

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS
CAMPUS DE BOTUCATU

Carlos Alberto Oliveira de Biagi Júnior

**PREDIÇÃO DE NOVOS ALVOS QUIMIOTERÁPICOS PARA O TRATAMENTO DE
CÂNCER DE MAMA BASEADO EM CENTRALIDADES DE REDES**

Trabalho de Conclusão de Curso apresentado
ao Instituto de Biociências, Campus de Botu-
catu, UNESP.

Botucatu, Novembro de 2014.

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
INSTITUTO DE BIOCÊNCIAS
CAMPUS DE BOTUCATU

Carlos Alberto Oliveira de Biagi Júnior

**PREDIÇÃO DE NOVOS ALVOS QUIMIOTERÁPICOS PARA O TRATAMENTO DE
CÂNCER DE MAMA BASEADO EM CENTRALIDADES DE REDES**

Trabalho de Conclusão de Curso apresentado
ao Instituto de Biociências, Campus de Botu-
catu, UNESP.

Orientador: Prof. Dr. José Luiz Rybarczyk Filho

Botucatu, Novembro de 2014.

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Biagi Júnior, Carlos Alberto Oliveira de.

Predição de novos alvos quimioterápicos para o tratamento de câncer de mama baseado em centralidades de redes / Carlos Alberto Oliveira de Biagi Júnior. - Botucatu, 2014

Trabalho de conclusão de curso (bacharelado - Física Médica) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: José Luiz Rybarczyk Filho
Capes: 20903006

1. Mamas - Câncer. 2. Câncer - Quimioterapia. 3. Análise de microarranjo. 4. Expressão gênica.

Palavras-chave: Câncer; Centralidades; Microarranjo.

Agradecimentos

- A Deus pelo dom da vida, proteção e cuidado;
- Ao Professor Dr. José Luiz Rybarczyk Filho pela excelente orientação, incentivos, e amizade. Meu muito obrigado por tudo;
- À Dra. Agnes Takeda pela ajuda na confecção desse trabalho;
- Aos amigos do laboratório, em especial à Tahila e Pedro;
- Aos meus amigos de sala Gerson, Mayara e Raphael pela amizade e companheirismo;
- Ao meu pai Carlos, minha mãe Selma, minha irmã Natália e minha namorada Karen pelo apoio irrestrito e incentivo incondicional em todos os aspectos;
- A todos os professores, amigos, servidores e familiares que de alguma forma contribuíram para que esse trabalho pudesse ser feito.
- À PROPE pela bolsa de Iniciação Científica oriunda do programa primeiros projetos (Processo 0028/015/13-PROPe/CDC), UNESP pela bolsa PIBIC/Reitoria, e ao apoio financeiro do CNPq (473789/2013-2 e 458810/2013-4);

Resumo

Um dos tratamentos mais utilizados no câncer de mama é a quimioterapia, que reduz a massa tumoral em aproximadamente em 80%. Ela também é recomendada para pacientes pós operados para a prevenção de metástases. Entretanto, as populações celulares podem desenvolver mutações quimioresistentes. Atualmente a combinação de drogas é muito eficiente no tratamento, mas a descoberta de novas drogas fará com que novas combinações sejam mais eficazes qpara casos específicos e possibilitará resultados mais rápidos no tratamento de câncer de mama.

O presente trabalho tem como objetivo encontrar novos alvos protéicos que possam ser utilizados para o desenvolvimento de novos quimioterápicos para o tratamento de câncer de mama.

Então, utilizou-se da biologia de sistemas para encontrar novos alvos protéicos. Foi construída uma rede baseada nos alvos protéicos já conhecidos, foram utilizadas análises de centralidades para evidenciar as proteínas mais importantes na rede. Correlacionamos os valores de centralidade com informação de expressão gênica para verificar se as proteínas mais importantes do ponto de vista topológico estão diferencialmente sub ou super expressas. A análise a convergiu para um bom resultado, foi possível evidenciar algumas proteínas que podem ser utilizadas como novos alvos quimioterápicos,pois alguns estudos atuais estão focados nestas proteínas.

Lista de Figuras

1.1	Etapas do ciclo celular. 2013 Nature Education	p. 1
1.2	Principais tipos de alterações genéticas que levam a ativação de oncogenes. O proto-oncogene (top) corresponde a uma região regulatória (RS) seguida de uma codificante (gene). 1. Indica mutação pontual ; 2. Indica translocação, em que há uma outra região regulatória; 3. Amplificação corresponde a presença de cópias múltiplas do gene, resultando na expressão em excesso (Adaptado de (KUFE; POLLOCK, 2003))	p. 2
1.3	Estágios da Mama mostrados esquematicamente. As Figuras foram retiradas do link http://www.cancerresearchuk.org	p. 5
4.1	Rede para o câncer de mama utilizando os dados de Acencio e colaboradores (ACENCIO et al., 2013) com 2297 nodos e 52459 interações.	p. 13
4.2	Exemplo de alguns gráficos comparativos entre as centralidades.	p. 15

Lista de Tabelas

4.1	Nomes dos quimioterápicos e seus respectivos alvos protéicos.	p. 12
4.2	Valores das centralidades utilizadas encontradas a partir do NetworkX.	p. 14
4.3	Resultado das expressões gênicas para os sete alvos conhecidos.	p. 17
4.4	Resultados das expressões gênicas dos possíveis novos alvos.	p. 17

Sumário

1	Introdução	p. 1
1.1	O ciclo celular e o câncer	p. 1
1.2	O Câncer de Mama	p. 2
1.2.1	Anatomia da Mama	p. 3
1.2.2	Tipos de Câncer de Mama	p. 3
1.2.3	Estágios	p. 4
1.2.4	Tratamento	p. 5
2	Objetivo	p. 8
3	Material e Métodos	p. 9
3.1	Redes	p. 9
3.2	DrugBank	p. 10
3.3	Expressão Gênica e <i>Gene Expression Omnibus</i> (GEO)	p. 10
4	Resultados e Discussão	p. 12
4.1	Alvos Proteicos	p. 12
4.2	Rede para o Câncer de Mama	p. 12
4.3	Expressão Gênica	p. 15
5	Conclusões	p. 18
	Referências Bibliográficas	p. 19

1 *Introdução*

1.1 O ciclo celular e o câncer

Em biologia, ciclo celular é o termo utilizado para um conjunto de processos necessários para que uma célula viva realize divisão celular. Esse processo é importante para a manutenção da vida e, por isso é necessário que o processo seja controlado. Em eucariontes, o ciclo é dividido em quatro etapas: G₁, S, G₂ e M (Figura 1.1). De modo simplificado, na fase G₁ ocorre a expressão de genes e síntese de proteínas responsáveis pela duplicação do DNA. Esse evento ocorre na fase S, que é denominada fase de síntese. Em seguida, na fase G₂ ocorre síntese de proteínas responsáveis pelo processo de divisão celular. Por fim há a fase M (mitose), na qual o material genético duplicado é separado e enviado para cada uma das células filhas formadas a partir da célula original (ALBERTS et al., 2002). Durante esse processo é essencial que

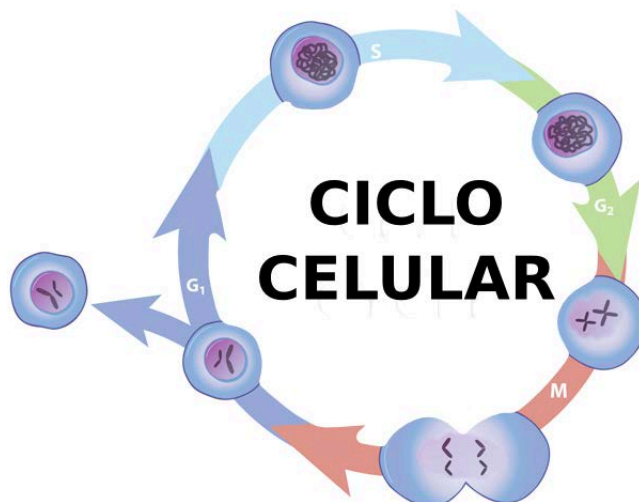


Figura 1.1: Etapas do ciclo celular. 2013 Nature Education

ocorra a verificação de que os eventos tenham ocorrido sem erros. A origem das células cancerosas está associada a anomalias na regulação do ciclo celular e à perda de controle da mitose. Quando os genes controladores do ciclo celular sofrem alterações (mutações) há o surgimento

de um câncer. O ciclo celular é regulado por duas classes de genes, os proto-oncogenes (são responsáveis pela produção de proteínas que atuam na estimulação do ciclo celular, quando ativos estimulam a ocorrência do ciclo celular) e os genes supressores (responsáveis pela produção de proteínas que atuam inibindo o ciclo celular, e quando ativos inibem a ocorrência do ciclo celular). Quando há mutações nos proto-oncogenes eles se transformam em oncogenes (genes causadores do câncer) (Figura 1.2). As mutações que afetam os genes supressores de tumor per-

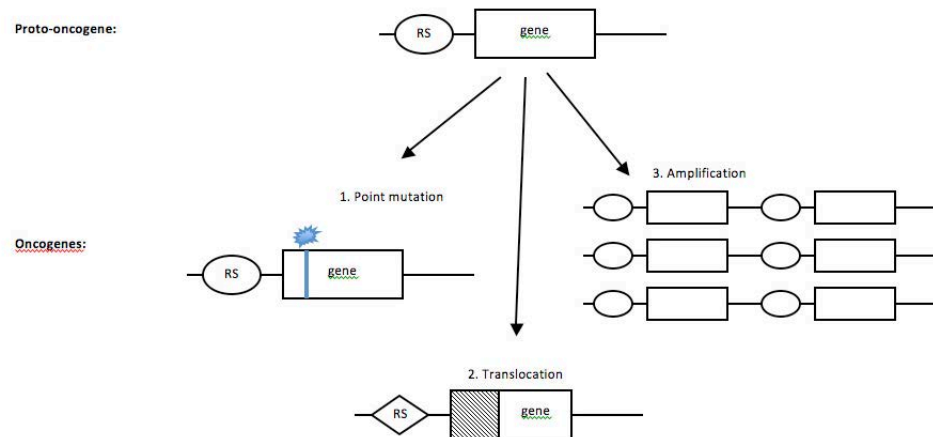


Figura 1.2: Principais tipos de alterações genéticas que levam a ativação de oncogenes. O proto-oncogene (top) corresponde a uma região regulatória (RS) seguida de uma codificante (gene). 1. Indica mutação pontual ; 2. Indica translocação, em que há uma outra região regulatória; 3. Amplificação corresponde a presença de cópias múltiplas do gene, resultando na expressão em excesso (Adaptado de (KUFE; POLLOCK, 2003))

turbam o sistema inibidor e o ciclo celular fica desregulado, promovendo a ocorrência desordenada de divisões celulares e o surgimento de células cancerosas (ALMEIDA et al., 2005). Essas células são resistente aos controles que mantêm um tecido normal e, como resultado dividem-se mais rápido que as células parentais e dependem menos da sinalização de outras células, inclusive esquivando-se da apoptose (morte celular programada), processo que normalmente seria aplicado a uma célula defeituosa (O'CONNOR; ADAMS, 2010). Essas células podem levar a formação de tumores, massas de células que crescem e dividem-se incontrolavelmente. Tumores malignos apresentam células diferentes das células do tecido de origem e podem adquirir habilidade de espalhar e migrar para outras partes do corpo, processo denominado metástase (CHOW, 2010).

1.2 O Câncer de Mama

O câncer de mama é o câncer mais comum entre as mulheres no mundo, tanto em países desenvolvidos ou em países em desenvolvimento. Estimou-se 1.67 milhões de novos casos até

o ano de 2012; as maiores taxas foram observadas na Europa Ocidental, com 96 novos casos a cada 100 mil mulheres, e as menores taxas na África Ocidental e na Ásia Oriental, com 27 novos casos a cada 100 mil mulheres. Na Inglaterra, por exemplo, 49.900 mulheres são diagnosticadas com câncer de mama por ano; 1 em 8 mulheres vão desenvolver câncer de mama durante suas vidas (LOPES et al., 1996). No Brasil as estatísticas dizem que haverá 57.120 novos casos de câncer de mama no ano de 2014; a incidência do mesmo é maior nas mulheres das regiões Sudeste, Sul, Centro Oeste e Nordeste (ESTIMATIVA, 2010).

1.2.1 Anatomia da Mama

As mamas são glândulas com a função principal de produção de leite materno para o recém-nascido. Elas são constituídas por várias estruturas (Figura 1.3): ácino (menor parte da glândula mamária, responsável pela produção de leite no período de lactação), lóbulo mamário (é constituído pelo conjunto de acinos), lobo mamário (conjunto de lóbulos mamário que se liga a papila por um ducto), ductos mamários (são 15 a 20 canais que conduzem o leite até a papila), tecido glandular (conjunto de lobos e ductos), papila (protuberância elástica onde desembocam os ductos mamários), aréola (estrutura central da mama onde se projeta a papila) e tecido adiposo (tecido gorduroso revestidor da mama, cuja característica advém do tipo físico, alimentação e idade da mulher)

1.2.2 Tipos de Câncer de Mama

Existem vários tipos de câncer de mama, mas alguns deles são bastante raros. Em alguns casos, um único tumor na mama pode ser uma combinação destes tipos (SMITH et al., 2003):

- **Carcinoma Ductal *In Situ*:** Também conhecido como carcinoma intraductal, é considerado não invasivo ou câncer de mama pré-invasivo. A diferença entre o carcinoma ductal *in situ* e carcinoma invasivo é que as células não se espalharam através dos ductos para o tecido mamário adjacente. O carcinoma ductal *in situ* é considerado um pré-câncer, pois em alguns casos pode se tornar um câncer invasivo.
- **Carcinoma Lobular *In Situ*:** As células se parecem com as células cancerosas que crescem nos lobos das glândulas produtoras de leite, mas não se desenvolvem através da parede dos lobos.
- **Carcinoma Ductal Invasivo:** Este é o tipo mais comum de câncer de mama. O carcinoma ductal invasivo (ou infiltrante) se inicia em um duto de leite, rompe a parede desse

duto e cresce no tecido adiposo da mama. Podendo se espalhar (metástase) para outras partes do corpo através do sistema linfático e da circulação sanguínea. Cerca de 80% dos cânceres de mama invasivos correspondem ao carcinoma ductal infiltrante.

- **Carcinoma Lobular Invasivo:** Inicia nas glândulas produtoras de leite (lobos) e espalha-se para outras partes do corpo. Cerca de 10% dos cânceres de mama invasivos correspondem ao carcinoma lobular invasivo. Este tipo pode ser mais difícil de ser detectado na mamografia do que o carcinoma ductal invasivo.

1.2.3 Estágios

O câncer de mama pode ser dividido em alguns estágios (BEAHR; HENSON, 1992) (Figura 1.3). Essa divisão é realizada para informar o grau de desenvolvimento do câncer. Essa avaliação é feita tomando por base a dimensão do tumor, a avaliação da extensão aos linfonodos e a presença ou não de metástases para outras partes do corpo.

- **Estágio 0:** É o chamado carcinoma *in situ* que não se infiltrou pelos dutos ou lóbulos, sendo um câncer não invasivo. O estágio zero significa que as células do câncer estão presentes ao longo da estrutura de um lóbulo ou um ducto, mas não se espalharam para o tecido gorduroso vizinho, isto é, as células cancerosas que ainda não invadiram os tecidos circundantes;
- **Estágio IA:** O tumor invasivo é pequeno (menos de 2.0 cm de diâmetro) e não se espalhou pelos linfonodos, isto é, o tumor que permaneceu no local no qual se originou sem disseminação para os linfonodos ou locais distantes;
- **Estágio IB:** Não há tumor na mama, porém existe a presença pequenos grupos de células cancerígenas - maiores que 0.2 mm e menores que 2.0 mm - estes grupos estão presentes nos gânglios linfáticos; ou existem tumor na mama, mas não é maior que 2.0 cm e existem grupos de células cancerígenas - maiores que 0.2 mm e menores que 2.0 mm - nos nódulos linfáticos;
- **Estágio IIA:** O tumor tem menos que 2.0 cm e infiltrou linfonodos axilares; ou o tumor tem entre 2.0 e 5.0 cm, mas não atinge linfonodos axilares;
- **Estágio IIB:** Não há evidência de tumor na mama, mas existe câncer nos linfonodos axilares. O tumor tem de 2.0 a 5.0 cm e atinge linfonodos axilares; ou O tumor é maior que 5.0 cm, mas não atinge linfonodos axilares.

- **Estágio IIIA:** O tumor é menor que 5.0 cm e se espalhou pelos linfonodos axilares que estão aderidos uns aos outros ou a outras estruturas vizinhas; ou O tumor é maior que 5.0 cm, atinge linfonodos axilares os quais podem ou não estar aderidos uns aos outros ou a outras estruturas vizinhas;
- **Estágio IIIB:** O tumor se infiltra na parede torácica ou causa inchaço ou ulceração da mama ou é diagnosticado como câncer de mama inflamatório. Pode ou não ter se espalhado para os linfonodos axilares, mas não atinge outros órgãos do corpo;
- **Estágio IIIC:** Tumor com qualquer tamanho que não se espalhou para partes distantes, mas que atinge linfonodos acima e abaixo da clavícula ou para linfonodos dentro da mama ou abaixo do braço;
- **Estágio IV:** É o câncer metastático. O tumor com qualquer tamanho que espalhou-se para outros locais do corpo, tais como: ossos, pulmões, fígado, rins, intestinos, cérebro.

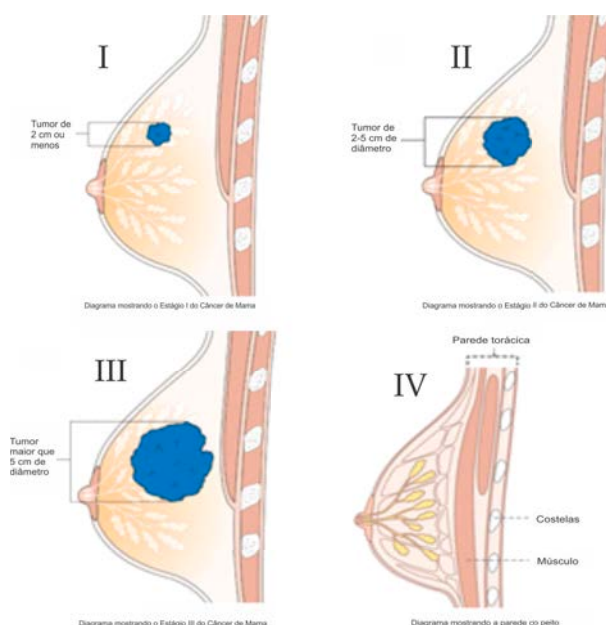


Figura 1.3: Estágios da Mama mostrados esquematicamente. As Figuras foram retiradas do link <http://www.cancerresearchuk.org>

1.2.4 Tratamento

Quando o paciente é diagnosticado com o câncer de mama há um grande leque de possibilidades de tratamentos, dependendo o tipo e o estágio que se encontra o tumor. Temos por exemplo a cirurgia, que não necessariamente é o primeiro tratamento, mas é muito indicado com finalidade de cura. Há também a indicação da radioterapia dependendo do estágio da doença,

sendo recomendada como complemento curativo cirúrgico ou como tratamento paliativo para diminuir os sintomas relacionados à doença. Por fim há a quimioterapia que pode ser indicado como medida pré operatória para diminuir o tamanho do tumor (cerca de 80% de diminuição da massa tumoral) e permitir a cirurgia, como complemento ao tratamento cirúrgico ou ainda como tratamento paliativo(SILVA, 2014).

A cirurgia é uma das modalidades mais antigas de tratamento para o câncer. Busca-se a retirada do tumor do local atingido pela doença. Quanto mais precocemente for feito o diagnóstico do câncer, maiores as chances da eficácia da cirurgia. É comum que, após a cirurgia, a paciente continue o tratamento com outros métodos para controlar a doença como, por exemplo, a quimioterapia.

Quimioterápicos são medicamentos que destroem células tumorais por interferir no processo de divisão celular. Células tumorais morrem devido à toxicidade dos quimioterápicos durante o seu processo de multiplicação. Os quimioterápicos se espalham por todo o organismo pela corrente sanguínea, e não matam apenas células tumorais que estejam se dividindo, mas também células normais do organismo. Devido a este efeito sobre as células normais, que acabam morrendo também, é que os quimioterápicos têm diversos efeitos colaterais como queda de cabelo, aftas (mucosite), diarreia, anemia, baixa da imunidade (baixa dos glóbulos brancos), risco de sangramento (baixa de plaquetas) entre outros(MEDICINE, 2009).

A quimioterapia pode ser combinada com a aplicação de um ou mais quimioterápicos. O uso de drogas isoladas (monoquimioterapia) é mais utilizado em pacientes com doença metastática e incurável, exceto em situações onde é urgente a resposta rápida ao tratamento, assim evitando interferir muito na qualidade de vida do paciente, sendo atualmente de uso muito restrito. A poliquimioterapia é de eficácia comprovada e tem como objetivos atingir populações celulares em diferentes fases do ciclo celular, utilizar a ação sinérgica das drogas, diminuir o desenvolvimento de resistência às drogas e promover maior resposta por dose administrada (ROSADO; HENRIQUES; BONATTO, 2011).

É muito provável que o paciente diagnosticado uma primeira vez com o câncer de mama desenvolva novamente o câncer após algum tempo do primeiro tratamento. Este retorno do câncer é denominada recorrente, e pode retornar na mama (chamada de recorrência local) na parede torácica ou em qualquer parte do corpo como ossos ou outros linfonodos (chamados de metástases à distância). O Hospital Johns Hopkins monitora de perto seus pacientes por até 2 anos após o primeiro diagnóstico e o tratamento completo para que eles possam acompanhar se o paciente vai desenvolver o câncer recorrente (MEDICINE, 2009).

Um detalhe importante é que o quimioterápico que for utilizado no primeiro tratamento do

primeiro diagnóstico não será mais eficaz no câncer recorrente, porque as células cancerígenas vão criar uma quimiorresistência a esses quimioterápicos. Então é necessário que haja outras alternativas de quimioterápicos para que o tratamento seja eficaz e resultante como o primeiro (KUO, 2007).

O desenvolvimento do câncer está relacionado com várias interações químicas entre pequenas moléculas e células neoplásicas, o que torna o estudo difícil já que os estudos das características dessas moléculas são realizados em ensaios isolados com a atividade de uma única droga. Usando as ferramentas de biologia de sistemas poderíamos estudar alvos protéicos para quimioterápicos, e sugerir novos alvos para tratamentos com o intuito de evitar que as células cancerígenas desenvolvam quimiorresistência.

2 *Objetivo*

O objetivo deste trabalho é utilizar a biologia de sistemas para eleger novos alvos protéicos para o câncer de mama. Assim, será necessário construir uma rede a partir de alvos de quimioterápicos utilizados no tratamento de câncer de mama e correlacionar as medidas de centralidades de redes com a expressão gênica.

3 *Material e Métodos*

3.1 Redes

Uma rede pode ser definida como um grafo, que é representado por uma coleção de nodos (vértices) ligados por arestas (links). Podemos encontrar redes nos mais variados contextos, como por exemplo: redes sociais, biológicas, tecnológicas, entre outras. As redes sociais são aquelas na qual as pessoas estão conectadas por relações profissionais, familiares, amizade, etc. As redes biológicas são redes que representam sistemas naturais, como por exemplo cadeias alimentares, redes de neurônios, caminhos metabólicos, etc. Redes de informação reúnem pessoas e organizações para o intercâmbio de informações, como por exemplo: redes de conceitos, preferências, grafo da Web, etc. E por fim temos as redes tecnológicas que são feitas pelo homem para distribuição de produtos ou recursos, por exemplo: redes de ligações telefônicas, internet, transporte, esgoto, água, luz, etc (NEWMAN, 2003, 2000).

Neste estudo iremos utilizar a rede de Acencio e colaboradores (ACENCIO et al., 2013) que é constituída por associações metabólicas, regulação transcricional e interação proteína proteína. Com base nestes dados iremos construir uma rede que represente as associações de alvos de quimioterápicos do cancer de mama. A rede será analisada através do pacote NetworkX (HAGBERG; SWART; CHULT, 2008), a fim de verificar quais são os pontos mais críticos à sua estabilidade. As centralidades utilizadas para esse tipo de análise são as seguintes (SCARDONI; PETTERLINI; LAUDANNA, 2009):

- **Grau de Intermediação:** Medida muito similar ao estresse, informando o quanto que um nodo específico está em relação aos demais nodos da rede. Auxilia na definição de módulos do sistema;
- **Grau de Proximidade:** É o nível de proximidade de um nodo relacionada aos outros nodos da rede;
- **Grau de Centralidade:** É definido como o número de ligações incidentes em um nodo. Esta medida também é conhecida como conectividade;

- **Radialidade:** É a probabilidade de uma proteína ser funcionalmente relevante para várias outras proteínas, mas com a possibilidade de ser irrelevante para algumas outras proteínas;
- **Estresse:** Conta o número de caminhos mais curtos que passam por um nó
- **Coeficiente de Clusterização:** determinado matematicamente pela equação 3.1:

$$C(i) = \frac{2n_i}{k_i(k_i - 1)} \quad (3.1)$$

onde n_i é o número de conexões dos vizinhos de um dado nodo i entre eles mesmos, e k_i , sua conectividade;

- **Coeficiente Topológico:** é uma medida relativa à extensão na qual um nodo compartilha vizinhos com outros nodos;
- **Comunicabilidade Central:** mede o quão fácil é para se enviar uma mensagem a partir de um nodo até outro [9].

Essas centralidades fornecem informações relevantes sobre o funcionamento de uma rede e nos permitirão inferir os nodos críticos à sua estabilidade e quais nodos possuem maior possibilidade de desestruturar o sistema como um todo.

3.2 DrugBank

O DrugBank (<http://www.drugbank.ca/>) (WISHART et al., 2006) é um banco de dados disponibilizado pela Universidade de Alberta (Canadá), sendo o único recurso de bioinformática e quimioinformática que combina informações detalhadas de medicamentos químicos e farmacológicos com informações dos alvos (sequência, estrutura química, via metabólica, etc.) de cada droga. O DrugBank contém atualmente o registro de 7740 drogas, dentre elas: drogas experimentais, nutracêuticos, pequenas moléculas e drogas biotecnológicas aprovadas pelo órgão governamental norte-americano de Administração de Alimentos e Medicamentos (*Food and Drug Administration* - FDA).

3.3 Expressão Gênica e *Gene Expression Omnibus* (GEO)

Os microarranjos são empregados na detecção e quantificação dos níveis de expressão de ácidos nucleicos (expressão gênica) provenientes de amostras biológicas, os quais são pos-

tas para hibridizar com o DNA fixado no arranjo. Estas amostras são o objeto de estudo do pesquisador, sejam elas preparadas a partir de células normais, cancerosas, tecido neuronal, etc.

Os dados de expressão gênica serão resgatados do repositório público GEO (BARRETT; EDGAR, 2006) (<http://www.ncbi.nlm.nih.gov/projects/geo/>), o qual possui uma enorme quantidade de experimentos com chips de cDNA sobre genes expressos em diferentes situações celulares. Atualmente existem cerca de 703566 experimentos para *Homo sapiens*.

Para a análise de expressão gênica será utilizado o programa Rstudio e o pacote LIMMA que permitir calcular os níveis de expressão gênica.

4 Resultados e Discussão

4.1 Alvos Proteicos

Primeiramente, foi necessário relacionar os quimioterápicos e os seus respectivos alvos protéicos no tratamento de câncer de mama. Esta busca foi realizada no site da *American Cancer Society*. Os quimioterápicos mais utilizados do tratamento de câncer de mama são: Fluorouracil, Doxorubicina, Epirubicina, Tamoxifeno, Trastuzumabe, Paclitaxel, Ciclofosfamida, Docetaxel, Raloxifeno e Lapatinib (WISHART et al., 2008). Com base na informação dos nomes dos quimioterápicos para o tratamento do câncer de mama, foi feita uma análise dentro do banco de dados *DrugBank* para relacionar drogas e alvo protéicos. A Tabela 4.1 apresenta os resultados desta análise.

Tabela 4.1: Nomes dos quimioterápicos e seus respectivos alvos protéicos.

Quimioterápico	Alvo Protéico
Fluorouracil	TYMS
Doxorubicina	TOP2A
Epirubicina	CHD1
Tamoxifeno	ESR1, ESR2
Transtuzumabe	ERBB2, EGFR, C1R, C1QA, C1QB, C1S, FCGR1A, FCGR2A, FCGR2B, FCGR2C, FCGR3B, FCGR3A
Paclitaxel	TUBB1, BCL
Ciclofosfamida	
Docetaxel	TUBB1, BCL2
Raloxifeno	ESR1, ESR2
Lapatinib	EGFR, ERBB2

4.2 Rede para o Câncer de Mama

Com base nos alvos protéicos da Tabela 4.1 foi construída uma rede que representasse a interação dos quimioterápicos e proteínas, para isso foi utilizada a rede de Acencio e colaboradores (ACENCIO et al., 2013). Buscou-se todos as proteínas que cada alvo protéico interagiu e

depois foi realizado um merge de todas as redes de cada alvo. As 21 proteínas alvo nos retornou uma rede com 2297 proteínas e 52459 interações. O interessante é que 90 % dessas interações são oriundas de interação proteína-proteína e o restante é oriundo de associações metabólicas e regulação transcricional. A figura 4.1 apresenta a rede do câncer de mama.

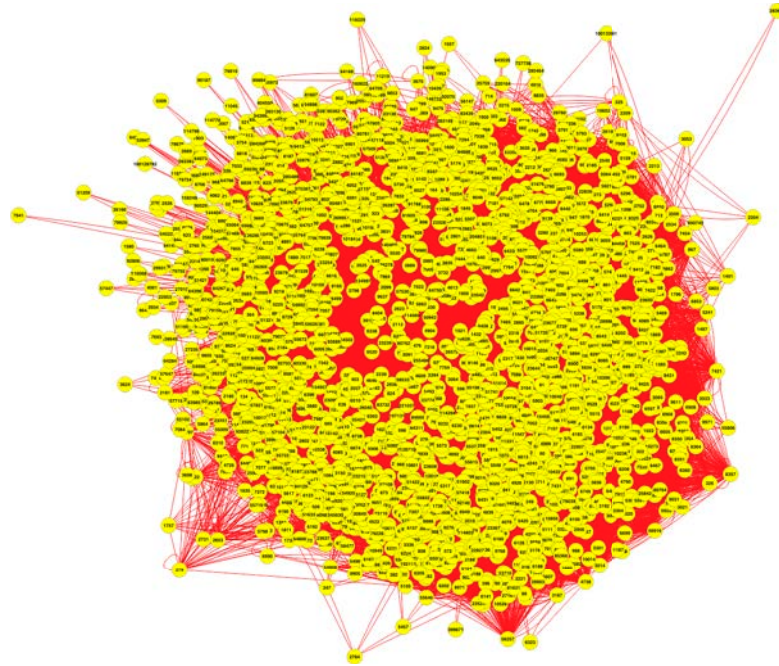


Figura 4.1: Rede para o câncer de mama utilizando os dados de Acencio e colaboradores (ACENCIO et al., 2013) com 2297 nodos e 52459 interações.

A partir do NetworkX (HAGBERG; SWART; CHULT, 2008) calculamos todas as centralidades da rede do câncer de mama. A análise dos resultados foram realizadas através de interpretação de gráficos (Figura 4.2). Cruzamos todas as centralidades e observamos os valores de centralidades que eram superiores aos valores da média de cada centralidade e comparamos com os valores de cada centralidade dos alvos protéicos já conhecidos no tratamento de câncer. A tabela 4.2 apresenta os valores das proteínas que apresentaram maiores valores de centralidades. O próximo passo é comparar a expressão gênica de câncer de mama com as proteínas que possuem altos valores de centralidade. A Tabela 4.2 mostra os resultados dessas centralidades que foram obtidas.

Tabela 4.2: Valores das centralidades utilizadas encontradas a partir do NetworkX.

proteína	Grau de Intermediação	Grau de Proximidade	Coefficiente de Clusterização	Grau de Centralidade	Radialidade	Estresse	Coefficiente Topológico	Comunicabilidade Central
UBC	0.20	0.56	0.09	172	0.94	1126108	0.06	5.29×10^{60}
ESR1	0.24	0.55	0.07	201	0.95	657750	0.05	4.12×10^{50}
EGFR	0.15	0.55	0.09	210	0.83	577916	0.05	2.62×10^{49}
TYMS	0.03	0.33	0.27	40	0.80	273782	0.16	2.62×10^{48}
SYK	0.02	0.42	0.28	66	0.78	204574	0.12	2.65×10^{48}
TK1	0.02	0.39	0.38	17	0.76	161346	0.14	1.96×10^{48}
EGF	0.03	0.46	0.18	103	0.77	158798	0.09	1.89×10^{47}
MYC	0.02	0.48	0.20	66	0.83	157900	0.08	4.79×10^{49}
IGF1R	0.03	0.48	0.14	97	0.80	154144	0.07	4.21×10^{48}
TP53	0.03	0.49	0.13	83	0.82	153530	0.06	2.36×10^{49}
SRC	0.02	0.50	0.24	102	0.81	153342	0.09	1.06×10^{49}
MET	0.02	0.46	0.14	104	0.79	149292	0.08	3.07×10^{48}
PIK3R1	0.02	0.50	0.27	104	0.81	126668	0.09	8.16×10^{48}
ZFYVE9	0.02	0.34	0.27	11	0.79	120608	0.16	6.71×10^{48}
MAPK1	0.02	0.49	0.23	74	0.80	115554	0.09	4.59×10^{48}
BCL2	0.02	0.43	0.12	59	0.81	113214	0.07	6.42×10^{48}
AKT1	0.01	0.47	0.20	79	0.81	109568	0.10	5.33×10^{48}
PCNA	0.01	0.43	0.36	42	0.80	106704	0.09	5.19×10^{48}
STAT3	0.01	0.45	0.30	47	0.80	106498	0.12	5.67×10^{48}

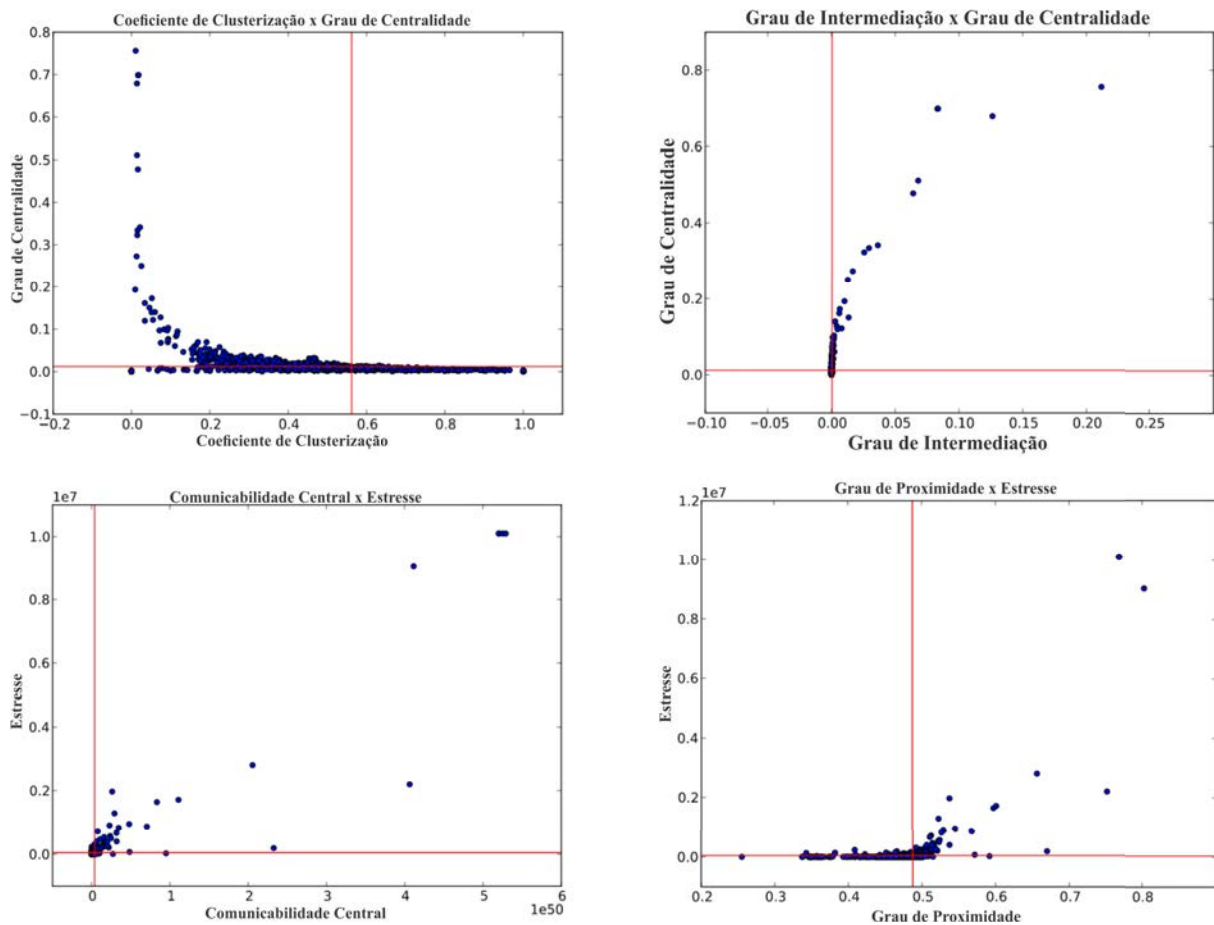


Figura 4.2: Exemplo de alguns gráficos comparativos entre as centralidades.

4.3 Expressão Gênica

Os dados referente a câncer de mama foram retirados do banco de dados GEO. O conjunto analisado está armazenado no GEO sob o código GSE7904 (RICHARDSON et al., 2006), esse conjunto é composto de 62 amostras, sendo 43 de tumor (Basal, Não BLC e BLC), 7 de tecido normal e 12 de organelas normais. Os tumores considerados BLC são aqueles que possuem o gene BRCA1 associado a atividade tumoral.

Os alvos protéicos da tabela 4.2 possuem níveis de expressão que podem ser considerados super ou sub expressos nos tecidos analisados. A correlação entre as centralidades e a análise de expressão diferencial indicou sete alvos protéicos (C1QA, C1QC, EGFR, ERBB2, ESR1, FCGR1A e FCGR1B) que apresentavam altos níveis de expressão no tecido tumoral. Estes alvos correspondem as drogas conhecidas: Lapatinib, Trastuzumabe, Tamoxifeno, Raloxifene. Adicionalmente, a análise da expressão gênica possibilitou detectar 16 possíveis alvos: AR, CYP2B6, SLC26A3, ETS1, GATA2, GATA3, FOXA1, MYC, YBX1, CXCL14, SGSM2, GA-

BARAPL1, REEP1, MAP1LC3B, CRIPAK, AREG E AREGB.

Estes alvos possuem níveis de expressão diferencial e valores de centralidades similares aos EGFR, ERBB2 e ESR1. A Tabela 4.3 apresenta os resultados das expressões gênicas dos sete alvos já conhecidos, e a Tabela 4.4 os resultados das expressões gênicas dos possíveis novos alvos, onde ***logFC*** nos dá o valor do contraste que normalmente representa a variação entre as condições experimentais, sendo que quando esse valor assume valores positivos representa que o gene está mais expresso, e quando assume valores negativos significa que o gene está menos expresso. Enquanto que ***AveExpr*** representa o valor da expressão média de todas os Arrays e canais (SMYTH, 2004).

Existem estudos que estão tentando comprovar a eficácia e a ação desses 16 novos alvos encontrados no presente trabalho (RONDE et al., 2013; JAMIESON et al., 2014; WANG et al., 2012; BOCKHORN et al., 2014; ZHANG; TONG, 2014; WIKMAN et al., 2014; BOYER-GUITTAUT et al., 2014; TALUKDER; MENG; KUMAR, 2005; LINDSTRÖM et al., 2014). Isto pode indicar que as proteínas com altos valores de centralidades tem um papel fundamental no câncer de mama, e que poderiam ser estudadas no desenvolvimento de novas drogas.

Tabela 4.3: Resultado das expressões gênicas para os sete alvos conhecidos.

proteína	Non-BLC-Basal logFC — AveExpr	Normal breast-Basal logFC — AveExpr	Normal breast-Normal.organelle logFC — AveExpr	Normal.organelle-Basal logFC — AveExpr	Normal.organelle-BRCA1 logFC — AveExpr	Normal.organelle-Non-BLC logFC — AveExpr
CIQA			3.083 — 8.965	-2.893 — 8.965	-3.208 — 8.965	-2.646 — 8.965
CIQC			3.445 — 8.951	-3.020 — 8.951	-3.001 — 8.951	-2.991 — 8.951
EGFR						1.118 — 6.834
ERBB2	2.292 — 9.780					
ESR1	3.817 — 7.984	4.238 — 7.984				-2.337 — 9.780
FCGR1A/B				-1.766 — 5.875		-1.699 — 5.875

Tabela 4.4: Resultados das expressões gênicas dos possíveis novos alvos.

Proteína	BRCA1-Não BLC logFC — AveExpr	Não BLC-Basal logFC — AveExpr	Normal-Basal logFC — AveExpr	Normal-Não BLC logFC — AveExpr	Normal.organelle logFC — AveExpr	organelle-Basal logFC — AveExpr	organelle-BRCA1 logFC — AveExpr	organelle-Não BLC logFC — AveExpr
AR		2.848 — 6.930	3.171 — 6.930					
CYP2B6		1.798 — 6.919	2.778 — 4.555	2.685 — 4.555				
SLC26A3						2.928 — 4.555		2.835 — 4.555
ETSI					-3.254 — 6.584	2.426 — 6.584	2.427 — 6.584	3.197 — 6.584
GATA2		1.024 — 6.346	1.186 — 6.346			1.037 — 6.346		
GATA3	-2.960 — 6.712	2.925 — 6.712	2.326 — 6.712		1.850 — 6.712			-2.449 — 6.712
FOXA1	-3.105 — 5.500	3.637 — 5.500	2.252 — 5.500					-2.858 — 5.500
MYC		-1.814 — 10.254				1.633 — 10.254		3.447 — 10.254
YBX1		-1.188 — 11.527						
CXCL14			3.922 — 9.024					
SGSM2			1.283 — 9.555					
GABARAPL1						1.493 — 7.834		1.449 — 7.834
REEP1		1.939 — 6.563						-2.046 — 6.563
MAP1LC3B					-1.039 — 4.845	1.301 — 4.845	1.362 — 4.845	1.220 — 4.845
CRIPAK		1.634 — 7.057	2.058 — 7.057					
AREGB			3.253 — 6.754			5.465 — 6.754	5.247 — 6.754	3.837 — 6.754

5 *Conclusões*

As análises mostram uma correlação da expressão gênica com os valores de centralidade da rede. As proteínas selecionadas como possíveis alvos quimioterápicos mostram que o método foi capaz de indicar alvos já conhecidos, como o caso das proteínas EGFR, ERBB2 e ESR1 e observa-se que outras proteínas podem ser possíveis novos alvos quimioterápicos, como é o caso do AR, CYP2B6, SLC26A3, ETS1, GATA2, GATA3, FOXA1, MYC, YBX1, CXCL14, SGSM2, GABARAPL1, REEP1, MAP1LC3B, CRIPAK, AREG E AREGB que são proteínas importante relacionadas a outros tumores. É necessária a aplicação desta metodologia em mais amostras de expressão gênica de câncer de mama para comprovação da funcionalidade da mesma.

Referências Bibliográficas

- ACENCIO, M. L. et al. Prediction of oncogenic interactions and cancer-related signaling networks based on network topology. *PloS one*, Public Library of Science, v. 8, n. 10, p. e77521, 2013.
- ALBERTS, B. et al. *Molecular Biology of the Cell*. 4th. ed. [S.l.]: Garland Science Taylor & Francis Group, 2002.
- ALMEIDA, V. d. et al. Câncer e agentes antineoplásicos ciclo-celular específicos e ciclo-celular não específicos que interagem com o dna: uma introdução. *Quim Nova*, SciELO Brasil, v. 28, n. 1, p. 118–29, 2005.
- BARRETT, T.; EDGAR, R. [19] gene expression omnibus: Microarray data storage, submission, retrieval, and analysis. *Methods in enzymology*, Elsevier, v. 411, p. 352–369, 2006.
- BEAHR, O. H.; HENSON, D. E. of the manual for staging of cancer. *Cancer*, Wiley Online Library, v. 69, n. 11, p. 2869–2869, 1992.
- BOCKHORN, J. et al. Acquired differentiation and loss of malignancy of spontaneous pulmonary metastases in human-in-mouse breast cancer models. *Cancer research*, AACR, p. canres–1188, 2014.
- BOYER-GUITTAUT, M. et al. The role of gabarapl1/gec1 in autophagic flux and mitochondrial quality control in mda-mb-436 breast cancer cells. *Autophagy*, Landes Bioscience, v. 10, n. 6, p. 0–1, 2014.
- CHOW, A. Y. Cell cycle control by oncogenes and tumor suppressors: Driving the transformation of normal ccell into cancancer cells. *Nature Education*, v. 3, n. 9, 2010.
- ESTIMATIVA, I. Incidência de câncer no brasil. www.inca.gov.br/estimativa/2010. Acesso em, v. 22, n. 02, p. 2010, 2010.
- HAGBERG, A.; SWART, P.; CHULT, D. S. *Exploring network structure, dynamics, and function using NetworkX*. [S.l.], 2008.
- JAMIESON, D. et al. Pharmacogenetics of adjuvant breast cancer treatment with cyclophosphamide, epirubicin and 5-fluorouracil. *Cancer chemotherapy and pharmacology*, Springer, v. 74, n. 4, p. 667–674, 2014.
- KUFE, D. W.; POLLOCK, R. E. *Holland-Frei Cancer Medicine*. [S.l.]: BC Decker, 2003.
- KUO, M. T. Roles of multidrug resistance genes in breast cancer chemoresistance. In: *Breast Cancer Chemosensitivity*. [S.l.]: Springer, 2007. p. 23–30.

LINDSTRÖM, S. et al. Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nature communications*, Nature Publishing Group, v. 5, 2014.

LOPES, E. R. et al. Câncer de mama: epidemiologia e grupos de risco. *Rev. bras. cancerol*, v. 42, n. 2, p. 105–16, 1996.

MEDICINE, J. H. *Breast Cancer Recurrence*. 2009. Disponível em:

<http://www.hopkinsmedicine.org/avon_foundation_breast_center/breast_cancers_other_conditions/breast.c

NEWMAN, M. E. Models of the small world. *Journal of Statistical Physics*, Springer, v. 101, n. 3-4, p. 819–841, 2000.

NEWMAN, M. E. The structure and function of complex networks. *SIAM review*, SIAM, v. 45, n. 2, p. 167–256, 2003.

O'CONNOR, C. M.; ADAMS, J. U. Essentials of cell biology. *NPG Educations*, 2010.

RICHARDSON, A. L. et al. X chromosomal abnormalities in basal-like human breast cancer. *Cancer cell*, Elsevier, v. 9, n. 2, p. 121–132, 2006.

RONDE, J. J. de et al. Serpina6, bex1, agtr1, slc26a3, and laptm4b are markers of resistance to neoadjuvant chemotherapy in her2-negative breast cancer. *Breast cancer research and treatment*, Springer, v. 137, n. 1, p. 213–223, 2013.

ROSADO, J. O.; HENRIQUES, J. A.; BONATTO, D. A systems pharmacology analysis of major chemotherapy combination regimens used in gastric cancer treatment: predicting potential new protein targets and drugs. *Current cancer drug targets*, Bentham Science Publishers, v. 11, n. 7, p. 849–869, 2011.

SCARDONI, G.; PETTERLINI, M.; LAUDANNA, C. Analyzing biological network parameters with centiscape. *Bioinformatics*, Oxford Univ Press, v. 25, n. 21, p. 2857–2859, 2009.

SILVA, I. N. do Cancer Jose Alencar Gomes da. Novembro 2014. Disponível em: <www.inca.gov.br>.

SMITH, R. A. et al. American cancer society guidelines for breast cancer screening: update 2003. *CA: a cancer journal for clinicians*, Wiley Online Library, v. 53, n. 3, p. 141–169, 2003.

SMYTH, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*, v. 3, n. 1, 2004.

TALUKDER, A.; MENG, Q.; KUMAR, R. Cripak, a novel endogenous pak1 inhibitor. *Oncogene*, Nature Publishing Group, v. 25, n. 9, p. 1311–1319, 2005.

WANG, Y. et al. Gata2 negatively regulates pten by preventing nuclear translocation of androgen receptor and by androgen-independent suppression of pten transcription in breast cancer. *Human molecular genetics*, Oxford Univ Press, v. 21, n. 3, p. 569–576, 2012.

WIKMAN, H. et al. Loss of cadm1 expression is associated with poor prognosis and brain metastasis in breast cancer patients. *Oncotarget*, 2014.

WISHART, D. S. et al. Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*, Oxford Univ Press, v. 34, n. suppl 1, p. D668–D672, 2006.

WISHART, D. S. et al. Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research*, Oxford Univ Press, v. 36, n. suppl 1, p. D901–D906, 2008.

ZHANG, Y.; TONG, T. Foxa1 antagonizes ezh2-mediated cdkn2a repression in carcinogenesis. *Biochemical and biophysical research communications*, Elsevier, 2014.