

2.02.05 - Genética / Genética Humana e Médica

ANÁLISE DE REDES DE PROTEÍNA EM GENES ASSOCIADOS À FENDA ORALAna Luiza Meneguci Moreira Franco^{1,2*}, Flávia Martinez de Carvalho^{1,2}

1. Programa de Pós-Graduação em Genética – UFRJ

2. Laboratório de Epidemiologia de Malformações Congênitas (LEMC) – Instituto Oswaldo Cruz | FIOCRUZ

Resumo

As fendas orais (FO) são anomalias congênitas complexas envolvendo fatores genéticos e ambientais. Trabalhos analisando redes de proteínas com genes relevantes na etiologia das FO já foram feitos, entretanto ainda há uma lacuna nesse conhecimento sobre a fenda labial com ou sem fenda palatina não síndrômica (NSFL/P). Nosso objetivo é compreender a interação entre os genes associados à NSFL/P descritos no OMIM através de análises de rede de proteína. A partir da construção de redes de interação proteína-proteína, identificamos as principais sub-redes e calculamos os “*hub genes*”. Entre os processos biológicos observados, destacamos o desenvolvimento do palato, embriogênese de membros e odontogênese, já a análise dos “*hub genes*” detectou os dez genes mais relevantes dentro da rede proposta, sendo eles: *SHH*, *BMP4*, *FGF8*, *FGFR1*, *FGFR2*, *CDH1*, *FGFR3*, *MSX1*, *PAX9* e *IRF6*. Assim, esperamos compreender melhor o papel desses genes e a interação entre eles no contexto multifatorial das FO.

Palavras-chave: Fenda labial com ou sem fenda palatina não síndrômica; anomalia congênita; genética humana

Apoio financeiro: CNPq, bolsista de mestrado ANMMM.

Introdução

As fendas orais (FO) são anomalias congênitas caracterizadas pela disrupção das estruturas faciais e geralmente são subdivididas em fenda labial com ou sem fenda palatina (FL/P) e fenda palatina (FP) e são classificadas como não síndrômica quando não estiverem associadas a outra anomalia¹. São classificadas como um fenótipo multifatorial², envolvendo fatores genéticos e ambientais³, entretanto a participação dos genes na etiologia das FO continua sendo, em grande parte, desconhecida. A prevalência das FO varia de acordo com a etnia sendo mais frequente em ameríndios e asiáticos e tem a prevalência global de aproximadamente 10 a cada 10.000 nascidos vivos⁴. Além disso, representam um grande problema de saúde pública, uma vez que, requerem um tratamento multidisciplinar envolvendo diversas especialidades como odontologia, nutrição, cirurgia, fonoaudiologia, entre outros⁵. Também já foi observado que podem afetar a interação social e a relação materno-fetal nos primeiros anos⁶. Recentemente um trabalho de Dong *et al.* 2022, utilizou uma metodologia de análise de redes de interação entre proteínas que é muito empregada em estudos com câncer, para relacionar genes já associados com FO, disponíveis nas bases de dados, a fim de delimitar os caminhos biológicos mais promissores. Entretanto, as palavras chaves utilizadas podem não ter incluído na sua totalidade os genes relacionados as FO, principalmente quando se trata das FO não síndrômicas. O objetivo desse trabalho é aprimorar essa metodologia, focando nas fendas labiais com ou sem fenda palatina não síndrômica (NSFL/P) a fim de lançar luzes sobre os caminhos biológicos mais promissores para futuras análises genéticas em pacientes com FO.

Metodologia**Coleta de dados**

Inicialmente foram feitas buscas na base de dados do OMIM⁹, as palavras chaves utilizadas foram: “non syndromic cleft lip”, “nonsyndromic cleft lip and palate” e “nosyndromic oral cleft”. Foram removidas as duplicatas e todos os genes encontrados adicionados à uma lista.

Análise de enriquecimento funcional

A ferramenta DAVID¹⁰ foi utilizada para realizar a análise de enriquecimento baseada no Gene Ontology¹¹, focando nos processos biológicos (PB). FDR (False discovery rate) menor que 0,05 foi considerado estatisticamente significativo. Finalmente foram selecionadas os dez termos relacionados a PB do Gene Ontology com menor valor de p.

Análise de rede de interação entre proteínas

A base de dados STRING¹² foi utilizada para a construção de redes de interação proteína-proteína. Foram selecionadas as interações com score, que são indicativos de confiança, maior que 0,4, nesse caso consideramos um nível de confiança médio com o objetivo de maximizar o número de interações. Em seguida, a aplicação MCODE¹³ para o software Cytoscape¹⁴ foi utilizada para a formação de clusters dentro da rede obtida. Foram selecionados os clusters com score de detecção de complexo molecular maior que 3.

Finalmente os *hub genes* foram identificados, a partir da rede gerada pela base de dados STRING¹², e ranqueados através da aplicação cytoHubba¹⁵, o método de ranqueamento escolhido foi o MCC (“Maximal

Clique Centrality”), que segundo o artigo de Chin *et al.* 2014 tem uma melhor desempenho na precisão de predição de proteínas essenciais.

Resultados e Discussão

Na etapa de busca, foram encontrados 67 genes associados com o fenótipo de interesse. A análise de enriquecimento apontou os dez processos biológicos (Gráfico 1) mais frequentes nesse conjunto de dados.

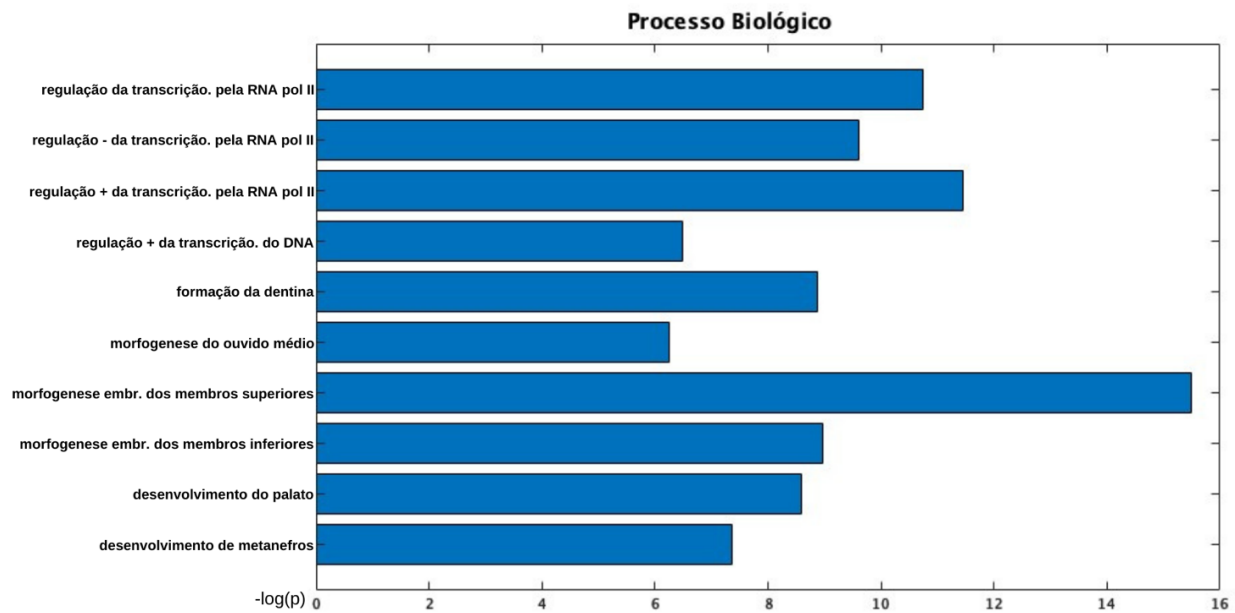


Gráfico 1: No eixo y: primeiros dez processos biológicos enriquecidos em nosso conjunto de dados, eixo x: menos log do valor de p. Imagem gerada no MATLAB¹⁶

Podemos ressaltar a presença de muitos genes envolvidos no desenvolvimento do palato (11,6% de todos os genes selecionados), o que era esperado devido ao fenótipo de interesse. Genes envolvidos na morfologia de membros superiores e inferiores também apresentaram valores de p relevantes na análise e é importante notar que alterações em dermatoglífos já foram relatadas em pacientes com fenda oral¹⁷. Outro processo biológico que merece destaque é a formação da dentina dos dentes visto que muitos familiares e portadores de FO apresentam anomalias dentárias, o que muitas vezes pode ser classificado como subfenótipo de FO¹⁷. Grande parte desses processos ocorre no mesmo período embriológico da formação da face e em regiões fisicamente próximas, e é esperado que alterações em genes chaves possam afetar diversas estruturas diferentes.

A construção de rede de interação proteína-proteína (Figura 1) permitiu que relacionássemos 57 genes dos 67 originais, sugeridos na listagem inicial de genes.

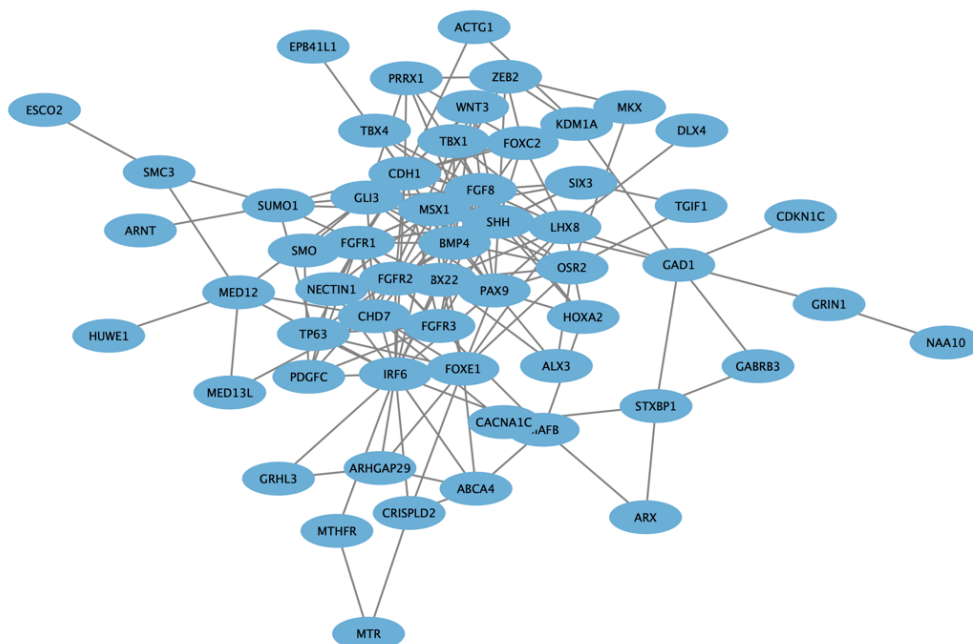


Figura 1: Rede de interação proteína-proteína. Figura produzida com o software Cytoscape¹⁴

A análise de clusters permitiu a identificação de 3 sub-redes (Figura 2).

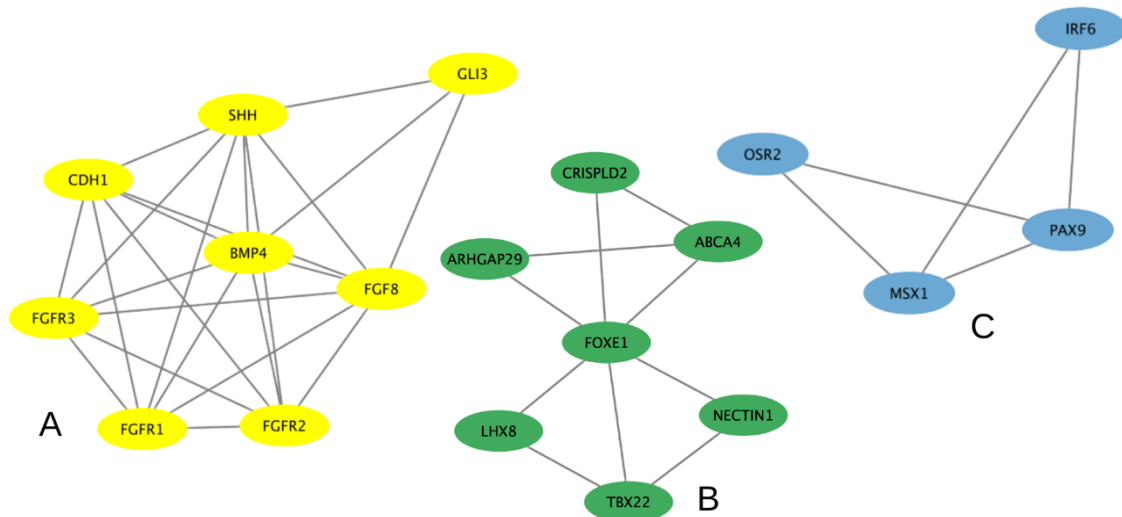


Figura 2: Sub-redes formadas a partir da análise de cluster. Os scores para as redes foram: A – 6,857, B – 3,333 e C – 3,333. Figura produzida com o software Cytoscape¹⁴

Na rede A podemos ver a presença de dois genes da via de sinalização de Sonic hedgehog (*GLI3*, *SHH*) conjuntamente com fator de crescimento de fibroblasto e alguns receptores (*FGF8*, *FGFR1*, *FGFR2* e *FGFR3*), durante o processo de formação da face há vários momentos onde a migração e crescimento celular são fatores chave, a presença desses genes evidencia esse aspecto. Os genes *BMP4*, *SHH*, *FGFR2* e *FGFR* e *CDH1* também estavam presentes na principal rede apresentada no artigo de Dong et al. 2022, mostrando a importância desses genes na etiologia das FO.

Na rede C encontramos 4 genes (*IRF6*, *OSR2*, *MSX1* e *PAX9*) que podem atuar como fatores de transcrição ou ativadores transcricionais, esses resultados convergem com os dez processos biológicos mais importantes listados no gráfico 1 e evidencia a importância da regulação gênica para a correta embriogênese.

Como resultado da análise dos “hub genes” temos os dez primeiros genes de acordo com o ranking calculado pelo aplicativo cytoHubba, sendo eles: *SHH*, *BMP4*, *FGF8*, *FGFR1*, *FGFR2*, *CDH1*, *FGFR3*, *MSX1*, *PAX9*, *IRF6* (Figura 3).

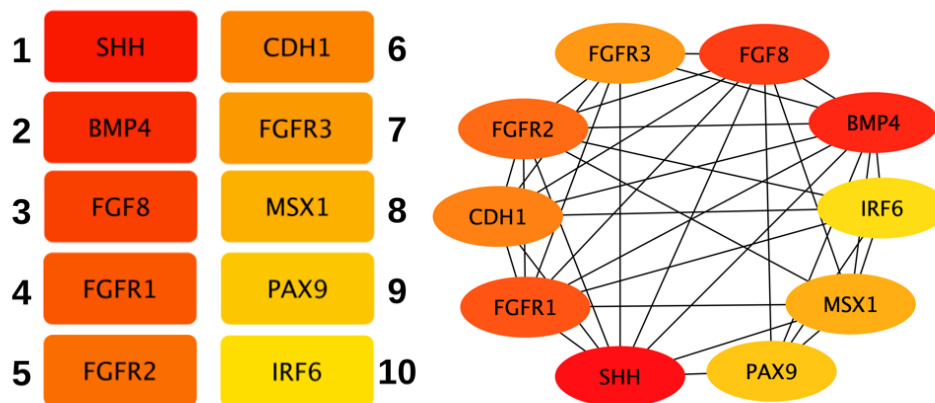


Figura 3: 10 primeiros “hub genes” identificado e a rede de interação proteína-proteína entre eles. Escala de cor representa o ranking dos genes. Figura produzida com o software Cytoscape¹⁴

Comparando com os dez “hub genes” identificada por Dong et al. (2022), o gene *SHH* permaneceu sendo o primeiro da lista, outros genes que também estiveram presentes foram: *FGFR1*, *BMP4* e *CDH1*. Um problema identificado por Dong et al. (2022) foi a ausência do gene *IRF6* nos “hub genes”, sendo esse um gene já associado com NSFL/P em diversas populações¹⁸. Na nossa análise é possível observar sua presença e interação com outros genes centrais.

Podemos observar a presença de genes de adesão como *CDH1* no sexto lugar, também na sub-rede B onde encontramos o gene *NECTIN1*, durante o processo de formação da face é necessária a correta fusão dos processos maxilares com os processos nasais e das prateleiras palatinas para que haja o desenvolvimento do lábio e palato respectivamente, evidenciando a importância das proteínas de adesão celular.

Corroborando com a sub-rede A podemos observar a presença um gene da via de sinalização de Sonic hedgehog (*SHH*) na primeira posição do ranking, essa via é uma das principais reguladoras do processo de desenvolvimento¹⁹. De forma geral podemos observar uma correspondência entre as sub-redes formadas, principalmente A e C e os 10 “hub genes”.

Conclusões

Com essas análises foi possível identificar os caminhos biológicos mais relevantes como por exemplo odontogênese e embriogênese de membros nos genes já associados com NSFL/P. Um ponto negativo desse trabalho é a utilização da base de dados OMIM que tem como foco genes ligados a doenças mendelianas, não sendo esse o caso das fendas orais, a incorporação do termo “nonsyndromic” tem como objetivo diminuir essa deficiência, entretanto não incorporamos genes candidatos ainda em fase de estudos à nossa lista final. Vale ressaltar que, após o advento do sequenciamento mássico, cada vez mais genes relacionados à FO sindrômica têm sido associados com NSFL/P.

Também foi possível relacionar os genes mais centrais o que auxilia o entendimento de sua importância dentro de um contexto multifatorial e pode guiar futuros estudos de associação em busca do melhor entendimento da etiologia genética das FO.

Referências bibliográficas

1. McInnes RR, Michaud J (2004) Taking it to the max: The genetic and developmental mechanisms coordinating midfacial morphogenesis and dysmorphology. *Clin Genet* 65:163-176. <https://doi.org/10.1111/j.0009-9163.2004.00225>.
2. Murthy J, Bhaskar L (2009). Current concepts in genetics of nonsyndromic clefts. *Int J PSurg* 42: 68-81.
3. Lorente C, Cordier S, Goujard J, Aymé S, Biachi F, Calzolari E, De Waller HEK, Knill-Jones R (2000) Tobacco and Alcohol Use During Pregnancy and Risk of Oral Clefts. *Am J of Public Health* 90(3):415-419.
4. Watkins SE, Meyer RE, Aylsworth AS (2014) Classification, Epidemiology, and Genetics of Orofacial Clefts. *ClinPlasticSurg* 41:149–163. <http://dx.doi.org/10.1016/j.cps.2013.12.003>
5. Strauss RP (1999) The Organization and Delivery of Craniofacial Health Services: The State of Art. *Cleft Palate-Cran J* 36(3):189-195. https://doi.org/10.1597/1545-1569_1999_036_0189_toadoc_2.3.co_2.
6. Scheller K, Urich J, Watzke S, Scheller C, Haase R (2020) Impact of cleft lip closure on the perception of the child by the mother and the social environment - A retrospective, self-reported, standardized pretest. *J CranioMaxilloSurg* 48:969-976. <https://doi.org/10.1016/j.jcms.2020.07.013>
7. McInnes RR, Michaud J (2004). Taking it to the max: The genetic and developmental mechanisms coordinating midfacial morphogenesis and dysmorphology. *Clin Genet* 65:163-176. <https://doi.org/10.1111/j.0009-9163.2004.00225>.
8. Dong X, Meng T *et al.* (2022). Uncovering the Pathogenesis of Orofacial Clefts Using Bioinformatics Analysis. *The Journal of Craniofacial Surgery*. 33:1-5
9. Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), (2022). World Wide Web URL: <https://omim.org/>
10. Jiao X, Sherman BT, Stephens R, *et al.* DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* 2012; 28:1805–1806
11. Ashburner *et al.* Gene ontology: tool for the unification of biology. *Nat Genet*. May 2000;25(1):25-9.
12. Szklarczyk D, Franceschini A, Wyder S *et al.* (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 43(D1): D447-D453. <https://doi.org/10.1093/nar/gku1003>
13. Bader GD, Hogue CWV (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4,2. <https://doi.org/10.1186/1471-2105-4-2>.
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks *Genome Research* 2003 Nov; 13(11):2498-504
15. Chin C, Chen S *et al.* (2014) cytoHubba: identifying hub objects and subnetworks from complex interactome. *BMC Systems Biology* 8:3-7.
16. MATLAB. (2022). *version 7.10.0 (R2022a)*. Natick, Massachusetts: The MathWorks Inc.
17. Kane, Alex A.; Wenger, Sharon L.; Losee, Joseph; Deleyiannis, Frederick; Ma, Lian; De Salamanca, Javier E.; Weinberg, Seth M.; Neiswanger, Katherine; Martin, Rick A.; Mooney, Mark P. (2006). The Pittsburgh Oral-Facial Cleft Study: Expanding the Cleft Phenotype. *Background and Justification*, 43(1), 7–20.
18. Wattanawong K, Rattanasiri S, McEvoy M, Attia J, Thakkinstian A (2016) Association between IRF6 and 8q24 polymorphisms and nonsyndromic cleft lip with or without cleft palate: Systematic review and meta-analysis. *Birth Defects Research Part A: Clinical and Molecular Teratology* 106(9): 773-788.
19. Choudhry Z, Rikani AA, Choudhry AM *et al.* (2014) Sonic hedgehog signalling pathway: a complex network. *Annals of Neurosciences* 21(1):28-31