



Evento	Salão UFRGS 2020: SIC - XXXII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2020
Local	Virtual
Título	Predição de candidatos a biomarcadores de câncer a partir da análise de redes de co-expressão gênica
Autor	LAURA GALANT SPEGGIORIN
Orientador	MARIANA RECAMONDE MENDOZA GUERREIRO

Predição de candidatos a biomarcadores de câncer a partir da análise de redes de co-expressão gênica

Laura Galant Speggorin, Profa. Dra. Mariana Recamonde Mendoza

Instituto de Informática, Universidade Federal do Rio Grande do Sul

A busca por genes associados a doenças complexas como câncer se torna cada vez mais importante para ajudar a melhor compreender e tratar tais doenças. Um dos objetivos é diminuir a quantidade de genes candidatos a biomarcadores, sendo métodos de aprendizado de máquina e teoria de redes complexas comumente adotados para esta tarefa. O objetivo deste estudo é utilizar medidas de centralidade para comparar a estrutura de redes de co-expressão geradas a partir de tecido tumoral, adjacente e saudável, e o potencial de uso de medidas de centralidades para identificação de biomarcadores. Utilizamos dados públicos de quatro tipos de câncer (mama, pulmão, próstata e tireoide), contendo amostras de tecido tumoral, adjacente ao tumor e saudável, sendo os dois últimos utilizados como controles. Os dados brutos foram pré-processados e filtrados para manter apenas genes que apresentam variância entre grupos. A correlação de Pearson intra-grupo foi calculada para pares de genes, e redes de co-expressão foram geradas utilizando um threshold $|\rho| > 0.6$. Sete métricas de centralidade (grau, grau ponderado, *betweenness*, *closeness*, *eigenvector*, *PageRank*, *Clustering Coefficient*) foram extraídas para os genes em cada rede, e sumarizadas através da média, mediana, desvio padrão. A significância estatística foi avaliada com um teste t de Student ($\alpha=95\%$). Nossos resultados apontam diferenças estruturais claras entre as três redes, o que reforça a suspeita de que a alteração da expressão dos genes entre os tecidos é significativa o bastante para ser usada para prever candidatos a biomarcadores. Levanta também a possibilidade de estudar mais a fundo as diferenças marcantes encontradas entre as redes de tecido adjacente e saudável. Foram usados também genes já associados a cada câncer para estudar o comportamento destes entre redes, e descobriu-se que estes apresentam diferença mais significativa em certas métricas, principalmente *eigenvector*, na rede de tecido adjacente.