

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

ONDE OS RAROS SÃO COMUNS:

Relatos, estratégias e reflexões acerca dos "clusters" em genética médica populacional

Augusto César Cardoso dos Santos

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Doutor em Ciências (Genética e Biologia Molecular).

Orientadora: Prof^a. Dr^a. Lavínia Schüler Faccini

Porto Alegre, março de 2022

Esse trabalho foi desenvolvido no Laboratório de Genética Médica e Evolução do Departamento de Genética da Universidade Federal do Rio Grande do Sul (UFRGS) e contou com o apoio financeiro e institucional do Instituto Nacional de Genética Médica Populacional (INAGEMP), da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e do Hospital Infantil Albert Sabin (HIAS, em Fortaleza - CE).

O aluno teve dedicação exclusiva ao doutorado e recebeu bolsa do CNPq por 18 meses.

DEDICATÓRIA

*Dedicado, in memoriam, à Gabi (Gabriela Cardoso),
por seu legado em Genética Médica Populacional e pelas doces lembranças.*

AGRADECIMENTOS

Lav, pela generosidade, leveza e sorriso largo.

PPGBM e Maria Cátira (à época, coordenadora), por permitir que parte do trabalho fosse realizado afastado de sua sede.

Banca, pelas considerações.

Erlane, pela acolhida no Ceará e trabalho duro.

Roberto Giugliani, pelo apoio constante.

Virginia Ramallo *y todo el grupo de Argentina, por su ayuda con los apellidos.*

Marcelo Zagonel, por tantos mapas.

Luiza e Guillermo, pela ajuda com os trabalhos.

Renata Canalle, por me apresentar à GEMEPO.

Giovanny Pinto e Hygor, por me apresentarem à genética no laboratório.

Mestres geneticistas da UFPI/UFDPar, pela inspiração.

Ursula, Nelson, Fernanda, Luciana, Sídia, Lúcia, Elmo, Grande Salzano e tantos outros profissionais do PPGBM, por serem uma referência.

Camila, Raíssa, Renata e Yelena, pela amizade fulminante.

Thayne, pela generosidade e dicas preciosíssimas.

Perp, Juliano, Zu, Bibi, Mari, Gabi e todo o 113, por tanta troca e café.

Julia, por me acompanhar.

Giovanny França, por tanta confiança e ensinamentos.

Equipe UT-VAC e DASNT/SVS/Ministério da Saúde, pelo acolhimento e crescimento.

Work Family de Brasília, por sobrevivermos à pandemia trabalhando com dados da pandemia.

Andrea Lobo, pelo cuidado e doçura.

Laís, por me lembrar dos ipês.

Alexandra Elbakyan, por ampliar o acesso ao conhecimento.

Lula e Dilma, pela expansão das universidades e institutos federais brasileiros que me permitiu sonhar em ser doutor.

Lívia, por confiar em mim quando nem eu mais confiava.

Breno, por tanto carinho e parceria.

Tia Elizete, por me apresentar à literatura.

Vovó Fransquinha e vovô Pedro, por amarem tanto esse netinho.

Todas as comunidades trabalhadas, por tanta receptividade e confiança na Ciência.

Mãe, pai, irmãos, cunhada, sobrinhos, família, pelo apoio irrestrito.

EPÍGRAFE

*Ponho o pé no caminho
Sigo cantando
Passo compasso passando vivendo
Sempre cantando
Assim como a cigarra
Que no seu canto se agarra
Não quer saber de parar
Só para quando o papo estourar
Ah! Ah! Ah! Ah! Ah!
Sonhos de flor e espinho
Sigo cantando
Passo compasso passando vivendo
Sempre cantando
Assim como um pássaro
Que tem lá no sertão
Que se chama sofrer
Mas que também de alegria
Sabe cantar
La ra ra iá!
De alegria cantar porque sei
Que a dor não tem lugar permanente no
 coração de ninguém, de ninguém*

Sempre Cantando, Moraes Moreira (1975)

"Tenha coragem; encare o mundo"
José de Arimatea dos Santos (meu pai)

SUMÁRIO

LISTA DE ABREVIATURAS	4
LISTA DE FIGURAS, TABELAS E QUADROS	6
RESUMO	7
ABSTRACT.....	9
RESÚMEN	11
APRESENTAÇÃO	13
<i>De onde saiu essa tese? Aproximações entre ciência e cinema</i>	13
<i>Uma consideração técnica</i>	15
CAPÍTULO I: REVISÃO DE LITERATURA.....	17
<i>Genética Médica Populacional</i>	17
<i>Doenças raras</i>	18
<i>Doenças genéticas raras</i>	20
<i>O caso das mucopolissacaridoses</i>	22
<i>Anomalias congênitas</i>	24
<i>Clusters, rumores, isolados, consanguinidade e outros conceitos</i>	27
<i>O pioneirismo brasileiro</i>	28
<i>Os isolados genéticos</i>	31
<i>O caso da Finlândia</i>	34
<i>O estudo dos sobrenomes e o método isonímico</i>	36
<i>O espaço da tese: Ceará, Região Nordeste e América do Sul</i>	38
<i>Região Nordeste</i>	38
<i>Ceará</i>	41
<i>América do Sul</i>	43
CAPÍTULO II: JUSTIFICATIVA.....	46
CAPÍTULO III: OBJETIVOS	48
CAPÍTULO IV: <i>An invincible memory: what surname analysis tells us about history, health and population medical genetics in the Brazilian Northeast</i>	49
CAPÍTULO V: <i>The importance of geographical space in the characterization of mucopolysaccharidoses: clinical and molecular profile of a case-series from Northeast, Brazil</i>	66
CAPÍTULO VI: <i>El mapa de todos: clusters of rare diseases and congenital anomalies in South America</i>	86

CAPÍTULO VII: DISCUSSÃO	112
<i>Onde os raros são comuns?</i>	112
<i>Sobrenomes e genética médica populacional no Nordeste do Brasil</i>	115
<i>Concentração de MPS VI em municípios cearenses</i>	117
<i>Clusters na América do Sul</i>	119
CAPÍTULO VIII: CONCLUSÕES	123
CAPÍTULO IX: PERSPECTIVAS	126
CAPÍTULO X: REFERÊNCIAS BIBLIOGRÁFICAS	129
CAPÍTULO XI: APÊNDICES	141
<i>APÊNDICE 1: Demais trabalhos que detectaram a variante ARSB, c.1143–8T>G)</i>	141
<i>APÊNDICE 2: Reportagem para a BBC</i>	142
<i>APÊNDICE 3: scientia brasilis</i>	150

LISTA DE ABREVIATURAS

AHDC1	<i>AT-Hook DNA Binding Motif Containing 1</i>
ARSB	<i>Arylsulfatase B</i>
ASGR1	<i>Asialoglycoprotein Receptor 1</i>
BA	Bahia
CENISO	Censo Nacional de Isolados
CHIT1	<i>Chitinase 1</i>
CID-10	Classificação Internacional de Doenças, 10 ^a Revisão
CNV	<i>Copy Number Variants</i>
DF	Distrito Federal
DNA	Ácido Desoxirribonucleico
ECLAMC	Estudo Latino-Americano de Malformações Congênitas
EUROCAT	Rede Europeia de Vigilância das Anomalias Congênitas
F	Coeficiente de Endogamia
FHD	Doenças de Herança Finlandesa
GAG	Glicosaminoglicanos
GEMEPO	Genética Médica Populacional
GJB2	<i>Gap Junction Protein Beta 2</i>
GMI	Índice Global de Moran
HBB	<i>Hemoglobin Subunit Beta</i>
IDH	Índice de Desenvolvimento Humano
INAGEMP	Instituto Nacional de Genética Médica Populacional
INS	Isonimia
KLC2	<i>Kinesin Light Chain 2</i>
LD	Desequilíbrio de Ligação
MPS	Mucopolissacaridose
OMIM	<i>Online Inheritance in Men</i>
PI	Piauí
qPCR	Reação em Cadeia da Polimerase em tempo real
RN	Rio Grande do Norte

RS	Rio Grande do Sul
Sinasc	Sistema de Informações sobre Nascidos Vivos
SIM	Sistema de Informações sobre Mortalidade
SNP	<i>Single Nucleotide Polymorphism</i>
SPOAN	<i>Spastic Paraplegia, Optic Atrophy, and Neuropathy</i>
STR	<i>Short Tandem Repeats</i>
SUS	Sistema Único de Saúde
TRE	Terapia de Reposição Enzimática
UF	Unidade da Federação
UFRGS	Universidade Federal do Rio Grande do Sul

LISTA DE FIGURAS, TABELAS E QUADROS

Figuras

Figura 1: Genética Médica Populacional entre outros.	18
Figura 2: Oportunidades e ações, ao longo da vida, para a prevenção de anomalias congênicas.	25
Figura 3: Direção das principais ondas de povoamento do território finlandês.	35
Figura 4: Níveis de consanguinidade no território brasileiro calculados a partir de registros de casamentos católicos.	40
Figura 5: Caminhos dos primeiros migrantes Baianos e Pernambucanos.	42
Figura 6: Diversidade de climas da América do Sul (de acordo com a classificação climática de Koppen-Geiger).....	45
Figura 7: Esquema-resumo das principais metodologias e resultados que compõem a tese.	113

Tabelas

Tabela 1: Visão geral sobre as mucopolissacaridoses.	23
Tabela 2: Grau de parentesco, proporção estimada de genes em comum e coeficiente de endogamia (F) da prole de diferentes uniões consanguíneas.	32

Quadros

Quadro 1: O caso de Xia-Gibbs 21	21
Quadro 2: As redes internacionais para a vigilância de anomalias congênicas..... 26	26
Quadro 3: O INAGEMP 29	29
Quadro 4: “Mutações”, entre aspas..... 30	30
Quadro 5: Um cluster de genodermatose no litoral cearense 43	43
Quadro 6: Anomalias congênicas e doenças raras..... 112	112
Quadro 7: Tu, PI 115	115
Quadro 8: Vigilância de anomalias congênicas no Brasil 121	121
Quadro 9: Saúde Brasil 122	122

RESUMO

Neste trabalho, utilizamos diferentes abordagens para mapear e caracterizar aglomerados geográficos de doenças raras e anomalias congênitas, os *clusters*, em diferentes perspectivas. Os *clusters* tratam-se do principal objeto de estudo da genética médica populacional (GEMEPO), um ramo que aplica a genética médica para o estudo e atendimento de comunidades que apresentam questões médicas com fundo genético em uma frequência acima da esperada.

O Nordeste do Brasil apresenta-se como uma região estratégica para a condução de estudos na área de GEMEPO, pois exhibe, em comparação com as demais macrorregiões brasileiras, altas taxas de consanguinidade e uma alta concentração de *clusters*. Além disso, trata-se de uma região territorialmente ampla, com uma população numerosa e diversa, além de apresentar uma baixa quantidade de serviços de referência e profissionais na área de genética médica. Neste sentido, foi utilizada a análise de sobrenomes de mais de 37 milhões de pessoas, através do método isonímico, em conjunto com indicadores de saúde e demais informações históricas e sociodemográficas, a fim de estudar, de modo abrangente e barato, a população dos 1.794 municípios nordestinos.

Os principais resultados deste trabalho apontaram para uma distribuição isonímica heterogênea (*Global Moran Index*, $GMI = 0.58$; $p < 0.001$), com as maiores taxas concentradas em um grupo de municípios circunscritos na região história do Quilombo dos Palmares, o maior conglomerado espacial de escravizados fugidos da América Latina. Nestes locais, foram relatados rumores e aglomerados geográficos de doenças genéticas (como albinismo oculocutâneo e síndrome de Verma-Naumoff) e anomalias congênitas. Além disso, foi encontrada uma correlação positiva entre o índice de isonímia e a frequência de nascidos vivos com anomalias congênitas ($r = 0.268$; $p < 0.001$), com uma correlação espacial de ambos os indicadores ($GMI = 0.50$; $p < 0.001$).

Em outro trabalho, foi investigado retrospectivamente o perfil clínico, geográfico e molecular de 76 indivíduos com mucopolissacaridoses (MPS) diagnosticados no Ceará, um dos nove estados nordestinos. Nesta série de casos, a MPS II foi o tipo mais comum, com a maioria dos indivíduos afetados apresentando variantes patogênicas do tipo *missense*. Casos com MPS I, por sua vez, apresentaram os sintomas, foram diagnosticados e, eventualmente, morreram mais cedo que os demais tipos de MPS. O principal resultado do trabalho apontou para uma concentração espacial de 13 indivíduos com MPS VI na região leste do Ceará

compartilhando a mesma variante patogênica (*ARSB*, c.1143-8T>G) em homozigose, em municípios relativamente pequenos e com histórias de formação em comum. Os resultados foram discutidos à luz dos processos evolutivos e biossociais possivelmente relacionados com tal achado, além das suas implicações em termos de GEMEPO e saúde pública.

Finalmente, uma revisão sistemática qualitativa da literatura científica permitiu avaliar o cenário dos *clusters* em todo o continente sul-americano. Todos os países sul-americanos, exceto Guiana e Paraguai, apresentaram comunidades com alta prevalência de doenças raras e anomalias congênitas. Esta paisagem integrativa e diversa possibilitou reflexões acerca da formação e organização da população da América do Sul e ampliou o nosso entendimento sobre os *clusters* em si. O trabalho, portanto, foi ao encontro de estudos prévios que classificam a América do Sul (e a América Latina, em geral) como um *hot spot* para a presença de comunidades isoladas e com problemas genéticos, e pode representar o primeiro passo em direção à construção de um censo continental para o registro de *clusters* e rumores, aos moldes do brasileiro Censo Nacional de Isolados (CENISO).

Historicamente, logramos avanços científicos importantes em diversas áreas do conhecimento humano a partir dos trabalhos com *clusters*, sobretudo para a área da genética, como a identificação de genes e variantes patogênicas relacionados a doenças, estudo de traços complexos, descrição de novas entidades patológicas, etc. O mapeamento dos locais "onde os raros são comuns" é o primeiro passo em direção a um atendimento especializado da comunidade, além do suporte técnico e metodológico a profissionais de saúde locais. Também permite advogar por medidas de prevenção, quando possível, em se tratando de mitigação dos fatores de risco e aconselhamento genético das famílias afetadas.

Palavras-chave: doenças raras; anomalias congênitas; isolamento; consanguinidade.

ABSTRACT

In this study, different approaches were used to map and characterize geographical clusters of rare diseases and congenital anomalies in different perspectives. Clusters are the main object of study of population medical genetics (PMG), a branch that applies medical genetics to the study and care of communities that present medical issues with a genetic background at a higher-than-expected frequency.

The Northeast of Brazil is a strategic region for conducting studies in the PMG field, as it exhibits, when compared to other Brazilian macro-regions, high rates of consanguinity and a high concentration of clusters. Furthermore, it is a territorially wide region, with a large and diverse population, in addition to having a low number of reference services and professionals in the field of medical genetics. In this regard, surname analysis of more than 37 million people was used through the isonymic method, together with health indicators and other historical and sociodemographic information, in order to study, in a comprehensive and inexpensive way, the population of the 1,794 northeastern counties.

The main results of this study pointed to a heterogeneous isonymic distribution (*Global Moran Index*, $GMI = 0.58$; $p < 0.001$), with the highest rates concentrated in a group of counties circumscribed in the historical region of *Quilombo dos Palmares*, the largest spatial conglomerate of escaped enslaved people in Latin America. In these places, rumors and geographical clusters of genetic diseases (such as oculocutaneous albinism and Verma-Naumoff syndrome) and congenital anomalies have been reported. Moreover, a positive correlation was found between the isonymy index and the frequency of live births with congenital anomalies ($r = 0.268$; $p < 0.001$) with both indicators being spatially correlated ($GMI = 0.50$; $p < 0.001$).

In another study, the clinical, geographical and molecular profile of 76 individuals with mucopolysaccharidosis (MPS) diagnosed in *Ceará*, one of the nine northeastern states, was retrospectively investigated. In this case series, MPS II was the most common type, with most affected individuals having missense-type pathogenic variants. Cases with MPS I, in turn, showed symptoms, were diagnosed and eventually died earlier than other types of MPS. The main result of the study pointed to a spatial concentration of 13 individuals with MPS VI in the eastern region of *Ceará* sharing the same pathogenic variant (*ARSB*, c.1143-8T>G) in homozygosis, in relatively small counties and with common formation histories. The results

were discussed in light of the evolutionary and biosocial processes possibly related to this finding, in addition to their implications in terms of PMG and public health.

Finally, a qualitative systematic review of the scientific literature allowed us to evaluate the cluster scenario throughout the South American continent. All South American countries, except Guyana and Paraguay, had communities with a high prevalence of rare diseases and congenital anomalies. This integrative and diverse landscape allowed reflections on the formation and organization of the population of South America and broadened our understanding of the clusters themselves. The study, therefore, was in line with previous ones that classify South America (and Latin America in general) as a hot spot for the presence of isolated communities with genetic problems and may represent the first step towards the creation of a continental census for the registration of clusters and rumors, along the lines of the Brazilian National Census of Isolates (CENISO).

Historically, we have achieved important scientific advances in various areas of human knowledge from the work with clusters, especially in the area of genetics, such as the identification of genes and pathogenic variants related to diseases, the study of complex traits, the description of new pathological entities, etc. Mapping the places "where the rare are common" is the first step towards specialized community care, as well as technical and methodological support for local health professionals. It also allows advocating for preventive measures, where possible, when it comes to mitigating risk factors and genetic counseling of affected families.

Keywords: rare diseases; congenital anomalies; isolation; consanguinity.

RESÚMEN

En este estudio se utilizaron diferentes enfoques para mapear y caracterizar grupos geográficos de enfermedades raras y anomalías congénitas, los clústeres, desde diferentes perspectivas. Los clústeres son el principal objeto de estudio de la genética médica de poblaciones, rama que aplica la genética médica al estudio y atención de comunidades que presentan problemas médicos con antecedentes genéticos con una frecuencia mayor a la esperada.

El Nordeste de Brasil se presenta como una región estratégica para la realización de estudios en el área, ya que exhibe, en comparación con otras macro regiones brasileñas, altas tasas de consanguinidad y alta concentración de clústeres. Encima, es una región territorialmente amplia, con una población numerosa y diversa, además de contar con un número bajo de servicios de referencia y profesionales en el área de la genética médica. En este sentido, se utilizó el análisis de apellidos de más de 37 millones de personas, a través del método isonímico, junto con indicadores de salud y otras informaciones históricas y sociodemográficas, con el fin de estudiar, de forma integral y económica, la población de los 1.794 municipios del noreste.

Los principales resultados de este estudio apuntaron a una distribución isonímica heterogénea (Índice Moran Global, $GMI = 0,58$; $p < 0,001$), con las tasas más altas concentradas en un grupo de municipios circunscritos en la región histórica del *Quilombo dos Palmares*, el mayor conglomerado espacial de personas esclavizadas fugitivas en América Latina. En estos lugares se han reportado rumores y conglomerados geográficos de enfermedades genéticas (como el albinismo oculocutáneo y el síndrome de Verma-Naumoff) y anomalías congénitas. Además, se encontró una correlación positiva entre el índice de isonimia y la frecuencia de nacidos vivos con anomalías congénitas ($r = 0,268$; $p < 0,001$), con una correlación espacial de ambos indicadores ($GMI = 0,50$; $p < 0,001$).

En otro estudio, se investigó retrospectivamente el perfil clínico, geográfico y molecular de 76 individuos con mucopolisacaridosis (MPS) diagnosticados en Ceará, uno de los nueve estados del nordeste. En esta serie de casos, MPS II fue el tipo más común, y la mayoría de los individuos afectados tenían variantes patogénicas de tipo *missense*. Los casos con MPS I, a su vez, presentaban síntomas, eran diagnosticados y eventualmente fallecían antes que otros tipos de MPS. El principal resultado del estudio apuntó a una concentración espacial de 13 individuos con MPS VI en la región este de Ceará compartiendo la misma variante

patogénica (ARSB, c.1143-8T>G) en homocigosis, en municipios relativamente pequeños y con antecedentes de formación en común. Los resultados se discutieron a la luz de los procesos evolutivos y biosociales posiblemente relacionados con este hallazgo, además de sus implicaciones en términos de genética médica poblacional y salud pública.

Finalmente, una revisión sistemática cualitativa de la literatura científica permitió evaluar el escenario de clústeres en todo el continente sudamericano. Todos los países de América del Sur, excepto Guyana y Paraguay, tenían comunidades con una alta prevalencia de enfermedades raras y anomalías congénitas. Este paisaje integrador y diverso permitió reflexiones sobre la formación y organización de la población de América del Sur y amplió nuestra comprensión de los propios clústeres. El estudio, por tanto, estuvo en línea con estudios previos que catalogan a América del Sur (y Latinoamérica, en general) como un punto de acceso para la presencia de comunidades aisladas y con problemas genéticos y puede representar el primer paso hacia la creación de un censo continental para registrar clústeres y rumores, como el Censo Nacional de Aislados de Brasil (CENISO).

Históricamente, hemos logrado importantes avances científicos en diversas áreas del conocimiento humano a partir del trabajo con clústeres, especialmente en el área de la genética, como la identificación de genes y variantes patogénicas relacionadas con enfermedades, el estudio de rasgos complejos, la descripción de nuevas entidades patológicas, etc. El mapeo de lugares “donde lo raro es común” es el primer paso hacia la atención comunitaria especializada, además del apoyo técnico y metodológico a los profesionales locales de salud. También permite abogar por medidas preventivas, cuando sea posible, en términos de mitigación de factores de riesgo y asesoramiento genético para las familias afectadas.

Palabras clave: enfermedades raras; anomalías congénitas; aislamiento; consanguinidad.

APRESENTAÇÃO

De onde saiu essa tese? Aproximações entre ciência e cinema

Essa tese saiu de um filme. De um não, de alguns. O primeiro deles eu vi no quinto semestre de faculdade, no litoral do Piauí. Ele se chamava *Quatro Heranças* (2012), o documentário que minha professora de genética médica da graduação em Biomedicina havia adquirido em seu último congresso. O filme conta a história de investigação e atendimento de quatro comunidades brasileiras com uma alta prevalência de questões com forte componente genético em sua etiologia, geralmente raras. O filme girava em torno de uma área conhecida como "genética médica populacional", que eu acabara de descobrir.

Pela genética eu já havia optado assim que iniciei a faculdade, por ser um campo no qual eu encontrava respostas (e, principalmente, perguntas não-respondidas) sobre a vida terrestre e suas implicações, obtidas de modo mensurável e reprodutível nos laboratórios do *campus*. A genética médica uniu este conhecimento ao estudo de doenças e cuidado em saúde, importantes para a minha formação em Biomedicina e para a minha vontade de compartilhar conhecimento com outras pessoas. A genética médica populacional, a partir do filme, revelou-se um meio pelo qual eu poderia aplicar tudo isso em prol do bem-estar de comunidades, aproximando meus estudos científicos a outras áreas que também me pareciam importantes, como a história, antropologia, sociologia, epidemiologia e, como eu viria a descobrir, o cinema. Eu achava que tinha descoberto um caminho a seguir.

A confirmação veio com outro filme. Ou melhor, com uma citação a uma escola de cinema. Com uma rápida busca na internet, identifiquei quem eram os "atores" que apareciam no *Quatro Heranças* e supus que boa parte deles estaria no próximo congresso nacional de genética médica. Dito e feito! Na grade do evento, dentre as várias apresentações sobre as potencialidades do exoma (que eu também havia descoberto na ocasião do que se tratava) na prática clínica, uma me chamou atenção por possuir outro foco e por ser conduzida por alguém que eu conhecera através do filme sobre genética médica populacional: Lavínia.

Ao falar sobre uma comunidade do sul do Brasil que apresentava uma alta taxa de nascimentos gemelares (Tagliani-Ribeiro et al. 2012; De Oliveira et al. 2013; Matte et al. 2019), Lavínia mostrou um trabalho intitulado *So Close, So Far Away* (que, na época, eu imaginei ser uma referência ao filme *Faraway, So Close* (1993), do alemão Win Wenders, traduzido para o português como *Tão Longe, Tão Perto*) (De Oliveira et al. 2013). Porém, o clímax viria logo em seguida, com uma paráfrase que direcionaria boa parte da minha curta trajetória científica:

“Para fazer genética médica populacional”, disse ela, “basta um questionário na mão e uma ideia na cabeça”.

"Uma câmera na mão e uma ideia na cabeça" é uma célebre frase creditada a um dos principais cineastas brasileiros de todos os tempos, Glauber Rocha, diretor de filmes como *Deus e o Diabo na Terra do Sol* (1964) e *Terra em Transe* (1967). Segundo ele, uma câmera e uma ideia bastariam para se fazer um filme. A frase também se tornou lema do Cinema Novo, movimento vanguardista que, na década de 1960, propunha um modo de fazer cinema engajado na realidade e voltado para a transformação da sociedade brasileira (Cantarino 2007). Ao aproximar a genética do cinema, Lavínia me apresentou a um modo de pensar e fazer ciência que eu (sobretudo, o "eu" que estava prestes a me graduar) julgava apropriado e atraente.

Concluída a faculdade, estabelecidos os devidos contatos e após aprovação no programa de pós-graduação em que alguns dos trabalhos apresentados no *Quatro Heranças* foram realizados, aproximava-se a hora de escolher qual projeto seria abordado durante o meu mestrado. Via e-mail, minha futura orientadora havia me "oferecido" algumas possibilidades, dentre as quais um projeto para continuar os trabalhos sobre a "cidade dos gêmeos" e a opção de trabalhar com o tema científico mais quente naqueles idos de 2016: Zika e microcefalia (Schuler-Faccini et al. 2016; BRASIL 2019a).

Eu ainda morava em Parnaíba (PI), minha cidade natal, e precisava tomar uma decisão antes de me mudar para Porto Alegre (RS), a qual também veio através de um filme. Em uma sala improvisada de cinema de um secular casarão público parnaibano, assisti *O Médico Alemão* (2013), da diretora argentina Lucía Puenzo, a qual eu já conhecia por sua película anterior, *XXY* (2007), em que genética e sexualidade ocupam posições centrais. No filme de 2013, Puenzo fala sobre os últimos dias do médico nazista Joseph Mengele na América do Sul, o mesmo fato que motivou a população de Cândido Godói (a "cidade dos gêmeos" mencionada anteriormente), através de um político local, a procurar o Serviço de Genética Médica do Hospital de Clínicas de Porto Alegre, em 1994. Sua principal solicitação era a investigação científica da recorrência de nascimentos gemelares na região e testar a hipótese de que ela seria causada por uma suposta passagem de Mengele na pequena cidade durante sua trajetória pelo continente sul-americano (Matte et al. 2019). Estava decidido o meu projeto de mestrado.

Do mestrado veio o doutorado, que se fez em três regiões diferentes do país (sul - Porto Alegre, RS; nordeste: Fortaleza, CE; e centro-oeste: Brasília, DF) e, de lá para cá, em muitos outros momentos, nossos trabalhos em genética se aproximaram do cinema, da música, da literatura e de tantas outras áreas do conhecimento humano. Como preciso encerrar esta pequena introdução, sair da primeira pessoa e iniciar, de fato, a revisão de literatura que baseia

os trabalhos científicos presentes nessa tese, utilizarei como exemplo final *Varda por Agnès* (2019), o último filme da cineasta belga, radicada na França, Agnès Varda.

Neste documentário autobiográfico, Varda compartilha seus processos de criação e reflexões acerca do cinema, com foco especial em como contar uma história (que ela chama de “*cine-writing*”). Para ela, existem três palavras que nortearam a construção de todos os seus filmes: inspiração, criação e compartilhar. Segundo Varda, em tradução livre:

A inspiração é a razão pela qual se faz um filme. As motivações, ideias, circunstâncias e eventualidades que acendem o desejo e te fazem trabalhar para fazer um filme. A criação é como fazemos o filme. Que meios você usa? Que estrutura? Sozinha ou não? Em cores ou não? A criação é o trabalho. A terceira é compartilhar. Não fazemos filmes para assistir sozinhos, fazemos para mostrá-los. No fundo, é preciso saber o porquê daquele trabalho.

Assim como para o cinema de Varda, inspiração, criação e compartilhar também foram processos fundamentais para a construção e comunicação dos trabalhos científicos que aqui serão mostrados. Afinal, todo trabalho científico parte da formulação de hipóteses sobre certas questões (“inspiração”) a serem testadas com rigor metodológico, para obter resultados reprodutíveis (“criação”) e descritos de modo a comunicar ciência e interagir com trabalhos prévios, através de artigos científicos, instrumentos acadêmicos, palestras, dentre outros (“compartilhar”). Ainda em sua obra, Varda justifica o seu processo de criação cinematográfica: “Não é a necessidade de fazer imagens, é que... tenho olhos curiosos”. De modo semelhante, essa tese não foi feita pela necessidade de fazer artigos científicos, é que... também temos olhos curiosos.

Uma consideração técnica

Nessa tese, além dos três trabalhos principais que serão apresentados nos capítulos IV, V e VI, serão destacadas outras produções científicas desenvolvidas pelo autor principal e sua orientadora, em parceria com grupos de pesquisa e assistência à saúde, e que foram publicadas durante o período correspondente ao doutorado. Assim, ao longo do texto, trechos marcados **em azul** farão correspondência com quadros informativos laterais, que trarão um breve resumo,

informações técnicas e link para acesso ao material já publicado. Uma lista de todos os quadros mencionados pode ser encontrada na seção “LISTA DE FIGURAS, TABELAS E QUADROS”.

CAPÍTULO I: REVISÃO DE LITERATURA

Genética Médica Populacional

As implicações da hereditariedade na saúde dos indivíduos e suas famílias constituem as bases da genética médica. Porém, certas doenças e demais questões com influência genética podem afetar não apenas indivíduos e seus familiares, mas também toda uma comunidade. A aplicação dos conhecimentos de genética médica para o estudo e a assistência à saúde de populações específicas constitui as bases da genética médica populacional, por vezes referenciada pelo acrônimo GEMEPO (Castilla 2005; Giugliani 2012a; Giugliani et al. 2019b).

Podemos dizer que, assim como a família está para a genética médica, a população está para a genética médica populacional. Deste modo, o foco de trabalho mais importante em GEMEPO são as comunidades que apresentam uma alta frequência de doenças genéticas ou anomalias congênitas (Castilla 2005; Giugliani et al. 2019b). Como mostrado na **Figura 1**, ela surge na fronteira entre outras subáreas da genética, como a já mencionada genética médica; a genética de populações, que investiga populações a partir de um enfoque evolutivo e cruzando com aspectos da antropologia; epidemiologia genética, tradicionalmente envolvida no estudo de doenças crônicas comuns de etiologia poligênica; e a genética comunitária, que está mais relacionada à saúde coletiva (Giugliani 2012a).

Portanto, GEMEPO é diferente de genética comunitária. Este último termo foi introduzido na literatura científica na década de 1990, constantemente definido como uma "ponte" entre a genética médica e a saúde pública (Ten Kate 2005; Aswini and Varun 2010). O principal objetivo da genética comunitária é o fornecimento de serviços de genética para a comunidade como um todo (seja ela um vilarejo isolado ou a população de um país), com um importante foco em programas de *screening* neonatal e aconselhamento genético (Ten Kate 1998; Ramalho and Silva 2000; Ten Kate 2005).

Além disso, é importante mencionar que "genética comunitária" pode ter um significado mais abrangente se estiver fazendo referência ao ramo das ciências biológicas que se dedica a entender como as dinâmicas entre ecologia e evolução moldam comunidades e ecossistemas, sem se limitar à espécie humana (Hersch-Green et al. 2011). Por sua vez, o termo "genética médica populacional" (ou suas variantes, como "genética médica em nível populacional") vem sendo amplamente utilizado em um contexto voltado mais especificamente para a saúde de comunidades específicas, sobretudo em trabalhos latino-americanos (Giugliani

2012a; Castilla and Schuler-Faccini 2014; Poletta et al. 2014; Cardoso et al. 2018; Giugliani et al. 2019b).

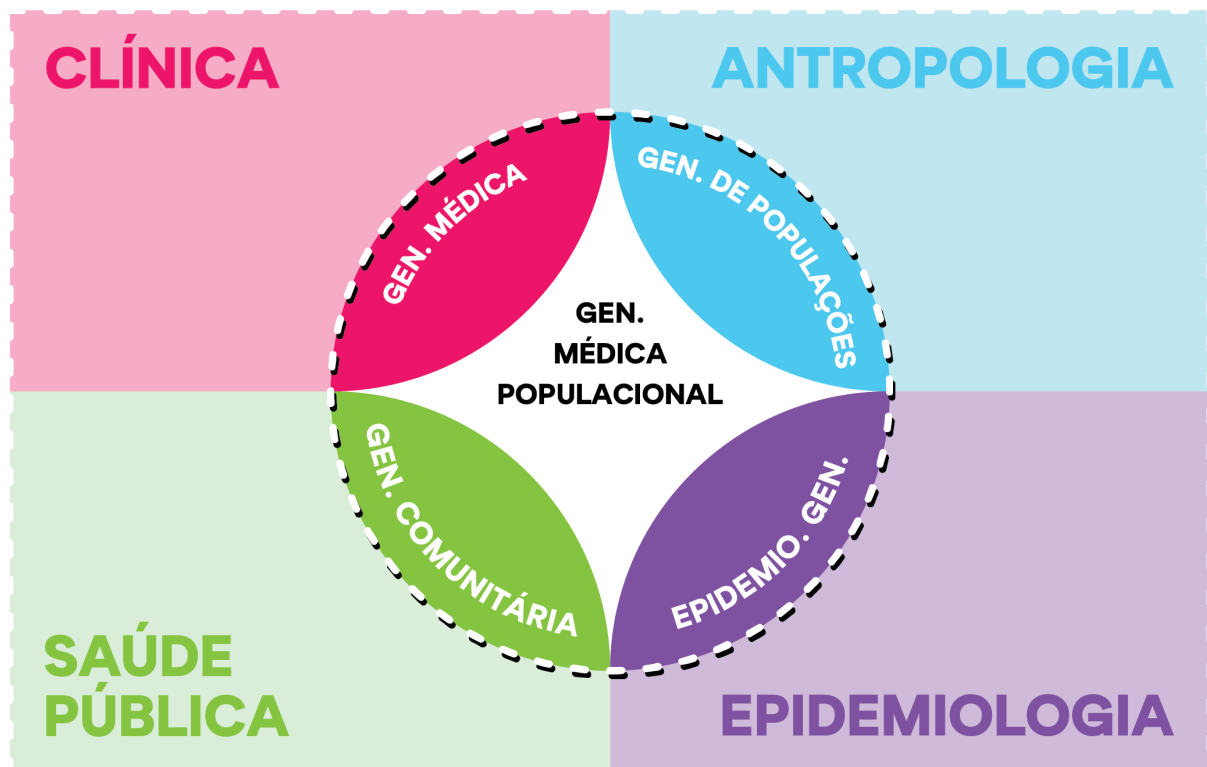


Figura 1: Genética Médica Populacional entre outros.

Fonte: Adaptado de Castilla e Schuler-Faccini, em comunicação pessoal.

O Brasil possui papel de pioneirismo no estudo de genética médica em nível populacional. Isto pode ser atestado através dos clássicos estudos conduzidos pelo geneticista paranaense Newton Freire-Maia que, a partir da década de 1950, investigaram múltiplos cenários e fatores envolvidos com a alta frequência de doenças genéticas em determinadas comunidades brasileiras (Freire-Maia 1957; Freire-Maia et al. 1978; Freire-Maia 1990). Além disso, foi criado em 2008, pelo Governo Federal, o Instituto Nacional de Genética Médica Populacional (INAGEMP), para apoiar o estudo e o atendimento de populações com alta prevalência de doenças raras (ou com um forte componente genético) e/ou de anomalias congênitas (Castilla and Schuler-Faccini 2014; Giugliani et al. 2019b), que serão abordadas a seguir.

Doenças raras

Doenças raras correspondem a um grupo numeroso de condições pouco frequentes, em comparação com outras doenças amplamente prevalentes na população em geral. Caracterizam-se como um grupo heterogêneo de doenças que geralmente se apresentam de maneira crônica ou progressiva e podem afetar qualquer sistema humano (Schieppati et al. 2008; Melnikova 2012; Richter et al. 2015).

Na União Europeia, uma doença é considerada rara quando afeta menos de 5 pessoas em 10.000 (WHO 2013). Nos Estados Unidos, é qualquer condição que afeta menos de 200.000 pessoas, o que corresponde, a uma prevalência de aproximadamente 1 em 1.630, considerando a população atual do país (USA 1992; Ferreira 2019). Já no Brasil, a frequência adotada pelo Ministério da Saúde é de 65 pessoas a cada 100.000, isto é, 1,3 afetados por 2.000 (Brasil 2014). Mundialmente, as estimativas podem variar de 5 a 80 por 100.000 (Ferreira 2019; Wakap et al. 2020).

Entretanto, se individualmente afetam poucas pessoas, coletivamente existem entre 5.000 e 8.000 diferentes tipos de doenças raras, que ocorrem entre 3,5% a 5,9% da população global (cerca de 263 a 446 milhões de indivíduos), elevando as doenças raras para uma questão de saúde pública (Boycott et al. 2019; Wakap et al. 2020). Esta prevalência global foi estimada a partir dos dados sobre 6.172 doenças raras únicas contidas no Orphanet, um portal sobre doenças raras apoiado pela Comissão Europeia que é uma importante referência para a temática em todo o mundo (Ferreira 2019; Orphanet 2022).

Entre 70% e 80% das doenças raras possuem causas genéticas reconhecidas – o restante é devido a causas infecciosas, imunológicas, desconhecidas, dentre outras (Wakap et al. 2020; Orphanet 2022). Entretanto, o diagnóstico e a pesquisa etiológica das doenças raras podem constituir tarefa complexa e desafiadora, e o cenário global destas doenças é dinâmico, visto que novas descobertas moleculares e clínicas são descritas regularmente, com potencial para alterar os números e estimativas mencionados há pouco (Boycott et al. 2019; Ferreira 2019; Lee et al. 2020).

Do ponto de vista clínico, a investigação genética destas doenças permite diagnóstico acurado, orientação terapêutica e aconselhamento genético; do ponto de vista científico, permite expandir o conhecimento não apenas sobre a patologia, mas sobre a fisiologia dos sistemas orgânicos e componentes genômicos que compõem nossa biologia (Giugliani et al. 2019a). Todavia, por muito tempo, o estudo das doenças raras manteve-se negligenciado (Federhen et al. 2014).

Para a maior parte das doenças raras não há cura ou biomarcadores validados e o pequeno número de casos e o conhecimento científico limitado dificultam o reconhecimento

clínico ou laboratorial de muitas delas (Melnikova 2012; WHO 2013). Também é importante destacar o baixo interesse da indústria farmacêutica em desenvolver e comercializar produtos terapêuticos inovadores voltados para o pequeno "mercado consumidor", motivo pelo qual os medicamentos para tais doenças são comumente conhecidos como "drogas órfãs" (USA 1992; WHO 2013; Lee et al. 2020).

O diagnóstico, assistência à saúde e pesquisa sobre doenças raras em um país emergente como o Brasil, constitui-se uma tarefa complexa (Federhen et al. 2014; Passos-Bueno et al. 2014; Giugliani et al. 2019a). No Brasil, a genética médica tem uma história recente, mas pioneira em se tratando de América Latina, com seus primeiros passos ocorrendo em meados da década de 1950. Embora a importância do campo tenha sido reconhecida pela comunidade médica já na década de 1960, a genética médica como disciplina permanece como opcional em muitas faculdades de medicina e demais de cursos na área de saúde (Novoa and Burnham 2011; Passos-Bueno et al. 2014).

Em âmbito governamental, a “Política Nacional de Atenção Integral às Pessoas com Doenças Raras”, ou Portaria N° 199/2014, estabeleceu diretrizes para o atendimento integral a indivíduos afetados por estas doenças no Sistema Único de Saúde (SUS) (Brasil 2014). Tal política surgiu como uma consequência de um grande esforço e pressão de parte da sociedade brasileira, objetivando reduzir a mortalidade e morbidade para indivíduos com doenças raras (Federhen et al. 2014). Todavia, existem apenas duas dezenas de centros de referência para doenças raras distribuídos de maneira irregular no território brasileiro, localizando-se sobretudo nas principais cidades do país, com vínculo a algum hospital universitário (Schwartz et al. 2021).

Com o objetivo de aumentar o conhecimento epidemiológico sobre as doenças raras no Brasil e ajudar os gestores em saúde a melhorar as políticas voltadas para a temática no país, encontra-se em fase de desenvolvimento um estudo multicêntrico envolvendo 40 instituições nacionais, dentre as quais 18 hospitais universitários, 17 serviços de referência para doenças raras e cinco serviços de triagem neonatal. Esta rede de colaboração é a primeira iniciativa de uma grande coleção de dados epidemiológicos sobre doenças raras na América Latina (Félix et al. 2022).

Doenças genéticas raras

Como exposto, a grande maioria das doenças raras possui causa genética. O estudo de (Wakap et al. 2020) mostrou que, dentre as doenças raras cadastradas na Orphanet, cerca de 71.9% apresentavam origem genética. Destas, cerca de 35% exibiam um padrão de herança autossômico recessivo, 24% autossômico dominante, 7% eram anomalias cromossômicas e 5% foram de doenças recessivas ligadas ao X.

A presença de raros trechos do genoma que divergem das sequências de referência, aliada à interpretação molecular e clínica dos pacientes, confirma o diagnóstico de doenças genéticas raras e também contribui de maneira importante para o prognóstico do paciente e para seus familiares. Isto porque a identificação de variantes genéticas patogênicas abre possibilidades de aconselhamento genético, investigação pré-natal de futuras gestações na família, dentre outros (Evans et al. 2017; Pogue et al. 2018). Estima-se que cerca de 50 a 60 novas doenças genéticas são registradas a cada ano nas bases de dados que exploram a temática, notavelmente a supramencionada Orphanet (mais especificamente, o Orphadata, que corresponde à base de dados do portal) e o OMIM (do inglês, *Online Inheritance in Men*), um catálogo de genes e doenças genéticas humanos (<https://www.omim.org/>) (Lee et al. 2020).

Avanços tecnológicos na área da biologia molecular, como o sequenciamento de **exoma** ou de genomas, vêm aumentando consideravelmente o nosso entendimento sobre doenças raras e sobre nossa própria biologia (Pogue et al. 2018; Boycott et al. 2019). Além das análises de sequenciamento de DNA, que identificam mudanças na sequência de nucleotídeos que resultam em um aminoácido alterado ou em uma parada prematura da tradução protéica, existem metodologias específicas para a identificação de variações citogenéticas, como a análise de microarranjos cromossômicos,

Quadro 1: O caso de Xia-Gibbs

Em um artigo científico publicado em 2020 na revista *Molecular Syndromology* (11:24-29), relatamos o primeiro caso brasileiro de síndrome de Xia-Gibbs, uma doença genética rara (pouco mais de 100 casos relatados mundialmente) autossômica dominante caracterizada por atrasos neurodesenvolvimentais globais, hipotonia, convulsões e pela presença de variantes patogênicas no gene *AHDC1*. Por meio da análise de **exoma** do caso e sequenciamento dos genitores, identificamos a variante c.451C>T; p.Arg151*, que correspondeu à variante mais próxima da região aminoterminal na proteína já descrita até então. Os resultados foram discutidos em termos de correlação genótipo-fenótipo e em contexto com todos os outros casos já relatados na literatura científica.

Acesso:

<https://bitly.com/bIzUh>



como no caso de doenças causadas por deleções e ampliações (como *copy number variants* ou apenas CNV) cromossômicas (Pogue et al. 2018).

De acordo com Giugliani et al. (2019b), o interesse em doenças genéticas raras vem aumentando nos últimos anos por alguns fatores principais: (1) novas doenças são continuamente descritas; (2) muitos indivíduos com doenças raras requerem atenção e políticas públicas dos governos; e (3) o estudo de doenças genéticas contribui para elucidação de processos fisiológicos ou relacionados a doenças não-genéticas. Sobre este último ponto, vale mencionar que o estudo de doenças genéticas raras vem ajudando na descoberta sobre o papel dos genes e de proteínas, na compreensão de aspectos relacionados a sequências genômicas não-codificantes, fatores modificadores, dentre outros (Wexler et al. 2004; Giugliani 2012a; Makrythanasis and Antonarakis 2013).

De fato, desde os estudos pioneiros de Archibald Garrod com erros inatos de metabolismo, percebeu-se que o estudo das doenças de origem genética pode oferecer uma ampla oportunidade para investigar fenômenos fisiológicos e evolutivos na nossa espécie (Pogue et al. 2018; Lee et al. 2020). Para (Lee et al. 2020), em tradução livre:

Doenças genéticas humanas raras podem ser vistas como resultados da triagem genética avançada que a natureza tem continuamente executado em nós desde o surgimento de nossa espécie.

A seguir, detalharemos o caso das mucopolissacaridoses (MPS) como um exemplo de grupo de doenças genéticas raras, as quais também serão abordadas no Capítulo V desta tese.

O caso das mucopolissacaridoses

As MPS são um grupo de doenças raras causadas pela deficiência de enzimas lisossomais que afetam a degradação dos glicosaminoglicanos (GAG). O acúmulo de GAG em diferentes tipos de células resulta em progressivo dano celular, o qual pode atingir múltiplos órgãos e sistemas e levar a falhas orgânicas, alterações cognitivas e redução na expectativa de vida entre os indivíduos com a condição (Muenzer 2011; Giugliani 2012b).

Existem sete tipos principais de MPS causados pelo defeito ou alteração em 11 enzimas distintas, a saber: MPS I (ou síndrome de Hurler, Hurler-Scheie ou Scheie), enzima α -L-iduronidase; MPS II (ou síndrome de Hunter), enzima iduronato-2-sulfatase; MPS III A-D

(síndrome de Sanfilippo), causada por deficiência em quatro diferentes tipos de enzimas (A - heparan N-sulfatase, B - α -N-acetilglicosaminidase, C - α -glicosaminida acetiltransferase e D - N-acetilglucosamina 6-sulfatase); MPS IV A e B (ou síndrome de Morquio), enzimas galactose 6-sulfatase (A) e β -galactosidase (B); MPS VI (ou síndrome de Maroteaux-Lamy), enzima arilsulfatase B; MPS VII (ou síndrome de Sly), enzima β -glucuronidase e MPS IX (ou síndrome de Natowicz), enzima hialuronidase (Muenzer 2011). As MPS são herdadas em um padrão autossômico recessivo, com exceção da MPS II, que apresenta um padrão de doença ligado ao X recessivo (Muenzer 2011; Giugliani 2012a; Celik et al. 2021). A **Tabela 1** resume as principais características bioquímicas e moleculares das MPS.

Tabela 1: Visão geral sobre as mucopolissacaridoses.

Tipo de MPS	GAG relacionado	Enzima deficiente	Locus (gene)	OMIM
I (síndrome de Hurler, Hurler-Scheie, Scheie)	dermatan sulfato, heparan sulfato	α -L-iduronidase	4p16.3 (<i>IDUA</i>)	# 607016
II (síndrome de Hunter) ^a	dermatan sulfato, heparan sulfato	iduronato-2-sulfatase	Xq28 (<i>IDS</i>)	# 309900
III A - D (síndrome de Sanfilippo)	heparan sulfato	A: heparan N-sulfatase B: α -N-acetilglicosaminidase C: α -glicosaminida acetiltransferase D: N-acetilglucosamina 6-sulfatase	A: 17q25.3 (<i>SGSH</i>) B: 17q21 (<i>NAGLU</i>) C: 8p11.1 (<i>HGSNAT</i>) D: 12q14.3 (<i>GNS</i>)	A: # 252900 B: # 252920 C: # 252930 D: # 252940
IV A e B (síndrome de Morquio)	A: queratan sulfato, condroitina sulfato B: queratan sulfato	A: galactose 6-sulfatase B: β -galactosidase	A: 16q24.3 (<i>GALNS</i>) B: 3p21.33 (<i>GLB1</i>)	A: # 253000 B: # 253010
VI (síndrome de Maroteaux-Lamy)	dermatan sulfato e sulfato de condroitina	arilsulfatase B	5q14.1 (<i>ARSB</i>)	# 253200

VII (síndrome de Sly)	dermatan sulfato, heparan sulfato e sulfato de condroitina	β -glucuronidase	7q11.21 (<i>GUSB</i>)	# 253220
IX (síndrome de Natowicz)	hialuronan	hialuronidase	3p21.31 (<i>HYAL1</i>)	# 601492

^a Único tipo que é transmitido de modo recessivo ligado ao X; todos os demais são transmitidos de modo autossômico recessivo. **Fonte:** Adaptado de (Muenzer 2011; Giugliani 2012a; Celik et al. 2021).

A incidência e prevalência ao nascimento das MPS varia amplamente dependendo do lugar ou background étnico (Khan et al. 2017; Kobayashi 2019). Analisando dados de diferentes países, Khan et al. (2017) estimaram prevalências ao nascimento das MPS entre 1,04 a 4,80/100.000 nascidos vivos. No Brasil, a prevalência ao nascimento das MPS foi estimada recentemente em 1,57/100.000 nascidos vivos, com a MPS II sendo a mais comum (0,48/100.000), seguida pela MPS VI (0,35/100.000) (Josahkian et al. 2021).

Em se tratando da distribuição geográfica das MPS, também é importante levar em consideração que algumas variantes patogênicas podem, por motivos que serão abordados mais adiante, ser mais comuns em determinada região ou grupo étnico, influenciando a epidemiologia da doença em tais situações (Khan et al. 2017). Neste sentido, o Brasil apresenta uma frequência alta de MPS VI, se comparado à maioria dos países europeus, o que pode ser parcialmente explicado pela presença de variantes patogênicas fundadoras em algumas regiões do país (Costa-Motta et al. 2011; Costa-Motta et al. 2014; Khan et al. 2017; Celik et al. 2021).

Anomalias congênitas

As anomalias congênitas (por vezes chamadas de defeitos congênitos ou malformações congênitas¹) correspondem às alterações estruturais ou funcionais do corpo humano que tiveram origem no período de desenvolvimento embrionário ou fetal, isto é, no pré-natal. Elas podem ser diagnosticadas antes, durante ou mesmo após o nascimento e afetam cerca de 3% a 6% de todos os nascimentos no mundo, sendo que uma parte importante será diagnosticada ao longo do desenvolvimento da criança (BRASIL 2021; WHO 2022).

¹ As malformações congênitas correspondem a um tipo específico de anomalias congênitas que são resultado de problemas no desenvolvimento de órgãos e estruturas que podem não se formar, formar-se de modo anormal ou incompleto (BRASIL, 2021).

Coletivamente, as anomalias congênitas constituem um grupo heterogêneo e diverso de alterações que podem ser muito graves (como defeitos de tubo neural e alguns tipos de cardiopatias congênitas) ou relativamente leves, com opções de correção terapêutica precoce (como fendas orais e polidactilia). Os principais fatores de risco relacionados às anomalias congênitas são de ordem genética (monogênicas ou cromossômicas), ambiental (agentes físicos, químicos, biológicos, mecânicos ou condições maternas, também chamados de teratógenos) ou multifatorial (Christianson et al. 2006; Mazzu-Nascimento et al. 2017; BRASIL 2021).

Estima-se que, anualmente, as anomalias congênitas sejam responsáveis por cerca de 300 mil óbitos no período neonatal em todo o mundo. Em muitos países, elas representam a principal causa de mortalidade em crianças com menos de um ano, impactando, para além do nível individual, as famílias, os sistemas de saúde e a sociedade em geral (Christianson et al. 2006; BRASIL 2021; WHO 2022). No entanto, cerca de 70% das anomalias podem ser evitadas ou terem seus efeitos minimizados com adoção de medidas de prevenção primária (como evitar a ingestão de álcool no período periconcepcional), secundária (como cirurgias intrauterinas para correção de defeitos de tubo neural) e terciária (como tratamento ortodôntico no caso de fendas orais), como mostrado na **Figura 2** (Christianson et al. 2006; BRASIL 2021).



Figura 2: Oportunidades e ações, ao longo da vida, para a prevenção de anomalias congênitas. **Fonte:** BRASIL (2021).

Fonte: BRASIL (2021).

Dado o impacto que representam na saúde global e seu potencial de prevenção, as anomalias congênitas são consideradas um importante problema de saúde pública global (Boyle et al. 2018; BRASIL 2021). Este impacto pode ser ainda maior em países de baixa e média renda, devido à falta de serviços adequados para o cuidado de crianças afetadas e uma maior taxa de exposição a infecções e desnutrição (Sitkin et al. 2014).

Portanto, assim como para as doenças raras, muitas anomalias congênitas podem ser individualmente raras, mas coletivamente representam uma significativa questão de saúde pública. Para a **Rede Europeia de Vigilância das Anomalias Congênitas** (EUROCAT), as anomalias constituem um importante grupo entre as doenças raras (EUROCAT 2012; Cardoso-dos-Santos et al. 2020b). No Brasil, as anomalias congênitas estão contidas no Eixo 1 (composto pelas "doenças raras de origem genética") da supramencionada Portaria N° 199/2014 (que instituiu a Política Nacional de Atenção Integral às Pessoas com Doenças Raras no âmbito do SUS) (Brasil 2014).

Ainda considerando a realidade brasileira, registram-se anualmente cerca de 25.000 anomalias congênitas visíveis ao nascimento no Sistema de Informações sobre Nascidos Vivos (Sinasc) (Cardoso-Dos-Santos et al. 2021). Embora o Sinasc possua abrangência nacional e captação universal em instituições públicas e privadas, é sabido que o seu registro de anomalias congênitas apresenta acentuada subnotificação, sobretudo para defeitos menos conspícuos e em determinadas regiões do país (BRASIL 2021).

Quadro 2: As redes internacionais para a vigilância de anomalias congênitas

Em um artigo de revisão narrativa publicado em 2020 no periódico *Epidemiologia e Serviços de Saúde* (29(4):1-13), identificamos e caracterizamos seis redes internacionais de colaboração para vigilância de anomalias congênitas. Além da mencionada EUROCAT, também foram identificadas outras cinco redes que cobriam populações da América Latina (ECLAMC e ReLAMC), Ásia (SEAR-NBBBD), Reino Unido (BINOCAR) e o *Clearinghouse* (ou ICBDSR), que cobre diferentes países do mundo. Estas redes contribuem para o entendimento sobre o impacto das anomalias congênitas nas populações, além de fortalecer programas individuais e criar iniciativas de vigilância em regiões desassistidas.

Acesso:

<https://bitly.com/fddIK>



Clusters, rumores, isolados, consanguinidade e outros conceitos

Doenças raras e anomalias congênitas podem ocorrer em uma frequência bem mais alta que o esperado em determinados locais. Trata-se da definição operacional de *clusters* (ou "aglomerados") geográficos no âmbito da GEMEPO (Castilla and Schuler-Faccini 2014; Gili et al. 2016; Cardoso et al. 2018). *Clusters* de certas doenças genéticas que são muito raras na população em geral ocorrem devido à maior prevalência de variantes patogênicas fundadoras em determinada comunidade ou por conta de variações socioculturais (Costa-Motta et al. 2011; Ferreira 2019). Seguindo o mesmo raciocínio, *clusters* de anomalias congênitas de origem ambiental ou multifatorial podem ocorrer pela concentração de fatores de risco em determinada população, geneticamente relacionada ou não (Castilla and Orioli 2003; Castilla and Schuler-Faccini 2014; Gili et al. 2016; Melo et al. 2021).

A noção de *cluster* é amplamente utilizada pela epidemiologia como definição de uma área e/ou período de tempo no qual a prevalência de uma patologia em particular é alta, superando os casos esperados da mesma (Elliott and Wakefield 2001). Ainda em um contexto epidemiológico, um *cluster* geográfico, portanto, é definido como a prevalência mais alta que a esperada (derivada de dados populacionais comparáveis) de determinada doença, em determinada população vivendo em uma área geográfica definida ao longo de um período de tempo (Castilla and Orioli 2003). Todavia, (Castilla and Orioli 2003) ressaltam que, no âmbito das doenças genéticas, este conceito precisa ser adaptado, de modo que o que realmente deve ser considerado no caso dos *clusters* de doenças genéticas é a frequência de genes e genótipos ao invés de fenótipos. Para eles, em tradução livre:

Esta definição pode ser diferente para doenças genéticas, uma vez que as doenças em questão são geralmente muito raras e as taxas de prevalência esperadas são desconhecidas, e pela presença de indivíduos heterozigotos assintomáticos para condições recessivas e não-penetrantes para dominantes.

Nesta área, outras definições operacionais importantes incluem os rumores, que correspondem a qualquer tipo de relato sobre a ocorrência não-usual de doenças raras e anomalias congênitas ou mesmo de fatores de risco associados (Castilla and Orioli 2003; Castilla and Schuler-Faccini 2014; Cardoso et al. 2018). Os alarmes, por sua vez, correspondem à

frequência incomum de anomalias congênitas em determinados lugar e tempo, os quais são atestados por algum teste estatístico. Finalmente, epidemias se referem à frequência incomum de determinada doença que se provou real. Portanto, as conclusões sobre epidemias iniciam a partir de uma hipótese de trabalho, seja ela um rumor ou um alarme (Castilla and Orioli 2003).

O pioneirismo brasileiro

Em se tratando do estudo e assistência de *clusters* geográficos na perspectiva da GEMEPO (de agora em diante, apenas *clusters*), o Brasil assume papel de destaque pelo pioneirismo e atual campo de atuação (Castilla and Schuler-Faccini 2014; Giugliani et al. 2019b). O pioneirismo é resultado de estudos encabeçados por grupos de pesquisa e assistência que vêm sendo desenvolvidos em todo o país, os quais possuem suas raízes ligadas aos trabalhos de Freire-Maia que, a partir da década de 1950, estudou comunidades brasileiras que concentram casos de doenças genéticas raras ou seus fatores de risco (como consanguinidade e isolamento geográfico, que serão explorados no tópico seguinte) (Freire-Maia 1957; Castilla and Schuler-Faccini 2014; Giugliani et al. 2019b).

Ainda no que se refere ao protagonismo brasileiro na investigação de *clusters*, (Castilla and Schuler-Faccini 2014) destacaram estudos pioneiros que descreveram comunidades com alta prevalência de doenças genéticas, como acondrogênese de Grebe no estado da Bahia, aqueiropodia em Minas Gerais e albinismo oculocutâneo na Ilha dos Lençóis, no Maranhão (Quelce-Salgado 1964; Freire-Maia et al. 1978; Freire-Maia and Opitz 1981). Além destes, outros trabalhos investigando *clusters* nacionais vêm contribuindo com a descrição de novas entidades patológicas e suas bases etiológicas, como a síndrome de Richieri-Costa-Pereira, síndrome de SPOAN e, mais recentemente, da síndrome congênita pelo vírus Zika, que se apresentou de maneira epidêmica em algumas localidades do país entre 2015 e 2017 (Macedo-Souza et al. 2005; Passos-Bueno et al. 2014; del Campo et al. 2017; BRASIL 2019a). Outros trabalhos mais recentes com ***clusters* brasileiros** também vêm logrando êxito na identificação das variantes patogênicas associadas a certas doenças genéticas, incluindo a caracterização fenotípica da população, análise de marcadores de ancestralidade, de fatores de risco não-genéticos, dentre outros (Macedo-Souza et al. 2005; Costa-Motta et al. 2011; Passos-Bueno et al. 2014; Cardoso et al. 2018; Giugliani et al. 2019b; Arruda et al. 2020)

Em um desses trabalhos, Macedo-Souza et al. (2005) descreveram uma grande família consanguínea no município de Serrinha dos Pintos (RN), da qual 25 indivíduos apresentavam uma doença neurológica caracterizada por atrofia óptica congênita, paraplegia espástica

progressiva de início precoce e neuropatia (que, no inglês formam o acrônimo SPOAN). Consistente com um padrão autossômico recessivo de transmissão, a síndrome é causada por deleções localizadas na região não-codificante *upstream* do gene *KLC2*, citogeneticamente localizada em 11q13. Após a descrição da SPOAN no Brasil, outros casos vêm sendo descritos em demais países, mormente Egito e Irã (Macedo-Souza et al. 2005; Bazvand et al. 2021).

Outro exemplo de *cluster* brasileiro bem caracterizado e atendido é o de MPS VI no município de Monte Santo, localizado no sertão da Bahia, onde foram identificados 13 indivíduos com a doença carregando a variante p.H178L, no gene *ARSB*. Esta concentração espacial dos casos levou o pequeno município a apresentar uma das maiores prevalências mundiais da doença, estimada em 1:5.000 nascidos vivos (Costa-Motta et al. 2011).

Em Monte Santo, todos os indivíduos afetados mostraram-se homozigotos para a mesma distribuição haplotípica de 10 SNPs (do inglês, *single nucleotide polymorphisms*) intragênicos, o que, junto com a análise dos heredogramas, sugeriu fortemente um efeito de fundador subjacente à alta frequência de MPS VI no município. Finalmente, também foi encontrada na população do local uma incidência aumentada de demais doenças genéticas, tais como fenilcetonúria e hipotireoidismo congênito, além de um alto índice de indivíduos heterozigotos para a variante patogênica responsável pelos casos de MPS VI na região (Costa-Motta et al. 2011; Costa-Motta et al. 2014; Giugliani et al. 2019b).

A importância do Brasil para a genética médica populacional também se deve ao Instituto Nacional de Genética Médica Populacional (INAGEMP), criado em 2008 para abarcar iniciativas nesta área em nível nacional (Giugliani et al.

Quadro 3: O INAGEMP

Em um artigo científico publicado na revista *Genetics and Molecular Biology* (42(1 suppl 1):312-320 – 2019), em alusão aos 10 anos de criação do INAGEMP, foram abordadas quatro iniciativas no âmbito do Inagemp, dentre as quais, o nosso relato em se tratando da investigação do *cluster* de gemelaridade em uma pequena cidade no Rio Grande do Sul. Além deste, também foram relatados outros projetos em GEMEPO, como o estudo piloto de *screening* neonatal para mucopolissacaridose VI em Monte Santo (BA) e os estudos com a doença de Machado-Joseph em famílias brasileiras com ancestralidade açoriana. Enfatizamos o processo de comunicação dos resultados e suas implicações para as comunidades envolvidas.

Acesso:

<https://bitly.com/xSrhl>



2019b). A identificação de grupos de risco para doenças raras por meio de estudos epidemiológicos é relevante para prevenção e tratamento das mesmas, os quais são de interesse de profissionais e autoridades de saúde, pacientes e suas associações e familiares, bem como de laboratórios envolvidos com diagnóstico e tratamento das patologias (Federhen et al. 2014; Cardoso et al. 2018). Todavia, a delimitação das populações em risco não é uma tarefa trivial.

Nosso grupo tem trabalhado com a “estratégia rumor”, isto é, a partir de um relato sobre possíveis isolados genéticos proveniente de fonte especializada ou não (os "rumores"), mas cuja validação segue um protocolo bem-delimitado e estabelecido. Estas informações, junto com a revisão constante da literatura especializada, alimentam o Censo Nacional de Isolados (CENISO), ligado ao INAGEMP, cujo principal objetivo é mapear e fornecer apoio para a identificação e caracterização dos *clusters* no Brasil (Castilla and Schuler-Faccini 2014; Cardoso et al. 2018). Os principais resultados do CENISO, assim como outras abordagens para mapeamento e caracterização de *clusters*, serão abordadas ao longo desta tese.

Muito mais que apenas procurar por "mutações", o trabalho com diferentes *clusters*, dentro e fora do Brasil, vêm ajudando na promoção à saúde dos indivíduos afetados e seus familiares (Giugliani et al. 2019b), no mapeamento de muitos genes e sua função na biologia humana (Wexler et al. 2004; Wexler 2012), investigação de traços complexos e doenças comuns (Villanueva et al. 2015; Nioi et al. 2016; Chaubey 2017; Matte et al. 2019; Mountford et al. 2020), investigação de fatos relacionados à história e estrutura genética das populações humanas (Panoutsopoulou et al. 2014; Matte et al. 2019; Anagnostou et al. 2019), dentre outros.

Quadro 4: “Mutações”, entre aspas

Em um trabalho de divulgação científica publicado na revista *Genética na Escola* (15(1):2-8), discutimos sobre a recomendação dos principais livros e manuais de Genética Médica para evitar o uso do termo “mutação” e “mutante” ao referir-se a seres humanos. Assim, discutimos o uso de tais termos na música e comportamento (a partir de artistas e obras da cultura pop, como David Bowie, Os Mutantes e o movimento Tropicalista), na literatura oriental (a partir do *I Ching*), em materiais audiovisuais (como filmes e séries retratando a temática dos *freak shows*), em histórias ficcionais em contexto com a teoria do pós-humanismo (como no universo dos *X-Men*) e, finalmente, no cenário evolucionista.

Acesso:

<https://bitly.com/LtIAx>



Os isolados genéticos

Uma comunidade humana pode ser definida como um grupo de pessoas que interagem entre si. Esta interação ocorre tipicamente dentro de um território específico e os membros dessa comunidade frequentemente compartilham valores, crenças ou comportamentos em comum (Neal 2020). Ocorre que o processo de formação e de organização social de determinadas comunidades pode deixar importantes marcas no "genoma populacional" e resultar na formação dos isolados genéticos, um importante tipo de *cluster* no âmbito da GEMEP (Castilla and Adams 1996; Arcos-Burgos and Muenke 2002; Giugliani et al. 2019b).

Os isolados populacionais podem ser definidos, do ponto de vista evolutivo, como subpopulações formadas a partir de um pequeno número de indivíduos que se isolaram em decorrência de algum evento fundador e ali permanecem por muitas gerações. Eventos fundadores na espécie humana incluem colonização de novos territórios, fome, guerras, desastres ambientais, epidemias, isolamento cultural, dentre outros. O estabelecimento de uma nova população por poucos fundadores originais, que contém somente uma pequena fração da variação genética total da população parental é definido como efeito do fundador (ou efeito-fundador) (Ridley 2006; Hatzikotoulas et al. 2014).

Na prática, os isolados representam comunidades que mantem um baixo fluxo gênico entre as populações ao redor por conta de barreiras geográficas ou culturais (por vezes, ambas) (Arcos-Burgos and Muenke 2002; Hatzikotoulas et al. 2014; Chaubey 2017). Barreiras geográficas estão presentes no isolamento de comunidades finlandesas, da Sardenha, no Mediterrâneo e de algumas populações que vivem em lugares de elevada altitude na América do Sul. Por sua vez, barreiras culturais estão implicadas com a origem de comunidades etno-religiosas, como o grupo menonita Amish (e mais especificamente, os *Old Order Amish*) e judeus Ashkenazi, na América do Norte, e a comunidade Parsi, na Índia e no Paquistão (Peltonen et al. 1995; Arcos-Burgos and Muenke 2002; Chaubey 2017).

Nestes locais, é comum observar consequências genéticas do isolamento, tanto em decorrência da consequente consanguinidade quanto por conta do baixo fluxo gênico com populações ao redor. A consanguinidade, isto é, a reprodução entre indivíduos aparentados, aumenta a chance de que ambos os genitores sejam portadores de uma variante patogênica herdada por um ancestral em comum (Nussbaum et al. 2008; Bittles 2010; Chaubey 2017).

Em outras palavras, a consanguinidade aumenta a probabilidade do aparecimento de indivíduos homocigotos para *loci* relacionados a doenças autossômicas recessivas (Bittles 2010). Logo, o risco de uma criança ser homocigota para um alelo recessivo raro é proporcional

ao grau de parentesco entre os genitores (Nussbaum et al. 2008; Bittles 2010). A **Tabela 2** apresenta a probabilidade de que a prole de diferentes uniões consanguíneas seja homocigota por descendência em determinado *locus*, a qual é conhecida por coeficiente de endogamia² (*F*) (Nussbaum et al. 2008).

Tabela 2: Grau de parentesco, proporção estimada de genes em comum e coeficiente de endogamia (*F*) da prole de diferentes uniões consanguíneas.

Tipo de união consanguínea	Grau de parentesco	Proporção de genes em comum	Coefficiente de endogamia (<i>F</i>)
Genitor(a) e filho(a)	1º	1/2	1/4
Irmãos ^a	1º	1/2	1/4
Irmão(a) – meio-irmão(a)	2º	1/4	1/8
Tio(a) – sobrinho(a)	2º	1/4	1/8
Primos em primeiro grau	3º	1/8	1/16
Primos duplos em primeiro grau	2º	1/4	1/8
Primos em segundo grau	5º	1/32	1/64

^a O mesmo para gêmeos dizigóticos. **Fonte:** Adaptado de (Nussbaum et al. 2008).

Como mencionado, além da consanguinidade, o baixo fluxo gênico com outras populações é outro evento que leva a consequências genéticas em isolados, pois o mesmo contribui para a redução da variabilidade genética de tais populações. A depender do tempo de isolamento e em conjunto com fatores como consanguinidade, número de fundadores e de “gargalos de garrafa”³, o fluxo gênico limitado ou ausente com populações vizinhas leva a níveis aumentados de homogeneidade genética e desequilíbrio de ligação (LD). Como consequência, os isolados costumam apresentar um reduzido tamanho efetivo populacional (*N_e*), ou seja, poucos indivíduos são necessários para explicar a variabilidade genética observada (Arcos-Burgos and Muenke 2002; Hatzikotoulas et al. 2014; Panoutsopoulou et al. 2014).

Outro fenômeno evolutivo importante a ser considerado é a deriva genética, isto é, as mudanças aleatórias nas frequências alélicas em gerações sucessivas. Assim, pelo acaso e

² Embora estejam intimamente relacionadas, endogamia e consanguinidade diferem em conceito. A endogamia consiste na união preferencial entre indivíduos pertencentes a um mesmo grupo por razões culturais/religiosas ou geográficas, envolvendo ou não a união entre pessoas aparentadas (consanguinidade) (Nussbaum et al. 2008).

³ Gargalo de garrafa (no inglês, *bottleneck*) diz respeito à redução drástica do tamanho de determinada população.

sobretudo em populações pequenas, alguns alelos podem ter sua frequência alterada, em consequência da redução significativa da variabilidade genética (Arcos-Burgos and Muenke 2002; Panoutsopoulou et al. 2014). De fato, deriva genética e consanguinidade possuem efeito sinérgico na maioria das populações isoladas resultantes de um gargalo de garrafa (Arcos-Burgos and Muenke 2002; Anagnostou et al. 2019).

De uma perspectiva paramétrica, os isolados genéticos podem ser determinados pela dosagem dos padrões de homoziguidade-heteroziguidade, pois populações altamente endocruzadas apresentam uma frequência aumentada de homoziguidade em detrimento ao número de heterozigotos (Arcos-Burgos and Muenke 2002; Anagnostou et al. 2019). Porém, alguns trabalhos vêm adicionando complexidades a este contexto (Garagnani et al. 2009; Anagnostou et al. 2019). Por exemplo, Anagnostou et al. (2019) investigaram SNPs a partir de abordagens de varredura genômica em populações isoladas da Europa e mostraram que os isolados não se estruturam como unidades uniformes do ponto de vista genômico. Ao invés disso, elas apresentaram altos níveis de heterogeneidade genômica inter-individual, comparáveis aos de comunidades abertas.

Independentemente do nível de homozigose, uma variante patogênica fundadora ou *de novo* pode aumentar rapidamente em frequências nessas pequenas comunidades, sobretudo na presença de consanguinidade, levando eventualmente ao surgimento de doenças genéticas (Peltonen et al. 1995; Arcos-Burgos and Muenke 2002; Charoute et al. 2015). Portanto, populações isoladas representam um desenho de estudo poderoso para a busca por variantes raras e de baixa frequência e sua associação com fenótipos complexos (Panoutsopoulou et al. 2014).

Também convém mencionar que, além de características genéticas, populações de isolados compartilham, em muitos exemplos, um conjunto homogêneo de exposições ambientais, como dieta e estilo de vida. Isto realça a potência destes lugares para estudos de mapeamento genômico, que ao longo do tempo vêm contribuindo para a descoberta de genes e genótipos responsáveis tanto pela origem de doenças mendelianas, quanto na investigação genética para traços complexos, conforme revisado por Kristiansson et al. (2008) e (Hatzikotoulas et al. 2014).

Para exemplificar algumas destas potencialidades que envolvem o trabalho com *clusters* e abordar as dinâmicas biossociais comumente implicadas na sua origem, detalharemos a seguir o caso da Finlândia, considerado um "isolado populacional de referência" (Mooney et al. 2018).

O caso da Finlândia

"Em populações humanas, o isolamento está relacionado a uma localização geográfica excepcional ou a algum fator cultural ou à prevalência de doenças genéticas relativamente raras". Esta afirmativa, em tradução livre, foi retirada do artigo *Messages from an Isolate: Lessons from the Finland Gene Pool* (Peltonen et al. 1995), que relata o caso singular da Finlândia, país que apresenta características geográficas, de organização social e dinâmicas populacionais que favorecem o surgimento de casos de doenças de base genética e multifatorial (Norio 2003; Locke et al. 2019).

Registros históricos e arqueológicos sustentam a tese de que o território finlandês (cerca de 337.030 km², sendo o oitavo maior país da Europa), recortado por centenas de lagos, foi fundado por pequenas populações ancestrais há cerca de 2.000 a 2.500 anos (Peltonen et al. 1995; Peltonen 2010). Sob influência também de questões de ordem natural e geográfica, intensificou-se um processo irregular de ocupação do país a partir do século XVI, com migrações internas saindo do sul em direção a terras mais ao centro, oeste e norte do território (**Figura 3**). Como resultado, ocorreu a formação de muitos grupos populacionais isolados, sobretudo na região nordeste do território finlandês, os quais cresceram em quantidade a partir de suas poucas dezenas de famílias fundadoras (FinDis 2022).

Além disso, a Finlândia recebeu uma pequena quantidade de novos imigrantes se comparada a demais países europeus, sobretudo em alguns pontos específicos do país, os quais poderiam adicionar novas informações ao *pool* genético da população. Esta situação também pode ter sido agravada pelo isolamento linguístico e cultural da população finlandesa, falante das línguas urálicas, com relação a demais povos europeus falantes das línguas indo-europeias (Arcos-Burgos and Muenke 2002). Em poucas palavras, a Finlândia, e sobretudo algumas comunidades internas, manteve-se isolada de outros países por razões geográficas e socioculturais.

Finalmente, cabe mencionar que a Finlândia possui longevos e abrangentes registros eclesiásticos que contêm informações demográficas que remontam ao século XVII e subsidiam estudos demográficos, de saúde e especialmente genéticos. A partir destes registros, por exemplo, constatou-se que a prática de casamentos consanguíneos vem se mostrando relativamente comum na história da população finlandesa. Todos os fatores supramencionados, em conjunto, contribuem para o aumento de frequência de variantes patogênicas responsáveis pela origem de doenças genéticas (especialmente aquelas de traço autossômico recessivo) que

são raras em todo o mundo, mas ocorrem com uma frequência notavelmente elevada na Finlândia (Peltonen et al. 1995).

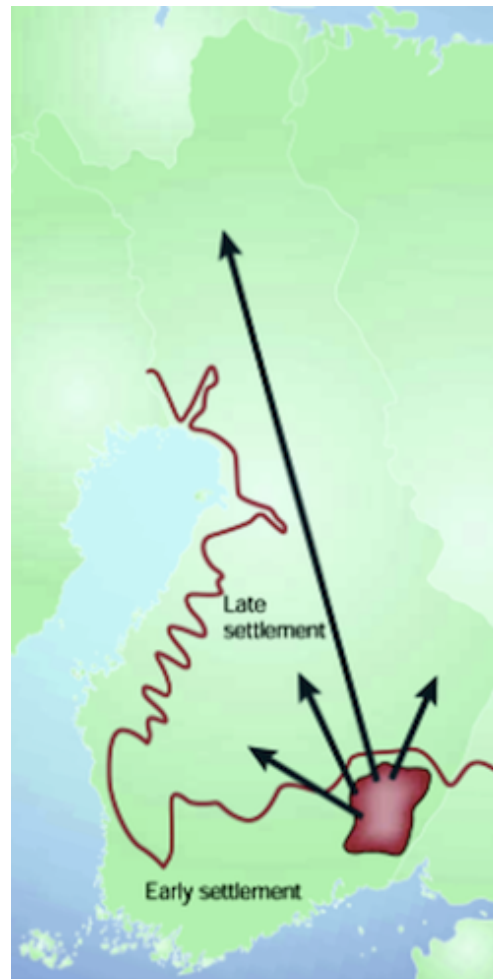


Figura 3: Direção das principais ondas de povoamento do território finlandês.

Fonte: Peltonen et al. (2000).

Ocorre muito na literatura científica o uso do termo “doença de herança finlandesa” (do inglês, *Finnish heritage disease* ou FHD) em alusão a cerca de 40 doenças que são significativamente mais comuns em pessoas de origem finlandesa (mais precisamente, pessoas cujos ancestrais são finlandeses étnicos, nativos da Finlândia ou de algumas regiões da Suécia e Rússia), se comparadas a demais populações mundiais (Norio 2003)⁴. De fato, o estudo de

⁴ Em seu trabalho de 1973, assinado também por Nevanlinna e Perheentupa, Reijo Norio foi pioneiro listar as primeiras 10 doenças que seriam mais frequentes na população finlandesa que em qualquer outra do mundo. Neste trabalho intitulado *Hereditary diseases in Finland*, os autores também apresentam a expressão “*rare flora in rare soul*” como metáfora para a combinação de fatores que estaria implicada em suas observações.

comunidades finlandesas representa o mais bem-sucedido esforço em mapear doenças genéticas em um isolado populacional (Arcos-Burgos and Muenke 2002).

Do ponto de vista de eventos biológicos, pode-se entender que o caso da Finlândia, assim como outros isolados genéticos, foi resultado da influência de poucas populações fundadoras (efeito do fundador), com o acaso e a reprodução entre indivíduos aparentados influenciando na frequência de certos genótipos (deriva genética e consanguinidade, respectivamente, atuando de maneira sinérgica), dentre os quais, aqueles responsáveis pelas FHD. Curiosamente, vale lembrar que a mesma combinação de fatores que faz aumentar a frequência de certas doenças genéticas entre a população finlandesa, também está implicada no fato de outras, como fenilcetonúria, fibrose cística e doença do xarope de bordo, serem extremamente raras (Peltonen et al. 1995; Arcos-Burgos and Muenke 2002; Peltonen 2010).

O estudo dos sobrenomes e o método isonímico

Na maioria das populações humanas, os sobrenomes, assim como os genes, são transmitidos de geração a geração. Portanto, os nomes familiares constituem uma ferramenta poderosa para a pesquisa populacional. Em estudos com enfoque genético, eles vêm sendo utilizados desde o século XIX⁵ como indicadores de estrutura populacional e eventos de migração (Jobling 2001; Redmonds et al. 2011).

Em sociedades patrilineares, como no caso do Brasil, onde os nomes familiares são geralmente passados dos pais para seus filhos e filhas, a transmissão dos sobrenomes acompanha um padrão muito semelhante ao de transmissão do cromossomo Y (Jobling 2001; Monasterio 2017). Muito semelhantes, porém não idêntico pois, enquanto todas as crianças recebem o nome familiar, apenas metade dos filhos (os do sexo masculino) recebe o cromossomo Y (Jobling 2001).

De um ponto de vista estatístico, a distribuição dos sobrenomes permite fornecer informações quantitativas sobre a estrutura de populações. É o caso do método isonímico ou isonímia, que corresponde à probabilidade de que duas pessoas compartilhem o mesmo sobrenome por ancestralidade comum (Rodríguez-Larralde et al. 1993). Este método foi desenvolvido por Crow and Mange (1965) para estimar a consanguinidade em sociedades em que os sobrenomes são transmitidos de modo regular e que apresentam uma forte relação com

⁵ Mais precisamente, a partir de 1875, com os estudos de George Darwin, filho de Charles Darwin, que utilizou os sobrenomes para estimar a frequência de casamentos consanguíneos na Inglaterra (Joblin, 2001).

a ancestralidade. Neste caso, a frequência de casamentos entre pessoas de mesmo sobrenome pode ser usada como medida de consanguinidade (Crow and Mange 1965).

Portanto, o método isonímico é baseado na premissa de que os sobrenomes são heterogeneamente distribuídos pelo espaço físico e social que constituem as populações humanas. Assim, seguindo o método descrito por Rodríguez-Larralde et al. (1993), a isonímia (ou I_{NS}) pode ser calculada como:

$$I_{NS} = \sum_k \left(\frac{N_{ki}}{N_i} \right)^2 - \left(\frac{1}{N_i} \right)$$

onde N_{ki} corresponde à frequência do sobrenome k na população i e N_i representa o tamanho da população em questão. Os valores de I_{NS} variam entre 0 e 1, de tal modo que quanto mais próximo a 0, maior será a variabilidade da população em se tratando de sobrenomes, visto que a probabilidade de que duas pessoas compartilhem um ancestral comum próximo será menor.

A partir da isonímia, podem ser calculados outros indicadores, como o índice α de Fisher, que representa a diversidade de sobrenomes de determinado local, calculado a partir do inverso da isonímia (isto é, $\alpha = 1/I_{NS}$). Assim, baixos valores de α são encontrados em populações mais isoladas, podendo refletir altas taxas de consanguinidade e deriva genética, enquanto valores altos são observados em localidades com maiores taxas de imigração. Seguindo o mesmo raciocínio, populações pequenas ou isoladas apresentam altos valores de B , que corresponde ao percentual da população coberto pelos sete sobrenomes mais frequentes (Rodríguez-Larralde et al. 1993; Dipierri et al. 2014).

Assim, unidades espaciais, como municípios e estados, podem representar um valor isonímico de sua população, possibilitando a criação de cenários comparativos. Análises estatísticas podem apoiar a detecção de aglomerados espaciais de unidades adjacentes que concentram altas ou baixas taxas de isonímia (ou qualquer outro indicador). É o caso do índice global de Moran (GMI), que testa a dependência espacial global entre todos os polígonos em estudo e é expressa por meio de um valor único para toda a região (Anselin 2010; Cardoso-Dos-Santos et al. 2018).

Como um método relativamente barato (se comparado a métodos de varredura genômica, por exemplo) e abrangente, a análise e distribuição dos sobrenomes por meio do método isonímico tem sido realizada em diversas populações ao redor do mundo, fornecendo informações importantes no que se refere à demografia, antropologia e dados de saúde (Dipierri et al. 2007; Liu et al. 2012; Barrai et al. 2012; Dipierri et al. 2016; Carrieri et al. 2019). Os sobrenomes vêm se mostrando importantes indicadores de migração, consanguinidade, mistura

genética e isolamento em níveis local, regional e nacional (Dipierrri et al. 2005; Scapoli et al. 2007; Tarskaia et al. 2009; Dipierrri et al. 2011; Liu et al. 2012; de Oliveira et al. 2013; Carrieri et al. 2019).

O espaço da tese: Ceará, Região Nordeste e América do Sul

Os trabalhos que serão produzidos a partir desta tese incluem o estudo de três diferentes espaços geográficos que se sobrepõem, mas que, cada um à sua maneira, podem fornecer uma oportunidade para ampliarmos nosso entendimento sobre os *clusters*, a saber: a região Nordeste do Brasil, o estado do Ceará e o continente sul-americano. Cada um deles será abordado com mais detalhes adiante.

Região Nordeste

O Nordeste é uma das cinco macrorregiões geográficas brasileiras e é formada por nove Unidades da Federação (UF), a saber: Alagoas, Bahia, Ceará, Maranhão, Paraíba, Piauí, Pernambuco, Rio Grande do Norte e Sergipe. De acordo com os resultados do último censo, tratava-se da segunda região mais populosa do Brasil (mais de 53 milhões de habitantes), possuindo a terceira maior área geográfica (mais de 1,5 milhões km²) e o mais baixo IDH (0,663) (IBGE 2010).

É da região Nordeste que começa a ser contada a história da colonização portuguesa no Brasil, a partir do século XVI. Essa região foi o principal centro econômico e administrativo durante o período colonial e onde ocorreram os primeiros processos de mistura étnica que dariam origem ao *background* genético da população brasileira, sobretudo entre europeus, pessoas escravizadas de origem africana e nativo-americanos (Pena et al. 2009; Manta et al. 2013; Ruiz-Linares et al. 2014).

De uma perspectiva histórica e sociológica, a formação do Nordeste brasileiro ocorreu a partir de duas atividades principais que refletem as características geológicas e ambientais encontradas na região: (1) o cultivo da cana-de-açúcar, que ocorreu em uma zona úmida e com densa vegetação, sobretudo nos atuais territórios de Pernambuco, Alagoas, Sergipe e Bahia, e que era baseada na concentração de grandes extensões de terra como a principal forma de propriedade e na escravização de pessoas de origem africana (sobretudo, da região central da África) como instituição de classe social (Arcanjo 1996; Freyre 2004); (2) delimitação de

territórios internos e pecuária extensiva, que ocorreu em um "outro Nordeste"⁶, representado sobretudo por regiões semiáridas e pastoril, equivalentes a algumas regiões do Ceará e Piauí, principalmente (Menezes 1970; Arcanjo 1996; Freyre 2004).

O Nordeste brasileiro é uma região estratégica para a condução de estudos em GEMEPO pois nela são encontrados os maiores índices de consanguinidade no país, como mostrado na **Figura 4**, extraída de um estudo clássico de Freire-Maia (1957), mas que também viria a ser replicada por estudos mais recentes (Azevêdo et al. 1969; Santos et al. 2010; Weller et al. 2012). Além disso, dados do CENISO indicam que essa região concentra a maior quantidade de rumores de isolados genéticos, e onde muitos *clusters* já foram relatados, sobretudo na região do Sertão, uma área marcada por um clima semi-árido e que apresentou historicamente frequentes secas, fome e vulnerabilidades sociais e de saúde (Menezes 1970; Cardoso 2007; Cardoso et al. 2018; Andrade et al. 2021).

Cardoso (2007) relata o modo pelo qual a região sertaneja era representada na metade do século XX:

Reaparecendo, mais uma vez de forma calamitosa na cena política, econômica e social em fins dos anos 1950 a seca nordestina seria a centelha que reacenderia o debate em torno da condição do sertão nordestino. Encravado na já então denominada “região-problema”, este célebre sertão iria ser tomado por muitos intelectuais e políticos quase como uma “região-problema dentro de uma região-problema”, onde clima, solo e homem estariam imbricados em uma estranha simbiose com a miséria. (...) Da poesia à economia, passando pela geografia, antropologia, sociologia e até nutrologia, se falava em Nordeste e, principalmente, em sertão do Nordeste Seja nas teorias do “dualismo” econômico, seja nas apreciações sociológicas ou antropológicas ou na sua particular conformação Geográfica e climato-botânica o “outro” é o Sertão, é o “Outro Nordeste”.

⁶ "O outro Nordeste" é como se intitula o livro do sociólogo cearense Djacir Menezes (1937), no qual o autor discorre sobre a evolução social e política do Nordeste da "civilização do couro" e suas implicações históricas. O título é uma referência a "Nordeste", do pernambucano Gilberto Freyre (1937), em que o mesmo aborda a região a partir da cultura da cana-de-açúcar.

Apesar de alguns avanços recentes no que se refere à melhoria de indicadores socioeconômicos e sociais no Sertão do Nordeste, as dinâmicas históricas e sociais de suas comunidades pode implicar em importantes implicações em genética médica (Laberge et al., 2005; Cardoso, 2007).



Figura 4: Níveis de consanguinidade no território brasileiro calculados a partir de registros de casamentos católicos.

Os diâmetros dos círculos são proporcionais aos valores dos coeficientes de consanguinidade (mostrados no trabalho de Freire Maia 1957). Estados: (1) Rio Grande do Sul; (2) Santa Catarina; (3) Paraná; (4) São Paulo; (6) Rio de Janeiro; (7) Minas Gerais; (8) Espírito Santo; (9) Bahia; (10) Sergipe; (11) Alagoas; (12) Pernambuco; (13) Paraíba; (14) Rio Grande do Norte; (15) Ceará; (16) Piauí; (17) Maranhão; (18) Pará; (19) Amazonas; (20) Goiás; (21) Mato Grosso; (22) Distrito Federal; (5) Cidade do Rio de Janeiro. Territórios: (22) Guaporé; (23) Acre; (24) Rio Branco; (25) Amapá. **Fonte:** Freire-Maia (1957).

Ceará

O Ceará é um estado localizado na região Nordeste do Brasil que possui uma população estimada (2021) de 9.240.580 pessoas, sendo a oitava maior população do país (IBGE 2022). O território cearense possui cerca de 184 municípios, distribuídos em uma área de 150 mil km², apresentando o 17º maior índice de desenvolvimento humano (IDH) entre os estados brasileiros (0,682) e, de acordo os resultados do último censo, 62,3% da população cearense se considerava parda, 31,6% branca, 4,6% preta, 1,2% amarela e 0,2% indígena. Além disso, cerca de 25% da população cearense residia em áreas rurais (IBGE 2010).

A formação do atual estado do Ceará ocorreu tardiamente, se comparado ao restante do Brasil. As primeiras ocupações do território para atender aos interesses da Coroa Portuguesa ocorreram a partir do século XVII, motivadas sobretudo pela pecuária e demarcação de território, com o conseqüente extermínio das populações nativas (Menezes 1970; Pontes et al. 2012; IBGE 2022). Como mostrado na **Figura 5**, os primeiros movimentos migratórios que dariam origem ao atual estado do Ceará partiram principalmente de Pernambuco e Bahia, visto que o principal objetivo destas expedições era subsidiar a zona dos engenhos, que cresciam de modo efervescente na região da Zona da Mata nordestina (Pontes et al. 2012).

Em um estudo sobre a formação do território cearense apoiado pelo Instituto de Pesquisa e Estratégia Econômica do Ceará, ligado ao governo estadual, Pontes et al. (2012) afirmam que:

Ao longo da história, desde a colonização [do Ceará], percebe-se um processo de fragmentação, resultado de um arranjo territorial político e seletivo. Inicialmente, as unidades de menor hierarquia na organização político-administrativa do país surgiram de uma evolução das comunidades a partir da ocupação territorial, com a expropriação de terras dos nativos, visando à reprodução do capital mercantilista europeu.

Do ponto de vista de Genética Médica, o estado conta com dois centros de referência em doenças raras, o Hospital Universitário Walter Cantídio e o Hospital Infantil Albert Sabin, sendo que este último é também um centro de referência para o tratamento de doenças metabólicas no estado, para onde são comumente encaminhados os pacientes com MPS e demais erros inatos de metabolismo (BRASIL 2019b).

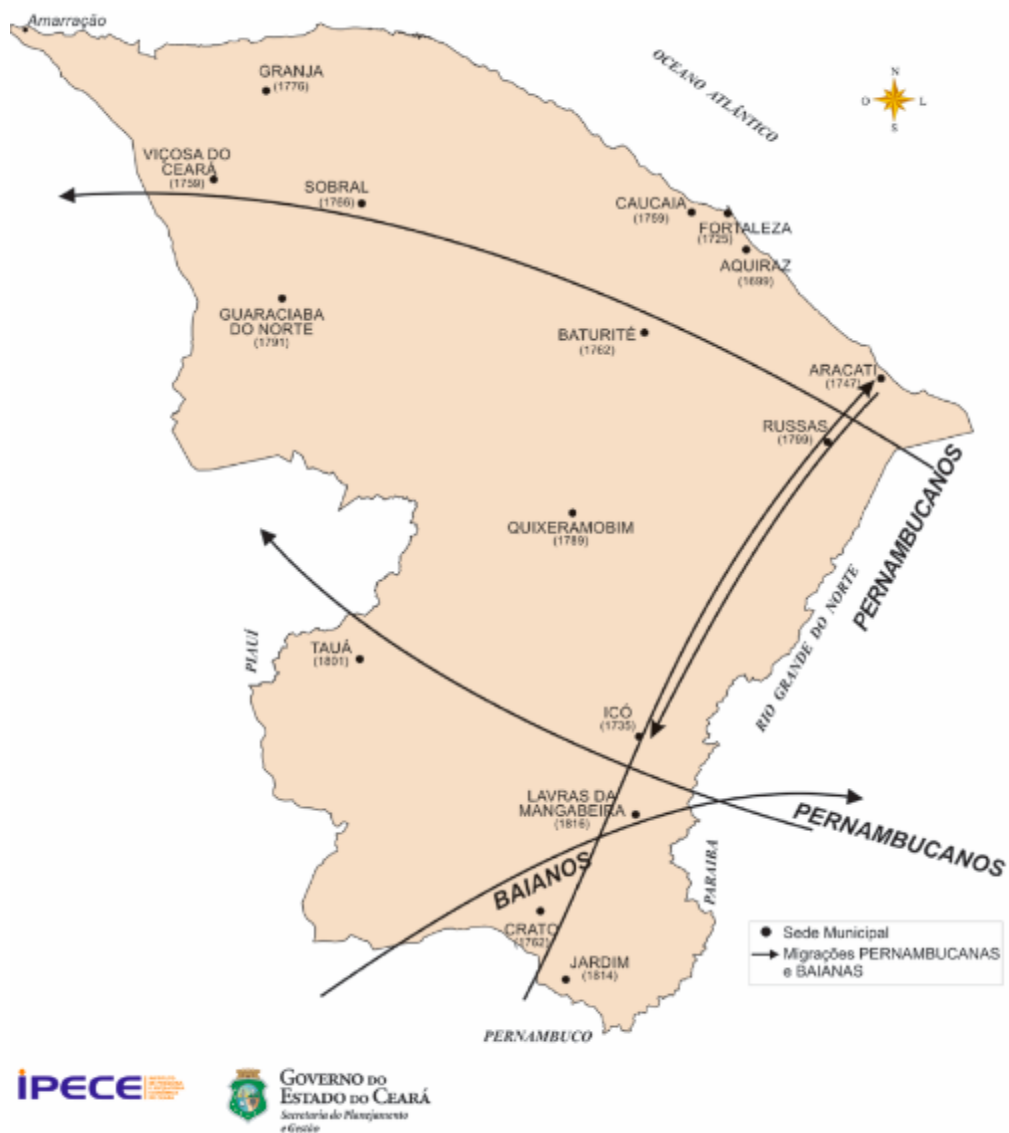


Figura 5: Caminhos dos primeiros migrantes Baianos e Pernambucanos.

Fonte: Pontes et al. (2012).

Em se tratando de GEMEPO, alguns estudos descreveram *clusters* de doenças genéticas no Ceará, como doença de Gaucher tipo I (Chaves et al. 2015) no noroeste do estado e picnodisostose (Araujo et al. 2016) geograficamente dispersa, ambas autossômicas recessivas e associadas a uma alta consanguinidade parental nas comunidades estudadas. O trabalho que descreveu 27 casos de picnodisostose em 22 famílias cearenses mostrou que "alta frequência de picnodisostose (...) é uma consequência da alta consanguinidade na região" e que "muitas mutações, provavelmente introduzidas há muito tempo no Ceará, devem ter se espalhado devido aos casamentos consanguíneos e à migração populacional interna" (Araujo et al. 2016).

Além disso, nosso grupo de pesquisa também trabalhou na descrição de aglomerados geográficos de doenças autossômicas dominantes no estado, como a **síndrome cutânea de CYLD** (Arruda 2016) na cidade de Aracati, no litoral cearense (localizada a 150 km da capital Fortaleza) e doença de Huntington na região noroeste do estado (Furtado et al. 2020). No CENISO, há o registro de 20 rumores cadastrados para a população do Ceará, além de cinco casos descartados (Cardoso et al. 2019).

Quadro 5: Um cluster de genodermatose no litoral cearense

Em um artigo publicado na revista *Journal of Community Genetics* (11:279-284), relatamos as principais características fenotípicas e moleculares de uma família de Aracati (no litoral do Ceará) cuja análise genealógica revelou pelo menos 48 pessoas com múltiplas lesões cutâneas transmitidas de modo autossômico dominante. A presença da variante c.2806C>T, p.Arg936, no éxon 20 do gene *CYLD* em todos os afetados e em nenhum dos não-afetados permitiu concluir o diagnóstico da doença (síndrome cutânea de CYLD), fornecer aconselhamento genético e assistência dermatológica e combater, de certa maneira, o forte estigma local associado à condição. A prevalência mínima da doença em Aracati foi estimada em 3,5/10.000.

Acesso:

<https://bitly.com/GUCEZ>



América do Sul

A América do Sul é o quarto maior continente do mundo, com uma extensão territorial de quase 18 milhões km², onde vivem mais de 400 milhões de pessoas. Fazem parte do continente 12 países (Argentina, Bolívia, Brasil, Chile, Colômbia, Equador, Guiana, Paraguai,

Suriname, Uruguai e Venezuela) e um território internacional (Guiana Francesa) (Ramos 2022; Wikipedia 2022).

A atual população sul-americana é o resultado de cinco séculos de mistura genética entre quatro componentes ancestrais principais: nativo-americanos, ibéricos e africanos, desde o século XV, além de ondas migratórias mais recentes. Como mostrado na **Figura 6**, o continente apresenta uma ampla diversidade de climas, bem como de ambientes naturais e geológicos, onde se instalaram comunidades com diferentes padrões culturais, componentes genéticos e ancestrais (Ruiz-Linares et al. 2014; Ramos 2022).

Castilla e Adams (1996) discutiram sobre o fato de que grande parte da população sul-americana encontra-se organizada em comunidades pequenas e semi-isoladas, com pouca imigração. Os autores compararam características genealógicas da população de Aicuña⁷, localizada em uma árida região da Argentina, com as populações de duas obras da literatura latino-americana (*Cem anos de Solidão*, Gabriel García Márquez, 1967 e a trilogia *O tempo e o vento*, Érico Veríssimo, 1949-1962). De acordo com os autores do manuscrito:

Todas as três populações exibiram muitas características em comum, tais como longos períodos de guerra civil e grandes heredogramas com caminhos complexos de herança resultando em padrões complexos de consanguinidade. Temas genéticos para todos os três são: (1) o uso de registros genealógicos para substanciar a propriedade da terra ou o poder político de uma família; (2) o registro genealógico de descendentes biológicos independente de seu status marital ou legal no grupo; (3) a existência de genealogias dos ramos aristocráticos nas mesmas famílias, o que ilustra o princípio da legal da primogenitura; (4) o valor dos sobrenomes como indicadores de parentesco e a extensão do isolamento genético, e (5) a consciência das consequências deletérias da consanguinidade.

⁷ População que concentra uma alta frequência de indivíduos com ataxia-telangiectasia (OMIM: # 208900) e albinismo oculocutâneo (OMIM: # 203200) (Castilla; Adams, 1996; Dipierri et al., 2013; Dipierri et al., 2014).

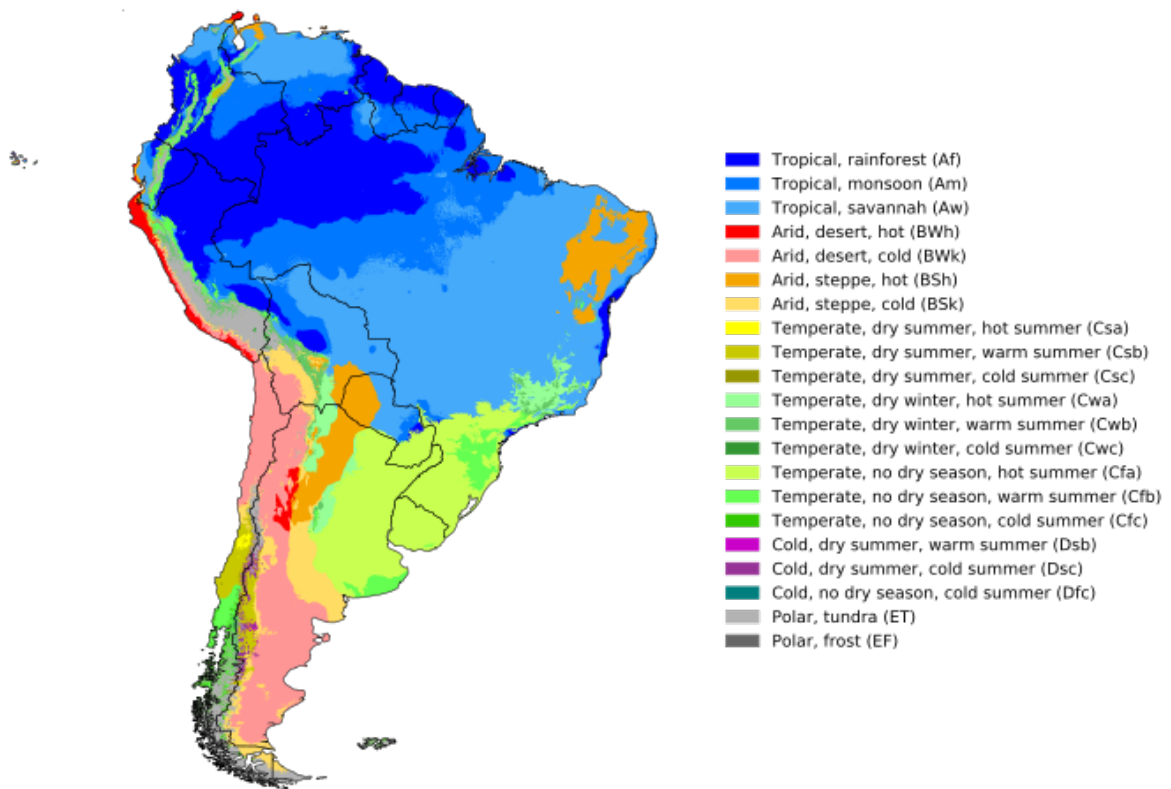


Figura 6: Diversidade de climas da América do Sul (de acordo com a classificação climática de Köppen-Geiger).

Fonte: Adaptado de Beck et al. (2018).

CAPÍTULO II: JUSTIFICATIVA

Os aglomerados geográficos de doenças raras e anomalias congênitas constituem um componente majoritário da genética médica populacional, que é um ramo da genética médica que se dedica ao estudo e cuidado médico de comunidades inteiras (e não apenas do indivíduo ou da família, como comum na prática clínica). O estudo dos *clusters* vem contribuindo com muitas descobertas científicas nas áreas de genética humana e médica, antropologia, sociologia, dentre outros (Castilla 2005; Giugliani 2012a; Panoutsopoulou et al. 2014).

No âmbito da genética, os *clusters* são considerados ferramentas poderosas para mapear variantes genéticas para o estudo de traços genéticos e complexos. Além disso, os *clusters* vem nos permitindo avançar no conhecimento sobre a história de formação e organização das sociedades, em certas modalidades terapêuticas, identificação de fatores de risco, dentre outros (Peltonen et al. 1995; Peltonen et al. 2000; Arcos-Burgos and Muenke 2002; Giugliani 2012a; Dipierri et al. 2014; Panoutsopoulou et al. 2014).

Em nível de saúde coletiva, o mapeamento e caracterização dos lugares "onde os raros são comuns" é estratégico para o desenho de estudos e políticas de assistência à saúde e sociais a comunidades que concentram indivíduos com doenças raras e anomalias congênitas (Castilla and Schuler-Faccini 2014; Costa-Motta et al. 2014; Cardoso et al. 2019). O estudo genealógico, a identificação da causa, o diagnóstico de heterozigotos e o compartilhamento da informação com base científica, em combinação com a história clínica de cada caso em particular, fornecem subsídios para um aconselhamento genético e estratégias de planejamento reprodutivo mais eficientes, além de oferecer uma oportunidade para combater mitos e preconceitos que permeiam o universo das doenças raras e anomalias congênitas (Penchaszadeh 2002; Castilla 2005; Schuler-Faccini and Giugliani 2012; Costa-Motta et al. 2014; Matte et al. 2019; Cardoso-Santos et al. 2020a; Arruda et al. 2020).

Portanto, a identificação e caracterização dos *clusters* representam um passo inicial em direção ao diagnóstico e atendimento adequados das comunidades afetadas; também permitem delimitar áreas estratégicas para o desenho de políticas de promoção à saúde para estes locais que concentram casos de doenças que, por vezes, cursam de modo crônico e podem exigir o cuidado multiprofissional. Isto é importante porque a presença de muitos casos de doenças raras (e, portanto, pouco estudadas) e complexas pode impactar de modo significativa a organização de toda uma rede de saúde, especialmente em localidades que apresentam baixos índices socioeconômicos (Castilla and Schuler-Faccini 2014; Giugliani et al. 2019b)

Vale ainda mencionar que doenças raras e anomalias congênitas emergem como tópicos importantes na atualidade, tanto em se tratando de avanço do conhecimento científico, quanto no que se refere a uma importante agenda em saúde pública. Dentro e fora do Brasil, a temática permeia debates ligados à promoção e vigilância em saúde, políticas de equidade, dentre outros (ten Kate 1998; ICBDSR 2014; Cardoso-dos-Santos et al. 2020b; BRASIL 2021; Cardoso-dos-Santos et al. 2021; WHO 2022).

CAPÍTULO III: OBJETIVOS

Objetivo Geral

Mapear e caracterizar regiões geográficas com uma frequência inesperadamente alta de doenças genéticas e anomalias congênitas (*clusters*).

Objetivos Específicos

- Avaliar a estrutura isonímica no Nordeste do Brasil como um *proxy* para isolamento e demais dinâmicas populacionais, no contexto de demais indicadores de saúde, históricos e socioeconômicos;
- Analisar o perfil clínico, molecular e geográfico dos indivíduos com mucopolissacaridoses do Ceará;
- Revisar os aglomerados geográficos de doenças raras e anomalias congênitas no continente sul-americano.

CAPÍTULO IV: *An invincible memory: what surname analysis tells us about history, health and population medical genetics in the Brazilian Northeast*

Artigo publicado na revista *Journal of Biosocial Science* (53(2)183-198 – 2020)

<https://doi.org/10.1017/S0021932020000127>

CAPÍTULO V: *The importance of geographical space in the characterization of mucopolysaccharidoses: clinical and molecular profile of a case-series from Northeast, Brazil*

Manuscrito em preparação.

CAPÍTULO VI: *El mapa de todos: clusters of rare diseases and congenital anomalies in South America*

Manuscrito em preparação.

CAPÍTULO VII: DISCUSSÃO

Onde os raros são comuns?

Esta tese intitula-se "Onde os raros são comuns" por dois motivos principais: (1) representa um sucinto conceito de *cluster*⁸, palavra mencionada logo em seguida, no subtítulo ("*relatos, estratégias e reflexões acerca dos "clusters" em Genética Médica Populacional*"); (2) representa a principal pergunta de pesquisa do presente trabalho, isto é, quais (e como) são os lugares que apresentam uma frequência inesperadamente elevada de **doenças raras e anomalias congênitas**? Aqui, foram apresentadas pesquisas que discutem diferentes estratégias de detecção e caracterização dos *clusters*, os quais figuram entre as principais frentes de trabalho da GEMEPO (Giugliani et al. 2019b).

Compõem o núcleo central desta tese três trabalhos que abordam os *clusters* a partir de diferentes estratégias e considerando diferentes espaços, conforme mostrado na **Figura 7**. Assim, a partir de dados de sobrenomes e indicadores de saúde, delineamos a paisagem isonímica para a região Nordeste do país e identificamos uma região específica que concentra agrupamentos e rumores de doenças genéticas e anomalias congênitas (Capítulo IV). No Capítulo V, identificamos um aglomerado geográfico de MPS VI em pequenos municípios da região leste do estado do Ceará, cujos indivíduos compartilhavam a mesma variante

⁸ Considerando especificamente aqueles *clusters* representados pelos isolados genéticos, por estarem mais intimamente relacionados à raridade.

Quadro 6: Anomalias congênitas e doenças raras

No terceiro capítulo do livro "Saúde Brasil 2020/2021: uma análise da situação de saúde e da qualidade da informação" (2021), publicado pelo Ministério da Saúde, exploramos a sobreposição entre os conceitos de anomalias congênitas e doenças raras. Neste trabalho, mostramos que os Sistemas de Informações sobre Nascidos Vivos (Sinasc) e Mortalidade (SIM) podem ser úteis para a obtenção de informações epidemiológicas sobre doenças raras no país, especialmente aquelas que fazem parte do Programa de Triagem Neonatal e que estão contidas no capítulo XVII da décima versão da Classificação Internacional de Doenças (CID-10; "Malformações congênitas, deformidades e anomalias cromossômicas).

Acesso:

<https://bitly.com/itMCK>



patogênica. Já no Capítulo VI, a partir de revisão sistemática da literatura, mapeamos e caracterizamos os *clusters* no continente sul-americano e seus prováveis aspectos de formação e implicações.

Termos como "isolamento geográfico", "efeito do fundador", "migração", "consanguinidade" e suas variantes podem ser encontrados em todos os três trabalhos previamente apresentados. De fato, estes fenômenos estão implicados na origem de grande parte dos isolados genéticos e sua presença em populações do Ceará, de locais do Nordeste brasileiro de outras regiões da América Latina já foi atestada por múltiplos trabalhos (Castilla and Adams 1996; Arcos-Burgos and Muenke 2002; Gomes 2011; Castilla and Schuler-Faccini 2014; Giugliani et al. 2019b).

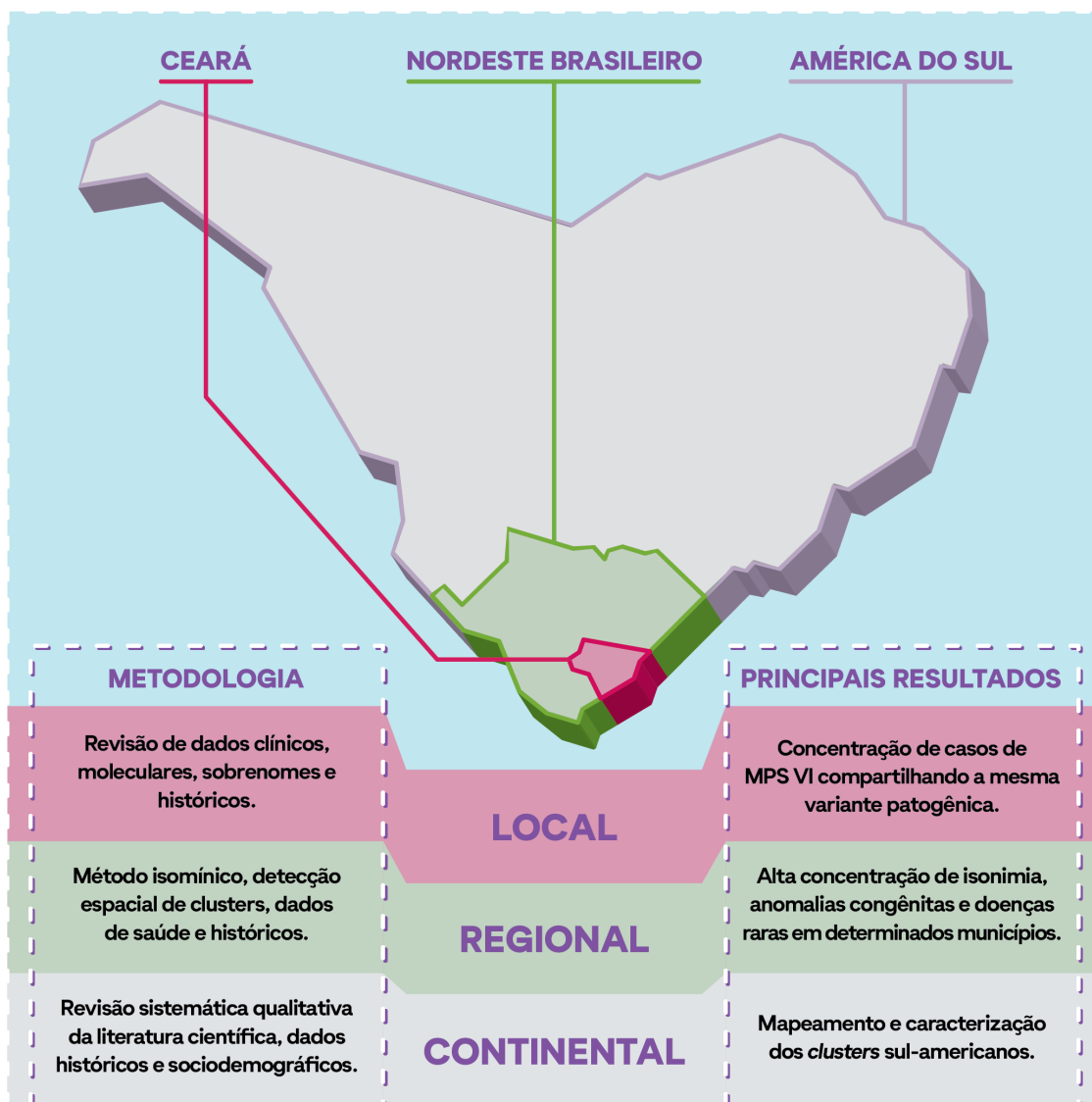


Figura 7: Esquema-resumo das principais metodologias e resultados que compõem a tese.

(a) Capítulo IV, (b) Capítulo V, (c) Capítulo VI. **Fonte:** Elaboração própria.

Castilla and Adams (1996) discutiram sobre o fato de que grande parte da população sul-americana encontra-se organizada em comunidades pequenas e semi-isoladas, com pouca imigração. Os estudos aqui apresentados fortalecem a noção de que os fenômenos acima mencionados (bem como outros previamente discutidos), que estão associados à frequência aumentada de doenças genéticas em delimitado espaço geográfico, encontram-se amplamente espalhados pela região sul-americana.

Os *clusters* em nível local, regional e continental originam-se como resultado das complexas e diversas dinâmicas biossociais estabelecidas intra e entre as complexas e diversas comunidades humanas. Conforme demonstrado por estudos prévios, a história e diversidade populacional possuem importantes implicações em Genética Médica (Laberge et al. 2005; Charoute et al. 2015).

Portanto, podemos entender os *clusters* como traços de origem multifatorial, com caráter biossocial. Origem multifatorial porque emerge da combinação de fenômenos em nível genético (como mutação, deriva genética a aumento de homozigidade) e ambiental (os fatores biogeográficos e dinâmicas sociais, históricas e culturais estabelecidas intra e entre comunidades). Caráter biossocial porque os *clusters* resultam da combinação entre características biológicas das populações com a maneira pela qual eles se organizam em sociedade e interagem com o meio.

Por envolver tamanha interdisciplinaridade, o trabalho com os *clusters* exige a integração entre diferentes áreas do conhecimento humano, para muito além das Ciências Biológicas e áreas da Saúde, atravessando terrenos como História, Sociologia, Antropologia, Geografia, dentre outras (Manno et al. 2014; Cardoso et al. 2019; Mountford et al. 2020; D'Angelo et al. 2020). Por vezes, os *clusters* podem revelar detalhes desconhecidos ou, porventura, esquecidos sobre a história de formação e organização social das populações (Mountford et al. 2020).

Além das metodologias que utilizamos neste trabalho para detectar e caracterizar os *clusters*, também convém mencionar outras abordagens metodológicas que vêm sendo utilizadas por outros grupos de pesquisa, como a análise de varredura espacial a partir de dados de prontuários clínicos e moleculares, amplamente utilizada para dados do Estudo Latino-Americano de Malformações Congênitas, ou ECLAMC (Poletta et al. 2007; Gili et al. 2015; Gili et al. 2016; Agost 2019) e de varredura genômica (Paradisi et al. 2015; Nioi et al. 2016; Anagnostou et al. 2019). Sobre este último, cita-se como exemplo o trabalho de Nioi et al. (2016), que, a partir do sequenciamento genômico (*microarrays* Illumina) de 2.636 islandeses,

mostrou uma associação entre uma variante patogênica rara no íntron 4 do gene *ASGR1* associados com risco reduzido para doença coronariana na população insular e isolada da Islândia.

A seguir, serão discutidos os principais resultados gerados por esta tese.

Sobrenomes e genética médica populacional no Nordeste do Brasil

Neste trabalho, mostramos que o estudo dos sobrenomes, em combinação com indicadores sociodemográficos e em saúde, pode representar uma alternativa de baixo custo para a obtenção de informações estratégicas no âmbito da Genética Médica Populacional. Estudamos os sobrenomes de mais 37 milhões de votantes de todos os municípios da região Nordeste do Brasil, representando um total de 74,714 sobrenomes diferentes.

A análise de sobrenomes revelou uma baixa diversidade de sobrenomes (representado pelo valor de α) e altos valores de isonimia, isto é, da probabilidade de que duas pessoas compartilhem o mesmo sobrenome por ancestralidade comum. Os valores de isonimia encontrados para a região Nordeste foram bem maiores que o de outras populações latino-americanas previamente relatadas (Rodríguez-Larralde et al. 1993; Dipierri et al. 2005; Barraí et al. 2012; Dipierri et al. 2014).

Este resultado atesta não apenas as altas taxas de consanguinidade documentadas historicamente para toda a região (Freire-Maia 1957; Azevêdo et al. 1969; Freire-Maia 1990; Weller et al. 2012; Santos et al. 2013), mas também (e talvez de maneira bem mais importante) o processo de transmissão polifilética ou horizontal dos sobrenomes que se estabeleceu de maneira muito acentuada durante a formação do Nordeste brasileiro, e que mantém seus reflexos até hoje (Rodrigues 2011a; Rodrigues 2011b; Chagas and Nunes 2016; Ribeiro 2018), como na predominância de sobrenomes de origem ibérica e **baixa representação daqueles de origem nativa.**

Quadro 7: Tu, PI

Em um poema publicado na revista acadêmica *Simbiótica* (7(3):430 – 2020), vinculada à Universidade Federal do Espírito Santo, o autor reflete sobre a ausência de sobrenomes de origem indígena no território do Piauí, conforme mostrado no trabalho Capítulo IV. Embora sem representação no sobrenome de sua atual população, a herança indígena está presente na origem do nome do estado, de muitas das suas cidades, rios e demais localidades.

Acesso:

<https://bityli.com/gukoH>



Analisado em conjunto com demais parâmetros (como altos valores do índice B, que corresponde ao percentual da população coberto pelos sete sobrenomes mais frequentes), nossos resultados apontaram para a presença de lugares pequenos e isolados. Este achado mostrou-se ainda mais significativo para alguns pontos dos estados de Sergipe, Pernambuco e Alagoas, notavelmente em um aglomerado de 37 municípios que se mostrou circunscrito à região histórica equivalente ao Quilombo dos Palmares, a maior e mais conhecida comunidade de pessoas escravizadas e seus descendentes nas Américas (Anderson 1996; Gomes 2011). De acordo com Gomes (2011):

O grande mocambo de Palmares não estava, no entanto, concentrado em um único lugar. Ao contrário, reuniu várias comunidades interdependentes e articuladas no nordeste açucareiro de Pernambuco e Alagoas.

Demais trabalhos científicos e material jornalístico vêm relatando o isolamento geográfico e um forte traço de consanguinidade entre comunidades que atualmente compõem os municípios identificados com os maiores valores de isonimia (Gomes 2011; Rodrigues 2011a; Rodrigues 2011b; Chagas and Nunes 2016; Ribeiro 2018). Nossos resultados também mostraram que esses municípios apresentaram, em média, uma prevalência maior de anomalias congênitas ao nascimento. Além disso, foram registrados três rumores de aglomerados geográficos de doenças autossômicas recessivas entre alguns dos municípios apontados: albinismo oculocutâneo (OMIM: 203100) em Santana do Mundaú (AL) e Quipapá (PE) e síndrome de Verma-Naumoff (OMIM: 613091) em Gameleira (PE) (Cardoso et al. 2018).

Ou seja, neste trabalho discutimos um Nordeste cuja origem está ligada ao cultivo da cana-de-açúcar, que originou comunidades que tinham em sua base, nas palavras de Gilberto Freyre (2004), "o latifúndio como forma de propriedade, a monocultura como forma de exploração econômica e a escravidão como instituição de classe social". Neste quesito, o Nordeste de Freyre contrapõe-se ao "outro Nordeste", de Djalma de Menezes (1970), que corresponde ao sertão pastoril, de clima semi-árido e que foi melhor abordado no trabalho que será discutido a seguir, sobre o *cluster* de MPS VI no Ceará. Sem desconsiderar as complexidades inerentes ao processo de formação histórico e social desta extensa macrorregião geográfica brasileira, é possível afirmar que ambos os "Nordestes" (tanto o de Freyre quanto o de Menezes) concentraram, durante sua história de formação, fenômenos biosociais (tais como casamentos consanguíneos, isolamento geográfico, consanguinidade, migração e provável

efeito fundador) como consequência de diferentes dinâmicas que por ali se estabeleceram (Menezes 1970; Arcanjo 1996; Freyre 2004)

Finalmente, como será mostrado mais adiante, a revisão sobre os *clusters* na América do Sul nos mostrou que existem pelo menos outros dois *clusters* diretamente com os eventos relacionados à escravidão de pessoas de origem africana no continente: um de ataxia espinocerebelar do tipo 7 em Nirgua, na Venezuela, e outro de β -talassemia, na fronteira entre Guiana Francesa e Suriname (Broquere et al. 2010; Paradisi et al. 2015). Neste último, os cinco probandos carregavam a variante *HBB*, c.315+1G>T e pertenciam ao grupo ("*maroon*") Bush Negroes. Assim, sugeriu-se que a variante teria sido originada no Oeste da África, com seu conseqüentemente espalhamento neste grupo étnico por conta de efeito fundador e/ou deriva genética (Broquere et al. 2010).

Concentração de MPS VI em municípios cearenses

Ao investigarmos as características clínicas, moleculares e geográficas de 76 indivíduos com MPS do Ceará, identificamos uma concentração espacial de MPS VI em pequenos municípios (Icó, Quixeré, Jaguaratama, Russas e Beberibe) na região leste do estado, os quais compartilhavam a mesma variante patogênica (*ARSB* c.1148-8T>G) em homozigose. Discutimos sobre alguns fatos que podem ajudar a explicar este achado.

Conforme mostrado no **Apêndice 1**, esta variante patogênica rara já foi relatada em outros estudos (Petry et al. 2005; Garrido et al. 2007; Karageorgos et al. 2007; Garcia et al. 2010; Leal et al. 2014; Horovitz et al. 2015; Franco et al. 2016; Tomanin et al. 2018), mas apenas em pacientes do Brasil, Argentina, Portugal e Espanha. Em um desses estudos, Garcia et al. (2010) descreveram as alterações esqueléticas de dois garotos portugueses (não aparentados e ambos filhos de pais não-consanguíneos) com a variante, um em homozigose e o outro em heterozigose. De acordo com os autores, a variante *ARSB* c.1148-8T>G é "comum em pacientes brasileiros e na nossa Unidade Metabólica no Hospital Pediátrico de Coimbra, Coimbra, Portugal (observação pessoal)".

É possível que essa mesma variante rara associada à MPS VI tenha surgido *de novo* nas diferentes populações dos quatro países mencionados, mas nós também podemos considerar um cenário hipotético, porém baseado em fatos reais, que resultaria na formação do *cluster* cearense. Assim, hipotetizamos que a origem dessa variante ocorreu em algum momento na Península Ibérica (considerando os casos de MPS VI relatados em Portugal e na Espanha), com

sua conseqüente chegada nas Américas (Brasil e Argentina, respectivamente) a partir do século XVI, através de indivíduos com a variante em heterozigose.

O início do processo de colonização europeia no caso particular do Brasil ocorreu através da região Nordeste e foi acelerado pelo cultivo e exportação da cana-de-açúcar que perdurou como principal atividade econômica da colônia até o século XVII, notavelmente nas capitãncias de Pernambuco e da Bahia. É principalmente destes lugares que partem os migrantes que iniciaram a ocupação da Colônia do Ceará, sobretudo no decorrer do século XVIII, inicialmente para defesa territorial e aldeamentos indígenas e mais intensamente por conta da criação de gado (Menezes 1970; IBGE 2010; Pontes et al. 2012). A propósito, o caminho entre Icó e Russas (atravessando Jaguaratama e Quixeré, todos municípios identificados no *cluster* cearense de MPS VI) corresponde a um dos principais caminhos de migração pernambucana no Ceará e, posteriormente, amplamente utilizada para fins comerciais, como pode ser visto na **Figura 4** dessa tese (Pontes et al. 2012; IBGE 2022).

Muitos municípios cearenses formaram-se a partir da expansão das fazendas de gado para núcleos populacionais maiores, incluindo as sesmarias (terras concedidas aos exploradores mediante requerimento de posse), vilas e municípios, sobretudo no interior do estado, em uma região marcada por um clima semi-árido e longos períodos de seca (Catunda 1955; Magalhães 1970; Menezes 1970). Nestes lugares, estabeleceu-se como prática comum a endogamia que era, nas palavras de Catunda (1955):

(...) na maioria dos casos forçada por insulamento feudal e, às vezes, espontânea, determinada pelo sentimento etnocêntrico do preconceito racial do sangue.

Neste mesmo estudo, Catunda (1955) discorre sobre um "clã parental" que teria se fixado em Icó e, através da obtenção de novas sesmarias, foram espalhando-se rumo ao norte acompanhados por inúmeras relações consanguíneas. Além do exemplo particular, outros trabalhos vêm relatando um importante componente de consanguinidade e certo grau de isolamento geográfico na formação de comunidades dos sertões do Ceará e do Nordeste, em geral (Santos et al. 2010; Gomes 2011; Weller et al. 2012; Santos et al. 2013; Chaves et al. 2015).

Logo, a variante patogênica detectada no *cluster* de MPS VI no Ceará pode ter surgido na população fundadora (colonizadores europeus), chegando ao território cearense por meio de processos migratórios internos ao Brasil-Colônia e aumentado em frequência com a formação

de comunidades com forte traço endogâmico, relativo isolamento e número reduzido de indivíduos. Neste cenário, a ação do acaso (deriva genética) pode ter levado ao surgimento de indivíduos homozigotos para a variante em algum lugar isolado, com consequente espalhamento para outras áreas em consequência dos processos migratórios. Assim, o *cluster* cearense ilustra outro fato importante: nem todos os *clusters* restringem-se aos limites que, atualmente, separam comunidades, cidades, províncias e podem envolver uma região geográfica mais abrangente (Manno et al. 2014).

Vale mencionar que já foram descritas outras duas populações nordestinas com alta frequência de MPS VI, curiosamente na Bahia (em um município sertanejo chamado Monte Santo) e em Pernambuco (localização não-informada), cuja alta frequência das variantes *ARSB*, c.533A>T (p.His178Leu) e *ARSB*, c.1143-1T>G, respectivamente, foi relacionado a um provável efeito fundador associado à formação de tais populações (Costa-Motta et al. 2014). Neste último caso, a variante patogênica identificada (c.1143-1T>G) encontra-se situada a apenas 7 pares de bases daquela encontrada no *cluster* cearense (Motta 2011; Costa-Motta et al. 2011).

Em Monte Santo (BA), a análise molecular dos familiares dos casos de MPS VI revelou uma alta taxa de homozigotos (cerca de 42%) e vem sendo desenvolvido um programa de *screening* neonatal em amostras de fácil coleta e transporte (Giugliani 2012a; Giugliani et al. 2019b). Além disso, a identificação dos clusters de MPS VI na Bahia e no Ceará contribuíram para a descentralização do processo terapêutico, principalmente no que se refere às infusões relacionadas à terapia de reposição enzimática, que rotineiramente ocorrem em centros de referência localizados em grandes cidades, permitindo o acesso à terapia em domicílio.

Além da descrição do aglomerado geográfico de MPS VI, nosso trabalho também discutiu informações clínicas e moleculares sobre demais casos de MPS no estado do Ceará, incluindo a descrição de variantes patogênicas possivelmente relatadas pela primeira vez, dentre as quais *IDS*, p.(Lys440Thrfs*15); *NAGLU*, p.(Glu28Lys) e *NAGLU*, p.(Tyr174=) e possíveis implicações do espaço geográfico no estudo desse grupo de doenças raras.

Clusters na América do Sul

O continente sul-americano é atualmente habitado por mais de 400 milhões de pessoas que refletem uma complexa histórica demográfica que inicia com o povoamento deste local há cerca de 14.000 anos por populações nativo-americanas. Do final do século XV adiante, múltiplas ondas de migração ocorreram ao continente a partir da colonização europeia (no

início, notavelmente por portugueses e espanhóis e, posteriormente, por demais populações) e do tráfico de pessoas escravizadas com forte componente africano. Os eventos de mistura genética ocorreram, portanto, em diferentes momentos e em um contexto de ampla diversidade geográfica, cultural e com histórica estratificação social (Ruiz-Linares et al. 2014; Salzano and Sans 2014; Homburger et al. 2015). Estes e tantos outros fenômenos que fogem ao escopo desta tese fazem a América do Sul (e a América Latina, em geral) um local único para a investigação de questões antropológicas e epidemiológicas (Salzano and Sans 2014).

Isto inclui o estudo dos *clusters*. O continente sul-americano vem sendo caracterizado como um local que concentra comunidades rurais semi-isoladas, com pouca imigração e onde a consanguinidade pode ser relativamente comum (Castilla and Adams 1996). Existem alguns *clusters* em populações da América do Sul bem caracterizados e conhecidos, mas a bibliografia sobre o assunto é multilíngue e dispersa. Os resultados da nossa revisão sistemática permitiram uma análise ampla e integrativa dos 122 *clusters* identificados na América do Sul e de demais aspectos relacionados à Genética Médica Populacional.

Inicialmente, a maior presença de *clusters* no Brasil evidenciou um papel de destaque na produção de estudos e na concentração de *clusters* em seu território, o que reflete sua maior densidade populacional, com forte influência de sua diversidade sociocultural, ambiental e populacional. Além disso, é também notável seu pioneirismo em estudos e iniciativas no campo da genética médica populacional, catalisados nos últimos anos pelo INAGEMP (Freire-Maia 1957; Freire-Maia 1990; Giugliani et al. 2019b).

Para o continente como um todo, ressaltamos alguns relatos de multi-*clusters*, isto é, regiões geográficas que concentraram mais de um grupo de doenças genéticas, com uma forte influência de variações de acordo com o componente ancestral de cada comunidade, além da presença de algumas comunidades insulares. A ilha de Providência, na Colômbia, representa um bom exemplo de ambas as características, visto que nesta comunidade insular foi registrada uma frequência elevada de dois tipos surdez: uma genética não-sindrômica associada com uma deleção no gene *GJB2*, encontrada entre indivíduos de origem europeia; e outra associada com a síndrome de Waardenburg, encontrada em famílias com ancestralidade africana (Lattig et al. 2008).

Outro exemplo de multi-*cluster* colombiano representa o exemplo mais extremo encontrado. Na comunidade de Antioquia, na região noroeste da Colômbia, já foram relatados agrupamentos espaciais de seis doenças genéticas: três autossômicas dominantes (síndrome de Lynch, doença de Alzheimer e síndrome de blefarofimose-ptose-epicantus) e três autossômicas recessivas (síndrome de Jarcho-Levin, acidose tubular renal com surdez e doença de Parkinson

juvenil) (Lopera et al. 1997; Pineda-Trujillo et al. 2001; Ramírez-Castro et al. 2002; Pineda-Trujillo et al. 2006; Nikali et al. 2008; Montoya and Morales 2009; Alonso-Espinaco et al. 2011). A ilha de Providência e Antioquia representam um peculiar caráter multi-étnico na origem destes *clusters* sul-americanos, pois, conforme assinalado por Mooney et al. (2018), “a maioria dos isolados examinados até o momento foram fundados a partir de uma única população ancestral”.

Também foram discutidos *clusters* com um forte traço étnico, como a alta prevalência de deficiência de quitotriosidase associada a uma variante patogênica no gene *CHIT1* em comunidades ameríndias pequenas e isoladas do Peru (Manno et al. 2014). Outros trabalhos mostraram a presença de *clusters* em comunidades cuja origem está ligada à escravização de pessoas de origem africana durante o regime colonial europeu na região (Broquere et al. 2010; Paradisi et al. 2015), como discutido no tópico anterior.

Por fim, destacamos um tipo de *cluster* que amplia o nosso principal objeto de trabalho para além dos isolados genéticos: os aglomerados geográficos de anomalias congênicas. Alguns tipos de anomalias congênicas podem ocorrer em uma frequência relativamente comum e, portanto, não se encaixam no conceito baseado em frequência de doença rara, e também podem ser causadas por fatores ambientais e multifatoriais (BRASIL 2021).

Para além da comparação com os isolados genéticos, *clusters* de anomalias congênicas fornecem uma janela de oportunidade para ações em Genética Médica Populacional e Saúde Coletiva. No Brasil, para o ano de 2018, elas foram a principal causa de morte em menores de 1 ano (23%), seguida por prematuridade (15%) e fatores maternos (15%) (BRASIL 2021). De fato, a detecção de aglomerados geográficos de certos tipos de anomalias congênicas faz parte do escopo de ações de programas locais, nacionais e supranacionais de **vigilância em saúde** (Castilla

Quadro 8: Vigilância de anomalias congênicas no Brasil

Em uma nota de pesquisa publicada na revista *Epidemiologia e Serviços de Saúde* (30(1):1-8 – 2021), documentamos o processo de definição da lista prioritária de anomalias congênicas para fortalecer o seu registro no Sistema de Informações sobre Nascidos Vivos (Sinasc). A lista foi elaborada por profissionais da Secretaria de Vigilância em Saúde (SVS/MS), em conjunto com representantes da Sociedade Brasileira de Genética Médica e consultada por demais sociedades médicas. Trata-se de uma definição operacional para a construção de um modelo nacional de vigilância de anomalias congênicas, que se encontra em fase de desenvolvimento no âmbito da SVS/MS.

Acesso:

<https://bityli.com/mLqSB>



and Orioli 2003; Cardoso-dos-Santos et al. 2020b; Melo et al. 2021). Em muitos casos é possível identificar prováveis fatores causais ou associados, subsidiando a tomada de decisão em se tratando de medidas de saúde pública de modo a prevenir novos casos e/ou mitigar seus efeitos deletérios.

Na América do Sul, assumiram papel de destaque os estudos conduzidos no âmbito do ECLAMC, um dos mais longevos e bem-estabelecidos programas de vigilância hospitalar de anomalias congênitas em todo o mundo (Cardoso-dos-Santos et al. 2020b). Dados coletados pelo ECLAMC têm subsidiado a detecção de *clusters* de anomalias congênitas em diversas regiões do continente sul-americano, dentre os podemos destacar o de microtia em Quito (Equador) e fenda labial em La Paz (Bolívia), pelo fato de ambas terem ocorrido em grandes centros urbanos (Castilla and Orioli 2003; Poletta et al. 2007; Gili et al. 2016).

Neste sentido, outro exemplo que vale menção foi a aglomeração espacial e temporal de microcefalia (e demais sinais e sintomas posteriormente descritos) em alguns pontos do Brasil e demais países sul-americanos (Ribeiro et al. 2017; Larrandaburu et al. 2017; BRASIL 2019a). Neste cenário epidêmico e de emergência em saúde pública, ocorreu a identificação de um novo teratógeno humano, o vírus Zika, e de uma nova condição que passaria a ser conhecida como síndrome congênita associada à infecção pelo vírus Zika (Schuler-Faccini et al. 2016; del Campo et al. 2017). Recentemente, em um livro intitulado **"Saúde Brasil: anomalias congênitas prioritárias para a vigilância ao nascimento"**, relatamos, pela perspectiva do Ministério da Saúde, a experiência acumulada durante o enfrentamento à situação de emergência a partir de 2015, e sobre como isso está relacionado à estruturação de um programa nacional de vigilância de anomalias congênitas para todo o Brasil (BRASIL 2021; Cardoso-dos-Santos et al. 2021).

Quadro 9: Saúde Brasil

No livro *"Saúde Brasil: anomalias congênitas prioritárias para a vigilância ao nascimento"* (2021), elaborado pelo Ministério da Saúde em parceria com profissionais da Universidade Federal do Rio Grande do Sul, Hospital de Clínicas de Porto Alegre e demais instituições, abordamos informações teóricas, epidemiológicas e práticas sobre as anomalias congênitas no Brasil, de modo a fortalecer o registro delas no Sistema de Informações Sobre Nascidos Vivos (Sinasc) e qualificar políticas de saúde. Exemplares físicos do livro foram enviados para as Secretarias Estaduais de Saúde de todo o Brasil, servindo de apoio para a estruturação do modelo nacional de vigilância em saúde.

Acesso:

<https://bityli.com/UkFbA>



CAPÍTULO VIII: CONCLUSÕES

Essa tese explorou diferentes estratégias para mapear e caracterizar os *clusters* em diferentes perspectivas geográficas (local, regional e multinacional). O estado do Ceará, a região Nordeste e a América do Sul, em geral, mostraram-se lugares estratégicos para o desenho de estudos e políticas de assistência à saúde e sociais a comunidades que concentram indivíduos com doenças raras e/ou anomalias congênitas. Cada um dos trabalhos aqui relatados pode ser fortalecido ou continuado com análises futuras, algumas das quais serão listadas no Capítulo IX.

De uma perspectiva local, utilizando como ponto de partida o estado do Ceará, caracterizamos uma série de casos de MPS em termos clínicos e moleculares. Nosso principal achado consistiu na presença de um aglomerado geográfico de municípios que concentraram casos de MPS VI, de diferentes famílias e que compartilhavam a mesma variante patogênica rara (*ARSB*, c.1143-8T>G). Discutimos alguns processos possivelmente envolvidos com a origem deste *cluster*, além das implicações deste e demais achados em termos de saúde individual e coletiva.

O trabalho com os sobrenomes no Nordeste do Brasil nos permitiu refletir sobre as altas taxas de isonímia e baixa diversidade de sobrenomes, fatores relacionados a isolamento e consanguinidade, e suas implicações para o aumento de frequências de anomalias congênitas e de *clusters*. Em um cenário em que toda a região apresentou altos índices de isonímia (se comparados a outros estudos já publicados), os maiores valores mostraram-se espacialmente concentrados em municípios que se originaram a partir de intenso isolamento geográfico e endogamia, e onde já foram relatadas concentrações geográficas de doenças autossômicas recessivas.

A revisão dos *clusters* na América do Sul possibilitou a criação de um fértil cenário comparativo que foi útil para refletir sobre aspectos relacionados às populações e à biogeografia do continente, os quais podem ajudar a explicar a origem de tais fenômenos e suas possíveis implicações. Este trabalho subsidia demais estudos e iniciativas nas áreas de Genética Médica Populacional, Saúde Coletiva, entre outros, e representa o primeiro passo em direção à construção de um censo continental de doenças raras e anomalias congênitas na América Latina, aos moldes do CENISO para a população brasileira.

Trabalhos com *clusters* vêm contribuindo para importantes descobertas científicas que beneficiam não apenas os indivíduos com determinada condição rara ou congênita, seus

familiares e comunidade, mas também ajudam a expandir consideravelmente o conhecimento sobre a biologia e comportamento humanos. Além disso, eles também levantam importantes questões de saúde comunitária pois tratam-se de lugares que concentram casos de doenças pouco conhecidas que geralmente cursam de maneira crônica e que, por vezes, exigem diagnóstico e cuidado multiprofissional. Isso pode ser um tópico especialmente importante em se tratando de regiões submetidas a diferentes fontes de vulnerabilidades sociais e de saúde.

Mesmo na ausência de tratamento específico para certo tipo de doença rara ou na impossibilidade de confirmar um diagnóstico molecular, a experiência empírica nos sugere a pensar que a informação, com base científica e quando repassada de maneira oportuna e apropriada, pode levar muitos benefícios a tais comunidades. Dentre eles, benefícios diretos à saúde, tais como: possibilidade de aconselhamento genético e planejamento reprodutivo, possibilidades terapêuticas locais ou em centros de referência, capacitação dos profissionais de saúde que atuam no atendimento destes indivíduos, dentre outros. Do ponto de vista social, a informação apropriada ajuda a derrubar mitos e preconceitos que cercam a dinâmica coletiva destes *clusters*.

Trabalhar com *clusters* também perpassa por questões bioéticas importantes, que não foram discutidas neste trabalho, mas que podem ser encontradas em outras publicações, como na de Matte et al. (2019), em que relatamos os principais resultados obtidos em mais de duas décadas de trabalhos com um *cluster* de gemelaridade no Sul do Brasil (que, voluntariamente, procurou os "cientistas" em busca de explicações para tal fenômeno). No trabalho, focamos nas particularidades do processo de comunicação dos resultados e aspectos éticos envolvidos com trabalhos em GEMEPO. Aliás, este *cluster* também foi tema da minha dissertação de mestrado em Genética e Biologia Molecular, pela UFRGS, intitulada "Twin Peaks": *investigando mistérios sobre a gemelaridade no Brasil* (disponível em: <https://www.lume.ufrgs.br/handle/10183/180672>).

Aliás, este *cluster* também foi tema do filme que eu assisti em sala de aula, em meados de 2013, em Parnaíba (PI). O mesmo filme que me apresentou a uma região de fronteira do conhecimento humano em que a genética e a biologia molecular encontram a saúde, sociologia, antropologia, história, geografia e tantas outras áreas que se dedicam ao estudo das dinâmicas populacionais e suas complexas relações com o meio.

O mesmo filme que me fez cruzar o país várias vezes, morar em três das suas cinco regiões, trabalhar no Ministério da Saúde durante uma pandemia, viver várias versões de mim mesmo e chegar até o momento de finalizar esta tese de doutorado. No início de tudo isso, na rodoviária de Parnaíba, meu pai me desejou coragem. Eu estava prestes a morar fora de casa,

pela primeira vez, para iniciar o mestrado em Porto Alegre (RS), uma cidade da qual minha família guardava pouca (se alguma) referência, além do fato de que era muito fria e distante.

Eu estava prestes a fazer o que fizeram tantos outros antes de mim: migrar. Assim como fizeram as famílias dos meus avós paternos e maternos, que se encontraram no litoral nordestino há menos de cinco décadas, seguindo o rastro dos seus antepassados que também migraram pelos estados do Ceará e do Piauí em busca de melhores condições de vida, desde que a história oral permite contar.

Na família da minha mãe, relata-se consanguinidade em vários pontos do nosso heredograma. Meus avós maternos são "primos legítimos". Ele, vaqueiro; ela, da roça. Ambos de sobrenome Cardoso. Ambos nascidos nos rincões entre Piauí e Ceará, isolados e castigados por longos períodos de estiagem, ocupando lugares ainda não arrendados à margem de cidades maiores. Minha avó teve onze gestações, uma delas minha mãe, noutras duas um natimorto e um "anjinho" (criança falecida), que "morreu de tanto sangrar". De onde venho, escuto relatos semelhantes ao da nossa família desde que nasci.

Termino, pois, este trabalho científico de maneira pessoal porque assim o comecei e também porque, em muitos momentos do meu processo de doutoramento, senti que, ao contar a história dos *clusters* abordados nesta tese, eu estava contando a minha própria história enquanto ser piauiense-cearense, nordestino e sul-americano. Independentemente da época e das motivações, as dinâmicas sociais, que nos fazem explorar e alterar o espaço ao nosso redor e por vezes resultam no aparecimento dos *clusters*, são o resultado da combinação de fatores biológicos e ambientais, de modo que estas transformações só podem ocorrer na presença de um fenômeno específico que nos afeta em nível individual e coletivo: a coragem. Foi preciso bastante dela para chegar até aqui.

CAPÍTULO IX: PERSPECTIVAS

Sobrenomes e genética médica populacional no Nordeste do Brasil

- Estudar cada cenário estadual individualmente, para identificar as maiores áreas de isonímia em cada local.
- Expandir as análises isonímicas para todo o Brasil, explorando o mesmo *dataset* que foi utilizado para a região Nordeste.
 - Comparar dados entre regiões.
 - Estudar marcadores de ancestralidade em nível isonímico.
- Estudo de características sociodemográficas e de saúde dos municípios (ou demais unidades espaciais) que concentram altos valores de isonímia e relatos de aglomeração espacial de doenças genéticas.
 - Índices de vulnerabilidade social e de saúde.
 - Discutir o impacto dos *clusters* na carga de saúde dos municípios, pensando em quão organizada e preparada encontram-se a rede e os serviços de saúde nestes locais que concentram uma alta frequência de indivíduos que possuem condições crônicas complexas.
- Investigar possíveis padrões fenotípicos ou genéticos entre as doenças genéticas relatadas nos *clusters*.
 - Biologia de sistemas.
 - *Human Phenotype Ontology Database*?

Concentração de MPS VI em municípios cearenses

As perspectivas sobre este trabalho incluem testar a hipótese de uma origem ancestral em comum da variante patogênica *ARSB* c.11431G>C nos casos do Ceará e possível relação com os demais casos já relatados. Para isso, podem ser considerados algumas das etapas listadas adiante:

- Levantar os grupos de pesquisa que já relataram a variante (**Apêndice 1**) e articular possibilidades de parcerias, o que inclui detalhar origem geográfica dos demais casos,

informações genealógicas, clínicas (severidade, tempos de diagnóstico e primeiros sintomas, TRE...).

- Estudo molecular para investigar origem ancestral e distância molecular entre os casos já relatados (Ceará e mundo).
 - Coleta de amostras ou possibilidade de sequenciamento em centros/instituições de conveniência e levantamento de informações genealógicas, clínicas (severidade, tempos de diagnóstico e primeiros sintomas, TRE...).
 - Avaliar LD de marcadores genéticos em *loci* adjacentes ao da variante.
 - Pensar na melhor técnica: genotipagem STR? marcadores? dosagem de SNPs por Sanger? painel? melhores algoritmos?
 - Coletar e analisar dados de sobrenomes e informações sociodemográficas.
- Determinação de heterozigotos na família e comunidade (considerar a possibilidade de ensaios moleculares mais custo-efetivos para determinação específica da variante - qPCR?).
 - Saídas de campo para coleta de amostras, informações históricas, sociodemográficas e genealógicas das famílias.
 - Parcerias locais (universidades? unidades de saúde?), com oficinas de capacitação e empoderamento da comunidade e de profissionais de saúde.
 - Aconselhamento genético e planejamento reprodutivo.
- Aprofundar a análise histórica dos casos do Ceará (novas saídas de campo? qual a relação com os casos de Pernambuco e outros possíveis casos brasileiros?).
- Cruzar dados de isonímia apenas para o estado do Ceará e cruzar com a distribuição de MPS e demais casos de doenças genéticas e anomalias congênitas.

Clusters na América do Sul

- Instituição de metodologia contínua de revisão de literatura científica para constante monitoramento e atualização para inclusão de novos *clusters* no continente.
 - Incluir o registro de rumores.
 - Divulgação do formulário de cadastro de rumores ou *clusters* entre profissionais da área de genética médica e populacional e vigilância de anomalias congênitas.
 - Qualificar os rumores por meio de metodologia que inclui revisão de literatura cinzenta.
- Inclusão de demais nações latino-americanas.

- Investigar possíveis padrões fenotípicos ou genéticos entre as doenças genéticas relatadas nos *clusters*.
 - Biologia de sistemas.
 - *Human Phenotype Ontology Database*?

CENISO

- Criação de um painel (Power BI? R?) para a atualização dos dados;
- Inserir o artigo/doi/autores de cada *cluster* (fases 3 e 4);
- Agrupar por macrorregião geográfica e inserir os geocódigos dos municípios e UF;
- Cruzar o centroide de cada *cluster* com as informações sobre distâncias aos centros de referência (doenças raras? APAEs?).

CAPÍTULO X: REFERÊNCIAS BIBLIOGRÁFICAS

- Agost L (2019) Tumores renales pediátricos y su posible asociación con poblaciones endogámicas en el centro de Argentina. *Pediatric renal tumors and their possible relation with endogamous populations in central Argentina*. *Revista Cubana de Pediatría* 91:
- Alonso-Espinaco V, Giráldez MD, Trujillo C, van der Klift H, Muñoz J, Balaguer F, Ocaña T, Madrigal I, Jones AM, Echeverry MM et al. (2011) Novel MLH1 duplication identified in Colombian families with Lynch syndrome. *Genetics in Medicine* 13:155–160. doi: 10.1097/GIM.0B013E318202E10B
- Anagnostou P, Dominici V, Battaglia C, Lisi A, Sarno S, Boattini A, Calò C, Francalacci P, Vona G, Tofanelli S et al. (2019) Inter-individual genomic heterogeneity within European population isolates. *PLOS ONE* 14:e0214564. doi: 10.1371/JOURNAL.PONE.0214564
- Anderson RN (1996) The Quilombo of Palmares: A New Overview of a Maroon State in Seventeenth-Century Brazil. *Journal of Latin American Studies* 28:545.
- Andrade L de MB, Guedes GR, de Souza Noronha KVM, Silva CMS, Andrade JP and Martins ASFS (2021) Health-related vulnerability to climate extremes in homoclimatic zones of Amazonia and Northeast region of Brazil. *PLOS ONE* 16:e0259780. doi: 10.1371/JOURNAL.PONE.0259780
- Anselin L (2010) Local Indicators of Spatial Association-LISA. *Geographical Analysis* 27:93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x
- Araujo TF, Ribeiro EM, Arruda AP, Moreno CA, de Medeiros PFV, Minillo RM, Melo DG, Kim CA, Doriqui MJR, Felix TM et al. (2016) Molecular analysis of the CTSK gene in a cohort of 33 Brazilian families with pycnodysostosis from a cluster in a Brazilian Northeast region. *European Journal of Medical Research* 21:33. doi: 10.1186/s40001-016-0228-7
- Aranjo JEM (1996) O gordo e o magro: o nordeste segundo Gilberto Freyre e Djacir Menezes. *Revista de Ciências Sociais* 27:73–83.
- Arcos-Burgos M and Muenke M (2002) Genetics of population isolates. *Clin Genet* 61:233–247. doi: 10.1034/J.1399-0004.2002.610401.X
- Arruda AP (2016) Tricoepitelioma Múltiplo Familiar: delineamento fenotípico e genotípico.
- Arruda AP, Cardoso-dos-Santos AC, Mariath LM, Feira MF, Kowalski TW, Bezerra KRF, da Silva LACT, Ribeiro EM and Schuler-Faccini L (2020) A large family with CYLD cutaneous syndrome: medical genetics at the community level. *Journal of Community Genetics* 11:279–284. doi: 10.1007/s12687-019-00447-2
- Aswini Y and Varun S (2010) Genetics in public health: Rarely explored. *Indian J Hum Genet* 16:47. doi: 10.4103/0971-6866.69326
- Azevêdo E, Morton NE, Miki C and Yee S (1969) Distance and kinship in northeastern Brazil. *American Journal of Human Genetics* 21:1–22.
- Barrai I, Rodriguez-Larralde A, Dipierri J, Alfaro E, Acevedo N, Mamolini E, Sandri M, Carrieri A and Scapoli C (2012) Surnames in Chile: A study of the population of Chile through isonymy. *American Journal of Physical Anthropology* 147:380–388. doi: 10.1002/ajpa.22000

- Bazvand F, Keramatipour M, Riazi-Esfahani H and Mahmoudi A (2021) SPOAN syndrome: a novel mutation and new ocular findings; a case report. *BMC Neurology* 21:1–5. doi: 10.1186/S12883-021-02051-9/FIGURES/2
- Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A and Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data* 2018 5:1 5:1–12. doi: 10.1038/sdata.2018.214
- Bittles AH (2010) Consanguinity, Genetic Drift, and Genetic Diseases in Populations with Reduced Numbers of Founders. *Vogel and Motulsky's Human Genetics: Problems and Approaches (Fourth Edition)* 507–528. doi: 10.1007/978-3-540-37654-5_19
- Boycott KM, Hartley T, Biesecker LG, Gibbs RA, Innes AM, Riess O, Belmont J, Dunwoodie SL, Jovic N, Lassmann T et al. (2019) A Diagnosis for All Rare Genetic Diseases: The Horizon and the Next Frontiers. *Cell* 177:32–37. doi: 10.1016/J.CELL.2019.02.040
- Boyle B, Addor MC, Arriola L, Barisic I, Bianchi F, Csáky-Szunyogh M, De Walle HEK, Dias CM, Draper E, Gatt M et al. (2018) Estimating Global Burden of Disease due to congenital anomaly: an analysis of European data. *Arch Dis Child Fetal Neonatal Ed.* doi: 10.1136/ARCHDISCHILD-2016-311845
- BRASIL (2019a) Secretaria de Vigilância em Saúde | Ministério da Saúde | Boletim Epidemiológico - Síndrome congênita associada à infecção pelo vírus Zika. Brasília - DF, pp 1–31
- Brasil (2014) PORTARIA Nº 199, DE 30 DE JANEIRO DE 2014. In: Ministério da Saúde. https://bvsms.saude.gov.br/bvs/saudelegis/gm/2014/prt0199_30_01_2014.html. Accessed 19 Feb 2020
- BRASIL (2021) Saúde Brasil 2020/2021 : anomalias congênitas prioritárias para a vigilância ao nascimento / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças não Transmissíveis, 1st ed. Brasília
- BRASIL (2019b) Doenças raras. In: Ministério da Saúde. <http://www.saude.gov.br/saude-de-a-z/doencas-raras>. Accessed 23 Jul 2019
- Broquere C, Brudey K, Hartevelde CL, Saint-Martin C, Elion J, Giordano PC and Romana M (2010) Phenotypic Expression and Origin of the Rare β -Thalassemia Splice Site Mutation HBB:c.315 + 1G>T. <https://doi.org/10.3109/036302692010484956> 34:322–326. doi: 10.3109/03630269.2010.484956
- Cantarino C (2007) Uma câmera na mão e uma idéia na cabeça! Era só isso mesmo? *Ciência e Cultura* 59:51.
- Cardoso GC, de Oliveira MZ, Paixão-Côrtes VR, Castilla EE and Schuler-Faccini L (2018) Clusters of genetic diseases in Brazil. *Journal of Community Genetics* 1–8.
- Cardoso GC, de Oliveira MZ, Paixão-Côrtes VR, Castilla EE and Schuler-Faccini L (2019) Clusters of genetic diseases in Brazil. *Journal of Community Genetics* 10:121–128. doi: 10.1007/s12687-018-0369-1
- Cardoso TT (2007) O novo nordeste: Celso Furtado, Josué de Castro e o debate sobre desenvolvimentismo e o sertão nordestino nos anos 50.
- Cardoso-dos-Santos A, Fagundes N and Schuler-Faccini L (2020a) Planeta Mutante ou análise comparativa do ser mutante da cultura pop à genética clínica. *Genética na Escola* 15:2–9.
- Cardoso-Dos-Santos AC, Alves RS de M, Medeiros-de-Souza AC, Bremm JM, Gomes J do A, Alves RFS, Araujo VEM de and França GVA de (2021) National congenital anomaly

- registers in the world: historical and operational aspects. *Epidemiologia e Serviços de Saúde* 30:e2021075. doi: 10.1590/S1679-49742021000400015
- Cardoso-Dos-Santos AC, Boquett J, Oliveira MZ de, Callegari-Jacques SM, Barbian MH, Sanseverino MTV, Matte U and Schuler-Faccini L (2018) Twin Peaks: A spatial and temporal study of twinning rates in Brazil. *PLoS One* 13:e0200885. doi: 10.1371/journal.pone.0200885
- Cardoso-dos-Santos AC, Magalhães VS, Medeiros-de-Souza AC, Bremm J, Alves R, Araujo V, Macario E, Oliveira W and França G (2020b) International collaboration networks for the surveillance of congenital anomalies: a narrative review. *Revista do Sistema Único de Saúde do Brasil* 29:14. doi: <https://bit.ly/3vjlScB>
- Cardoso-dos-Santos AC, Medeiros-de-Souza AC, Bremm JM, Alves RFS, Araújo VEM de, Leite JCL, Schuler-Faccini L, Teresa Vieira Sanseverino M, de Menezes Karam S, Maria Félix T et al. (2021) List of priority congenital anomalies for surveillance under the Brazilian Information System on Live Births. *Epidemiologia e Serviços de Saúde* 30:1–8. doi: <https://bit.ly/2QTAj8t>
- Carrieri A, Sans M, Dipierri JE, Alfaro E, Mamolini E, Sandri M, Rodríguez-Larralde A, Scapoli C and Barraí I (2019) The structure and migration patterns of the population of Uruguay through isonymy. *Journal of Biosocial Science* 1–15. doi: 10.1017/s0021932019000476
- Castilla EE (2005) On being a medical geneticist. *American Journal of Medical Genetics* 138 A:197–198. doi: 10.1002/ajmg.a.30956
- Castilla EE and Adams J (1996) Genealogical Information and the Structure of Rural Latin-American Populations: Reality and Fantasy. *Human Heredity* 46:241–255. doi: 10.1159/000154361
- Castilla EE and Orioli IM (2003) ECLAMC: The Latin-American Collaborative Study of Congenital Malformations. *Community Genetics* 7:76–94. doi: 10.1159/000080776
- Castilla EE and Schuler-Faccini L (2014) From rumors to genetic isolates. *Genetics and Molecular Biology* 37:186–193. doi: 10.1590/S1415-47572014000200005
- Catunda H (1955) O grupo Feitosa na formação social do Nordeste. *Revista do Instituto do Ceará* 96–99.
- Celik B, Tomatsu SC, Tomatsu S and Khan SA (2021) Epidemiology of Mucopolysaccharidoses Update. *Diagnostics* 11:273. doi: 10.3390/DIAGNOSTICS11020273
- Chagas M and Nunes P (2016) Quilombolas vivem em situação de miséria em comunidades de Alagoas. In: G1 . <http://g1.globo.com/al/alagoas/noticia/2016/11/quilombolas-vivem-em-situacao-de-miseria-em-comunidades-de-alagoas.html>. Accessed 2 May 2019
- Charoute H, Bakhchane A, Benrahma H, Romdhane L, Gabi K, Rouba H, Fakiri M, Abdelhak S, Lenaers G and Barakat A (2015) Mediterranean Founder Mutation Database (MFMD): Taking Advantage from Founder Mutations in Genetics Diagnosis, Genetic Diversity and Migration History of the Mediterranean Population. *Human Mutation* 36:E2441–E2453. doi: 10.1002/HUMU.22835
- Chaubey G (2017) Piecing together genetic histories of isolated populations - On Biology. In: *On Biology*. <https://blogs.biomedcentral.com/on-biology/2017/06/21/piecing-together-genetic-histories-of-isolated-populations/>. Accessed 1 Mar 2022

- Chaves RG, da Veiga Pereira L, de Araújo FT, Rozenberg R, Carvalho MDF, Coelho JC, Michelin-Tirelli K, de Freitas Chaves M and Cavalcanti GB (2015) Consanguinity and founder effect for Gaucher disease mutation G377S in a population from Tabuleiro do Norte, Northeastern Brazil. *Clinical Genetics* 88:391–395. doi: 10.1111/cge.12515
- Christianson A, Howson CP and Modell B (2006) *March of Dimes - Global Report on Birth Defects*. New York
- Costa-Motta FM, Acosta AX, Abé-Sandes K, Bender F, Schwartz IVD, Giugliani R and Leistner-Segal S (2011) Genetic studies in a cluster of Mucopolysaccharidosis Type VI patients in Northeast Brazil. *Molecular Genetics and Metabolism* 104:603–607. doi: 10.1016/j.ymgme.2011.09.017
- Costa-Motta FM, Bender F, Acosta A, Abé-Sandes K, Machado T, Bomfim T, Boa Sorte T, da Silva D, Bittles AH, Giugliani R et al. (2014) A Community-Based Study of Mucopolysaccharidosis Type VI in Brazil: The Influence of Founder Effect, Endogamy and Consanguinity. *Human Heredity* 77:189–196. doi: 10.1159/000358404
- Crow JF and Mange AP (1965) Measurement of inbreeding from the frequency of marriages between persons of the same surname. *Eugenics Quarterly* 12:199–203. doi: 10.1080/19485565.1965.9987630
- D'Angelo CS, Hermes A, McMaster CR, Prichep E, Richer E, van der Westhuizen FH, Repetto GM, Mengchun G, Malherbe H, Reichardt JKV et al. (2020) Barriers and Considerations for Diagnosing Rare Diseases in Indigenous Populations. *Frontiers in Pediatrics* 8:797. doi: 10.3389/FPED.2020.579924/BIBTEX
- de Oliveira MZ, Schüler-Faccini L, Demarchi DA, Alfaro EL, Dipierri JE, Veronez MR, Colling Cassel M, Tagliani-Ribeiro A, Silveira Matte U and Ramallo V (2013) So Close, So Far Away: Analysis of Surnames in a Town of Twins (Cândido Godói, Brazil). *Annals of Human Genetics* 77:125–136. doi: 10.1111/ahg.12001
- del Campo M, Feitosa IML, Ribeiro EM, Horovitz DDG, Pessoa ALS, França GVA, García-Alix A, Doriqui MJR, Wanderley HYC, Sanseverino MVT et al. (2017) The phenotypic spectrum of congenital Zika syndrome. *American Journal of Medical Genetics, Part A* 173:841–857. doi: 10.1002/ajmg.a.38170
- Dipierri J, Rodríguez-Larralde A, Alfaro E, Scapoli C, Mamolini E, Salvatorelli G, Caramori G, de Lorenzi S, Sandri M, Carrieri A et al. (2011) A Study of the Population of Paraguay through Isonymy. *Annals of Human Genetics* 75:678–687. doi: 10.1111/j.1469-1809.2011.00676.x
- Dipierri J, Rodríguez-Larralde A, Barraí I, Camelo JL, Redomero EG, Rodríguez CA, Ramallo V, Bronberg R and Alfaro E (2014) Random inbreeding, isonymy, and population isolates in Argentina. *J Community Genet* 5:241–8. doi: 10.1007/s12687-013-0181-x
- Dipierri JE, Alfaro EL, Scapoli C, Mamolini E, Rodríguez-Larralde A and Barraí I (2005) Surnames in Argentina: A population study through isonymy. *American Journal of Physical Anthropology* 128:199–209. doi: 10.1002/ajpa.20027
- Dipierri JE, Gomez ELA, Rodríguez-Larralde A and Ramallo V (2016) Isonymic Relations in the Bolivia-Argentina Border Region. *Hum Biol* 88:191–200.
- Dipierri JE, Rodríguez-Larralde A, Alfaro EL and Barraí I (2007) Isonymic structure of the Argentine Northwest. *Annals of Human Biology* 34:498–503. doi: 10.1080/03014460701425427

- Elliott P and Wakefield J (2001) Disease clusters: should they be investigated, and, if so, when and how? *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 164:3–12. doi: 10.1111/1467-985X.00180
- EUROCAT (2012) Special Report: Congenital Anomalies are a Major Group of Mainly Rare Diseases . Newtownabbey, Co Antrim
- Evans JP, Powell BC and Berg JS (2017) Finding the Rare Pathogenic Variants in a Human Genome. *JAMA* 317:1904–1905. doi: 10.1001/JAMA.2017.0432
- Federhen A, Vairo FP, Vanzella C, Boer AP, Baldo G and Giugliani R (2014) Clinical research and rare diseases: the Brazilian situation. *J Bras Econ Saúde Supl.*(1):17–23.
- Félix TM, de Oliveira BM, Artifon M, Carvalho I, Bernardi FA, Schwartz IVD, Saute JA, Ferraz VEF, Acosta AX, Sorte NB et al. (2022) Epidemiology of rare diseases in Brazil: protocol of the Brazilian Rare Diseases Network (RARAS-BRDN). *Orphanet Journal of Rare Diseases* 2022 17:1 17:1–13. doi: 10.1186/S13023-022-02254-4
- Ferreira CR (2019) The burden of rare diseases. *American Journal of Medical Genetics Part A* 179:885–892. doi: 10.1002/AJMG.A.61124
- FinDis (2022) The Finnish Disease Heritage. <http://www.findis.org/index.php>. Accessed 1 Mar 2022
- Franco JF, Soares DC, Torres LC, Leal GN, Cunha MT, Honjo RS, Bertola DR and Kim CA (2016) Impact of early enzyme-replacement therapy for mucopolysaccharidosis VI: results of a long-term follow-up of Brazilian siblings. doi: 10.4238/gmr.15017850
- Freire-Maia A and Opitz JM (1981) Historical note: The extraordinary handless and footless families of Brazil – 50 years of acheiropodia. *American Journal of Medical Genetics* 9:31–41. doi: 10.1002/AJMG.1320090108
- Freire-Maia N (1957) Inbreeding in Brazil. *Am J Hum Genet* 9:284–98.
- Freire-Maia N (1990) Genetic effects in Brazilian populations due to consanguineous marriages. *American Journal of Medical Genetics* 35:115–117. doi: 10.1002/ajmg.1320350121
- Freire-Maia N, de Andrade L, de Athayde-Neto A, Cavalli IJ, Oliveira JC, Marcedillo FA and Coelho A (1978) Genetic Investigations in a Northern Brazilian Island. *Human Heredity* 28:401–410. doi: 10.1159/000152991
- Freyre G (2004) *Nordeste : aspectos da influência da cana sobre a vida e a paisagem do nordeste do Brasil*, 7th ed. Global Editora, São Paulo
- Furtado LET de A, Carneiro SAF, Cardoso-Dos-Santos AC, Magalhaes SC and Filho JRMC (2020) Geographic Clusters of Huntington’s Disease in Northwestern Ceará: a Preliminary Analysis. *Neurotherapeutics* 17:1–41.
- Garagnani P, Laayouni H, González-Neira A, Sikora M, Luiselli D, Bertranpetit J and Calafell F (2009) Isolated populations as treasure troves in genetic epidemiology: the case of the Basques. *European Journal of Human Genetics* 2009 17:11 17:1490–1494. doi: 10.1038/ejhg.2009.69
- Garcia P, Sousa SB, Ling TP, Conceição M, Seabra J, White KK and Diogo L (2010) Skeletal complications in mucopolysaccharidosis VI patients: Case reports. *J Pediatr Rehabil Med* 3:63–69. doi: 10.3233/PRM-2010-0108
- Garrido E, Chabás A, Coll MJ, Blanco M, Domínguez C, Grinberg D, Vilageliu L and Cormand B (2007) Identification of the molecular defects in Spanish and Argentinian

- mucopolysaccharidosis VI (Maroteaux-Lamy syndrome) patients, including 9 novel mutations. *Mol Genet Metab* 92:122–130. doi: 10.1016/J.YMGME.2007.06.002
- Gili JA, Poletta FA, Giménez LG, Pawluk MS, Campaña H, Castilla EE and López-Camelo JS (2016) Descriptive analysis of high birth prevalence rate geographical clusters of congenital anomalies in South America. *Birth Defects Res A Clin Mol Teratol* 106:257–266. doi: 10.1002/BDRA.23481
- Gili JA, Poletta FA, Pawluk M, Gimenez LG, Campaña H, Castilla E and López-Camelo JS (2015) High Birth Prevalence Rates for Congenital Anomalies in South American Regions. *Epidemiology* 26:e53–e55. doi: 10.1097/EDE.0000000000000345
- Giugliani L, Vanzella C, Zambrano MB, Donis KC, Wallau TKW, Costa FM Da and Giugliani R (2019a) Clinical research challenges in rare genetic diseases in brazil. *Genetics and Molecular Biology* 42:305–311. doi: 10.1590/1678-4685-gmb-2018-0174
- Giugliani R (2012a) Newborn screening for lysosomal diseases: current status and potential interface with population medical genetics in Latin America. *J Inherit Metab Dis* 35:871–877. doi: 10.1007/S10545-011-9436-Z
- Giugliani R (2012b) Mucopolysaccharidoses: From understanding to treatment, a century of discoveries. *Genetics and Molecular Biology* 35:924–931. doi: 10.1590/S1415-47572012000600006
- Giugliani R, Bender F, Couto R, Bochernitsan A, Brusius-Facchin AC, Burin M, Amorim T, Xavier A, Acosta, Purificação A et al. (2019b) Population Medical Genetics: Translating Science to the Community. *Genet Mol Biol*. doi: 10.1590/1678-4685-GMB-2018-0096
- Gomes F (2011) *Palmares*, 2nd ed. Contexto, São Paulo
- Hatzikotoulas K, Gilly A and Zeggini E (2014) Using population isolates in genetic association studies. *Briefings in Functional Genomics* 13:371. doi: 10.1093/BFGP/ELU022
- Hersch-Green EI, Turley NE and Johnson MTJ (2011) Community genetics: what have we accomplished and where should we be going? *Philos Trans R Soc Lond B Biol Sci* 366:1453–1460. doi: 10.1098/RSTB.2010.0331
- Homburger JR, Moreno-Estrada A, Gignoux CR, Nelson D, Sanchez E, Ortiz-Tello P, Pons-Estel BA, Acevedo-Vasquez E, Miranda P, Langefeld CD et al. (2015) Genomic Insights into the Ancestry and Demographic History of South America. *PLOS Genetics* 11:e1005602. doi: 10.1371/JOURNAL.PGEN.1005602
- Horovitz DDG, Acosta AX, de Rosso Giuliani L and Ribeiro EM (2015) Mucopolysaccharidosis type VI on enzyme replacement therapy since infancy: Six years follow-up of four children. *Mol Genet Metab Rep* 5:19–25. doi: 10.1016/J.YMGMR.2015.09.002
- IBGE (2010) Censo Demográfico 2010. In: Instituto Brasileiro de Geografia e Estatística. <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/>. Accessed 13 Apr 2017
- IBGE (2022) IBGE CIDADES. In: IBGE CIDADES. <https://cidades.ibge.gov.br/>. Accessed 1 Mar 2022
- ICBDSR (2014) INTERNATIONAL CLEARINGHOUSE FOR BIRTH DEFECTS SURVEILLANCE AND RESEARCH.
- Jobling MA (2001) In the name of the father: surnames and genetics. *Trends Genet* 17:353–7.
- Josahkian JA, Trapp FB, Burin MG, Michelin-Tirelli K, Magalhães APPS de, Sebastião FM, Bender F, Mari JF de, Brusius-Facchin AC, Leistner-Segal S et al. (2021) Updated birth

- prevalence and relative frequency of mucopolysaccharidoses across Brazilian regions. *Genetics and Molecular Biology* 44:1–6. doi: 10.1590/1678-4685-GMB-2020-0138
- Karageorgos L, Brooks DA, Pollard A, Melville EL, Hein LK, Clements PR, Ketteridge D, Swiedler SJ, Beck M, Giugliani R et al. (2007) Mutational analysis of 105 mucopolysaccharidosis type VI patients. *Hum Mutat* 28:897–903. doi: 10.1002/HUMU.20534
- Khan SA, Peracha H, Ballhausen D, Wiesbauer A, Rohrbach M, Gautschi M, Mason RW, Giugliani R, Suzuki Y, Orii KE et al. (2017) Epidemiology of mucopolysaccharidoses. *Molecular Genetics and Metabolism* 121:227–240. doi: 10.1016/j.ymgme.2017.05.016
- Kobayashi H (2019) Recent trends in mucopolysaccharidosis research. *Journal of Human Genetics* 64:1–11. doi: 10.1038/s10038-018-0534-8
- Kristiansson K, Naukkarinen J and Peltonen L (2008) Isolated populations and complex disease gene identification. *Genome Biology* 9:1–9. doi: 10.1186/GB-2008-9-8-109/TABLES/2
- Laberge AM, Michaud J, Richter A, Lemyre E, Lambert M, Brais B and Mitchell GA (2005) Population history and its impact on medical genetics in Quebec. *Clinical Genetics* 68:287–301. doi: 10.1111/J.1399-0004.2005.00497.X
- Larrandaburu M, Vianna FS, Anjos-daSilva A, Sanseverino MT and Schuler-Faccini L (2017) Zika virus infection and congenital anomalies in the Americas: opportunities for regional action. *Revista Panamericana de Salud Pública* 41:1–8. doi: 10.26633/rpsp.2017.174
- Lattig M, Gelvez N, Plaza S, Tamayo G, Uribe J, Salvatierra I, Bernal J and Tamayo M (2008) Deafness on the island of Providencia - Colombia: different etiology, different genetic counseling. *Genet Couns* 19:403–12.
- Leal GN, de Paula AC, Morhy SS, Andrade JL and Kim CA (2014) Advantages of early replacement therapy for mucopolysaccharidosis type VI: echocardiographic follow-up of siblings. *Cardiol Young* 24:229–235. doi: 10.1017/S1047951113000152
- Lee CE, Singleton KS, Wallin M and Faundez V (2020) Rare Genetic Diseases: Nature's Experiments on Human Development. *iScience*. doi: 10.1016/J.ISCI.2020.101123
- Liu Y, Chen L, Yuan Y and Chen J (2012) A study of surnames in china through isonymy. *American Journal of Physical Anthropology* 148:341–350. doi: 10.1002/ajpa.22055
- Locke AE, Steinberg KM, Chiang CWK, Service SK, Havulinna AS, Stell L, Pirinen M, Abel HJ, Chiang CC, Fulton RS et al. (2019) Exome sequencing of Finnish isolates enhances rare-variant association power. *Nature* 2019 572:7769 572:323–328. doi: 10.1038/s41586-019-1457-z
- Lopera F, Ardilla A, Martínez A, Madrigal L, Arango-Viana JC, Lemere CA, Arango-Lasprilla JC, Hincapié L, Arcos-Burgos M, Ossa JE et al. (1997) Clinical Features of Early-Onset Alzheimer Disease in a Large Kindred With an E280A Presenilin-1 Mutation. *JAMA* 277:793–799. doi: 10.1001/JAMA.1997.03540340027028
- Macedo-Souza LI, Kok F, Santos S, Amorim SC, Starling A, Nishimura A, Lezirovitz K, Lino AMM and Zatz M (2005) Spastic paraplegia, optic atrophy, and neuropathy is linked to chromosome 11q13. *Annals of Neurology* 57:730–737.
- Magalhães J (1970) O vaqueiro na história do Ceará. *Revista do Instituto do Ceará* 94–103.
- Makrythanasis P and Antonarakis S (2013) Pathogenic variants in non-protein-coding sequences. *Clin Genet* 84:422–428. doi: 10.1111/CGE.12272

- Manno N, Sherratt S, Boaretto F, Coico FM, Camus CE, Campos CJ, Musumeci S, Battisti A, Quinnell RJ, León JM et al. (2014) High prevalence of chitotriosidase deficiency in Peruvian Amerindians exposed to chitin-bearing food and enteroparasites. *Carbohydrate Polymers* 113:607–614. doi: 10.1016/J.CARBPOL.2014.07.011
- Manta FS de N, Pereira R, Vianna R, Rodolfo Beuttenmüller de Araújo A, Leite Góes Gitaí D, Aparecida da Silva D, de Vargas Wolfgramm E, da Mota Pontes I, Ivan Aguiar J, Ozório Moraes M et al. (2013) Revisiting the Genetic Ancestry of Brazilians Using Autosomal AIM-Indels. *PLoS ONE* 8:e75145. doi: 10.1371/journal.pone.0075145
- Matte U da S, Cardoso-dos-Santos AC, Rodrigues G, Oliveira MZ de, Tagliani-Ribeiro A, Heck S, Dresch V, Schosler M and Schuler-Faccini L (2019) Decifrando o “mistério dos gêmeos”: Vinte anos de pesquisa em Cândido Godói, Rio Grande do Sul. *Clinical and Biomedical Research* 39:107–115.
- Mazzu-Nascimento T, Melo DG, Morbioli GG, Carrilho E, Vianna FSL, Da Silva AA and Schuler-Faccini L (2017) Teratogens: A public health issue – A Brazilian overview. *Genetics and Molecular Biology* 40:387–397. doi: 10.1590/1678-4685-GMB-2016-0179
- Melnikova I (2012) Rare diseases and orphan drugs. *Nature Reviews Drug Discovery* 11:267–268. doi: 10.1038/nrd3654
- Melo DG, Sanseverino MTV, Schmalfluss T de O and Larrandaburu M (2021) Why are Birth Defects Surveillance Programs Important? *Frontiers in Public Health* 9:1670. doi: 10.3389/FPUBH.2021.753342/BIBTEX
- Menezes D (1970) *O outro nordeste: formação social do nordeste*, 2nd ed. Arte Nova, Rio de Janeiro
- Monasterio L (2017) Surnames and ancestry in Brazil. *PLOS ONE* 12:e0176890. doi: 10.1371/journal.pone.0176890
- Montoya JH and Morales OL (2009) Four cases of Jarcho-Levin’s syndrome in the province of Antioquia, Colombia. *Biomédica* 29:25–32. doi: 10.7705/BIOMEDICA.V29I1.38
- Mooney JA, Huber CD, Service S, Sul JH, Marsden CD, Zhang Z, Sabatti C, Ruiz-Linares A, Bedoya G, Fears SC et al. (2018) Understanding the Hidden Complexity of Latin American Population Isolates. *American Journal of Human Genetics* 103:707–726. doi: 10.1016/J.AJHG.2018.09.013/ATTACHMENT/3251C1AD-D240-4FC0-9655-172D03133BAA/MMC1.PDF
- Motta FMMC (2011) Análise de mutações no gene arilsulfatase B em pacientes com mucopolissacaridose tipo VI do Brasil: definição de uma possível origem comum em Monte Santo/BA.
- Mountford HS, Villanueva P, Fernández MA, Jara L, de Barbieri Z, Carvajal-Carmona LG, Cazier JB and Newbury DF (2020) The Genetic Population Structure of Robinson Crusoe Island, Chile. *Frontiers in Genetics* 11:669. doi: 10.3389/FGENE.2020.00669/BIBTEX
- Muenzer J (2011) Overview of the mucopolysaccharidoses. *Rheumatology*. doi: 10.1093/rheumatology/ker394
- Nikali K, Vanegas JJ, Burley MW, Martinez J, Lopez LM, Bedoya G, Wrong OM, Povey S, Unwin RJ and Ruiz-Linares A (2008) Extensive founder effect for distal renal tubular acidosis (dRTA) with sensorineural deafness in an isolated South American population. *American Journal of Medical Genetics Part A* 146A:2709–2712. doi: 10.1002/AJMG.A.32495

- Nioi P, Sigurdsson A, Thorleifsson G, Helgason H, Agustsdottir AB, Norddahl GL, Helgadottir A, Magnusdottir A, Jonasdottir A, Gretarsdottir S et al. (2016) Variant ASGR1 Associated with a Reduced Risk of Coronary Artery Disease . *New England Journal of Medicine* 374:2131–2141. doi: 10.1056/NEJMOA1508419/SUPPL_FILE/NEJMOA1508419_DISCLOSURES.PDF
- Norio R (2003) The Finnish disease heritage III: the individual diseases. *Human Genetics* 2003 112:5 112:470–526. doi: 10.1007/S00439-002-0877-1
- Novoa MC and Burnham TF (2011) Challenges for the universalization of clinical genetics: the Brazilian case. *Revista panamericana de salud publica = Pan American journal of public health* 29:61–8.
- Nussbaum RL, Mcinnes RR and Willard HF (2008) *Thompson & Thompson Genética Médica*, 7th ed. Elsevier, São Paulo
- Orphanet (2022) The portal for rare diseases and orphan drugs. <https://www.orpha.net/consor/cgi-bin/index.php>. Accessed 2 Mar 2022
- Panoutsopoulou K, Hatzikotoulas K, Xifara DK, Colonna V, Farmaki AE, Ritchie GRS, Southam L, Gilly A, Tachmazidou I, Fatumo S et al. (2014) Genetic characterization of Greek population isolates reveals strong genetic drift at missense and trait-associated variants. *Nature Communications* 2014 5:1 5:1–11. doi: 10.1038/ncomms6345
- Paradisi I, Ikonomu V and Arias S (2015) Spinocerebellar ataxias in Venezuela: genetic epidemiology and their most likely ethnic descent. *Journal of Human Genetics* 2016 61:3 61:215–222. doi: 10.1038/jhg.2015.131
- Passos-Bueno MR, Bertola D, Dain D, Horovitz G, Evangelista De Faria Ferraz V and Brito LA (2014) Genetics and genomics in Brazil: a promising future. *Molecular Genetics & Genomic Medicine* 2:280–291. doi: 10.1002/mgg3.95
- Peltonen L (2010) Molecular Background of the Finnish Disease Heritage. <https://doi.org/10.3109/07853899709007481> 29:553–556. doi: 10.3109/07853899709007481
- Peltonen L, Palotie A and Lange K (2000) Use of population isolates for mapping complex traits. *Nature Reviews Genetics* 2000 1:3 1:182–190. doi: 10.1038/35042049
- Peltonen L, Pekkarinen P and Aaltonen J (1995) Messages from an Isolate: Lessons from the Finnish Gene Pool. *Biol Chem* 376:697–704.
- Pena SDJ, Bastos-Rodrigues L, Pimenta JR and Bydlowski SP (2009) DNA tests probe the genomic ancestry of Brazilians. *Brazilian Journal of Medical and Biological Research* 42:870–876. doi: 10.1590/S0100-879X2009005000026
- Penchaszadeh VB (2002) Preventing congenital anomalies in developing countries. *Community Genetics* 5:61–69. doi: 10.1159/000064632
- Petry MFG, Nonemacher K, Sebben JC, Schwartz I v.d., Azevedo ACM, Burin MG, de Rezende AR, Kim CA, Giugliani R and Leistner-Segal S (2005) Mucopolysaccharidosis type VI: Identification of novel mutations on the arylsulphatase B gene in South American patients. *J Inher Metab Dis* 28:1027–1034. doi: 10.1007/S10545-005-0020-2
- Pineda-Trujillo N, Apergi M, Moreno S, Arias W, Lesage S, Franco A, Sepulveda-Falla D, Cano D, Buritica O, Pineda D et al. (2006) A genetic cluster of early onset Parkinson's disease in a Colombian population. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 141B:885–889. doi: 10.1002/AJMG.B.30375

- Pineda-Trujillo N, Carvajal-Carmona LG, Buriticá O, Moreno S, Uribe C, Pineda D, Toro M, García F, Arias W, Bedoya G et al. (2001) A novel Cys212Tyr founder mutation in parkin and allelic heterogeneity of juvenile Parkinsonism in a population from North West Colombia. *Neuroscience Letters* 298:87–90. doi: 10.1016/S0304-3940(00)01733-X
- Pogue RE, Cavalcanti DP, Shanker S, Andrade R V., Aguiar LR, de Carvalho JL and Costa FF (2018) Rare genetic diseases: update on diagnosis, treatment and online resources. *Drug Discovery Today* 23:187–195. doi: 10.1016/J.DRUDIS.2017.11.002
- Poletta FA, Castilla EE, Orioli IM and Lopez-Camelo JS (2007) Regional analysis on the occurrence of oral clefts in South America. *Am J Med Genet A* 143A:3216–3227. doi: 10.1002/AJMG.A.32076
- Poletta FA, Orioli IM and Castilla EE (2014) Genealogical data in population medical genetics: field guidelines. *Genetics and Molecular Biology* 37:171–185. doi: 10.1590/S1415-47572014000200004
- Pontes LMV de, Albuquerque ELS and Medeiros CN de (2012) A QUESTÃO DOS LIMITES MUNICIPAIS DO ESTADO DO CEARÁ - IPECE. Fortaleza
- Quelce-Salgado A (1964) A New Type of Dwarfism with Various Bone Aplasias and Hypoplasias of the Extremities. *Human Heredity* 14:63–66. doi: 10.1159/000151832
- Ramalho AS and Silva RB (2000) Community Genetics: a new discipline and its application in Brazil. *Cadernos de Saúde Pública* 16:261–263. doi: 10.1590/S0102-311X2000000100029
- Ramírez-Castro JL, Pineda-Trujillo N, Valencia A v., Muñetón CM, Botero O, Trujillo O, Vásquez G, Mora BE, Durango N, Bedoya G et al. (2002) Mutations in FOXL2 underlying BPES (types 1 and 2) in Colombian families. *American Journal of Medical Genetics* 113:47–51. doi: 10.1002/AJMG.10741
- Ramos VA (2022) South America . In: *Britannica*. <https://www.britannica.com/place/South-America>. Accessed 1 Mar 2022
- Redmonds G, King Turi and Hey David (2011) *Surnames, DNA, and family history*, 1st ed. Oxford University Press, Oxford
- Ribeiro EM, Lopes TF, Kerbage SC, Pessoa ALS and de Góes Cavalcanti LP (2017) From the perception of a cluster of cases of children with microcephaly to congenital Zika syndrome in Brazil: The lessons we have learned and the challenges that lie ahead of us. *Journal of Venomous Animals and Toxins Including Tropical Diseases* 23:1–3. doi: 10.1186/S40409-017-0107-X/METRICS
- Ribeiro LF (2018) *Território e memória: uma etnografia na comunidade remanescente quilombola do Muquém em União dos Palmares – Alagoas*. Universidade Federal de Alagoas
- Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E and Hughes DA (2015) Rare Disease Terminology and Definitions—A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. *Value in Health* 18:906–914. doi: 10.1016/J.JVAL.2015.05.008
- Ridley M (2006) *Evolução - 3ª Ed. 2006*, 3rd ed. Artmed, São Paulo
- Rodrigues MFF (2011a) Paisagens, geossímbolos e dimensões da cultura em comunidades quilombolas. *Mercator* 10:103–121.

- Rodrigues R (2011b) Casamentos entre primos ameaçam saúde de quilombolas em Alagoas. In: O Estado de S.Paulo. <https://www.estadao.com.br/noticias/geral,casamentos-entre-primos-ameacam-saude-de-quilombolas-em-alagoas-imp-,660626>. Accessed 2 May 2019
- Rodríguez-Larralde A, Barraí I and Alfonzo JC (1993) Isonymy structure of four Venezuelan states. *Ann Hum Biol* 20:131–45.
- Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, Fuentes M, Pizarro M, Everardo P, de Avila F et al. (2014) Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet* 10:e1004572. doi: 10.1371/journal.pgen.1004572
- Salzano FM and Sans M (2014) Interethnic admixture and the evolution of Latin American populations. *Genetics and Molecular Biology* 37:151–170. doi: 10.1590/S1415-47572014000200003
- Santos SC dos, Melo US, Lopes SS dos S, Weller M and Kok F (2013) A endogamia explicaria a elevada prevalência de deficiências em populações do Nordeste brasileiro? *Ciência & Saúde Coletiva* 18:1141–1150. doi: 10.1590/S1413-81232013000400027
- Santos S, Kok F, Weller M, Paiva FRL de and Otto PA (2010) Inbreeding levels in Northeast Brazil: strategies for the prospecting of new genetic disorders. *Genetics and Molecular Biology* 33:220–223.
- Scapoli C, Mamolini E, Carrieri A, Rodríguez-Larralde A and Barraí I (2007) Surnames in Western Europe: A comparison of the subcontinental populations through isonymy. *Theoretical Population Biology* 71:37–48. doi: 10.1016/J.TPB.2006.06.010
- Schieppati A, Henter J-I, Daina E and Aperia A (2008) Why rare diseases are an important medical and social issue. *The Lancet* 371:2039–2041. doi: 10.1016/S0140-6736(08)60872-7
- Schüler-Faccini L and Giugliani R (2012) Population Medical genetics in Brazil. *Genomics and Health in the Developing World*. Oxford University Press, New York, pp 1231–1236
- Schuler-Faccini L, Sanseverino M, Vianna F, da Silva A, Larrandaburu M, Marcolongo-Pereira C and Abeche A (2016) Zika virus: A new human teratogen? Implications for women of reproductive age. *Clinical Pharmacology & Therapeutics* 100:
- Schwartz IVD, Randon DN, Monsorens N, Moura de Souza CF, Horovitz DDG, Wilke MVMB and Brunoni D (2021) SARS-CoV-2 pandemic in the Brazilian community of rare diseases: A patient reported survey. *Am J Med Genet C Semin Med Genet* 187:301–311. doi: 10.1002/AJMG.C.31883
- Sitkin NA, Ozgediz D, Donkor P and Farmer DL (2014) Congenital Anomalies in Low- and Middle-Income Countries: The Unborn Child of Global Surgery. *World Journal of Surgery* 2014 39:1 39:36–40. doi: 10.1007/S00268-014-2714-9
- Tarskaia L, El'chinova GI, Scapoli C, Mamolini E, Carrieri A, Rodríguez-Larralde A and Barraí I (2009) Surnames in Siberia: A study of the population of Yakutia through isonymy. *American Journal of Physical Anthropology* 138:190–198. doi: 10.1002/ajpa.20918
- Ten Kate LP (2005) Community Genetics: A Bridge between Clinical Genetics and Public Health. *Public Health Genomics* 8:7–11. doi: 10.1159/000083330
- ten Kate LP (1998) Editorial. *Public Health Genomics* 1:1–2. doi: 10.1159/000016137

- Tomanin R, Karageorgos L, Zanetti A, Al-Sayed M, Bailey M, Miller N, Sakuraba H and Hopwood JJ (2018) Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. *Hum Mutat* 39:1788–1802. doi: 10.1002/HUMU.23613
- USA (1992) Orphan Drug Act - Public Law 97-414.
- Villanueva P, Nudel R, Hoischen A, Fernández MA, Simpson NH, Gilissen C, Reader RH, Jara L, Echeverry MM, Francks C et al. (2015) Exome Sequencing in an Admixed Isolated Population Indicates NFXL1 Variants Confer a Risk for Specific Language Impairment. *PLOS Genetics* 11:e1004925. doi: 10.1371/JOURNAL.PGEN.1004925
- Wakap SN, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y and Rath A (2020) Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics* 28:165–173. doi: 10.1038/s41431-019-0508-0
- Weller M, Tanieri M, Pereira JC, Almeida EDS, Kok F and Santos S (2012) Consanguineous unions and the burden of disability: A population-based study in communities of Northeastern Brazil. *American Journal of Human Biology* 24:835–840. doi: 10.1002/ajhb.22328
- Wexler NS (2012) Huntington's Disease: Advocacy Driving Science. <https://doi.org/10.1146/annurev-med-050710-134457> 63:1–22. doi: 10.1146/ANNUREV-MED-050710-134457
- Wexler NS, Lorimer J, Porter J, Gomez F, Moskowitz C, Shackell E, Marder K, Penchaszadeh G, Roberts SA, Gayán J et al. (2004) Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A* 101:3498–503. doi: 10.1073/pnas.0308679101
- WHO (2013) Priority diseases and reasons for inclusion - Rare diseases.
- WHO (2022) Birth Defects. <https://www.who.int/news-room/fact-sheets/detail/birth-defects>. Accessed 1 Mar 2022
- Wikipedia (2022) South America. https://en.wikipedia.org/wiki/South_America. Accessed 1 Mar 2022

CAPÍTULO XI: APÊNDICES

APÊNDICE 1: Demais trabalhos que detectaram a variante *ARSB*, c.1143-8T>G (foram inseridos todos os estudos relatados no ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/variation/888/>) e em uma busca não-sistemática no Pubmed/Medline, acessados em fevereiro de 2022).

N	Work	ID	Gender	Age at publication	Origin	Allele 1	Allele 2	Clinical information?	Age at onset/diagnosis	Parental consanguinity
1	Petry et al., 2005	7	-	-	Brazil	1533del23	c.1143-8T>G	yes	-	-
2	Petry et al., 2005	12	-	-	Brazil	c.1143-8T>G	c.1143-8T>G	yes	-	-
3	Karageorgos et al., 2007	86	-	7	Brazil	c.1143-1G>C	c.1143-8T>G	no	-	-
4	Karageorgos et al., 2007	87	-	7	Brazil	c.1143-1G>C	c.1143-8T>G	no	-	-
5	Karageorgos et al., 2007	88	-	12	Brazil	c.1143-8T>G	c.1143-8T>G + p.R315Q	no	-	-
6	Karageorgos et al., 2007	91	-	5	Brazil	c.1143-8T>G	c.1534_1556del	no	-	-
7	Karageorgos et al., 2007	104	-	NA	Brazil	c.1143-8T>G	p.R315Q	no	-	-
8	Garrido et al., 2007	ML2	-	-	Spain	c.1143-8T>G	c.1143-8T>G	yes	Birth/2y	-
9	Garrido et al., 2007	ML7	-	-	Spain	c.[237-243delGGTGCTC]	c.1143-8T>G	yes	4y/11y	-
10	Garrido et al., 2007	ML11	-	-	Argentina	c.1143-8T>G	c.1143-8T>G	yes	3m/7y	-
11	García et al., 2010*	1	M	9	Portugal	c.1143-8T>G	c.1143-8T>G	yes	6m/3y	no
12	García et al., 2010*	2	M	10	Portugal	c.149T>A	c.1143-8T>G	yes	13m/1.6y	no
13	Leal et al., 2013; Franco et al., 2016	1**	F	14	Brazil	c.1143-8T>G	c.1143-8T>G	yes	2y/3y	yes
14	Leal et al., 2013; Franco et al., 2016	2**	F	7	Brazil	c.1143-8T>G	c.1143-8T>G	yes	7m/7m	yes
15	Horovitz et al., 2015	P1	M	-	Brazil	c.1143-8T>G	c.1143-8T>G	yes	1y	yes

*Notam que se trata de "a common mutation in Brazilian patients (Petry et al., 2005) and in our Metabolic Unit in Hospital Pediatrico de Coimbra, Coimbra, Portugal", em observação pessoal. **Irmãs.

APÊNDICE 2: Reportagem para a BBC.

04/03/2022 11:45

Onde estão os pacientes com doenças raras no Brasil? – BBC News Brasil

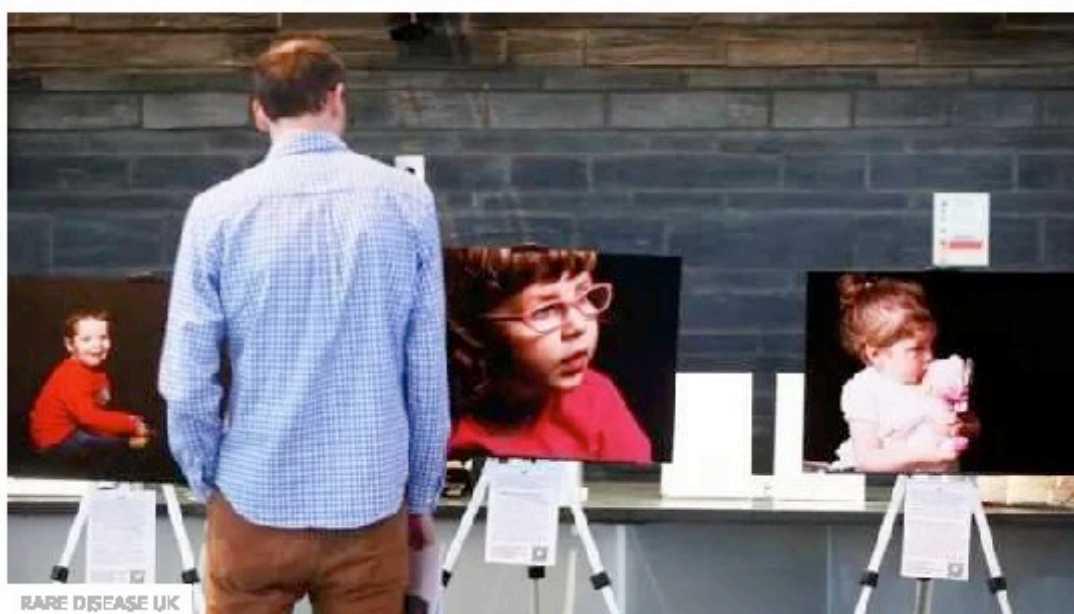
BBC NEWS BRASIL

Notícias Brasil Internacional Economia Saúde Ciência Tecnologia

Onde estão os pacientes com doenças raras no Brasil?

Evanildo da Silveira
De São Paulo para a BBC News Brasil

16 fevereiro 2019



A organização Rare Disease UK fez eventos em toda a Grã-Bretanha para conscientizar o público sobre estas doenças

Aos três anos de idade, a engenheira civil Bruna Galvão Marchesano de Freitas, hoje com 26, sofreu uma pequena fratura no pé. Nada demais. Teve o membro engessado e se curou com facilidade. Quatro anos mais tarde, no entanto, teve nova fratura, no fêmur. Dessa vez, a lesão chamou a atenção do ortopedista. Uma criança de sete anos não deveria ter os ossos tão densos – o que os torna inflexíveis e quebradiços. A investigação médica descobriu que a menina possui uma condição rara, de origem genética, chamada picnodisostose.

<https://www.bbc.com/portuguese/geral-47234287>

1/13

Bruna foi identificada durante o mais recente Censo Nacional de Isolados (CENISO), cujo objetivo é identificar populações brasileiras com alta frequência de enfermidades genéticas ou anomalias congênitas (causadas por fatores de risco genéticos, como casamentos consanguíneos), ou ambientais (talidomida, vírus Zika).

- [As crianças que vivem com as doenças mais raras do mundo](#)
- [A rara doença que faz com que pessoas parem de ouvir vozes masculinas](#)
- [Clique para assinar o canal da BBC News Brasil no YouTube](#)

Criado e realizado desde 2014 pelo Instituto Nacional de Ciência e Tecnologia de Genética Médica Populacional (Inagemp), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), o levantamento busca identificar "clusters" (aglomerados) ou comunidades com prevalência mais alta de uma doença rara.

"O objetivo é orientar quanto a políticas de saúde, aconselhamento genético, tratamento e prevenção nesses locais", explica a médica geneticista Lavinia Schuler Faccini, do Departamento de Genética da Universidade Federal do Rio Grande do Sul (UFRGS), coordenadora do CENISO.

A também médica geneticista Denise Cavalcanti, da Faculdade de Ciências Médicas da Universidade Estadual de Campinas (FCM-Unicamp), explica que doença rara, também conhecida como órfã, é aquela que acomete uma pequena proporção de indivíduos na população.

Talvez também te interesse



O que mudaria na prática com o 'PL dos agrotóxicos'



Coronavírus: A longa lista de possíveis sequelas da covid-19

Como a fome afeta a saúde física - e a mental - no longo prazo

Por que comer menos pode ser o segredo para uma vida mais longa e saudável

No Brasil, em geral, se adota os mesmos parâmetros de frequência da Europa, cerca de 1 pessoa em cada 2 mil. "Seja como for, em sua grande maioria (cerca de 80%) são de origem

genética.”

Como são descobertos os grupos com doenças raras

Apesar de cada doença rara afetar um número relativamente pequeno de pessoas, no conjunto elas acometem um grande número, pois existem de 5 a 8 mil doenças deste tipo conhecidas no mundo.

“Estima-se que pelo menos 6% da população tenha uma delas”, diz Lavínia. “Por isso, espera-se que existam 13 milhões de portadores no Brasil. Mas chamo a atenção para uma particularidade do CENISO: o objetivo não é mapear todas as pessoas com uma dessas enfermidades, mas comunidades onde uma delas pode não ser rara localmente e necessitar de uma política de saúde individualizada em nível de município ou estado, por exemplo.”

ARQUIVO PESSOAL

| Família que tem duas filhas, Bruna e Laura, com a condição rara plicnosistose

Nesse sentido, o biomédico Augusto César Cardoso dos Santos, doutorando em Genética e Biologia Molecular pela UFRGS, que também participa do CENISO, acrescenta que algumas doenças que são raras na população em geral podem concentrar-se com frequência inesperadamente elevada em determinados locais ou grupos populacionais. “O principal objetivo do Censo é mapear estes lugares (ou populações)”, explica.

 Podcast

Isso é feito de duas maneiras. A primeira é por meio de revisão da literatura científica, para identificar possíveis *clusters* (aglomerados, em Inglês), que já foram ou estão sendo estudados por algum grupo de pesquisa. "A outra é pela comunicação oral ou escrita de um 'rumor', isto é, qualquer pessoa (pesquisador da área ou não) pode informar a existência de um aglomerado geográfico de doença rara", explica Santos.

"O nosso trabalho é investigar essas informações, avaliar se elas consistentemente caracterizam-se como *clusters*, se a causa responsável pela enfermidade foi encontrada e se a população está sendo atendida, por exemplo."

BBC Lê

A equipe da BBC News Brasil lê para você algumas de suas melhores reportagens

Episódios

Onde estão as doenças raras no Brasil

O CENISO não é como um censo do IBGE, que rastreia e cadastra toda a população em anos determinados. O levantamento do Inagem vai sendo atualizados continuamente. Desde a sua criação, ela já recebeu centenas de rumores de *clusters*, dos quais 86 foram confirmados até em 2014, número que subiu para 144 em 2018, últimos dados publicados.

Santos diz que é interessante notar que a maior parte dos rumores está concentrada na região Nordeste do Brasil. "Embora cada caso seja único e necessite de investigação específica, podemos listar diversos fatores que podem nos ajudar a entender esta distribuição anormal", diz. "Entre eles, casamentos consanguíneos por motivos culturais ou religiosos, populações geograficamente isoladas, efeito genético de fundador (quando uma mutação surge num determinado local) e famílias com grande número de filhos."

ARQUIVO PESSOAL

Laura, 10 anos, portadora de piodisostose; um aglomerado de pessoas com a condição foi descoberto no Ceará

Ele conta que entre os *clusters* descobertos estão o de aniridia congênita (enfermidade ocular que leva à ausência de íris), em Alagoas; de albinismo (caracterizado pela ausência de pigmento na pele), no Maranhão; e de Doença de Machado-Joseph (caracterizado por perda de controle muscular e coordenação motora), no Rio Grande do Sul.

Há ainda o de mucopolissacaridose 6 (doença causada por deficiência de enzimas, que leva a disfunção na lubrificação dos órgãos, causando danos progressivos que podem afetar o cérebro, olhos, ouvidos, coração, fígado, ossos e articulações), em Monte Santo, na Bahia. "Ela tem uma prevalência mundial de uma pessoa afetada a cada 350 mil", diz Lavínia. "Na cidade baiana é em torno de um em cada 5 mil. Ainda é rara, mas proporcionalmente é muito alta (30 vezes mais alta)."

Apesar de doença, vida normal

O *cluster* de plicnodisostose foi descoberto no Ceará por Denise, que não participa do CENISO. "Em 2013 visitei, a convite da colega médica geneticista Erlane Ribeiro, o Hospital Infantil Albert Sabin em Fortaleza", conta. "Nessa visita de dois dias colaborando com Erlane para fazer o diagnóstico preciso de pacientes com suspeita de displasias esqueléticas (enfermidades que afetam o crescimento), nos deparamos com um grande número de pessoas com plicnodisostose. Identificamos mais de 25 pacientes em 22 famílias atendidas por ela. Visto se tratar de uma condição rara, cuja frequência estimada era de 1 (um) paciente a cada 1-1,7 milhão de indivíduos, propusemos um estudo inicial para identificar as bases moleculares daqueles pacientes."

De acordo com Denise, que também descobriu casos em outros Estados, como o da família de Bruna, que mora em Formosa (GO), a plicnodisostose é uma condição genética com padrão de herança autossômico recessivo, ou seja, mais frequente em filhos de casais consanguíneos. Ela começa a se manifestar depois do 1º ao 2º ano de vida.

"Seus portadores se caracterizam pela baixa estatura, uma face característica, que é pequena, com uma região frontal proeminente, olhos protrusos (mais para frente, em relação à posição que seria normal), uma hipoplasia (desenvolvimento deficiente ou incompleto de tecido ou órgão) maxilar e queixo pequenos", explica. "As mãos também são muito pequenas, com as unhas curtas e alargadas."

ARQUIVO PESSOAL

Bruna, 26 anos, também é portadora de plicnodisostose. A doença genética com padrão de herança autossômico recessivo, ou seja, mais frequente em filhos de casais consanguíneos

Embora não sejam consanguíneos, os pais de Bruna têm outra filha, Laura, de 10 anos, com plicnodisostose. "Descobrimos a condição genética da segunda pela experiência que tivemos com a primeira", conta a mãe delas, a dentista Iesa Galvão Lisboa Marchesano de Freitas. "Elas têm baixa estatura (em torno de 1,35 m), ângulo mandibular reto, o que provoca pouco

crescimento da mandíbula e apinhamento dentário, além de falanges terminais dos pés e mãos hipoplásicas, provocando um encurtamento dos dedos."

Segundo Iesa, isso não impede suas filhas de levarem uma vida normal. "Em primeiro lugar, a plicnodisostose não é uma doença, e sim uma condição genética, que felizmente não afeta em nada a vida das duas", diz. "Claro que em caso de fratura é necessário mais repouso, mas elas têm vida normal, estudam, trabalham, fazem academia, têm vida social. Todos nós convivemos muito bem com a condição genética das meninas, elas têm total consciência do problema e sabem na nossa ausência procurar socorro em caso de alguma fratura."

Já assistiu aos nossos novos vídeos no [YouTube](#)? Inscreva-se no nosso canal!

APÊNDICE 3: scientia brasilis.

Poema publicado na revista Aurora: Revista de Arte, Mídia e Política (13(37):177). Disponível em: <https://revistas.pucsp.br/aurora/article/view/46436>

POESIA ■ você viu?
mais uma ameaça de corte orçamentário
para a Ciência do Brasil

scientia brasilis
Augusto César
Cardoso-dos-Santos¹

como cientista que sou,
ando questionando-me quando foi que deixamos minha profissão
parecer tão desimportante à Nação

como cientista que sou,
aprofundo-me no questionamento, duvidando se minha profissão
é realmente tão importante à Nação

de que importa saber
que a pele da tilápia cicatriza
que a goma do cajueiro ameniza
que o leite materno previne
que Manuel Bandeira fez revolução

para que saber o que causa microcefalia
se ao cérebro brasileiro nunca foi dada tão pouca valia
ou de onde veio o homem americano
se páginas de horóscopo nos dizem para onde vamos
ou para que dizer à família do sertão
que o mau que a acomete é genética e não maldição

cinema marginal
esquistossomose
inteligência artificial
perna mecânica
Hermeto Pascoal
produto vetorial
plano real
saúde mental:
problema real

PROBLEMA REAL!

o cientista brasileiro precisa ocupar-se de algum problema real,
tal qual: como convencer o seu povo
que conhecimento é poder, liberdade e cura
sem que seja preciso dizer de novo
de novo
e de novo?

177

¹ Biomédico pela UFPI, mestre em Genética e Biologia Molecular pelo Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS, onde atualmente cursa doutorado. Orcid: 0000-0002-1499-9105