



| | |
|-------------------|--|
| Evento | Salão UFRGS 2018: SIC - XXX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS |
| Ano | 2018 |
| Local | Campus do Vale - UFRGS |
| Título | Transcriptogramer aplicado à oncologia infantil: Sarcoma de Ewing |
| Autor | MAURICIO GOMES DE QUEIROZ |
| Orientador | RITA MARIA CUNHA DE ALMEIDA |

Transcriptogramer aplicado à oncologia infantil: Sarcoma de Ewing

Maurício Gomes de Queiroz, Rita M.C. de Almeida - UFRGS

O sarcoma de Ewing (SE) é um tumor agressivo que acomete ossos e tecidos moles e afeta principalmente crianças e adultos jovens. Na sua maioria são tumores indiferenciados geneticamente caracterizados pela presença da translocação cromossômica envolvendo *EWSR1* e *FLII*. Apesar dos avanços no entendimento das bases biológicas do sarcoma de Ewing, ainda não está definido qual sua célula de origem. Uma maior compreensão da identidade e origem celular é necessária para a identificação de marcadores e alvos terapêuticos. Uma fonte de informação sobre o estado metabólico das células são transcriptomas, que podem ser obtidos por micro-arranjos ou por RNA-Seq. Essas técnicas fornecem uma medida da expressão gênica em escala genômica. Em virtude da escala desses dados, técnicas de bioestatísticas mais avançadas tornam-se imprescindíveis, além disso, dados de transcriptomas apresentam grande ruído estocástico, o que reduz muito o poder dos testes estatísticos.

Amostras de Sarcoma de Ewing disponibilizadas na literatura, no repositório Gene Expression Omnibus (GEO) sob o código GSE21511 foram analisadas [1]. Esse conjunto de dados compara amostras de biópsias de SE com amostras de tecidos normais e de linhagens transformadas obtidas de células tronco de Crista Neural, com intuito de investigar as células de origem dos tumores de Ewing.

De acordo com análise de von Levetzow^[1], as amostras primárias de SE são mais parecidas com uma linhagem de célula tronco embrionária humana derivada da crista neural (hNCSC), indicando assim sua origem. No entanto, devido à razão sinal-ruído nos dados durante a análise PCA conduzida, esta afirmação representa apenas 34,9% das variações entre as amostras.

Uma técnica estatística para a análise de dados de expressão gênica foi desenvolvida na UFRGS [2] e nomeada Transcriptogramer, que apresenta um aumento importante na razão sinal-ruído, melhorando portanto o poder do teste.

A pesquisa aqui desenvolvida visa aprofundar esta análise filtrando o ruído com a nova técnica, oferecendo como alternativa um olhar mais representativo. Primeiramente, foram replicados os resultados do artigo referenciado, e verificamos que, ao considerar apenas as três primeiras componentes principais no PCA tendo por entrada os dados de expressão, os autores desprezaram uma contribuição importante das outras componentes, modificando o resultado obtido. Em um segundo momento, utilizamos como entrada o resultado do transcriptograma de raio 30 e levando em conta 100% das variações, observamos que as amostras de SE são mais parecidas com linhagens mesenquimais, e não com as linhagens hNCSC, como sugere o artigo. Portanto, ao considerar apenas 34,9% das variações, deixa-se de lado muitas informações importantes que são evidenciadas após tratamento dos dados com o Transcriptogramer.

A utilização da ferramenta Transcriptogramer diminui o ruído dos dados, de tal forma que ao considerarmos as 3 componentes principais corresponde a 70% da variação enquanto a metodologia tradicional corresponde por 44%.

[1] von Levetzow C, Jiang X, Gwyne Y, von Levetzow G et al., ‘Modeling initiation of Ewing sarcoma in human neural crest cells’. *PLoS One* (2011)

[2] da Silva, S.R.M., Perrone, G.C., Dinis, J.M. and de Almeida, R.M.C., ‘Transcriptograms: Reproducibility enhancement and differential expression of non predefined functional gene sets in human genome’. *BMC Genomics*, 15, 1181 (2014).