



CIÊNCIA E TECNOLOGIA:  
IMPLICAÇÕES NO ENSINO, PESQUISA E EXTENSÃO

# FEPEG

F Ó R U M  
ENSINO • PESQUISA • EXTENSÃO • GESTÃO

REALIZAÇÃO:



APOIO:



ISSN: 1806-549X

## ANÁLISE GÊNICA DO CARCINOMA BASOCELULAR, CARCINOMA ESPINOCELULAR E CERATOSE ACTÍNICA PELO PARTICLE SWARM CLUSTERING MODIFICADO

**Autores:** DÉBORA GONÇALVES PEREIRA GUIMARÃES, EMISAEEL STÊNIO BATISTA GOMES, SABRINA FERREIRA DE JESUS, MARCELA GONÇALVES SOUZA, ELOÁ MANGABEIRA SANTOS, ANGELINY TAMIARANA LIMA TABOSA, ANDRÉ LUIZ SENA GUIMARÃES

### Introdução

O câncer de pele não melanoma constitui o grupo de neoplasias mais frequente em humanos, responsável por cerca de 30% de todos os tumores malignos registrados no Brasil. (INCA, 2017) Destes, o carcinoma basocelular (CBC) e o carcinoma espinocelular (CEC) são os mais comuns, ambos associados a fatores de risco como exposição à luz ultravioleta, pele clara e imunodepressão. (INCA, 2017) O CBC possui maior incidência, porém é menos agressivo. Já o CEC apresenta maiores índices de metástase e comumente surge a partir de uma lesão precursora, como a ceratose actínica (AK, do inglês *actinic keratosis*). (CHETTY et al., 2015) AK é uma lesão consequente da exposição solar cumulativa que pode evoluir tanto para a remissão espontânea quanto para a transformação maligna em CEC. (BERMAN; COCKERELL, 2013)

Neste contexto, a bioinformática surge como um importante mecanismo para identificar e comparar genes líderes envolvidos em sua patogênese. O reconhecimento de genes relacionados ao comportamento distinto entre CBC e CEC, além da diferença entre os genes expressos pelo CEC e AK abre portas para novas abordagens que podem contribuir para a prevenção e tratamento desses tipos de câncer. (POSWAR et al., 2015)

Este trabalho propõe uma abordagem para o agrupamento de dados e identificação de genes líderes baseada no algoritmo *Particle Swarm Clustering* (PSC), utilizado para caracterizar a expressão gênica de CBC, CEC e AK.

### Material e métodos

#### A. Bioinformática e análise de biologia de sistemas

Os principais genes envolvidos no CBC, CEC e AK foram identificados por meio das bases de dados: NCBI Gene, OMIM e Genecards. Utilizou-se a nomenclatura da Organização Genoma Humano (*Human Genome Organization*, HUGO). As palavras-chave para CEC foram: “*cutaneous squamous cell carcinoma*”, “*squamous cell carcinoma of the skin*”, “*skin squamous cell carcinoma*”, “*cutaneous squamous cell cancer*”, “*skin squamous cell cancer*” e “*squamous cell cancer of the skin*”; para CBC: “*basal cell carcinoma*”, “*basocellular carcinoma*” e “*basal cell epithelioma*”; para AK, “*actinic keratosis*”.

Dessa forma, gerou-se uma lista de genes potencialmente relacionados ao CBC, CEC e AK. Essa lista foi então expandida por meio do programa STRING (versão 9.05), considerando apenas interações baseadas em observações experimentais descritas em domínio público com alto grau de confiança (acima de 0,9, intervalo 0-0,99). Com esse processo foram identificados novos genes diretamente ligados ao CBC, CEC e AK. Evidências científicas da PubMed foram utilizados para evitar falsos positivos.

Para cada gene obtido foi calculado o *Weighted Number of Links* (WNL), além do valor total de interações (*total Interaction Score*, TIS), que representa todas as interações gênicas da base de dados do STRING. Esses valores são importantes para identificar os genes líderes, isto é, os que possuem maior grau de associações. Genes sem interações são chamados de genes órfãos. A análise topológica foi realizada por meio das ferramentas Cytoscape e FANMOD, enquanto a análise ontológica foi feita com o programa BinGO.

#### B. Métodos de agrupamento

Baseados no WNL e TIS, os genes foram agrupados pelo método K-Means Clustering. As categorias com maior WNL e menor TIS foram consideradas genes líderes. As demais categorias não foram levadas em conta na análise. Os testes ANOVA e Tukey-Kramer foram realizados para certificar os resultados. Significância estatística foi definida em  $P < 0,001$ . Utilizou-se também uma versão modificada do algoritmo PSC para combinar os dados gerados por bases distintas. Para permitir que o PSC encontre automaticamente o número de grupos, duas mudanças foram feitas no algoritmo: operador de crescimento e as partículas de remoção. Na execução do algoritmo, se o nível de concentração da partícula analisada for maior do que zero e sua similaridade com o padrão de entrada mais próximo for maior que um limiar, a partícula será clonada. Já o operador de remoção elimina as partículas em níveis de concentração iguais a zero.

### Resultados e discussão

Foi encontrado um total de 382 genes relacionados ao CBC, 178 relacionados ao CEC e 56 à AK. A análise ontológica e de regressão demonstra diferentes mecanismos associados a cada tipo de lesão.

Comparando-se os métodos de agrupamento para CBC, o mecanismo K-means demonstrou cinco *clustering groups* e indicou os genes líderes CTNNA1 e TP53. Por outro lado o PSC detectou dez *clustering groups* e os principais genes relacionados foram: CTNNA1, EGFR, JUN, MYC, SFRP2, PRKCA, BIRC5, MCM2, POMC. Para o CEC, K-means indicou três *clustering groups* e os genes líderes TP53, EGFR, CTNNA1, MAPK1, MAPK3 e MYC. Com PSC, foram obtidos seis *clustering groups* e os seguintes genes líderes: SP1, E2F1, STAT3, CDKN1A, ERBB2, RB1, CDK4, CCND1, CAV1, EGF, CTNNA1, IGF1R, CASP3, PARP1, MSH2, NOTCH1, CKS1B, EIF4E. Por fim, para AK, o TP53 foi sugerido como gene líder tanto pelo K-means quanto pelo PSC. Considerando o CBC e CEC, o K-means apresentou genes com maior TIS do que o PSC.

Este estudo foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG).



CIÊNCIA E TECNOLOGIA:  
IMPLICAÇÕES NO ENSINO, PESQUISA E EXTENSÃO

# FEPEG

F Ó R U M  
ENSINO • PESQUISA • EXTENSÃO • GESTÃO

REALIZAÇÃO:



APOIO:



ISSN: 1806-549X

A identificação de genes líderes se provou útil para determinar alvos terapêuticos e marcadores prognósticos, sendo inicialmente descrita com o K-means. (ORLANDO et al., 2013) Estudos subsequentes apontaram o agrupamento como um ponto chave a ser melhorado. (GUIMARAES et al., 2016)

O valor de TIS traduz interações extensas, enquanto o WNL se relaciona a interações específicas. Visto que o PSC apresentou menores valores de TIS e sugeriu diferentes genes líderes para CBC e CEC, o PSC pode ser útil na definição de alvos terapêuticos, enquanto o K-means pode ser direcionado para melhor compreensão da patogênese das neoplasias. Além disso, K-means pode ser mais indicado para identificar mecanismos biológicos similares em lesões menos frequentes. (POSWAR et al., 2015)

Este estudo demonstra, portanto, a importância de ferramentas de agrupamento na bioinformática para identificar alvos terapêuticos. Além disso, os genes relacionados a cada uma das lesões estudadas neste estudo são consistentes com a literatura, visto que mutações no gene TP53 são encontradas em grande parte dos casos de neoplasias malignas da pele e conferem resistência de células pré-neoplásicas à apoptose. (MADAN et al., 2010) O CTNNB1 foi reconhecido com gene líder tanto no CBC quanto no CEC. De acordo com o PSC, CEC se relaciona ao gene CTNNA1, ambos associados à adesão celular. Há evidências de que o *knockout* do CTNNA1 promove CEC *in situ*. (VASIOUKHIN et al., 2001) Por outro lado, observou-se o acúmulo nuclear de proteínas CTNNB1 no CBC. (SALDANHA et al., 2004)

Dessa forma, é clara a importância da bioinformática na identificação de genes com alto potencial de biomarcadores e alvos terapêuticos.

## Conclusões

Conclui-se que este estudo sugere o PSC como uma ferramenta interessante para agrupamento de genes na análise bioinformática de doenças prevalentes. Já o K-means deve ser utilizado em doenças menos frequentes. Além disso, os achados sugerem que o TP53 pode ter um papel central na AK, o CTNNB1 parece estar relacionado ao CBC, enquanto o CTNNA1 se associa ao CEC. Portanto, estes achados podem incentivar futuros estudos sobre o potencial desses genes como biomarcadores e alvos farmacológicos.

## Agradecimentos

Este trabalho foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e pela Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG).

## Referências bibliográficas

- BERMAN, B.; COCKERELL, C. J. Pathobiology of actinic keratosis: ultraviolet-dependent keratinocyte proliferation. **Journal of the American Academy of Dermatology**, v. 68, n. 1, p. S10-S19, 2013.
- CHETTY, P.; CHOI, F.; MITCHELL, T. Primary care review of actinic keratosis and its therapeutic options: a global perspective. **Dermatology and therapy**, v. 5, n. 1, p. 19-35, 2015.
- GUIMARAES, T. A. et al. Evaluation of the antineoplastic activity of gallic acid in oral squamous cell carcinoma under hypoxic conditions. **Anti-cancer drugs**, v. 27, n. 5, p. 407-416, 2016.
- INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA (INCA); COORDENAÇÃO DE PREVENÇÃO E VIGILÂNCIA. **Estimativa 2018: incidência de câncer no Brasil**. Rio de Janeiro: INCA, 2017.
- MADAN, V.; LEAR, J. T.; SZEIMIES, R. M. Non-melanoma skin cancer. **The Lancet**, v. 375, n. 9715, p. 673-685, 2010.
- ORLANDO, B.; BRAGAZZI, N.; NICOLINI, C. Bioinformatics and systems biology analysis of genes network involved in OLP (Oral Lichen Planus) pathogenesis. **Archives of oral biology**, v. 58, n. 6, p. 664-673, 2013.
- POSWAR, F. DE O. et al. Bioinformatics, interaction network analysis, and neural networks to characterize gene expression of radicular cyst and periapical granuloma. **Journal of endodontics**, v. 41, n. 6, p. 877-883, 2015.
- SALDANHA, G. et al. Nuclear  $\beta$ -catenin in basal cell carcinoma correlates with increased proliferation. **British Journal of Dermatology**, v. 151, n. 1, p. 157-164, 2004.
- VASIOUKHIN, V. et al. Hyperproliferation and defects in epithelial polarity upon conditional ablation of  $\beta$ -catenin in skin. **Cell**, v. 104, n. 4, p. 605-617, 2001.



CIÊNCIA E TECNOLOGIA:  
IMPLICAÇÕES NO ENSINO, PESQUISA E EXTENSÃO

# FEPEG

F Ó R U M  
ENSINO • PESQUISA • EXTENSÃO • GESTÃO

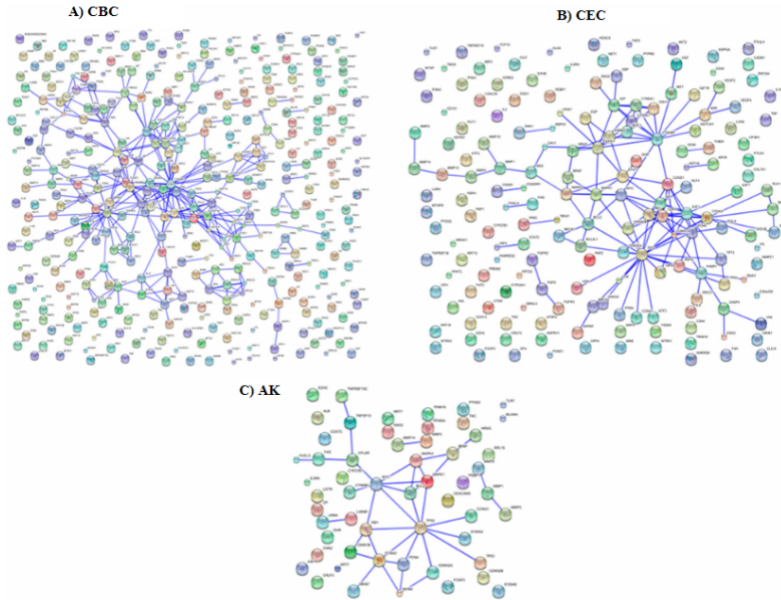
REALIZAÇÃO:



APOIO:



ISSN: 1806-549X



**Figura 1.** Mapa da interação entre genes possivelmente relacionados às lesões (nível de confiança > 0,90). Rede gerada no software STRING (versão 9.05). A) CBC. B) CEC. C) AK.