AVALIAÇÃO DA FREQUÊNCIA E REPRESENTATIVIDADE DE BIOBANCOS EM POPULAÇÕES SUB-REPRESENTADAS PARA ESTUDOS DE ASSOCIAÇÃO GENÔMICA AMPLA RELACIONADOS AO GLAUCOMA PRIMÁRIO DE ÂNGULO ABERTO

Palavras-Chave: Glaucoma Primário de ângulo Aberto (GPAA); Estudo de Associação Genômica Ampla (*Genome Wide Association Studies - GWAS*); BioBancos/Biorepositórios.

Kenji William Onaga Mees, **LEGH / CBMEG - UNICAMP**

Thiago Adalton Rosa Rodrigues, **LEGH / CBMEG - UNICAMP**

José Paulo Cabral de Vasconcellos, **Departamento de Oftalmologia - UNICAMP**

Mônica Barbosa de Melo **(Orientadora), LEGH / CBMEG - UNICAMP**

1. INTRODUÇÃO:

1.1. Glaucoma Primário de Ângulo Aberto

O glaucoma se refere a um grupo heterogêneo de neuropatias ópticas, caracterizado por provocar uma degeneração lenta, progressiva e irreversível do nervo óptico, favorecendo defeitos característicos como a perda das células ganglionares da retina, afilamento da camada de fibras nervosas da retina e aumento da escavação do disco óptico, com perda correspondente do campo visual[1,2].

O Glaucoma Primário de Ângulo Aberto (GPAA) é o subtipo mais prevalente de glaucoma no mundo, sendo caracterizado pelo ângulo do seio camerular aberto à gonioscopia, além de uma patogênese idiopática e acometimento de pacientes jovens ou adultos[3,4,5].

A etiologia do GPAA ainda é pouco compreendida, havendo conhecimento limitado acerca de seus fatores de risco[1,2,4]. Dentre tais precursores, a literatura cita: idade avançada, pressão intraocular elevada, histórico familiar e ancestralidade africana[5,6,7]. Desta forma, a etnia e base genética de cada coorte ressaltam a disparidade de sua prevalência e incidência em diferentes regiões e/ou grupos étnicos[8].

1.2. Genética do GPAA

Sua transmissão genética se manifesta majoritariamente através do padrão de herança complexo[9], onde uma intricada rede de interação envolvendo diferentes variantes de pequeno efeito (muitas vezes em múltiplos genes), somados a fatores ambientais e sistêmicos, podem modular a suscetibilidade para a expressão de um traço fenotípico (Herança poligênica/multifatorial)[10].

Desta forma, Estudos de Associação Genômica Ampla (*Genome-Wide Association Studies – GWAS*) representam uma abordagem eficaz baseada na genotipagem de variantes em larga escala, viabilizando a identificação de regiões genômicas associadas a um traço complexo e/ou endofenótipo[9,11].

Contudo, apesar da disseminação da técnica de *GWAS* ter gerado notáveis avanços no conhecimento acerca das bases genéticas do GPAA, é evidente que estes estudos ainda se mantêm limitados ao uso de coortes majoritariamente homogêneas, sobretudo compostas por ancestralidades europeias e asiáticas, com cobertura de ~93,73% dos participantes em estudos até o momento[12,13]. Como reflexo de tal ausência, indivíduos de ancestralidade miscigenada (afro-americanos, afro-caribenhos, hispânicos e latino-

americanos), e/ou minoritárias, permanecem amplamente sub-representados (6,27%), evidenciando a necessidade de novas investigações nestas populações[12,13]

Desta maneira, a fim de aumentar melhor elucidar as bases genéticas envolvidas no GPAA, é primordial que haja um esforço para a inclusão de mais populações sub-representadas nos estudos de associação genômica ampla.

1.3. Biobancos

Biobancos são coleções que realizam a organização, estoque e processamento de material biológico, adicionalmente a informações clínicas e/ou epidemiológicas dos participantes. Os biobancos possuem função central na realização de pesquisas em diversas áreas do conhecimento, ao viabilizar estudos em uma grande quantidade de amostras, sendo um requerimento crítico, por exemplo, em estudos para uma medicina personalizada.[14,15].

O termo "Biobanco" é frequentemente utilizado de forma equivocada para categorizar todo e qualquer repositório amostral de bioespécimes, desconsiderando implicações éticas e legais que cerceiam tal estrutura. Portanto, no Brasil, estas coleções são definidas tanto pela Resolução CNS nº 441/2011[16], quanto pela Portaria nº 2201/2011[17] do Ministério da Saúde como uma: "Coleção organizada de material biológico humano e informações associadas, coletado e armazenado para fins de pesquisa, conforme regulamento ou normas técnicas, éticas e operacionais pré-definidas, sob responsabilidade e gerenciamento institucional, sem fins comerciais". Tal definição difere da diretriz de "Biorrepositório", cuja definição também consta nos documentos supracitados, mas que, todavia, se diferencia entre outros parâmetros, por ser uma: "Coleção organizada para a execução de um projeto de pesquisa específico e estando sob o gerenciamento apenas do respectivo pesquisador, não estando sob jurisdição do Estado".

Tendo em vista a importância dessas coleções para a medicina personalizada, se torna evidente que uma composição multiétnica representativa de diversas etnias garante que tratamentos e intervenções médicas sejam efetivas em diferentes populações, além de, no contexto de um *GWAS*, aumentar o poder de descrição de novas associações[19].

Contudo, até o momento, múltiplas populações permanecem sub-representadas nestes acervos, o que transparece a necessidade de expansão não apenas em diversidade, mas em robustez[20]. Deste modo, a presente revisão busca descrever os principais Biobancos em *GWAS*s de populações sub-representadas para o GPAA, explorando sua hipotética capacidade de representar tais coortes.

MÉTODOS:

A pesquisa dos artigos foi conduzida através do *NHGRI-EBI GWAS Catalog,* descrito como o principal repositório público de *GWAS*s, concatenando e articulando importantes informações literárias depositadas no *PubMed*, além de dados biomédicos genômicos[21]. Vale mencionar que artigos com meta-análises em coortes multiétnicas foram ramificados em estudos independentes pelo catálogo, aumentando a acurácia descritiva de cobertura. Dito isto, a seleção dos trabalhos para análise ocorreu em quatro fases, conforme os critérios: **1º)** Seleção de trabalhos rotulados com o fenótipo: "*open-angle glaucoma*", sob registro de identificação EFO_0004190; **2º)** Especificidade para análises para GPAA e/ou seus endofenótipos, excluindo investigações em outros traços fenotípicos; **3º)** Inclusão de trabalhos nos últimos 10 anos; **4º)** Exclusão de estudos oriundos de coortes europeias (*"European")* e/ou asiáticas orientais (*"East Asian")*, tanto em caráter de descoberta, quanto de replicação, com a categorização de ancestralidade sendo realizada pelo próprio catálogo[12]. Subsequentemente, os estudos selecionados foram analisados e categorizados manualmente, visando a identificação de coortes sub-representadas (não provenientes da

Europa nem Ásia oriental) e miscigenadas. A definição de Biobanco, se deu com base nos seguintes critérios: **1º)** Deve ser associado a uma instituição/consórcio e não a uma pesquisa específica; **2º)** Deve possuir domínio na internet para acesso de informações sobre a coleção. Mais do que uma caracterização de biobancos, essa abordagem buscou revelar a sub-representação sistemática de populações diversas na literatura genômica atual, enfatizando a importância de ampliar a diversidade nas pesquisas. Os achados que se seguem reforçam essa necessidade e apontam potenciais variantes de interesse em contextos populacionais historicamente negligenciados.

RESULTADOS E DISCUSSÃO:

Para a seleção dos trabalhos a serem avaliados, os critérios de inclusão produziram os seguintes resultados em cada fase: 1º) Dos 7.255 estudos contidos no *GWAS Catalog,* 60 foram classificados como *"open-angle glaucoma"*; 2º) Selecionaram-se 42 estudos, pois eram específicos para o GPAA e/ou seus endofenótipos; 3º e 4º) Selecionaram-se 29 estudos, por conterem populações sub-representadas em suas coortes (não são europeias nem asiáticas orientais). Os trabalhos que se enquadravam nos critérios acima eram provenientes de 13 artigos, dos quais foi verificada a origem de suas coortes, permitindo a identificação de 14 Biobancos conforme a definição que consta nos métodos.

Em sequência, foi ponderada a ancestralidade e país de recrutamento das coortes que compunham estes acervos e que foram utilizadas em cada estudo **(Tabela 1)**.

BioBancos	Ancestralidade	País de Recrutamento
BioVU (Vanderbilt University Medical Center's biorepository)	Afro-americano ou afro-caribenho [22]	EUA [23]
BioME (Charles Bronfman Institute for Personalized Medicine Biobank)	Afro-americano ou afro-caribenho / Hispânico ou Latino-americano [22]	EUA [24]
Kaiser Permanente GERA (Genetic Epidemiology Research on Adult Health and Aging)	Afro-americano ou afro-caribenho [22]	EUA [25]
Women's Health Initiative	Afro-americano ou afro-caribenho [22]	EUA [26]
All of Us (AOU)	Africano não especificado [22]	EUA [27]
Penn Medicine BioBank (PMBB)	Afro-americano ou afro-caribenho [22]	EUA [27]
Million Veteran Program (MVP)	Afro-americano ou afro-caribenho [22]	EUA [27]
eMERGE	Africano não especificado [22]	EUA [28]
BBofA (Biobank of the America (US))	Afro-americano ou afro-caribenho [22]	EUA [29]
CCPM (Colorado Center for Personalized Medicine (US)	Afro-americano ou afro-caribenho / Hispânico ou Latino-americano [22]	EUA [30]
GNH (Genes & Health (UK)	Sul asiático [22]	Reino Unido [31]
MGI (Michigan Genomics Initiative (US)	Africano não especificado [22]	EUA [32]
UCLA ATLAS Community Health Initiative	Afro-americano ou Afro-caribenho / Africano não especificado / Sul asiático [22]	EUA [33]
UK Biobank	Afro-americano ou Afro-caribenho / Africano não especificado / Sul asiático [22]	Reino Unido [34]

***Tabela 1**: Tabela dos BioBancos selecionados, com a ancestralidade e país de recrutamento de suas coortes*

Assim, a partir dos resultados, é possível observar uma notável prevalência da ancestralidade "afro-americana ou afro-caribenha" (71,5%) e da "africana não especificada" (35,7%) nas coortes, e que, sobretudo, houve um recrutamento majoritariamente nos EUA, seguido pelo Reino Unido, indicando a falta de diversidade nestes grupos de estudo. Nenhum Biobanco foi identificado em países Africanos.

Tal panorama se torna ainda mais evidente ao se considerar a origem diaspórica destas populações recrutadas nos EUA. Atualmente o continente africano é divido em 5 principais grupos etnolinguísticos, mas apenas um deles é representado de forma predominante na diáspora para a América do Norte (família linguística Níger-Congo)[35]. Em paralelo, a população recrutada no Reino Unido é principalmente

composta por afro-caribenhos e de uma variedade de populações africanas, especialmente advindas de Gana e Nigeria, indicando que compartilham um histórico diaspórico semelhante ao dos EUA[36,37].

Portanto, o estudo de pequenos grupos diaspóricos africanos residindo em países como EUA e Reino Unido, somado a sua categorização em grupos de definição ampla, vão continuar falhando em representar a alta diversidade genética do continente, persistindo a dificuldade de mitigar as desigualdades no acesso à medicina personalizada.

Os resultados também evidenciaram uma marcante ausência de populações sub-representadas, como populações sul-asiáticas (21,4%), da Ásia Central, do Oriente Médio e hispânicas ou latino-americanas (14,28%), não somente em caráter de ancestralidade, mas também como países de recrutamento.

Neste contexto, vale ressaltar a escassez na representação das populações Hispânicas ou Latino-Americanas, além de miscigenadas o que inclui o Brasil, que revela um importante perfil genético para identificação de novas associações, sendo a população mais miscigenada do mundo segundo recente publicação de Nunes, K. *et al.* (2025)[38]. Tal composição é oriunda de diversos processos migratórios, incluindo uma base que permeia populações africanas, europeias e ameríndios, esta última, em particular, sendo uma das menos estudadas do mundo[38].

CONCLUSÃO:

Em suma, essa investigação identificou e categorizou os principais Biobancos utilizadas nos *GWAS*s para GPAA mundialmente. Há uma predominância das populações "Afro-americana ou afro-caribenha" e "Africano não especificado" recrutadas nos EUA, sendo priorizadas coortes ainda homogêneas e menos descritivas em diversidade genética, quando comparadas à África continental. Ademais, é notável a falta de Biobancos com amostragem de ancestralidades "Hispânica ou Latino-americana", resultando na carência de elucidação em suas bases genéticas, exibindo uma valiosa oportunidade para o estudo da população brasileira nesta área do conhecimento. São necessários novas iniciativas e cooperações internacionais para introduzir uma maior cobertura de etnias sub-representadas nos Biobancos para GPAA, permitindo sua maior representação e, consequentemente, acesso a uma medicina de precisão.

REFERÊNCAIS BIBLIOGRÁFICAS:

1. Weinreb, R. N.; Khaw, P. T. Primary open-angle glaucoma. The Lancet, 363:1711-20, 2004.
2. European Glaucoma Society Terminology and Guidelines for Glaucoma, 5th Edition. The British Journal of Ophthalmology, 105:1-169, 2021.
3. Quigley, H. A.; Broman. A. T. The number of people with glaucoma worldwide in 2010 and 2020. The British Journal of Ophthalmology, 90:262-7, 2006.
4. Coleman AL, Miglior S. Risk factors for glaucoma onset and progression. Survey of Ophthalmology, 53:S3-10, 2008.
5. Schuster AK, Wagner FM, Pfeiffer N, Hoffmann EM. Risk factors for open-angle glaucoma and recommendations for glaucoma screening. Ophthalmologe, 118:145-152, 2021.
6. Boland, M. V., Quigley HA. Risk factors and open-angle glaucoma: classification and application. Journal of Glaucoma, 16:406-18, 2007.
7. Weinreb R.N. et al. Primary open-angle glaucoma. Natural Reviews. Disease Primers, 2:16067, 2016.
8. Tham, Y.C. et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology, 121:2081-90, 2014.
9. Janssen, S. F. et al. The vast complexity of primary open angle glaucoma: disease genes, risks, molecular mechanisms and pathobiology. Progress in Retinal and Eye Research. 37: 31-67, 2013.
10. Nussbaum RL, McInnes RR, Willard HF. Thompson & Thompson – Genética Médica. 7ª Ed. Rio de Janeiro: Editora Guanabara Koogan S.A, 2008, pp. 183-229.
11. Tam, V., *et al.* Benefits and limitations of genome-wide association studies. Nature Review Genetics. 20, 467–484, 2019.
12. Morales J, *et al.*. A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. Genome Biology. 19(1):21, 2018.

13. Mills, M. C., Rahal C. The GWAS Diversity Monitor tracks diversity by disease in real time *Nature Genetics.* 52 242–3, 2020.

14. Annaratone, L. *et al.* Pathology and Biobanking Working Group. Basic principles of biobanking: from biological samples to precision medicine for patients. Virchows Arch. 479(2):233-246,2021.

15. Harati, M. D. *et al.* An Introduction to Starting a Biobank. Methods in Molecular Biology. ;1897:7-16, 2019.

16. BRASIL. Ministério da Saúde. Conselho Nacional de Saúde. Resolução CNS nº 441, de 12 de maio de 2011. Diretrizes para análise ética de projetos de pesquisas que envolvam armazenamento de material biológico humano ou uso de material armazenado em pesquisas anteriores. **Diário Oficial União**. 18 de julho de 2011; Seção 1:60-61. Disponível em: http://conselho.saude.gov.br/resolucoes/2011/Reso441.pdf.

17. BRASIL. Ministério da Saúde. Portaria nº 2.201, de 14 de setembro de 2011. Diretrizes Nacionais para Biorrepositório e Biobanco de Material Biológico Humano com Finalidade de Pesquisa. **Diário Oficial União**. 15 de setembro de 2011; Seção 1:40-42. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt2201_14_09_2011.html.

18. Mendy, M. *et al*. Biospecimens and Biobanking in Global Health. Clinics in Laboratory Medicine .38(1):183-207, 2018.

19. Peterson, R. E., *et al.*. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. Cell.179:589-603, 2019.

20. Lee, Y. S., *et al*. Ethical considerations for biobanks serving underrepresented populations. Bioethics. 39:240-249, 2025.

21. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014 Jan;42(Database issue): D1001-6.

22. Cerezo, M. *et al.* The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. Nucleic Acids Res. 53:D998-D1005, 2025.

23. Restrepo, N. A., *et al*. Local genetic ancestry in CDKN2B-AS1 is associated with primary open-angle glaucoma in an African American cohort extracted from de-identified electronic health records. BMC Med Genomics.11:70, 2018.

24. NHLBI TOPMed - NHGRI CCDG: The BioMe Biobank at Mount Sinai. **dbGaP Study**. Disponível em: https://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs001644.v1.p1. Acesso em: 02/08/2025.

25. Resource for Genetic Epidemiology Research on Aging (GERA). **dbGaP Study**. Disponível em: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000674.v2.p2. Acesso em: 02/08/2025.

26. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials. 19:61-109, 1998.

27. Verma, S. S. *et al* A multi-cohort genome-wide association study in African ancestry individuals reveals risk loci for primary open-angle glaucoma. Cell. 187:464-480.e10, 2024.

28. Verma, S. S. *et al.* Epistatic Gene-Based Interaction Analyses for Glaucoma in eMERGE and NEIGHBOR Consortium. PLoS Genet. 12(9):e1006186, 2016.

29. Zhou, W., *et al.* Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. Cell Genom. 2:100192, 2022.

30. Wiley, L. K., *et al.* Building a vertically integrated genomic learning health system: The biobank at the Colorado Center for Personalized Medicine. Am J Hum Genet.111:11-23, 2024.

31. About our study. **GENESANDHEALTH**. Disponível em: https://www.genesandhealth.org/volunteer/. Acesso em: 02/08/2025

32. Zawistowski, M. *et al*. The Michigan Genomics Initiative: A biobank linking genotypes and electronic clinical records in Michigan Medicine patients. Cell Genom. 3(2):100257, 2023.

33. Johnson R., *et al.* Leveraging genomic diversity for discovery in an electronic health record linked biobank: the UCLA ATLAS Community Health Initiative. Genome Med. 14:104, 2022.

34. Our participants.**UK Biobank.** Disponível em: https://www.ukbiobank.ac.uk/about-our-data/our-participants/. Acesso em: 02/08/2025.

35. Fatumo, S. *et al.* A roadmap to increase diversity in genomic studies. Nat Med. 2022 Feb;28(2):243-250.

36. Office for National Statistics (ONS), released 29 November 2022, ONS website, statistical bulletin, Ethnic group, England and Wales: Census 2021.

37. Good, J. "Niger-Congo, with a special focus on Benue-Congo". In Vossen, Rainer; Gerrit J. Dimmendaal (eds.). *The Oxford Handbook of African Languages*. Oxford University Press. pp. 139–160, 2020.

38. Nunes, K. *et al*. Admixture's impact on Brazilian population evolution and health. Science. 388:3564, 2025.