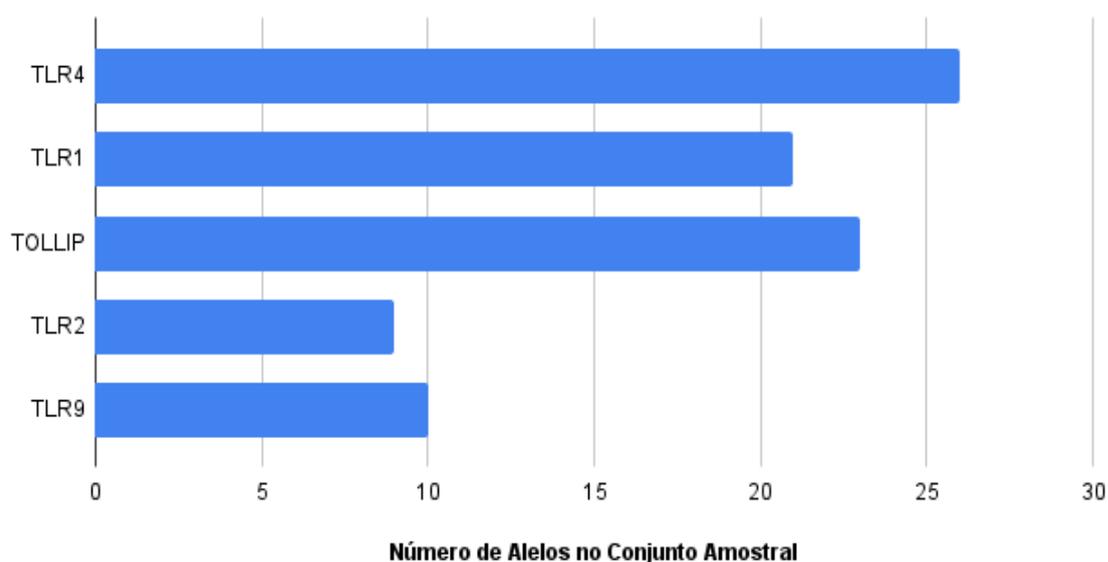


Fonte: Autoria própria (2024)

Partindo disso, ao direcionarmos atenção a América do Sul, nota-se constância do Gene TOLLIP e sua variável, com baixa incidência de alterações nos genes TLR e baixa presença de polimorfismos associados à resistência na região. Tal predominância, pode estar associada a: ancestralidade, aleatoriedade e/ou contextos epigenéticos de adaptação a um determinado ambiente e estilo de vida pretérito. Para elucidar essa, e outras questões, novos trabalhos são necessários, mas é possível pontuar a possibilidade de parte da população do continente ser mais suscetível à infecção, o que pode ter sido um fator que aumentou as probabilidades de morte por Tuberculose com a chegada dos colonizadores, e que a TB pode ter atuado como uma pressão seletiva nas sociedades agrícolas, como as dos Andes (Joseph *et al.*, 2023), antes do contato europeu.

Para o Norte da América (Gráfico 7) o contexto se inverte, uma vez que existe a predominância das alterações no gene TLR4, principalmente da sua variável **rs5743810** que oferece uma maior resistência à infecção, sendo tal panorama mais comum no Canadá.

Gráfico 7 - Proporção de genes predominantes na América do Norte antes do contato.



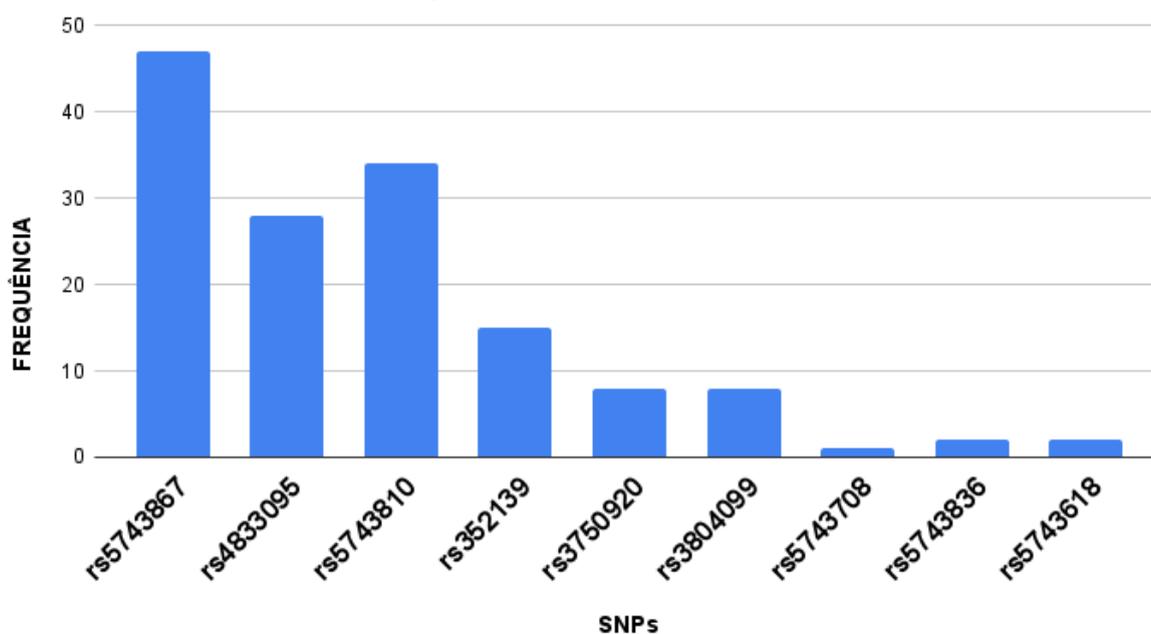
Fonte: Autoria própria (2024)

Tal contexto pode ter fornecido, em especial aos grupos do Canadá, uma maior resistência à patologia quando a linhagem europeia foi trazida para o continente americano. Todavia, tal possibilidade não é fortemente corroborada pelos dados levantados - em virtude da maior frequência de suscetibilidade em um contexto geral, o que pode ter culminado em mortes mais rápidas pela doença com o advento da colonização - e pelo ponto de não ter sido

possível, por razões da própria degradação dessas partes do genoma nas amostras canadenses, leituras mais complexas para esses indivíduos.

Ao falarmos dos SNP ao invés dos genes, existe uma forte predominância, em toda América, do polimorfismo **rs5743867** que ocorre no gene TOLLIP e oferece uma maior suscetibilidade à infecção por Tuberculose, o que corrobora o indicativo do gene TOLLIP ser o mais predominante no contexto geral da América pré-contato. Esse SNP foi encontrado em populações na Uganda (África Oriental), mas não em populações Tibetana e Han (Wu *et al.*, 2020). Para além deste, os polimorfismos que aparecem com mais frequência é o **rs5743810**, associado ao TLR4 e oferece uma maior resistência à infecção, e o **rs4833095**, relacionado ao TLR1 e associado a uma maior suscetibilidade à infecção.

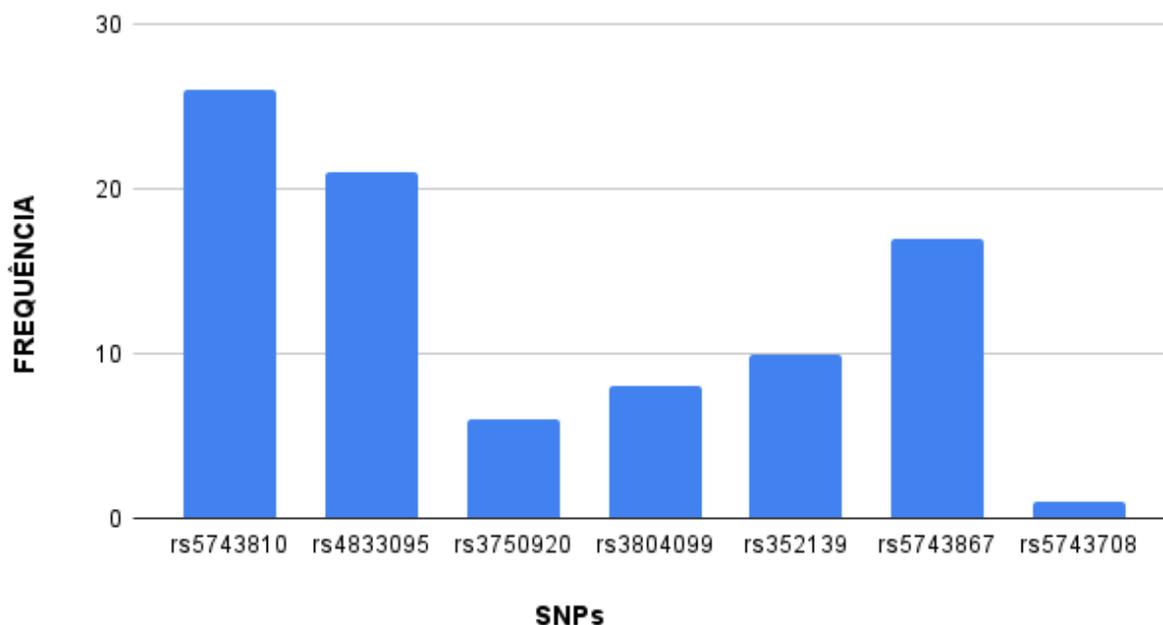
Gráfico 8 - Proporção das variantes (SNP) associados à TB presentes na América pré-contato, com a nítida predominância do **rs5743867**.



Fonte: Autoria própria (2024)

Apesar dessa predominância em um contexto geral, a América do Norte (Gráfico 9), de forma oposta, apresenta uma maior frequência do **rs5743810**, associado ao TLR4, que oferece uma maior resistência à doença, corroborando o que foi apontado anteriormente.

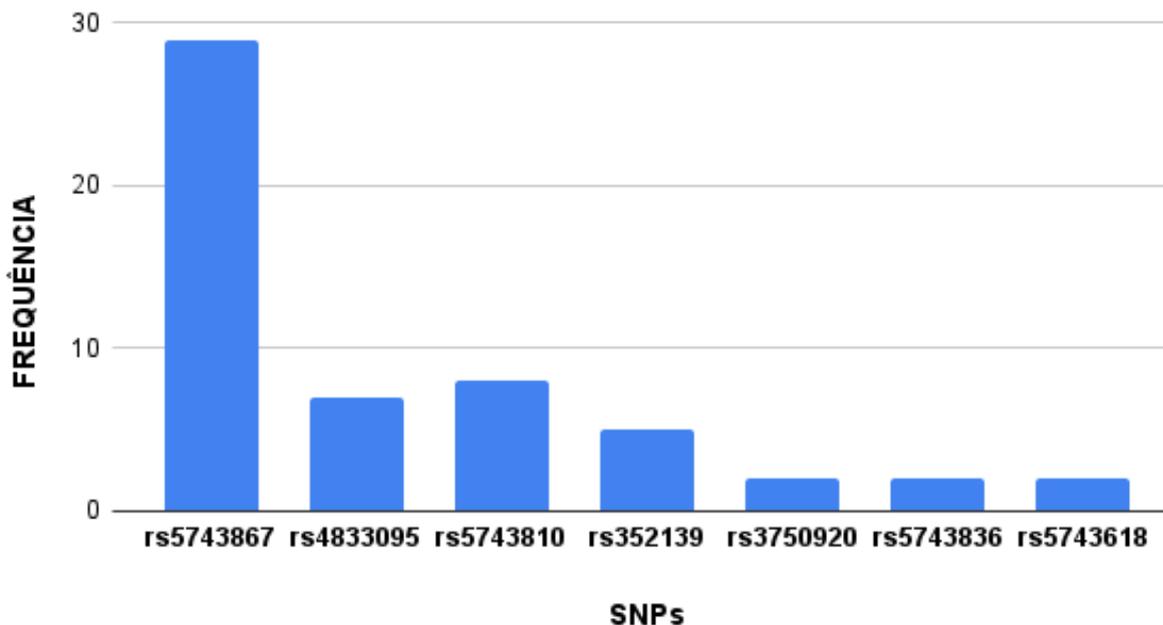
Gráfico 9 - Proporção das variantes (SNP) associados à TB presentes na América do Norte, com a nítida predominância do **rs5743810**.



Fonte: Autoria própria (2024)

A América do Sul (Gráfico 10), por sua vez, segue o panorama geral e apresenta, em grande frequência, o **rs5743867**, associado ao gene TOLLIP, oferecendo uma maior suscetibilidade à infecção, corroborando a discussão anterior.

Gráfico 10 - Proporção das variantes (SNP) associados à TB presentes na América do Sul, com a nítida predominância do **rs5743867**.



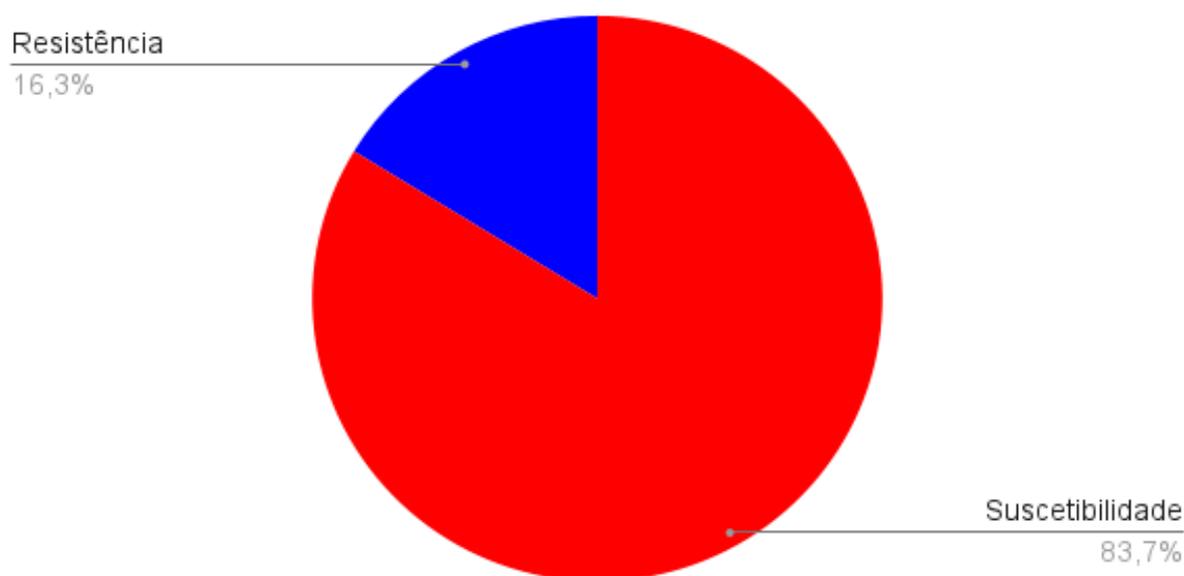
Fonte: Autoria própria (2024)

Dessa forma, nota-se uma predominância de alterações no grupo dos TLR na parte norte do continente, às vezes em um mesmo sítio, como: GbTo 31 (Canada), Lovelock (Estados Unidos); enquanto a parte sul sofre com uma maior modificação no gene TOLLIP, em indivíduos do mesmo sítio, como: Lapa do Santo (Brasil), Rio Uncallane (Peru); o que é um possível indicativo de uma maior suscetibilidade da população nesta parte do continente.

Esse contexto, assim como a predominância dos marcadores associados à suscetibilidade, são corroborados quando analisamos os dados por um viés cronológico de dois intervalos de tempo: Intervalo 1 (> 12.000 a 6.000 anos AP) e Intervalo 2 (5.000 a 500 anos AP); sendo estes escolhidos em função da maior frequência de indivíduos nesses espaços.

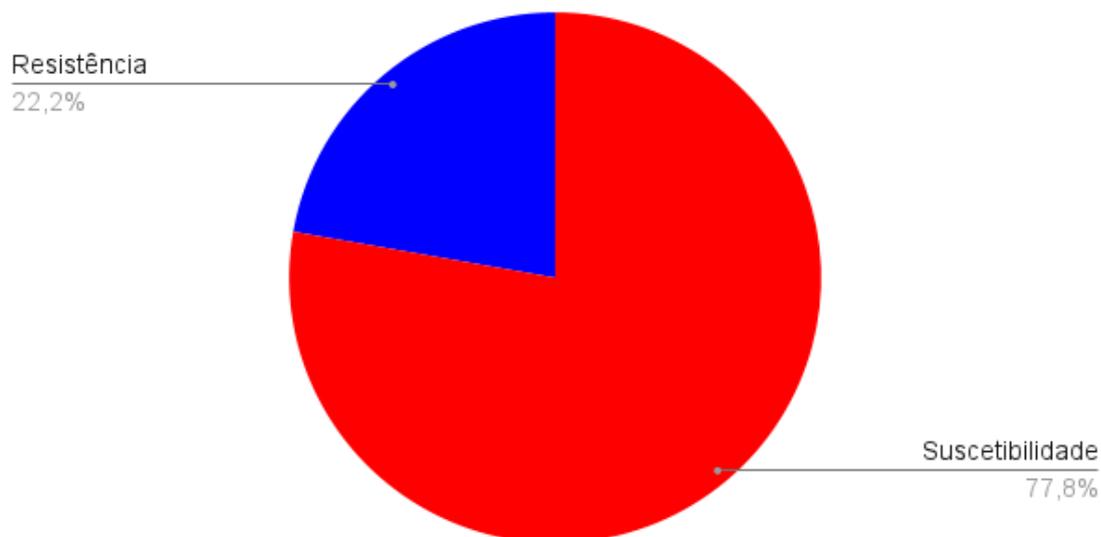
À vista disso, os gráficos seguintes (11, 12 e 13) apresentam a predominância dos marcadores associados à suscetibilidade à TB nas Américas no primeiro intervalo escolhido, sendo essa predominância levemente maior na parte Sul do Continente.

Gráfico 11- Proporção de variantes associadas à Suscetibilidade versus Resistência nas Américas no Intervalo 1.



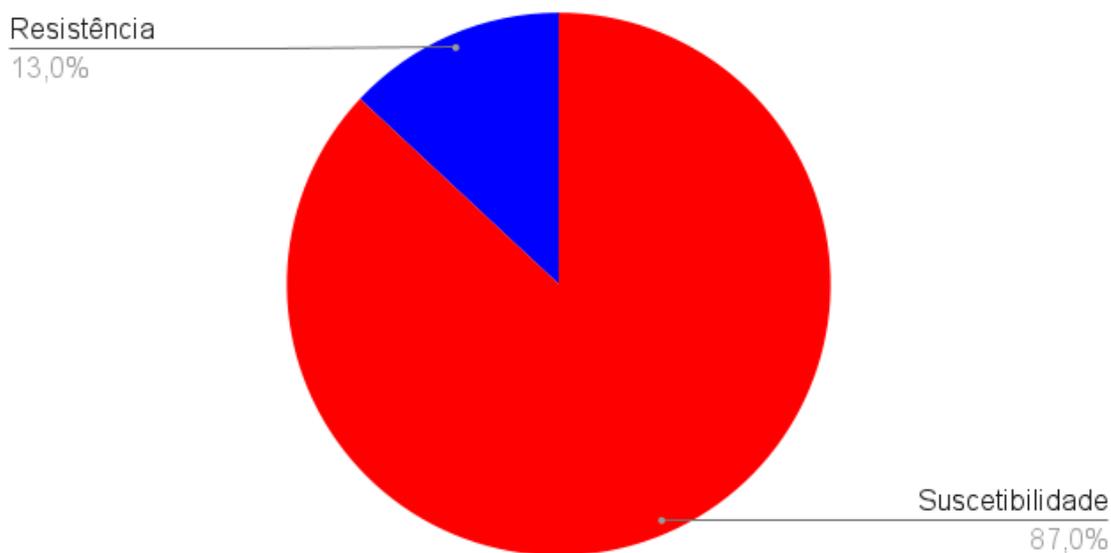
Fonte: Autoria própria (2024)

Gráfico 12 - Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Norte no Intervalo 1.



Fonte: Autoria própria (2024)

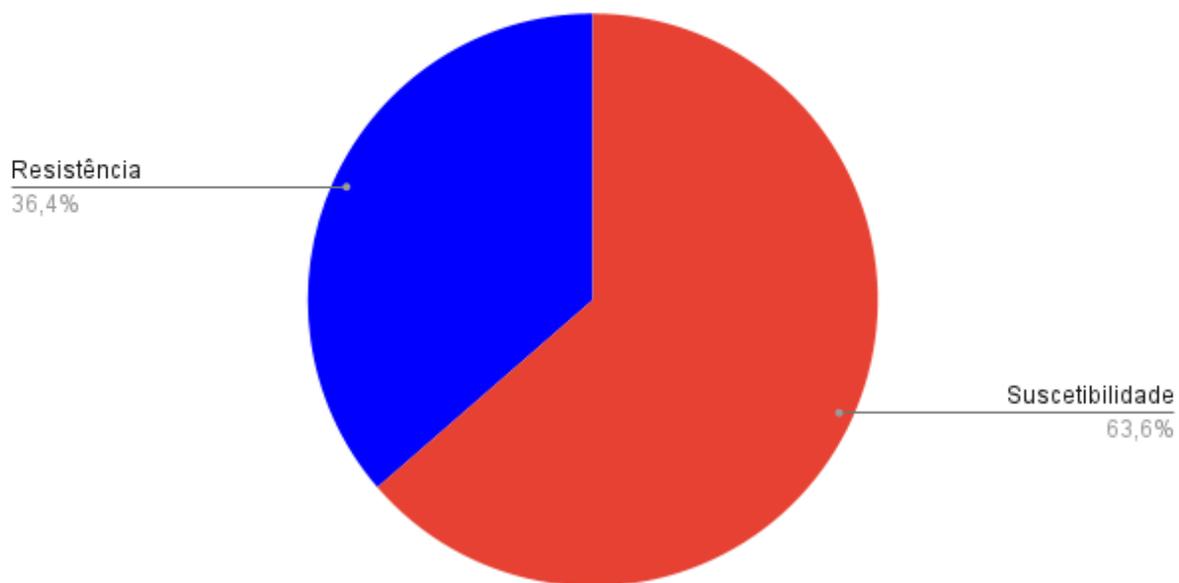
Gráfico 13 - Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Sul no Intervalo 1.



Fonte: Autoria própria (2024)

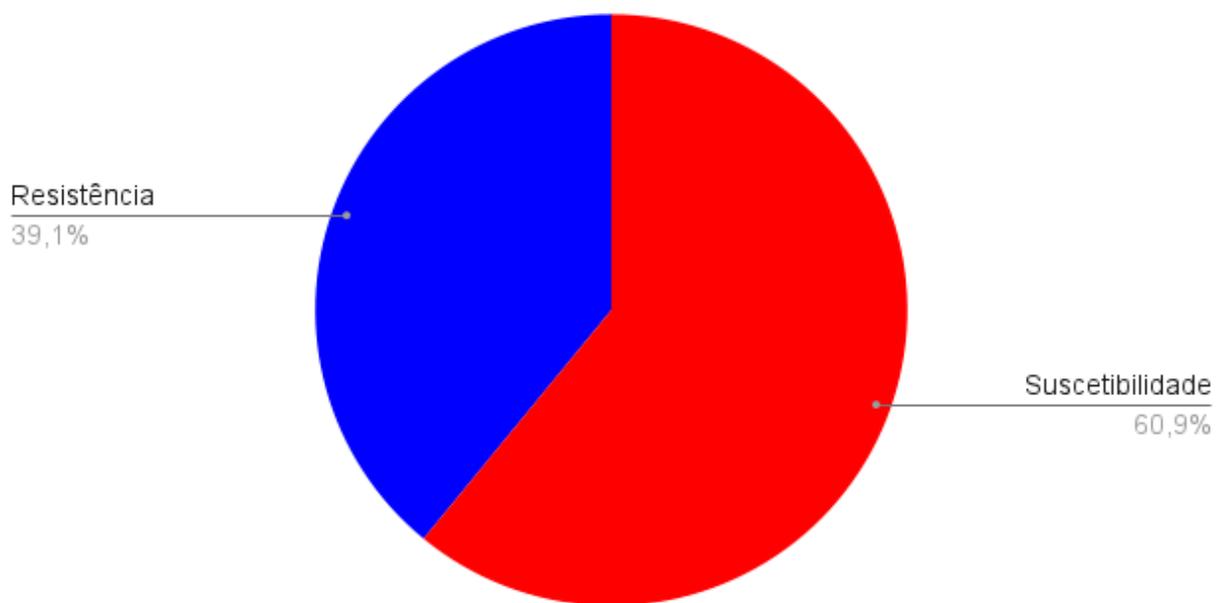
O cenário se perpetua no segundo intervalo, o que corrobora a predominância da suscetibilidade de uma forma geral, todavia, no segundo intervalo é possível visualizar, principalmente na América do Norte, um aumento associado à resistência contra à infecção.

Gráfico 14 - Proporção de variantes associadas à Suscetibilidade versus Resistência nas Américas no Intervalo 2.



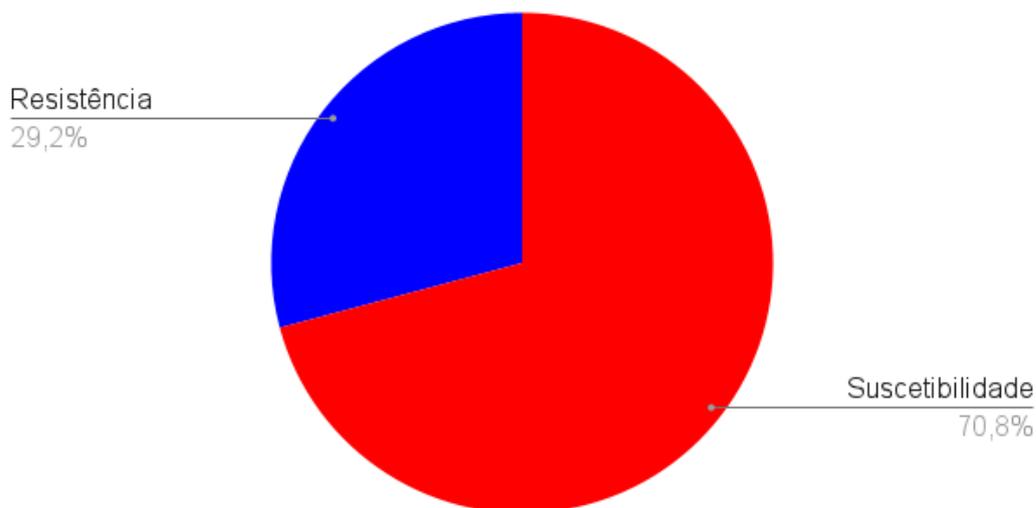
Fonte: Autoria própria (2024)

Gráfico 15 - Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Norte no Intervalo 2.



Fonte: Autoria própria (2024)

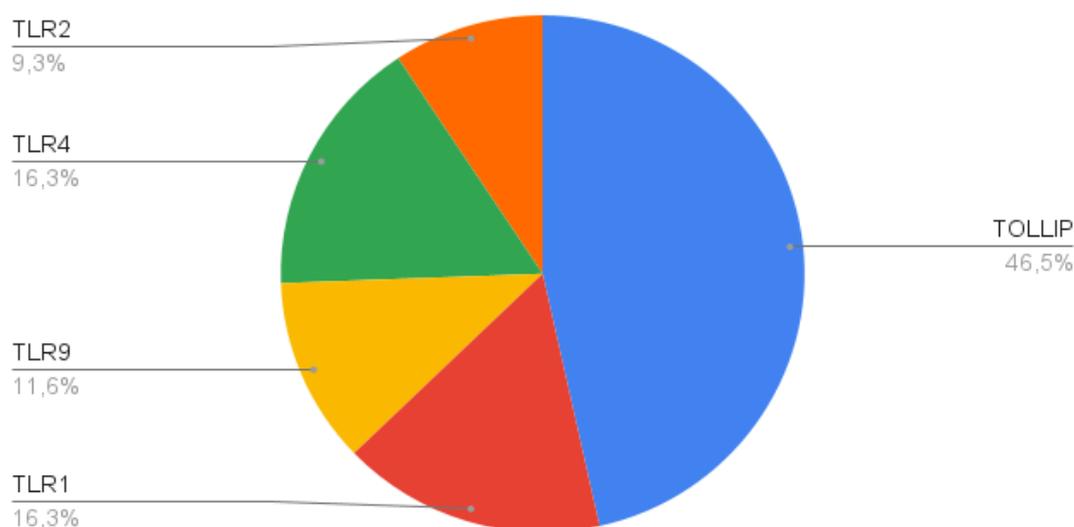
Gráfico 16 - Proporção de variantes associadas à Suscetibilidade versus Resistência na América do Sul no Intervalo 2.



Fonte: Autoria própria (2024)

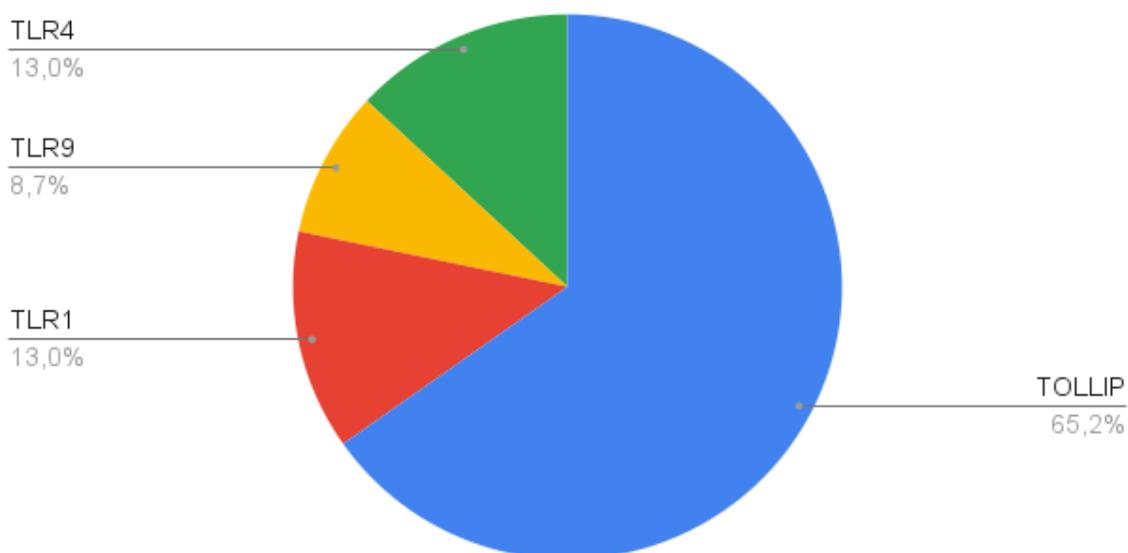
Esse aumento da resistência, em contextos mais recentes, acompanha uma mudança nos genes dominantes nas Américas nos distintos intervalos de tempo, com variações entre a parte Norte e Sul da região. De forma geral, corroborando os resultados anteriores, no intervalo 1 existe a predominância das variantes no gene TOLLIP nas Américas (gráfico 17), especialmente na América do Sul (gráfico 18). Todavia, não se observa essa predominância na América do Norte (gráfico 19); região em que predomina as alterações no gene TLR como apontado nos resultados anteriores.

Gráfico 17 - Proporção dos genes associados à TB nas Américas no Intervalo 1.



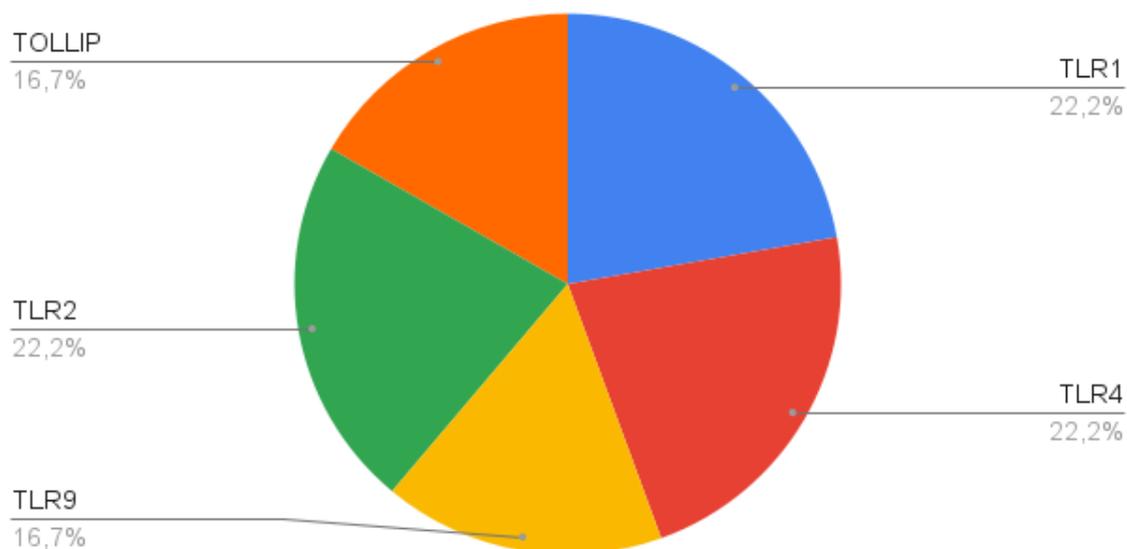
Fonte: Autoria própria (2024)

Gráfico 18 - Proporção dos genes associados à TB na América do Sul no Intervalo 1.



Fonte: Autoria própria (2024)

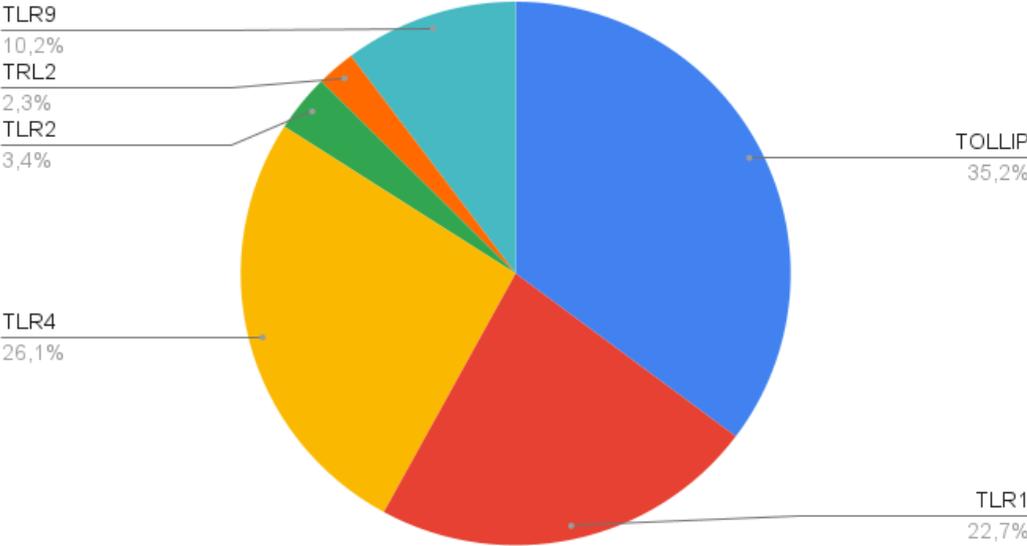
Gráfico 19 - Proporção de genes associados à TB na América do Norte no Intervalo 1.



Fonte: Autoria própria (2024)

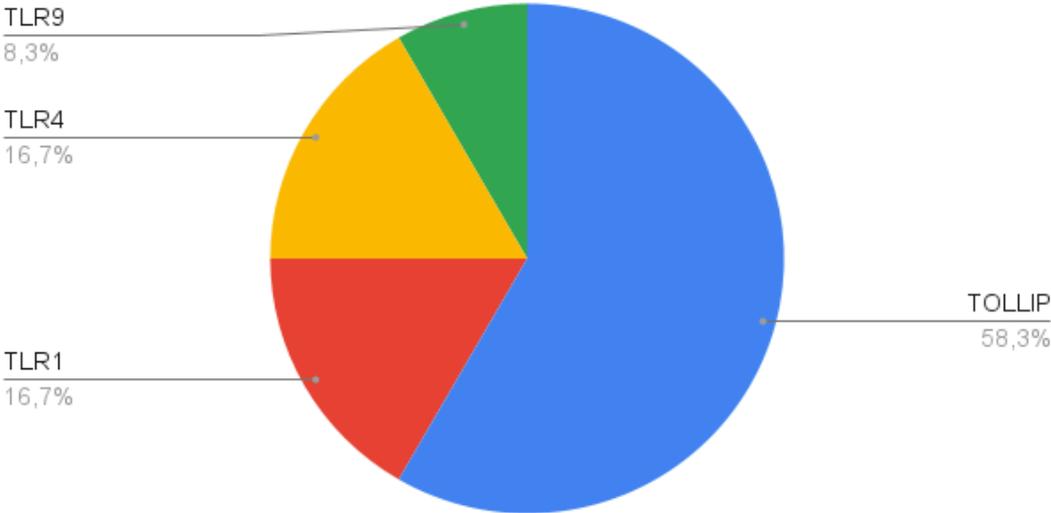
No segundo intervalo cronológico ainda é possível observar a predominância do gene TOLLIP, principalmente, na América do Sul. Entretanto, nesse intervalo mais recente, da mesma forma em que é possível visualizar um aumento nas alterações de resistência, também é possível observar um aumento na frequência dos gene TLR, principalmente do TLR4, especialmente, na América do Norte.

Gráfico 20 - Proporção de genes associados à TB nas Américas no Intervalo 2.



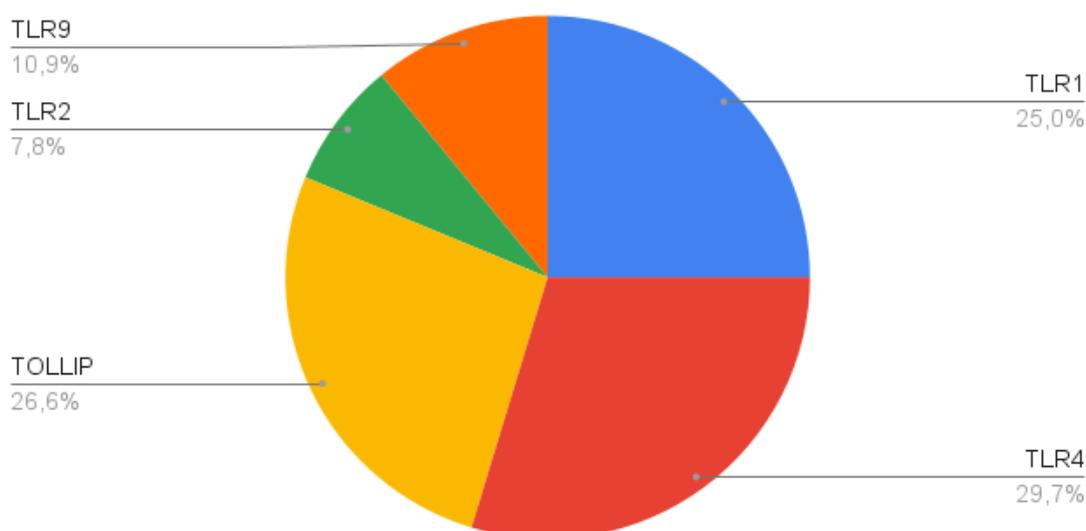
Fonte: Autoria própria (2024)

Gráfico 21 - Proporção de genes associados à TB na América do Sul no intervalo 2.



Fonte: Autoria própria (2024)

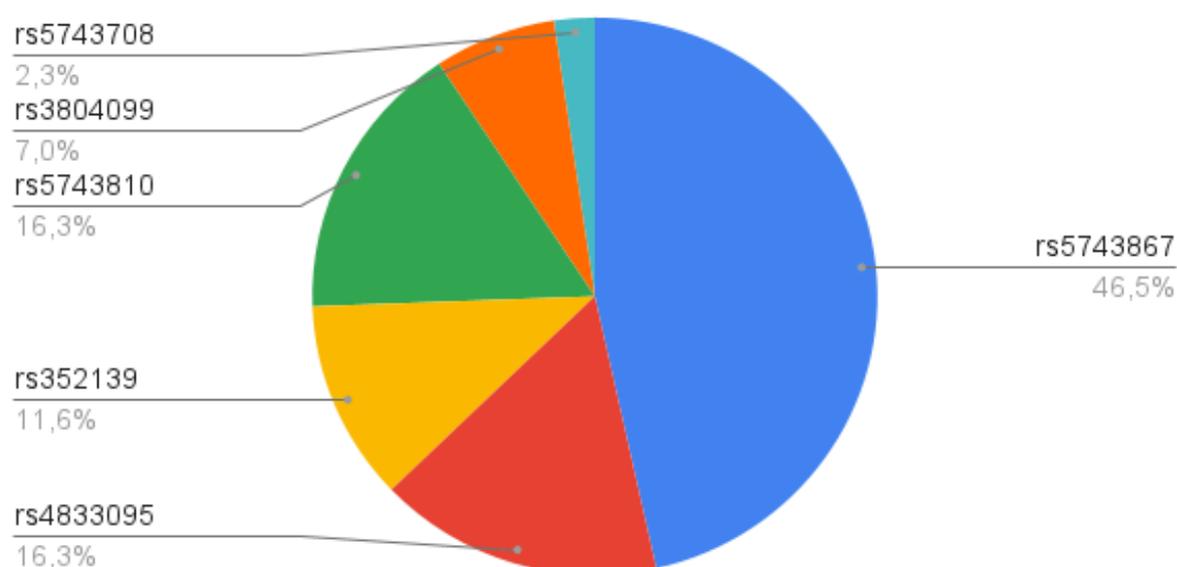
Gráfico 22 - Proporção de genes associados à TB na América do Norte no Intervalo 2.



Fonte: Autoria própria (2024)

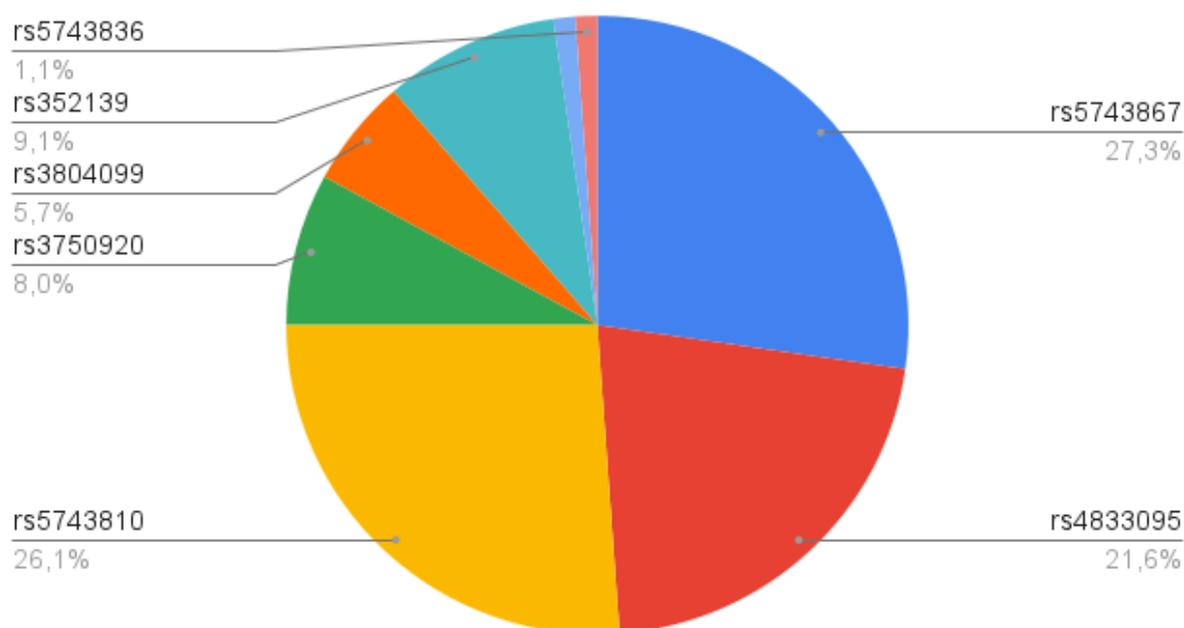
Ademais, finalizando o esboço cronológico, ao analisar a variação (SNP) predominante nas Américas nos distintos intervalos de tempo já discutidos, chegou-se aos seguintes resultados: Predominância da alteração **rs574367** – no gene TOLLIP – associada à Suscetibilidade, porém, existe um aumento das variações associadas ao gene TLR e o surgimento de uma nova variante, a **rs3750920**, também no gene TOLLIP, no intervalo 2 (gráfico 24).

Gráfico 23 - Proporção dos SNP associados à TB na América no Intervalo 1.



Fonte: Autoria própria (2024)

Gráfico 24 - Proporção dos SNP associados à TB na América no Intervalo 2.

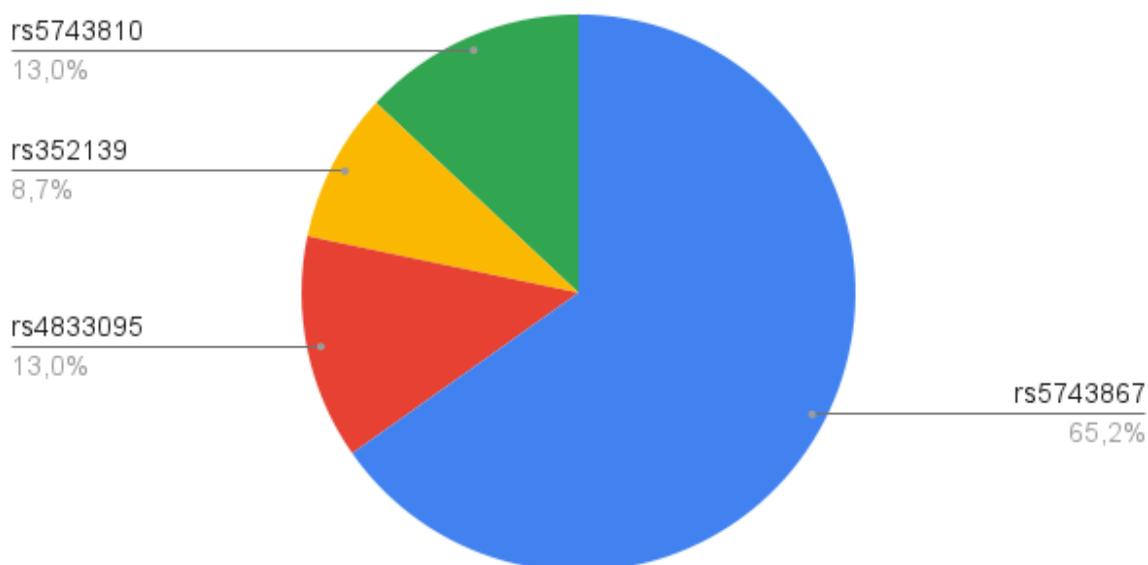


Fonte: Autoria própria (2024)

O surgimento dessa variante, que acomete o gene TOLLIP e é associada à resistência contra Tuberculose, acontece no mesmo cenário do aumento de resistência contra à infecção no intervalo de 5.000 a 500 anos AP; sua presença sendo mais constante na América do Norte e em contextos mais recentes.

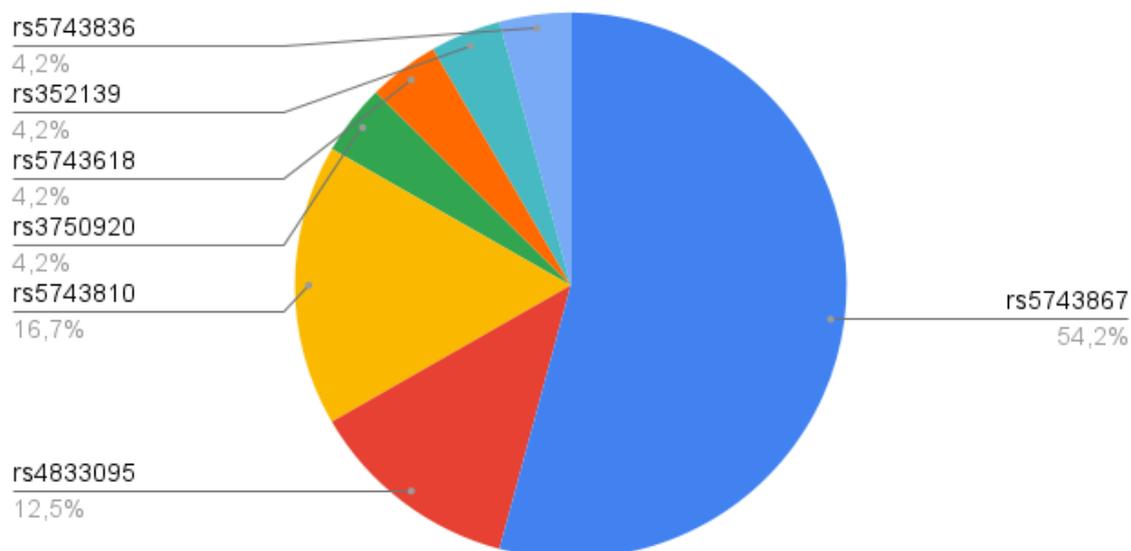
O cenário difere um pouco quando analisamos separadamente cada parte do continente. Na parte Sul, observamos a predominância da alteração **rs5743867** no gene TOLLIP associada à suscetibilidade, corroborando os resultados anteriormente apresentados. Essa predominância se perpetua em ambos os intervalos, entretanto, no segundo intervalo ela está associada a nova variante que surge nesse período – o SNP **rs3750920** – que, de acordo com os dados analisados, teve sua primeira aparição em um indivíduo do Chile de aproximadamente 4.500 anos AP.

Gráfico 25 - Proporção dos SNP associados à TB na América do Sul no Intervalo 1.



Fonte: Autoria própria (2024)

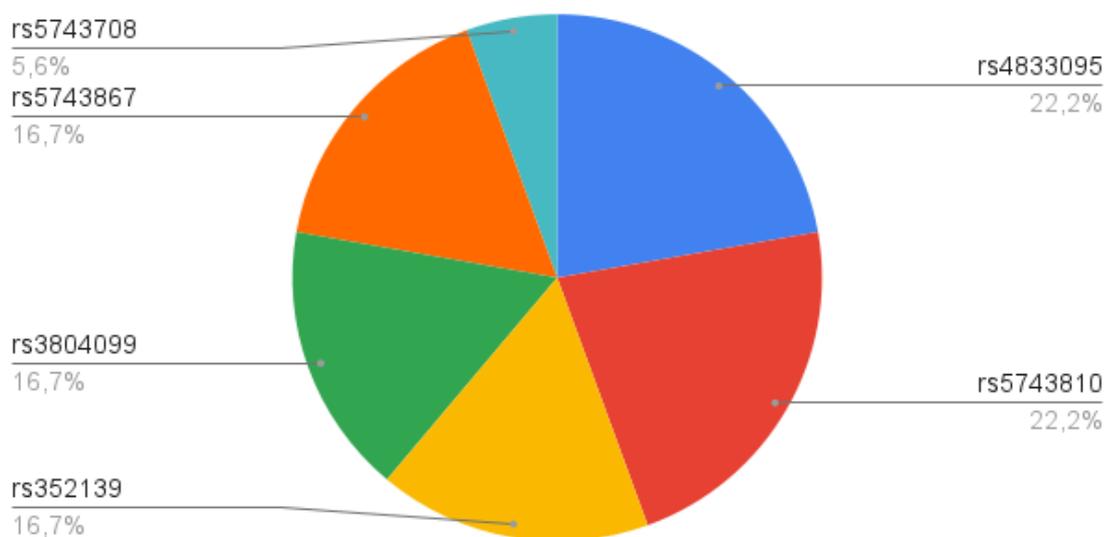
Gráfico 26 - Proporção dos SNP associados à TB na América do Sul no Intervalo 2.



Fonte: Autoria própria (2024)

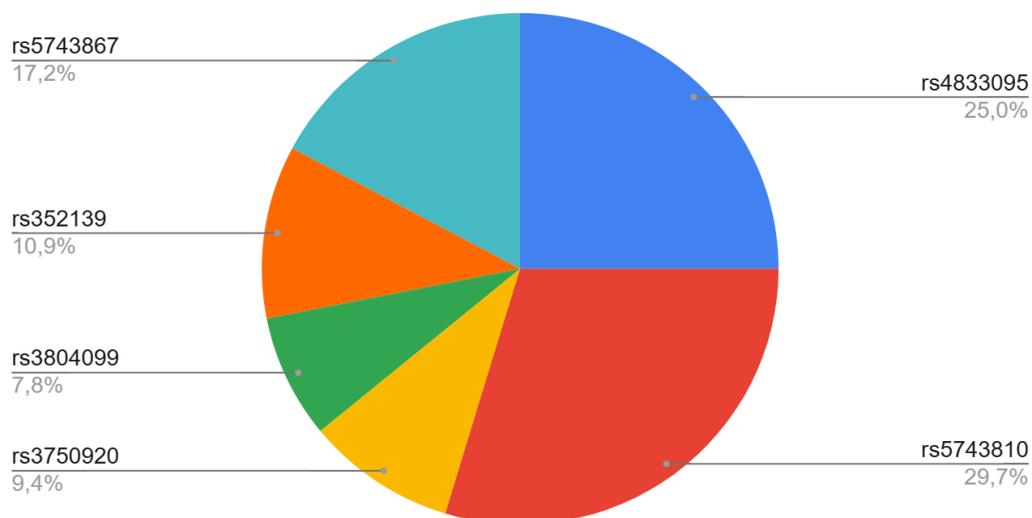
Para a parte Norte o cenário é um pouco distinto, como apontado pelos resultados anteriores. As variações predominantes são aquelas presentes no gene TLR, apresentando uma maior diversidade que o Sul. Além disso, a dominância TOLLIP apontada nos resultados anteriores em um contexto geral sem distinção cronológica, ocorre em função do surgimento da alteração **rs3750920**; esta que é muito mais comum na região Norte do que Sul, especialmente, em contextos mais próximos da colonização.

Gráfico 27 - Proporção de SNP associados à TB na América do Norte no Intervalo 1.



Fonte: Autoria própria (2024)

Gráfico 28 - Proporção dos SNP associados à TB na América do Norte no Intervalo 2.



Fonte: Autoria própria (2024)

Por fim, as análises cronológicas corroboram os resultados anteriormente observados e evidenciam uma maior variedade de alterações na parte Norte do continente. Além disso, evidenciam o surgimento de uma nova variante, em um contexto cronológico específico, o que pode sugerir que essa aparição esteja ligada com o aumento da resistência à infecção observado no mesmo período de tempo. Essa perspectiva pode nortear discussões sobre

possíveis cenários que conduziram a esse contexto, tais como: 1) Variação ambiental; 2) Mudança Populacional; 3) Nova variante do patógeno; dentre outros.

Nesse contexto, os presentes resultados apoiam a premissa da existência de marcadores associados à Tuberculose na América pré-contato, respondendo assim ao questionamento levantado. Dentre os polimorfismos encontrados, há uma predominância daqueles associados à suscetibilidade à infecção, o que levanta questões acerca de uma maior vulnerabilidade, a priori, por parte dos povos originários. Em razão disso, o gene que mais sofre alteração no contexto das Américas, incluindo a Mesoamérica observada pelas duas amostras do Belize, é o gene TOLLIP, mais especificamente alteração conhecida como **rs5743867** que oferece uma maior suscetibilidade à infecção, estando presente em diferentes indivíduos dentro de um mesmo sítio em países como: Saki Tzul (Belize), Lapa do Santo (Brasil), Gruta do Sumidouro (Brasil), Lucier (Canadá), Lovelock (Estados Unidos), Rio Uncallane (Peru), dentre outros.

Dito isso, cabe salientar que os SNP levantados são associados a L4 – linhagem europeia predominante nas Américas após a colonização (Joseph et al, 2023) – ou seja, pode-se não ter a mesma associação, de suscetibilidade ou resistência, a cepa existente nas Américas antes da colonização. Nessa perspectiva, é preciso pontuar que são conhecidas até o presente momento sete espécies de *Mycobacterium* que causam a tuberculose, sendo que esse conjunto constitui o chamado Complexo *Mycobacterium tuberculosis* (MTBC). Dentro desse complexo, existe a espécie *Mycobacterium tuberculosis* (M.tb), ou bacilo de Koch, que é a bactéria responsável pelo maior número de casos de tuberculose em humanos (Biblioteca Virtual de Saúde, 2007). Essa espécie, por sua vez, apresenta, até onde se sabe, sete distintas linhagens, diferenciadas por pequenas alterações em seu genoma, específicas de determinadas regiões. Dentre elas está a L4, ou linhagem Europeia, que é a variante do M.tb mais comum no mundo hodierno, tendo se tornado dominante nas Américas após a colonização.

Todavia, de acordo Joseph e colegas (2023), existia nas Américas antes da colonização uma cepa de Tuberculose com provável origem zoonótica. Para Krause e Trappe (2019), essa cepa, muito provavelmente, teria vindo da África para as Américas pelo mar, ou seja, uma transmissão zoonótica pelos animais marinhos. Isso porque, em concordância com os autores já citados, a cepa pré-colombiana (encontrada em múmias peruanas), como eles se referem em seu livro: “A jornada dos nossos genes: Uma história da Humanidade”, teria se originado do mesmo ramo da Tuberculose bovina, tendo o seu ancestral comum divergido a 5.000 mil anos em algum lugar da África.

Nesse sentido, as comparações genéticas entre esse patógeno encontrado nas múmias peruanas e a cepa atual da TB, assim como essa cepa pré-colombiana ainda ser encontrada em animais marinhos presentes no hemisfério sul (Krause e Trappe, 2019), corroboram uma possível transmissão dessa cepa por esse animais que teriam vindo da África para as Américas pelo atlântico (Krause e Trappe, 2019). Nesse panorama, esses dados vão contra a hipótese, por algum tempo aceita, de que a Tuberculose teria chegado nas Américas pela passagem do Estreito de Bering, a aproximadamente 15.000 anos. Todavia, ainda é incerto como se deu a chegada da doença nas Américas antes da colonização, da mesma forma que não é claro como a mesma chegou na Europa (Krause e Trappe, 2019).

Em contrapartida, observando as amostras analisadas e o indício de uma mesma alteração em indivíduos de um mesmo sítio com datações anteriores a 5.000 anos antes do presente, tais como: Brasil (Gruta do Sumidouro e Lapa do Santo) e do Peru (Lauricocha), se questiona se existiria outra cepa de Tuberculose, mais antiga que a pré-colombiana apresentada por Krause e Trappe (2019), visto que a repetição de alterações em indivíduos de um mesmo sítio pode ser um indicativo da doença anterior a 5 mil anos do presente. Esse cenário é corroborado pelos trabalhos abordados por Ujvari (2021), estes que encontraram a bactéria da Tuberculose, ou uma forma muito semelhante à mesma, em animais americanos há 17.000 anos no estado de Wyoming nos Estados Unidos.

Além disso, cabe destacar a oposição de genes e SNP encontrados nas distintas Américas. Tal cenário, em campo hipotético, pode ser enxergado como fruto de distintas rotas migratórias. Isso porque, as modificações no gene TOLLIP foram associadas, em alguns trabalhos, a populações africanas, enquanto os TLR são mais discutidos em populações asiáticas (Wu *et al.*, 2020). Logo, o estudo desses genes e suas variações, pode oferecer um novo caminho para se conhecer mais sobre o povoamento das Américas e as distintas rotas migratórias, ou serem um reflexo das épocas mais recentes – algumas de início da colonização – a que pertencem os indivíduos da América do Norte.

Tal conjectura é corroborada pelos resultados apresentados de uma perspectiva cronológica; esta que contorna a distinção e diversidade das variações no continente. Além disso, a interpretação por um olhar temporal, permitiu trazer a tona a presença de uma variação que só veio a se tornar comum na América em contextos mais recentes, tendo sua primeira aparição em um indivíduo a 4.500 anos AP. Essa variante, que acomete o gene TOLLIP e oferece uma maior resistência à infecção por Tuberculose, aparece no mesmo momento em que é possível visualizar um aumento na resistência dos indivíduos americanos, especialmente, os da América do Norte a priori. Ademais, da mesma forma que outras

variantes do gene TOLLIP associados à Tuberculose, esse SNP só foi encontrado em populações da Africanas.

Nesse sentido, partindo de uma data, torna-se possível traçar hipóteses sobre possíveis rotas migratórias humanas, ou do patógeno da Tuberculose, que podem ter levado ao surgimento dessa variação; bem como a possibilidade da mesma ser um cenário de adaptação a uma nova condição ambiental. Além disso, a perspectiva cronológica também permite esboçar cenários sobre a origem da Tuberculose em contexto americano, bem como abordar a hipótese de distintas bactérias do complexo MTBC conviverem com os nativos americanos antes do contato; o que os conduziu a distintas adaptações e variabilidades.

Por conseguinte, os polimorfismos levantados são estudados hodiernamente ligados à linhagem 4 (L4) da Tuberculose, que só chegou na América com a colonização (Joseph *et al.*, 2023), substituindo as prováveis linhagens existentes no continente. Esse panorama, nos leva a questionar se as alterações que foram encontradas nos indivíduos americanos, para a cepa pré-colombiana ou outro tipo, teriam as mesmas funções e consequência. Dessa forma, para esse contexto, propõem-se dois cenários possíveis: **1)** Em virtude da imunidade não ser tão robusta frente à diversidade de cepas existentes da Tuberculose (Tientcheu, *et al.*, 2017 *apud* Joseph *et al.*, 2023), os SNP encontrados também estariam associados à suscetibilidade e/ou resistência nesses indivíduos, mesmo sendo uma cepa diferente da L4, o que teria os deixado mais vulneráveis à nova variante, mais agressiva (Krause e Trappe, 2019), trazida pelos colonizadores europeus. **2)** Por se tratar de cepas distintas, os SNP encontrados poderiam ter associação a outros fatores, como outras doenças ou uma adaptação a um determinado ambiente ou estilo de vida, tal como ambiente de elevadas altitudes, visto que alguns sítios (Lovelock-EUA, Rio Uncallane-Peru, Gruta do Sumidouro e Lapa do Santo-Brasil, dentre outros) estão em ambientes elevados. Nesse sentido, a adaptação ao contexto em homeostase foi rompida com a chegada da nova variante, à qual eles não eram adaptados.

5 CONSIDERAÇÕES FINAIS

A área da Arqueogenética e da Arqueologia da Doença conduz ao limiar entre cultura e biologia revelando o quão interdisciplinar é a ciência arqueológica. As pesquisas iniciais dessa vertente lançaram luz aos estudos evolutivos e abriram um leque de possibilidades que se estende até os dias atuais.

Sendo uma das partes mais cativantes nas pesquisas de Arqueogenética, a evolução humana não é a única vertente sobre a qual essa abordagem se debruça. Em sua história, foi possível elencar pesquisas voltadas para rotas migratórias, ancestralidade e modelos de povoação e adaptação, sendo este último bastante proposto nos trabalhos desenvolvidos na América por meio dessa perspectiva.

Todavia, apesar das importantes contribuições adquiridas por essas análises, ainda são embrionários os trabalhos na região, em especial quando se fala de estudar o DNA antigo para inferir sobre doenças no passado e com isso, se possível, lançar luzes para o presente e o futuro. Hodiernamente, existem muitos trabalhos voltados para o estudo do DNA e a sua relação com as doenças infecciosas, como é o caso da Tuberculose, mas poucas ainda são realizadas com DNA antigo.

Apesar das inúmeras hipóteses que tratam parte do DNA como parte de antigos patógenos (Zimmerman, Barry; Zimmerman, David, 2022), como vírus e bactérias, pouco ainda se pesquisa, mesmo com as ferramentas existentes, sobre a interação entre esses patógenos e o ser humano, bem como tal relação afetou a adaptação humana e a própria evolução. Afinal as doenças infecciosas foram, e ainda são, uma das principais causas de mortalidade humana e, portanto, representam uma forte pressão seletiva na evolução humana (Murci, 2019).

No contexto da Tuberculose, inúmeros são os trabalhos, nos últimos cem anos, que estudam as alterações genéticas associadas à infecção como uma forma de compreender uma das doenças mais antigas ligada à espécie do *Homo sapiens*. Nessa perspectiva, a identificação e levantamento dessas alterações, mais especificamente polimorfismos de nucleotídeo único que acarretam em uma maior suscetibilidade ou resistência à infecção, junto com o aparato metodológico da bioinformática que permitiu a identificação dessas alterações nos indivíduos analisados e conduziu à elaboração de gráficos para uma visualização dos resultados – evidenciaram a presença dos polimorfismos associados à Tuberculose e auxiliaram na compreensão da distribuição e prevalência da doença nas distintas partes do continente americano.

Destarte, a presente pesquisa evidencia a presença de alterações relacionadas à Tísica pulmonar nos grupos pretéritos que ocupavam a América antes do contato, bem como a prevalência de uma suscetibilidade à infecção por parte desses indivíduos. Em contrapartida, os dados também revelam certa oposição, a priori, entre os genes e SNP mais observados na parte norte e sul do continente. Tal oposição, pode ser um reflexo das datações mais recentes para os indivíduos norte americanos, alguns datando do início da colonização, em comparação com os sulistas, ou ser uma nova possibilidade de estudo para rotas migratórias de povoamento das Américas, visto que o gene e SNP mais predominante na parte norte apresenta uma maior prevalência em grupos asiáticos enquanto os encontrados no sul foram observados, a priori, em populações da África.

À vista disso, os resultados levam a outros questionamentos, sendo um deles a origem da Tuberculose nas Américas. Isso porque, os resultados obtidos na presente pesquisa mostraram a repetição de uma mesma alteração em indivíduos de um mesmo sítio com datas referentes ao final do Pleistoceno (> 12.000 AP), o que permite a possibilidade de inferência da TB já existir no continente nesse período. Em contrapartida, os trabalhos referenciados, discorrem sobre uma cepa pré-colombiana no continente pré-contato que teria surgido na África, tendo origem no mesmo ramo que a Tuberculose bovina, e que teria chegado ao continente americano, provavelmente, pelo mar (Krause e Trappe, 2019) há aproximadamente 4.000 anos. Tais dados, por sua vez, não oferecem luz perante a possibilidade de presença de Tuberculose em indivíduos mais antigos a essa data, como pode ter acontecido no Brasil, no Chile, nos Estados Unidos e no Peru, sendo novos trabalhos e sequenciamento necessários para a melhor compreensão das diferentes rotas tomadas pela infecção na América.

Portanto, salienta-se a necessidade de mais sequenciamentos genéticos e pesquisas que se voltem para os estudos do patógeno da Tuberculose e sua interação hospedeiro-patógeno. Isso porque, existe uma grande diversidade dentro do complexo do *Mycobacterium tuberculosis* (MTBC) e essa diversidade pode ser manifesta em distintas formas de transmissão da doença, manifestações clínicas e respostas imunes, ou seja, isso sugere uma forte interação entre o hospedeiro e o patógeno que irá depender do genótipo do MTBC que irá influenciar diretamente na resposta imune da infecção (Tientcheu *et al.*, 2017), sendo este um campo frutífero para futuras pesquisas.

Destarte, os resultados obtidos ratificam a hipótese da presente pesquisa uma vez que as análises apontam para a presença de alterações genéticas associadas à Tuberculose nas

Américas antes do contato. Entretanto, a investigação acabou conduzindo a novos questionamentos propícios para pesquisas e trabalhos futuros. Dentre estes, salienta-se compreender a chegada, ou origem, da Tuberculose e como ela teria influenciado – exercido pressão seletiva – sobre os grupos humanos do continente Americano antes e após a colonização. Portanto, aponta-se a necessidade de mais sequenciamentos na área estudada; não só do DNA antigo humano como também o DNA do patógeno em questão, afinal, esses pequenos microrganismos não foram – não são – pequenos coadjuvantes no processo histórico e evolutivo humano.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, Ana Cristina Paes Barreto; PRADO, Alex Isidoro Ferreira; TAKENAMI, Lukary. Tuberculosis immunology: a narrative literature review. **Arq Asma Alerg Imunol** – Vol. 6, Nº 2, 2022.

BAKER, Oussama; LEE, Oona Y-C; *et al.* Human tuberculosis predates domestication in ancient Syria. **PubMed**, 1:S4-S12, junho, 2015. Doi: 10.1016/j.tube.2015.02.001.

BAYAT, Ardeshir. Bioinformatics. **BMJ**, v.324, doi 1018, abril, 2002.

Biblioteca Virtual em Saúde (Ministério da Saúde). Tuberculose. Novembro de 2007. Disponível em: <https://bvsms.saude.gov.br/tuberculose-21/#:~:text=O%20que%20%C3%A9%3F,descobridor%20da%20causa%20da%20doen%C3%A7a>). Acesso em: 16 de novembro de 2024.

BROSCH, Roland; *et al.* A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. **Proceedings of the National Academy of Sciences of United States of America (Proc Natl Acad Sci USA)**, 2002.

BUIKSTRA, Jane; ROBERTS, Charlotte. **The Archaeology of Disease**. The History Press, terceira edição, 2005.

BUIKSTRA, Jane; ROBERTS, Charlotte. **The Global History of Paleopathology - Pioneers and Prospects**. Oxford University Press, 15 de junho de 2012.

CAPELLUTO, Daniel G.S. Tollip: a multitasking protein in innate immunity and protein trafficking. **Microbes and Infection**, vol. 14, p. 140-147, 2012. Doi: <https://doi.org/10.1016/j.micinf.2011.08.018>

CASSANEGO, Samuel. O que é Bioinformática?. **PET Sistemas de Informação**, 2020. Disponível em: <https://www.ufsm.br/pet/sistemas-de-informacao/2020/12/03/o-que-e-a-bioinformatica>. Acesso em: 12 de janeiro de 2024.

CHEN, Ruifeng; *et al.* Human Toll-like receptor 2 genetics polymorphisms with tuberculosis susceptibility: A systematic review and meta-analysis. **Cytokine**, v. 172, doi 156405, dezembro, 2023.

COMAS, Iñaki; *et al.* Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. **Nature genetics**, v.45, p.1176-1182, 2014.

DANECEK, Petr; *et al.* Twelve years of SAMtools and BCFtools. **GigaScience**, 10 (2), 1–4, 2021. <https://doi.org/10.1093/gigascience/giab008>

FERNANDES, Nicollas Michell Brandim . Identificação De Ancestralidade Neandertal Nas Linhagens Genéticas De Indivíduos Antigos Do Brasil. **Monografia (Trabalho de Conclusão de Curso Graduação em Arqueologia)** - Centro de Filosofia e Ciências Humanas, Universidade Federal de Pernambuco, Recife, p.84, 2021.

FURIN, Jennifer; *et al.* Tuberculosis. **The Lancet**, v.393, i.10181, p.1642-1656, abril de 2019.

JOSEPH, Sophie, *et al.* Genomic evidence for adaptation to tuberculosis in the Andes before European contact. **iScience**, ScienceDirect, v.26, issues 2, 17 de fevereiro de 2023, 106034.

KERNER, Gaspard; *et al.* Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2000 years. **ScienceDirect**, 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0002929721000513>>. Acesso em: 10 de maio de 2023.

KRAUSE, Johannes; TRAPPE, Thomas. **A jornada dos nossos genes**. SEXTANTE, 2019.

LEAL, Rafael. Um estudo arqueogenético comparado de linhagens mitocondriais na América do Sul: espacialidade, derivação de haplogrupos e rotas migratórias entre 11.000 e 500 anos AP. **Monografia (Trabalho de Conclusão de Curso Graduação em Arqueologia)** - Centro de Filosofia e Ciências Humanas, Universidade Federal de Pernambuco, Recife, p.84, 2021.

LINDO, John; *et al.* The genetic prehistory of the Andean highlands 7000 years BP through European contact. **Sci. Adv.**, 4 (11), eaau4921. DOI: 10.1126/sciadv. aau4921, 2018.

LIU, CW., LIN, CJ., HU, HC. *et al.* The association of inflammasome and TLR2 gene polymorphisms with susceptibility to tuberculosis in the Han Taiwanese population. **Sci Rep** 10, 2020. <https://doi.org/10.1038/s41598-020-67299-6>

LONGHI, Renata Maronna Praça. Genetic polymorphism and immune response to tuberculosis in indigenous populations: a brief review. **SciELO Brasil**, junho de 2013. <https://doi.org/10.1016/j.bjid.2012.11.001> .

LUNA, Leandro Hernán; ARANDA, Claudia Marcela; *et al.* Oldest evidence of tuberculosis in Argentina: A multidisciplinary investigation in an adult male skeleton from Saujil, Tinogasta, Catamarca (905-1030 CE). **ScienceDirect**, v.125, dezembro, 2020.

MINISTÉRIO DA SAÚDE. **Brasil livre da Tuberculose: Plano Nacional pelo fim da Tuberculose como Problema de Saúde Pública**. Brasília, DF, 2017.

MURCI, Lluís Quintana. Human Immunology through the lens of evolutionary genetics. **Cell**, ScienceDirect, v.177, issues1, p.184-199, 21 de março, 2019. <https://doi.org/10.1016/j.cell.2019.02.033>

NELSON, Elizabeth; *et al.* Advances in the molecular detection of tuberculosis in pre-contact Andean South American. **International Journal of Paleopathology**, v.29, p.128-140, junho de 2020.

NOGUEIRA, Salvador; ZORZETTO, Ricardo. A Saúde nos Tempos do Imperador. **Pesquisa FAPESP**, 2013.

PÁLFI, György; *et al.* Tuberculosis in evolution. **Tuberculosis**, v.95, Supplement 1, junho, 2015.

OnlineBioinfo Bioinformática. O que é Bioinformática (e Biologia Computacional e Biocomputação)|Para que serve?. **Youtube**, 5 de julho de 2023.

RASMUSSEN, Morten *et al.* Ancient human genome sequence of an extinct PalaeoEskimo. **Nature**, v. 463, n. 7282, p. 757-762, 2010.

RENFREW, Colin; BAHN, Paul (Ed.). **Archaeology: The Key Concepts**. Abingdon: Routledge, 2005.

REINHARD, Andrew. Uma conversa sobre Archaeogaming com Andrew Reinhard. [Entrevista concedida a] Moura, Cleberson Henrique de *et al.* **Aedos**, Porto Alegre, v. 12, n. 26, p. 721-739, ago. 2020

REIS, Leonardo Mariano; RODRIGUES, Francisco W.; *et al.* Análise da relação de ocorrência de polimorfismo de nucleotídeo único do gene DOCK9 ceratocone. **Revista Brasileira de Oftalmologia**, 2016. Doi: <https://doi.org/10.5935/0034-7280.20160046>

ROBERTS, Charlotte; MANCHESTER, Keith. The study of Paleopathology.*In*: ROBERTS, Charlotte; MANCHESTER, Keith. **The Archeology of Disease**. The History Press, 2005, p. 23-70.

SABIN, Susan, *et al.* A seventeenth-century Mycobacterium tuberculosis genome supports a Neolithic emergence of the Mycobacterium tuberculosis complex. **Genome Biology**, 2020.

SANTOS, André Luis Campelo dos; SULLASI, Henry Lavalley; *et al.* Spatiotemporal fluctuations of population structure in the Americas revealed by a meta-analysis of the first decade of archeogenomes. **American Journal of Biological Anthropology**, Wiley Library, v.180, Issues 4, p. 703-714, dezembro de 2022.

SANTOS, André Luis Campelo dos. Paleomigrações na América: Uma Abordagem Arqueogenética. 2020. **Tese (Doutorado)** - Arqueologia, Universidade Federal de Pernambuco, Recife, 2020.

SHAH, J.A., VARY, J.C., CHAU, T.T., BANG, N.D., YEN N.T., FARRAR, J.J. *et al.* Human TOLLIP regulates TLR2 and TLR4 signaling and its polymorphisms are associated with susceptibility to tuberculosis. **J. Immunol**, v.189, p.1737–1746, 2012

STRAUSS, André Menezes. Arqueogenética: um olhar para o passado para garantir o futuro. **Jornal da USP**, 27 de outubro de 2022. Disponível em: <https://jornal.usp.br/artigos/arqueogenetica-um-olhar-para-o-passado-para-garantir-o-futuro/> . Acesso em: 24 de janeiro de 2024.

SILVA, Ilana Elisa Chaves. **Arqueologia da doença no cemitério histórico do Pilar - PE**. Dissertação (Mestrado em Arqueologia) - Centro de Filosofia e Ciências Humanas, Universidade Federal de Pernambuco. Recife, p.144. 2015.

SOUZA, Karolayne Silva, *et al.* **Polimorfismo em genes associados à suscetibilidade para a tuberculose**. Revista de Casos e Consultoria, v. 13, n. 1, e 13128387, p. 1 a 19, agosto de 2022.

STACIE D. A. BURKE. Tuberculosis: Past and Present. **Reviews in Anthropology**, v.40,issues 1, p. 27-52, março 2011. DOI: 10.1080/00938157.2011.547754

TEIXEIRA, Daniel de Azevedo. **Patologia Geral**. Teófilo Otoni (Minas Gerais), 2020.

TIENCHEU, Leopoldo D; *et al.* Immunological consequences of strain variation within the Mycobacterium tuberculosis complex. **European Journal of Immunology**, p. 432-445, 2017.

UJVARI, Stefan Cunha. **A História da Humanidade Contada pelos Vírus, bactérias, parasitas e outros microrganismos**. Editora Contexto, 2021.

VARSHNEY, Deepika; *et al.* Systematic review and meta-analysis of human Toll-like receptors genetic polymorphisms for susceptibility to tuberculosis infection. **Cytokine**, v.152, doi 155191, abril, 2022.

WU, Shouguan; LIU, Xiangmin *et al.* Polymorphisms of *TLR2*, *TLR4* and *TOLLIP* and tuberculosis in two independent studies. **PubMed Central**,2020. Disponível em:<<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7403954/>>. Acesso em: 5 de maio de 2023.

ZIMMERMAN, Barry E.; ZIMMERMAN, David J. **Medicina Macabra 3**. Ed. 1. Darkside, 17 de maio de 2022.