

Ciência e Visualização de Dados em Saúde

MO413-MC936 - 2024/1

Relatório do trabalho T2 - Integrando Redes

Equipe Get Out Zika

Nome	RA (e-mail)	Especialização
Bruno Brito Pereira da Silva	b213947 (@dac.unicamp.br)	Biologia
Márcio Silva Cruz	m290232 (@g.unicamp.br)	Computação
Rafael Simionato	r066505 (@g.unicamp.br)	Computação

Metodologia

Utilizamos a funcionalidade do *Multiple proteins* do STRING para encontrar as interações entre as proteínas disponíveis no arquivo *most-differentially-expressed-genes.csv*. Com isso descarregamos o arquivo de arestas do grafo gerado com informações a respeito das interações físicas documentadas:

network type: physical subnetwork
meaning of network edges: confidence
active interaction sources: experimental e database

Realizamos um procedimento similar buscando pela funcionalidade *Pathway* a rede de interações disponível na base de dados KEGG para o *Wnt Signaling Pathway*. Neste caso descarregamos os arquivos de nós e arestas do grafo gerado contendo informações a respeito das interações físicas documentadas por *confidence*, *experimental* e *database*, e por fim, combinamos todos os dados no Neo4j.

No Neo4j descrevemos um modelo lógico de interação Proteína-Proteína para realizar a importação das redes. A fim de evitar sobrescrever-se os valores das propriedades *experimental* e *database* nas arestas criadas na importação do primeiro arquivo, acrescentamos um sufixo *_wnt* às propriedades citadas na importação do segundo arquivo. Em seguida obtivemos uma rede com 2.129 nós *Protein* e 2.306 relacionamentos *Interacts* conforme a figura 1.

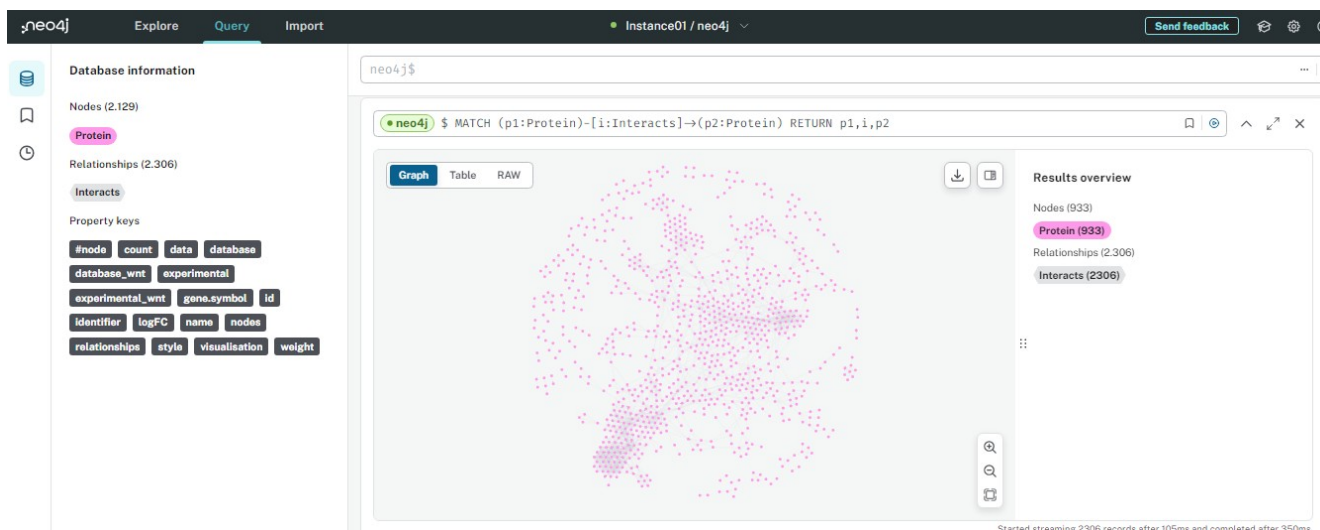


Figura 1 – Rede de interação Proteína-Proteína no Neo4j

Uma vez juntados os grafos, exportamos a rede consolidada para um arquivo CSV a ser utilizado no Cytoscape. **Importante** mencionar que neste arquivo constam apenas as proteínas que interagem entre si (933 proteínas - figura 1) na rede agregada.

Em seguida, no Cytoscape utilizamos a funcionalidade para importação de redes a partir de arquivos. Assim definimos os nós *source* e *target* atribuindo a eles as propriedades *logFC* e *wnt*.

A fim de se evidenciar as proteínas (nós) participantes do *Wnt Signaling Pathway*, unificamos as propriedades *experimental_wnt* e *database_wnt* em um tipo booleano. Utilizamos essa nova propriedade *wnt* para diferenciar os genes do *pathway* por forma (triangular) e tamanho. Evidenciou-se também os nós com valor de expressão diferencial (propriedade *logFC*) através de uma escala de cores variando de azul (baixa expressão) a vermelho (alta expressão) conforme figura 2A.

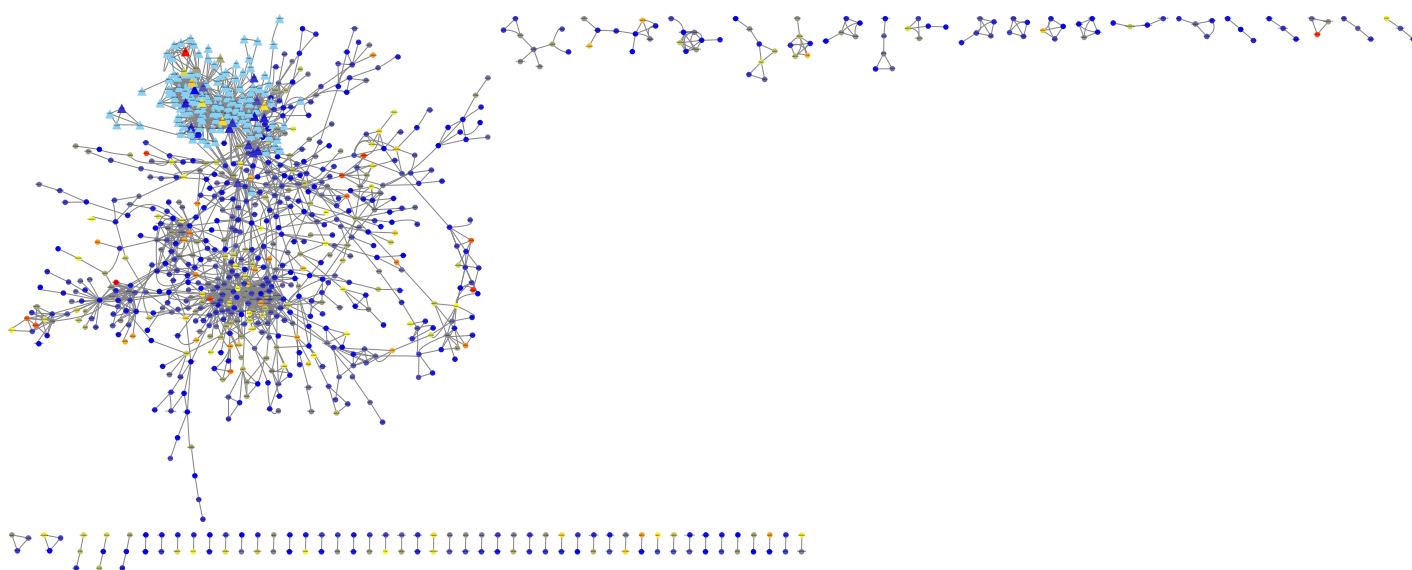


Figura 2A – Rede de interação Proteína-Proteína no Cytoscape

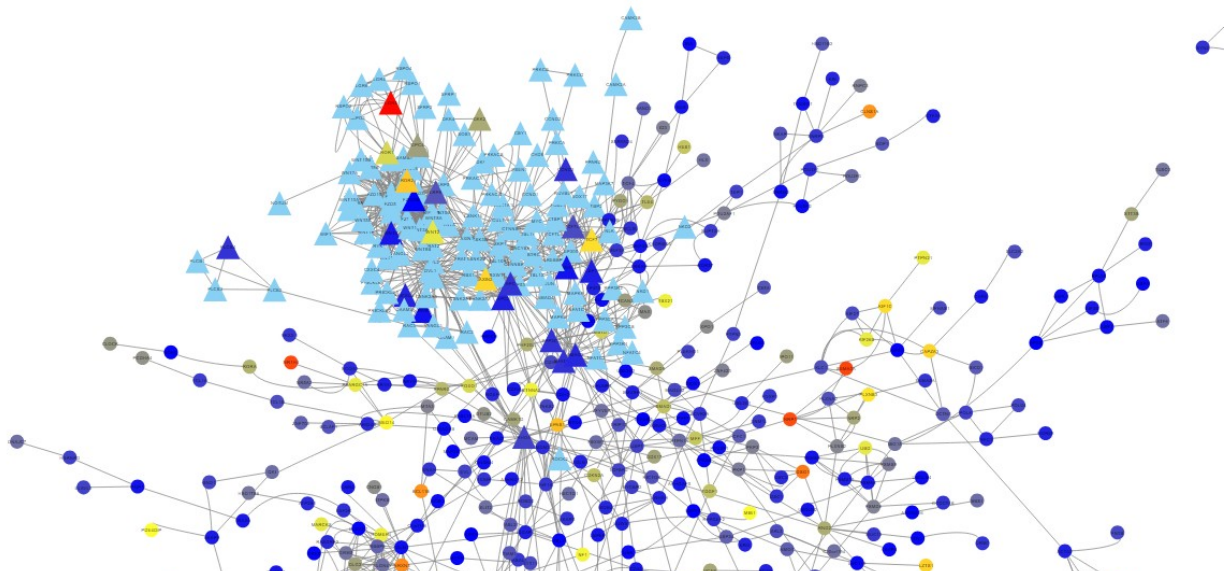


Figura 2B – Figura 2Aem zoom