Identificação de Variantes Genéticas Associadas ao Ceratocone por Meio do Sequenciamento de Exomas e Análises Funcionais em Família Brasileira

Palavras-chave: Ceratocone, Sequenciamento de Exomas, Biologia Molecular.

Lislie Vitória Rodrigues Guimarães, **LEGH / CBMEG – UNICAMP**
Thiago Henrique Paes, **LEGH / CBMEG – UNICAMP**
Me. Thiago Adalton Rosa Rodrigues, **LEGH / CBMEG – UNICAMP**
Ma. Júlia Nicoliello Pereira de Castro, **LEGH / CBMEG – UNICAMP**
Prof. Dr. Daniel de Almeida Borges, **Departamento de Oftalmologia - UNICAMP**
Profª. Drª. Mônica de Cassia Alves de Paula, **Departamento de Oftalmologia - UNICAMP**
Prof. Dr. Renato Nallin Montagnolli **(Coorientador), LMAM - UFSCar**
Profª. Drª. Mônica Barbosa de Melo **(Orientadora), LEGH / CBMEG – UNICAMP**

INTRODUÇÃO:

O ceratocone, do inglês *keratoconus (KC)*, é uma patologia oftalmológica degenerativa, caracterizada por uma ectasia progressiva de caráter não inflamatório, bilateral e multifatorial, a qual afeta patologicamente a curvatura corneana[1,2,3]. Tal protusão da córnea assume uma forma cônica, gerando um astigmatismo irregular[4]. Com desenvolvimento inicial assintomático, está relacionado a fatores sistêmicos, ambientais e genéticos, sendo manifesta majoritariamente por um padrão de herança mendeliana autossômica dominante[2,4]. Dessa maneira, clinicamente é frequentemente equivocado com casos de astigmatismo, sendo a única possibilidade de diagnóstico precoce mediante exames oftalmológicos de imagem[1,2]. Na ausência desses exames, o quadro clínico pode avançar com prejuízo significativo à acuidade visual do portador.

Trata-se de uma patologia que ocorre entre homens e mulheres de todas as etnias, apresentando uma incidência mundial de 1 portador a cada 2.000 nascidos por ano[2]; e de 1,4 a cada 2.000 no Brasil. Vale destacar que a coorte brasileira possui a heterogeneidade étnica mais diversa do mundo, sendo este o resultado direto de diversos eventos migratórios com a predominância de 3 principais etnias (ameríndia, europeia e africana), constituindo um importante modelo genotípico e fenotípico para estudos sobre o *KC*. Apesar dos avanços nas

investigações moleculares, populações latinas e miscigenadas ainda são raramente incluídas nesses estudos, especialmente quando comparadas às amplamente representadas coortes europeias e asiáticas. Nesse contexto, a elucidação do diverso perfil genético brasileiro representa uma oportunidade única, ampliando significativamente o horizonte para a identificação de variantes de nucleotídeo único *(Single Nucleotide Variants – SNVs)* potencialmente associadas ao desenvolvimento do *KC*. Tal diversidade genética exige a realização de estudos mais abrangentes, que contribuam para uma compreensão mais profunda da fisiopatologia e dos mecanismos moleculares envolvidos na doença.

Neste sentido, o objetivo deste estudo é investigar as bases genéticas envolvidas na etiologia do KC por meio dos sequenciamentos de exomas e análise de predição de risco em uma família brasileira **(Figura 1)**, visando a identificação e caracterização de genes candidatos envolvidos em vias biológicas relacionadas à patogênese e fisiopatologia dessa condição ocular. Desta forma, pretende-se desenvolver uma maior compreensão acerca das bases genéticas envolvidas na etiologia do ceratocone no Brasil.

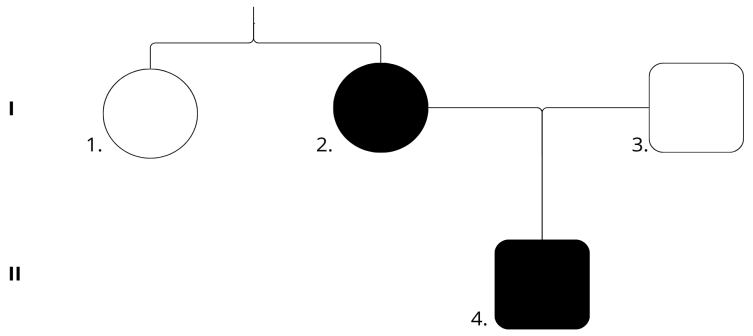

Figura 1 - Heredograma da família estudada.

MÉTODOS:

O sequenciamento de exoma foi realizado em uma família com padrão de herança autossômica dominante, composta por mãe e filho afetados, além de tia e pai não afetados. Foram extraídos os materiais genéticos a partir de sangue periférico dos pacientes, sendo subsequentemente preparada a biblioteca genômica e com respectivo sequenciamento de nova geração. Em seguida, foi realizada uma varredura das variantes já reportadas anteriormente em artigos científicos, tanto nacionais quanto internacionais, por meio de uma revisão sistemática na base de dados *PUBMed*, utilizando os descritores: *"keratoconus"; "genetics"; "review"; "Brazil"*.

Com isso, foi elaborado um *pipeline* para avaliação através de ferramentas de bioinformática. Neste sentido, os dados brutos do sequenciamento passaram por etapas de controle de qualidade utilizando o *FastQC* (v0.12.0); sendo as sequências (*reads*) com cobertura igual ou superior a Q20 mantidas para análises posteriores. O software *BWA* (versão 0.6) foi empregado no alinhamento das *reads* ao genoma de referência humano *GRCh38*. Frente as etapas pós-alinhamento, como marcação de duplicatas e recalibração da qualidade, foram executados comandos utilizando as ferramentas *Picard* e *GATK*, dentro do ambiente *Miniconda3* (versão 24.5.0). Vale destacar que as etapas posteriores de filtragem e controle de qualidade foram conduzidas no ambiente *R Studio* (v.4.1.2), utilizando os pacotes *readxl*, *dplyr* e *stringr*.

À priori, foi feita uma padronização do conjunto de dados, corrigindo valores nulos por inteiros e dados numéricos por caracteres, visando a execução de boas práticas de bioinformática[5,6,7]. Nesta etapa o *dataset* continha ~5.500 linhas e 102 colunas, porém, apenas algumas colunas foram utilizadas como parâmetros de filtragem das SNVs: *SYMBOL* (gene); *ZYG* (zigosidade) = heterozigose; *MAX_AF* (frequência máxima alélica de uma variante na base de dados gnomAD) < 0,05, *QUAL* (qualidade); DP (cobertura do sequenciamento) > 10; *Consequence* (categoria da mutação) = *missense*; exclusão de variantes consideradas toleradas/benignas por *SIFT* (preditor *in silico*) ; *PolyPhen* (preditor *in silico*); *Existing_variation* (variante previamente relatada), além da exclusão de variantes presentes nos não afetados.

Finalmente, a anotação funcional foi realizada com o auxílio do *Ensembl Variant Effect Predictor* (VEP), com o intuito de priorizar variantes com potencial relevância patogênica, baseado em seu impacto funcional prévio. Após a consolidação da lista de genes, foi realizada uma análise de enriquecimento funcional utilizando a ferramenta digital *Enrichr*, com o objetivo de identificar vias biológicas, processos celulares e funções moleculares potencialmente associadas ao conjunto de variantes identificado. Como forma de validação experimental dos achados será realizado o sequenciamento de Sanger. Será empregada a ferramenta *Primer3Plus* para a elaboração dos *primers* das sequências de interesse, além do *software GeneRunner* (v6.5.52x64) para controle de qualidade dos oligonucleotídeos.

RESULTADOS:

Esta análise molecular e de bioinformática filtrou variantes homozigóticas, identificando SNVs segregantes entre afetados, excluindo as presentes em não afetados.

Como resultado, foram encontradas 574 variantes em comum nos pacientes 2 e 4, sendo que 22 *SNVs* não estavam presentes nos não portadores. Entre as variantes *missense* identificadas, a rs201226914 (gene *PIEZO1)* e a rs185364182 (gene *ZSCAN10)* foram selecionadas como as mais relevantes para o contexto de desenvolvimento de *KC* nesta coorte da população brasileira, considerando inicialmente os aspectos biológicos, seguidos de predições *in silico* e diretrizes de classificação da *ACMG*[8].

Então, foram buscadas as vias biológicas em que estes genes agem a partir da ferramenta *EnrichR*. O banco de dados *Reactome Pathways* 2024 relacionou o gene *PIEZO1* com respostas celulares a estímulos em geral, principalmente mecânicos, além de respostas endoteliais ao estresse de cisalhamento, ativando a sinalização de integrinas, moléculas previamente descritas na literatura como participantes da patogênese do *KC*. Já o *ZSCAN10*, possui achados descrevendo mecanismos associados a ações de células-tronco embrionárias humanas. Ainda, o *ARCHS4 Kinases Coexp* relacionou ambos com o funcionamento de 10 quinases humanas, um grupo de enzimas responsáveis pela fosforilação, com habilidades de ativação ou inativação de moléculas-alvo, regulando diversos processos celulares. O *Kinase Perturbations from GEOdown* também associou ambos com o *knockout* do gene *MAP3K1*, fenômeno com potencial de afetar o desenvolvimento ocular, podendo resultar no fenótipo de "olhos abertos ao nascer". Por último, o *The Kinase Library* 2024 relacionou o *PIEZO1* com as espécies reativas de oxigênio (EROs), as quais, quando em desequilíbrio com as defesas antioxidantes, geram estresse oxidativo, já relacionado na literatura como fator significante para o desenvolvimento do *KC*.

DISCUSSÃO:

A literatura descreve que a resposta à estímulos mecânicos é crucial na patogênese desta afecção, visto que a má cicatrização na camada de *Bowman* resulta na ectasia corneana. Outro fator de risco é a fosforilação, desempenhando um papel crucial na regulação do ciclo e funções celulares que, quando alteradas, podem gerar severas disfunções oculares. Por fim, é importante realçar que a relação dos dois genes com o *knockout* do gene *MAP3K1*, associado ao desenvolvimento ocular, e do *PIEZO1* com as EROs, são relevantes no processo fisiopatológico do KC.

CONCLUSÕES:

Neste sentido, ao interpretar a análise funcional em um contexto biológico, as variantes nos genes *PIEZO1* e *ZSCAN10* podem contribuir como fator de risco para o desenvolvimento de *KC* nesta amostra da população brasileira. Vale enfatizar que análises adicionais em outros membros da família e coortes, além de validações experimentais, são necessárias para confirmação de sua patogenicidade.

BIBLIOGRAFIA

1. Harminder SD, Darren SJ, Ting MA, Dalia GS. Chapter 5 - Pathophysiology of Keratoconus. In: Izquierdo L, Henriques M, Mannis M. Keratoconus: Diagnosis and Management. *Elsevier,* 2022; 51-64
2. Rabinowitz YS. Keratoconus. *Surv Ophthalmol*. 1998;42(4):297-319.
3. Schor P. Ceratocone. *Arquivos Brasileiros de Oftalmologia.* 1998, 235-237.
4. Mas Tur V, MacGregor C, Jayaswal R, O'Brart D, Maycock N. A review of keratoconus: Diagnosis, pathophysiology, and genetics. *Surv Ophthalmol*. 2017;62(6):770-783.
5. Ren S, Yang K, Fan Q, et al. Bioinformatics analysis of key candidate genes and pathways in Chinese patients with keratoconus. *Exp Eye Res*. 2023;231:109488.
6. Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, Feng G. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. Cancer Inform. 2014 Sep 21;13(Suppl 2):67-82.
7. Hu D, Lin Z, Jiang J, Li P, Zhang Z, Yang C. Identification of Key Genes and Molecular Pathways in Keratoconus: Integrating Text Mining and Bioinformatics Analysis. *Biomed Res Int*. 2022;2022:4740141. Published 2022 Aug 23.
8. Richards, S., et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, vol. 17, no. 5, May 2015, pp. 405-24.