



UFC

**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA
GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS**

ELTON ALVES SANTOS

**CONSTRUÇÃO E ANÁLISE DE REDE DE INTERAÇÃO PROTEÍNA-PROTEÍNA
DE LEUCEMIA MIELOIDE AGUDA: BUSCA POR NOVOS ALVOS PARA A
TERAPIA ANTICÂNCER**

FORTALEZA

2021

ELTON ALVES SANTOS

CONSTRUÇÃO E ANÁLISE DE REDE DE INTERAÇÃO PROTEÍNA-PROTEÍNA DE
LEUCEMIA MIELOIDE AGUDA: BUSCA POR NOVOS ALVOS PARA A TERAPIA
ANTICÂNCER

Trabalho de Conclusão de Curso apresentado à coordenação do Curso de Ciências Biológicas do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Bacharel em Ciências Biológicas. Área de concentração: Biologia.

Orientador: Prof. Dr. Nicholas Costa Barroso Lima.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

S234c Santos, Elton Alves.

Construção e análise de rede de interação proteína-proteína de leucemia mieloide aguda: busca por novos alvos para a terapia anticâncer / Elton Alves Santos. – 2021.
55 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências, Curso de Ciências Biológicas, Fortaleza, 2021.

Orientação: Prof. Dr. Nicholas Costa Barroso Lima.

1. Leucemia mieloide aguda. 2. Câncer. 3. Redes de interação proteína-proteína. I. Título.

CDD 570

ELTON ALVES SANTOS

CONSTRUÇÃO E ANÁLISE DE REDE DE INTERAÇÃO PROTEÍNA-PROTEÍNA DE
LEUCEMIA MIELOIDE AGUDA: BUSCA POR NOVOS ALVOS PARA A TERAPIA
ANTICÂNCER

Trabalho de Conclusão de Curso
apresentado a Graduação em Ciências
Biológicas da Universidade Federal do
Ceará, como requisito parcial à obtenção
do título de Bacharel em Ciências
Biológicas. Área de concentração:
Biologia.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Nicholas Costa Barroso Lima (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Denise Cavalcante Hissa
Universidade Federal do Ceará (UFC)

Profa. Dra. Erika Freitas Mota
Universidade Federal do Ceará (UFC)

Aos meus pais, amigos e familiares

AGRADECIMENTOS

Ao Prof. Dr. Nicholas Costa Barroso Lima, pela excelente orientação.

Às professoras participantes da banca examinadora Denise Cavalcante Hissa e Erika Freitas Mota pelo tempo, pelas valiosas colaborações e sugestões.

Aos colegas da turma, pelas reflexões, críticas e sugestões recebidas.

A minha mãe, Joelina Alves dos Santos, por ter sido a melhor mãe do mundo enquanto viva.

A minha família, por todo o apoio nos momentos de dificuldade.

“A convergência da tecnologia da informação e da biologia permite aos cientistas traduzir o genoma humano em dados digitais que podem acelerar diagnósticos e curas.” Scott Gottlieb

RESUMO

As leucemias estão entre os tipos de câncer mais complexos, sendo bastante estudadas para melhora do prognóstico, diagnóstico e tratamento, visando uma maior sobrevivência dos pacientes. Com os artifícios computacionais e genômicos, é possível analisar como as proteínas relacionadas às leucemias linfoides interagem a fim de prever quais delas poderiam se tornar alvo de medicamentos. Utilizando mecanismos de busca em bancos de dados e na literatura, foi feita uma rede de leucemia mieloide aguda para posterior análise com o aplicativo cytoscape. Foram obtidas 249 proteínas para a LMA. Dentre elas, 45 proteínas de interesse, com duas não apresentando medicamentos registrados, segundo o banco de dados Drug Gene Interaction Database (DGIdb). Concluiu-se que a metodologia utilizada é de interesse científico para aprofundar os resultados obtidos na pesquisa com análises de maior complexidade.

Palavras-chave: leucemia mieloide aguda; câncer; redes de interação proteína-proteína.

ABSTRACT

Leukemias are among the most complex types of cancer, being widely studied to improve the prognosis and diagnosis, treatment, aiming at greater patient survival. With computational and genomic artifices, it is possible to analyze how proteins related to myeloid leukemias interact in order to predict which ones might become drug targets. Using search engines in databases and in the literature, an acute myeloid leukemia network was made for further analysis with the cytoscape application. 249 AML proteins were obtained. Among them, 45 proteins of interest, with two not having registered drugs, according to the Drug Gene Interaction Database (DGIdb). It was concluded that the methodology used is of scientific interest to deepen the results obtained in the research with more complex analyses.

Keywords: acute myeloid leukemia; cancer; protein-protein interaction networks.

LISTA DE FIGURAS

- Figura 1 – Representação do grafo apresentado por Leonhard Euler 29
- Figura 2 – Representação de rede de interações proteína-proteína de levedura..... 31
- Figura 3 – Rede de interação proteína-proteína primária feita a partir do banco de dados My Cancer Genome. As arestas indicam o tipo de informação que comprova a interação entre duas proteínas, amarela para evidência por mineração de texto, azul para evidência de banco de dados e rosa para dados experimentais..... 35
- Figura 4 – Rede de interação proteína-proteína consenso feita a partir dos bancos de dados my cancer genome e uniprot e pesquisa literária. As arestas indicam o tipo de informação que comprova a interação entre duas proteínas, sendo a amarela para evidência por mineração de texto, azul para evidência por banco de dados e rosa dados experimentais..... 36
- Figura 5 – Rede consenso com layout circular feito a partir do Cytoscape. Vermelho: Proteínas com maiores graus e mais mutadas; Roxo: Proteínas mais mutadas de acordo com o MCG; Turquesa: Proteínas com maiores graus; Verde: Proteínas do MCG; Rosa: Busca “acute+myeloid+leukemia” Azul: Busca “acute+myelogenous+leukemia” Laranja: Literatura ; Amarelo: Proteínas encontradas em mais de uma busca..... 37
- Figura 6 – Comparativo das proteínas presentes nas leucemias mieloide aguda e linfóide aguda. Em vermelho estão destacadas as proteínas presentes apenas na rede da LMA, em azul as proteínas presentes apenas na rede da LLA e em roxo as proteínas compartilhadas entre as duas condições..... 43
- Figura 7 – Comparativo das proteínas presentes nas leucemias mieloide aguda e linfóide aguda. Em vermelho estão destacadas as proteínas presentes apenas na rede da LLC, em azul as proteínas

presentes apenas na rede da LMA e em roxo as proteínas 43
compartilhadas entre as duas condições.....

LISTA DE QUADROS

Quadro 1 – Genes significativamente mutados de acordo com o mycancergenome.org.....	23
--	----

LISTA DE TABELAS

Tabela 1 – Proteínas mais mutadas de acordo com o My Cancer Genome e proteínas com maiores graus. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.....	38
Tabela 2 – Proteínas da rede PPI-LMA com maiores graus. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.....	39
Tabela 3 – Proteínas de interesse relacionadas às proteínas das tabelas 1 e 2. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.....	41

LISTA DE ABREVIATURAS E SIGLAS

C.I	Centralidade de Intermediação
C.P	Centralidade de Proximidade e
C.C	Coeficiente de Clusterização
DGIdb	The Drug Gene Interaction Database
LMA	Leucemia Mieloide Aguda
LLA	Leucemia Linfoblástica Aguda
LLC	Leucemia Linfocítica Crônica
MCG	My Cancer Genome
NIH	Instituto Nacional do Câncer Estadunidense
PPI	Interação Proteína-Proteína
TGF β	Fator de Transformação do Crescimento Beta
RAS	Família dos homólogos de oncogenes virais de sarcoma de rato
MYC	Homólogo de oncogene viral da mielocitomatose aviária

SUMÁRIO

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	16
2.1	CÂNCER	16
2.1.1	<i>Tratamentos</i>	21
2.2	LEUCEMIA MIELOIDE AGUDA	22
2.2.1	<i>Proteínas</i>	23
2.2.2	<i>Tratamentos</i>	27
2.3	BIOLOGIA DE SISTEMAS	28
2.3.1	<i>Grafos</i>	29
2.3.2	<i>Redes de Interação Proteína-Proteína</i>	31
3	METODOLOGIA	33
3.1	COLETA DE DADOS PRIMÁRIOS	33
3.2	COLETA DE DADOS ADICIONAIS	33
3.3	RENDERIZAÇÃO DA REDE E ANÁLISE TOPOLÓGICA	33
3.4	BUSCA NO DRUG GENE INTERACTION DATABASE	34
3.5	COMPARATIVO COM LEUCEMIA LINFOBLÁSTICA AGUDA E LEUCEMIA LINFOCÍTICA CRÔNICA	34
4	RESULTADOS E DISCUSSÃO	35
4.1	REDE PRIMÁRIA E CONSENSO	35
4.2	RENDERIZAÇÃO E ANÁLISE TOPOLÓGICA	37
4.3	BUSCA NO DRUG GENE INTERACTION DATABASE	42
4.4	COMPARATIVO COM LEUCEMIA LINFOBLÁSTICA AGUDA E LEUCEMIA LINFOCÍTICA CRÔNICA	42
5	CONCLUSÃO	45
	REFERÊNCIAS	46

1 INTRODUÇÃO

O avanço da biologia molecular, da bioinformática e das ômicas nos últimos anos possibilitou um maior entendimento das interconexões entre as mutações que contribuem para a progressão dos mais diversos tipos de câncer, dentre eles a Leucemia Mieloide Aguda (LMA). Com esse avanço, a farmacogenética e a farmacogenômica possibilitaram com que terapias personalizadas pudessem ser desenvolvidas e aplicadas, o que deu início a medicina personalizada (MA, LU, 2011).

A Farmacogenética é o estudo de como as diferenças genéticas influenciam a variabilidade nas respostas dos pacientes aos medicamentos (ROSES, 2000, p.857). Já a Farmacogenômica é o estudo da relação entre variantes em uma grande coleção de genes, até todo o genoma, e efeitos variáveis de drogas (RODEN *et al.*, 2006, p.750). Podemos, a partir disso, não só catalogar mutações em genes e proteínas de interesse, mas também prever interações entre estes genes e proteínas que podem indicar quais medicamentos utilizar para determinado tratamento.

Para que o estudo dessas interações fosse desenvolvido, surgiu a medicina de sistemas, um ramo da biologia de sistemas com foco em doenças humanas (FLORES *et al.*, 2013, p.566). Todas essas áreas são pilares da medicina p4 (preditiva, preventiva, personalizada e participatória) (HOOD, FLORES, 2012).

À medida que os avanços tecnológicos foram surgindo, as informações, sejam elas biológicas ou não, passaram a ser catalogadas em bancos de dados por grupos de cientistas, o que possibilitou maior facilidade de acesso e disseminação dos dados (BERG *et al.*, 2013). Dentre os bancos de dados biológicos, encontram-se o Gene ontology, o Uniprot, o Genecards o String o My cancer genome e o The drug gene interaction database (ASHBURNER *et al.*, 2000; APWEILER, 2004; STELZER *et al.*, 2016; SZKLARCZYK *et al.*, 2019; JAIN *et al.*, 2020; COTTO *et al.*, 2018).

Com o entendimento de que a biologia de sistemas utiliza de uma perspectiva holística para decifrar os mecanismos relacionados às doenças, a principal metodologia adotada pela área consiste em produzir redes de interações que comportam as mais variadas moléculas biológicas (BARABÁSI, OLTVAI, 2004). Dentre os softwares utilizados para a análise e estudo de redes de interação

proteína-proteína (PPIs) estão o Cytoscape, o ClueGo e o MCODE (SHANNON *et al.*, 2003; BINDEA *et al.*, 2009; BADER, HOGUE, 2003).

Tendo em mente todos esses conceitos, o objetivo deste trabalho se concentra na aplicação da teoria de sistemas para a corroboração de dados já existentes sobre a leucemia mieloide aguda e a busca por possíveis novos alvos moleculares para a terapia anticâncer.

2 REFERENCIAL TEÓRICO

2.1 CÂNCER

O nome câncer se refere ao conjunto de doenças caracterizadas pela expansão desordenada de células com sucessivas mutações que apresentam capacidade de invadir outras regiões do corpo. Segundo Foulds (1958), o estudo ocupacional data de 1775, com o reconhecimento do carcinoma de células escamosas do saco escrotal por Percivall Pott. Seus estudos levaram a proposição do modelo de iniciação, promoção e progressão da doença.

Ao longo de sua trajetória, células cancerosas induzem a formação de tumores, que são tecidos de alta complexidade que propiciam um ambiente favorável para a proliferação celular (HANAHAN, WEINBERG, 2011). Para que a proliferação continue estável, é necessária a aquisição de diferentes competências por parte das células cancerosas, sendo elas adquiridas ao curso de sucessivas mutações genéticas, culminando em fenótipos malignos.

Hanahan e Weinberg (2011) descrevem 6 competências, sendo elas a sustentação de sinais proliferativos, evasão de supressores do crescimento, resistência a morte celular, indução de angiogênese, imortalidade replicativa e ativação da invasão e metástase. Cada uma dessas competências possui mecanismos de atuação específicos, com participação de diversos agentes moleculares, mas todas convergem para a evolução da doença.

Grande parte dos agentes moleculares responsáveis por essas competências são oncogenes e genes supressores de tumor. Oncogenes codificam para proteínas que controlam a proliferação celular, apoptose ou ambos (CROCE, 2008, p.503), classificados entre fatores de transcrição, remodeladores de cromatina, fatores de crescimento e seus receptores, reguladores de apoptose e transdutores de sinal.

Weinberg (1991) descreve genes supressores de tumor como componentes celulares capazes de processar informação de inibição do crescimento, cuja alteração possibilita proliferação de natureza neoplásica. Dentre eles encontra-se o gene tp53, descrito por Strachan e Read (1996) como guardião do genoma, responsável tanto pelo retardo do ciclo celular para reparos no genoma

quando sinalização para a apoptose e autofagia. (GREEN, KROEMER, 2009; ZILFOU, LOWE, 2009).

As interações entre muitos oncogenes e genes supressores de tumor elucidam a complexidade do interactoma celular, possibilitando que diferentes alterações em diferentes proteínas desestabilizam a rede de comunicação responsável pela homeostase. Rotas de sinalização como a do fator de transformação do crescimento beta ($TGF\beta$) demonstram que modificações nos componentes, sejam elas pré ou pós-transcricionais, traducionais ou epigenéticas, direcionam o papel que ele possui, podendo agir tanto na supressão da proliferação celular quanto na promoção da transição epitelial mesenquimal, necessária para a invasão tecidual e metástase (IKUSHIMA, MIYAZONO, 2010; KATSUNO, LAMOUILLE, DERYNCK, 2013).

Como a imortalidade replicativa é uma marca central na tumorigênese, as vias relacionadas a genes pró e anti-apoptóticos estão estreitamente relacionadas ao desequilíbrio homeostático, com papel central da mitocôndria nesse processo (GREEN, REED, 1998).

Emily *et al.* (2001) apontam que a efetivação do apoptose permeia a liberação do citocromo c da membrana interna da mitocôndria para o citosol por meio da permeabilização da membrana externa mediada pelos genes pró-apoptóticos BAK e BAX, culminando na ativação de caspases efetoras. Para que essa cascata seja evitada, genes anti-apoptóticos BCL2 e BCL_{XL} se ligam à BAX e BAK, impedindo a sucessão de eventos.

Outra importante função mitocondrial, a respiração celular, contribui para o aumento de espécies reativas de oxigênio (ROS), culminando na inativação de supressores de tumor como a proteína PTEN e aumento de função em oncogenes como a tirosina-quinase Src, que promove invasão e metástase (PORPORATO *et al.*, 2014; SULLIVAN, CHANDEL, 2014).

Com o crescimento tumoral, North, Moenner e Bikfalvi (2005) atentam que a formação de vasos sanguíneos é imprescindível para a manutenção do microambiente, tendo em vista que o metabolismo acelerado das células cancerígenas reduz os níveis de oxigênio, causando um quadro de hipóxia. Juntamente com os níveis de ROS, a hipóxia atua por meio da indução do fator de

hipóxia (HIF), que regula a atividade celular e atua aumentando os níveis do fator de crescimento endotelial vascular (VEGF) (FORSYTHE *et al.*, 1996; LIAO, JOHNSON, 2007).

A sinalização celular depende de receptores de fatores de crescimento e transdutores de sinal presentes nas membranas plasmáticas, proteínas responsáveis por ativar cascatas de sinalização que culminam na expressão gênica. De acordo com Fantl, Jhonson e Williams (1993), transdutores de sinal são proteínas compostas por um domínio quinase intracelular responsável pela fosforilação de aminoácidos, sendo a tirosina o mais comum, geralmente agindo por meio de dimerização.

Segundo Ruan e Lai (2004), muitas células cancerígenas, na tentativa de burlar esse sistema de sinalização, adquirem a capacidade de auto estimulação, produzindo seus próprios fatores de crescimento. Para além disso, Hanahan e Weinberg (2011) citam a estimulação de células do estroma para a suplementação de fatores de crescimento, a hiper expressão de receptores em caso de baixa concentração extracelular e a ativação constitutiva de componentes das cascatas de sinalização.

Com o microambiente do tumor em desenvolvimento, o remodelamento dos seus componentes por células em proliferação, que por meio de proto-oncogenes da família dos homólogos de oncogenes virais de sarcoma de rato (RAS) e do homólogo de oncogene viral da mielocitomatose aviária (MYC) induzem um estado inflamatório no tecido, é essencial para o recrutamento de células imunes, que por sua vez providenciam fatores de crescimento (GRIVENNIKOV, GRETEN, KARIN, 2010; GRIVENNIKOV, KARIN, 2010).

De acordo com Sparmann e Bar-Sagi (2004), o oncogene RAS induz a inflamação associada a tumores por meio da interleucina 8 (IL8), que também auxilia na angiogênese. Grivennikov, Greten e Karin (2010) também atentam para a estreita relação entre diferentes origens da inflamação e o desenvolvimento de neoplasias, como inflamação crônica por infecção ou autoimunidade, exposição a agentes ambientais e até mesmo induzida por terapias anticâncer.

Baylin e Jones (2011) apontam que o estudo da epigenética atrelado a técnicas de sequenciamento de nova geração (NGS) tem aumentado gradativamente a importância da área no desenvolvimento de neoplasias, dando

origem ao termo “epigenoma do câncer”. O termo epigenética se refere ao estudo de alterações hereditárias na expressão gênica que ocorrem independentemente de alterações na sequência primária de DNA (SHARMA, KELLY, JONES, 2010, p.27).

Ainda segundo Sharma, Kelly e Jones (2010), as principais alterações epigenéticas permeiam a metilação do DNA, modificações de histonas e microRNAs. Estudos apontam que a metilação do DNA ocorre principalmente em ilhas CpG presentes em repetições centroméricas, retrotransposons, rDNA, promotores e extremidades 5' de genes (BIRD, 2002; TAKAI, JONES, 2002), servindo como mecanismo de silenciamento genético (BIRD, 1992; KESHET, LIEMAN-HURWITZ, CEDAR, 1986).

Histonas são proteínas básicas e os principais componentes dos cromossomos em todos os organismos eucarióticos (DELANGÉ, SMITH, 1971, p.279). Suas estruturas consistem em um domínio globular c-terminal e uma cauda n-terminal interagindo em nucleossomos para formar a cromatina, regiões condensadas do núcleo, sendo a complexidade de organização relacionada a quantidade de nucleossomos (LUGER *et al.*, 1997; JENUWEIN, ALLIS, 2001).

Jenuwein e Allis (2001) propõem que a modificação das caudas n-terminais das histonas e posição dos nucleossomos por oncogenes remodeladores da cromatina constituem um código relacionado a estrutura desta, possibilitando ou não o acesso a genes por fatores de transcrição, estando relacionados a progressão do câncer.

Todas as características relacionadas ao desenvolvimento de neoplasias culminam com a metástase. A compreensão da dinâmica desse processo ajudará a identificar alvos para terapias moleculares que podem interromper ou possivelmente reverter o crescimento e as metástases do câncer (FARES *et al.*, 2020, p.01).

Fares *et al.* (2020) descrevem 5 etapas presentes no desenvolvimento de metástases, sendo eles invasão tecidual, intravasão sanguínea, circulação, extravasão e colonização tecidual e escape imunológico.

A invasão tecidual envolve a transição epitelial-mesenquimal:

A transição epitelial-mesenquimal (EMT) é um processo celular durante o qual as células epiteliais perdem sua organização polarizada e junções célula-célula, sofrem alterações no formato celular e na organização citoesquelética e adquirem características mesenquimais, como morfologia celular semelhante a fibroblastos e aumento da migração e invasão celular (TIWARI *et al.*, 2012, p. 194)

Lamouille, Xu e Derynck (2014) apontam que esse processo se inicia com a desconstrução de junções celulares, com a diminuição dos níveis das proteínas das famílias claudina e ocludina, e degradação da E-caderina, não permitindo sua interação com a β -catenina e desestabilizando a polaridade celular. Posteriormente, as células passam por uma reorganização do esqueleto de actina e liberam fatores de degradação da matriz extracelular (MEC), permitindo a invasão tecidual.

Dentre as proteínas responsáveis por degradar a MEC estão as metaloproteinases. A função básica dessas enzimas é a degradação de componentes como o colágeno, e seus mecanismos são importantes para muitas funções como a angiogênese, crescimento celular, diferenciação, apoptose e vigilância imunológica (CURRAN, MURRAY, 2000; EGEBLAD, WERB, 2002). Moss, Jensen-Taubman e Stetler-Stevenson (2012) também discutem o papel dessas proteínas no estabelecimento do nicho pré-metastático.

Segundo Reymond, D'Água e Ridley (2013), a intravasão sanguínea pode ocorrer pela neovasculatura tumoral ou vasos subjacentes. O processo é similar ao de leucócitos, com rolagem pelo endotélio e transmigração que pode ser de maneira paracelular, com a participação da metaloproteinases 1 (MMP1) e macrófagos, que secretam moléculas como fator de necrose tumoral 1^a (TNF1A) e fator de crescimento epidermal (EGF) ou transcelular, com a criação de um poro por meio de complexos calmodulina-Ca²⁺ que ativam cadeias leves de miosina. (WYCKOFF *et al.*, 2007; KHUON *et al.*, 2010).

De acordo com Fares *et al.* (2020), as células metastáticas, após entrarem na circulação sanguínea, podem circular sozinhas ou em clusters. Duda *et al.* (2010) demonstram que, entre as células presentes em clusters, encontram-se fibroblastos do sítio primário, estando associados a maior eficiência do processo por auxiliarem na angiogênese do sítio de colonização ao secretarem fatores como VEGF.

Ao longo da sua trajetória, células e clusters metastáticos também podem interagir com leucócitos, o que pode acarretar em maior eficácia do processo, como exemplificado por Szczerba *et al.* (2019), ao elucidar que neutrófilos podem escoltar células metastáticas ao sítio secundário e promover o ciclo celular enquanto circulam pelo organismo. Coffelt *et al.* (2015) também elucidam que a indução da expressão de IL17 de células T $\gamma\delta$ por IL1 β promove a polarização de neutrófilos

dependentes de fator estimulador de colônias granulocitárias (G-CSF) que inibem células T CD8⁺ citotóxicas, o que facilita o escape imunológico.

A etapa de extravasão ocorre em capilares similarmente a de leucócitos *in vitro*, com rolagem pelo endotélio, aderência primária promovida por selectina ou caderina neuronal (N-caderina) e aderência estável por integrinas. A saída ocorre de maneira paracelular (REYMOND, D'ÁGUA, RIDLEY, 2013; LÄUBLI, BORSIG 2010; MILES *et al.*, 2008).

Para que a colonização ocorra, tumores propiciam o desenvolvimento de nichos pré-metastáticos (PMNs) em órgãos alvo. PMNs são o resultado de efeitos sistêmicos combinados de fatores secretados por tumores e vesículas extracelulares (EVs) eliminadas por tumores que promovem uma sequência temporal de eventos (PEINADO *et al.*, 2017, p. 302). Kaplan *et al.* (2005) demonstraram que o desenvolvimento de PMNs atrai células progenitoras hematopoiéticas da medula óssea (HPCs) positivas para VEGFR1 que auxiliam na chegada de células metastáticas e na neovascularização.

Giancotti (2013) enfatiza que o processo de aclimatação de células metastáticas após a chegada em sítios secundários pode desencadear um período de dormência devido a incapacidade de sair da aquiescência proliferativa ou de fugir da vigilância imunológica, culminando em relapsos anos depois do tratamento e cura da doença.

2.1.1 Tratamentos

Todos os processos pontuados acima são possíveis alvos para o tratamento anticâncer. De acordo com o Instituto Nacional do Câncer Estadunidense (NIH) (2020), as terapias existentes consistem em Cirurgia, radioterapia, quimioterapia, imunoterapia, hormonioterapia, transplante de células tronco e medicina personalizada.

NIH (2020) descreve a cirurgia como procedimento de remoção do câncer, funcionando melhor em tumores sólidos localizados e a radioterapia como um procedimento que utiliza altas doses de radiação para matar células cancerígenas e reduzir tumores. Quando o tumor permanece localizado no momento do diagnóstico, obtém-se uma cura em cerca de um terço dos pacientes por meio de

estratégias de tratamento local, como cirurgia ou radioterapia (CHU, SARTORELLI, 2004, p. 949).

Chu e Sartorelli (2004) apontam três tratamentos distintos envolvendo quimioterapia, sendo eles a quimioterapia primária, utilizada para cânceres avançados sem tratamentos alternativos, a quimioterapia neoadjuvante, utilizada para cânceres com tratamentos alternativos, mas não efetivos e a quimioterapia adjuvante, realizada após cirurgia para reduzir os riscos de reincidência local.

As imunoterapias contra os cânceres existentes incluem várias abordagens, que vão desde a estimulação de mecanismos efetores até a neutralização de mecanismos inibitórios e supressivos (FARKONA, DIAMANDIS, BLASUTIG, 2016, p.02). Farkona, Diamandis e Blasutig (2016) descrevem a utilização de citocinas, vacinas, terapia celular adotiva, anticorpos monoclonais e anticorpos anti PD1 E PD-L1 na clínica, suas vantagens e desvantagens.

2.2 LEUCEMIA MIELOIDE AGUDA

A leucemia mieloide aguda (LMA) é uma doença heterogênea caracterizada pela expansão clonal de progenitores mieloides (blastos) na medula óssea e sangue periférico (SAULTZ, GARZON, 2016, p.01). Atualmente, os subtipos de LMA são classificados de acordo com a Organização mundial da saúde, que subdivide a doença de acordo com características moleculares, morfológicas e associação com outras doenças (Vardiman, Thiele, Arber *et al.*, 2009).

De acordo com Siegel, Miller e Jemal (2020) o número de novos casos estimados da doença nos Estados Unidos em 2020 chegaria a 19.940, sendo 11.090 em homens e 8.850 em mulheres, com prováveis 11.180 mortes. No Brasil, o Instituto Nacional do Câncer (INCA) (2020) estima que para o triênio 2020-2022 serão diagnosticados com leucemia, por ano, cerca de 5.920 casos em homens e de 4.890 em mulheres, sem especificação detalhada para leucemia mieloide aguda. Esses dados demonstram que ainda há uma alta taxa de mortalidade da doença, mesmo com os últimos avanços da oncologia e novas terapias disponibilizadas.

Morfologicamente, os blastos da LMA variam em tamanho um pouco maiores do que os linfócitos até o tamanho de monócitos ou maiores. Os núcleos são grandes em tamanho, de forma variada e geralmente contêm vários nucléolos (SAULTZ, GARZON, 2016, p.01). A caracterização morfológica é muito importante

ao avaliar o esfregaço de um paciente, podendo ser comparada ao histórico de exames do mesmo, possibilitando um diagnóstico precoce e o acompanhamento preciso da progressão da doença.

Um fator que torna a LMA uma doença de difícil prognóstico são as inúmeras origens que ela possui. Shlush *et al.* (2014) apontam a existência de células tronco hematopoiéticas pré-leucêmicas com mutações características da LMA que são herdadas por blastos como sendo a via principal do início da leucemogênese. Dentre as mutações discutidas, a que ocorre na proteína DNMT3A é tida como a mais ancestral. Essa proteína pertence à família das DNA metiltransferases, responsáveis pela metilação de ilhas CpG, processo epigenético relacionado ao silenciamento e ativação de genes (CHALLEN *et al.*, 2012).

Raza e Galili (2012) elucidaram a contribuição das proteínas isocitrato desidrogenases 1 e 2 (IDH1/IDH2), em conjunto com a janus kinase 2 (JAK2V617F) para a transformação leucêmica de pacientes com mielofibroses. Já as proteínas metilcitosina dioxigenase (TET2) e proteína do grupo Polycomb ASLX1, em conjunto com JAK2 e IDH1 contribuem para a transformação leucêmica em neoplasmas mieloproliferativos (MPNs) (TEFFERI *et al.*, 2012). Shih *et al.* (2012) destacam que dentre essas proteínas, TET2, ASLX1 e DNMT3A, por serem fatores epigenéticos, estão associadas com baixa sobrevida de pacientes.

Concomitante a mutações em genes, Kvinlaug *et al.* (2011) ressaltam que a LMA também apresenta fusões genéticas de oncogenes que conferem capacidade de alteração dos padrões de transcrição e sinalização, como as proteínas fusionadas AML1-ETO e NUP98-HOXA9, sendo também associadas ao prognóstico bom ou ruim da doença (GRIMWADE *et al.*, 2010; CHOU *et al.*, 2009).

2.2.1 Proteínas

De acordo com o banco de dados My Cancer Genome, iniciativa do Vanderbilt-Ingram Cancer Center, cerca de 30 genes são significativamente mutados na LMA e estão listados no quadro 1 com seus respectivos nomes e funções.

Quadro 1 — Genes significativamente mutados de acordo com o mycancergenome.org

Gene	Nome	Função	ID
------	------	--------	----

ASLX1	ASXL Transcriptional Regulator 1	Necessário para a determinação normal da identidade de segmentos no embrião em desenvolvimento	GC20P032359
BCOR	BCL6 Corepressor	Interage seletivamente com o domínio POZ de BCL6	GC0XM040049
BCOR L1	BCL6 Corepressor Like 1	Interage com várias histonas desacetilases de classe II para reprimir a transcrição	GC0XP129980
CBL	Cbl Proto-Oncogene	Medeia a transferência de ubiquitina de enzimas de conjugação de ubiquitina (E2) para substratos específicos	GC11P119206
CEPBA	CCAAT Enhancer Binding Protein Alpha	Modula a expressão de genes envolvidos na regulação do ciclo celular, bem como na homeostase do peso corporal	GC19M033299
DNMT 3A	DNA Methyltransferase 3 Alpha	Metilação de ilhas CpG	GC02M025228
ETV6	ETS Variant Transcription Factor 6	Necessário para hematopoiese e manutenção da rede vascular em desenvolvimento	GC12P011649
EZH2	Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit	Pode desempenhar um papel nos sistemas hematopoiético e nervoso central	GC07M148807

FLT3	Fms Related Receptor Tyrosine Kinase 3	Regulação da Hematopoiese	GC13M 028003
GATA2	GATA Binding Protein 2	Papel essencial na regulação da transcrição de genes envolvidos no desenvolvimento e proliferação de linhagens celulares hematopoiéticas e endócrinas	GC03M 128479
IDH1	Isocitrate Dehydrogenase (NADP(+)) 1	Catalisa a descarboxilação oxidativa do isocitrato para 2-oxoglutarato	GC02M 208236
IDH2	Isocitrate Dehydrogenase (NADP(+)) 2	Catalisa a descarboxilação oxidativa do isocitrato para 2-oxoglutarato	GC15M 090083
JAK2	Janus Kinase 2	Envolvida em um subconjunto específico de vias de sinalização do receptor de citocinas	GC09P 004985
KIT	KIT Proto-Oncogene, Receptor Tyrosine Kinase	Fosforilação de múltiplas proteínas intracelulares que participam da proliferação, diferenciação migração e apoptose	GC04P 054657
KMT2A	Lysine Methyltransferase 2A	Papel essencial na regulação da expressão gênica durante o desenvolvimento inicial e hematopoiese	GC11P1 18436
KRAS	KRAS Proto-Oncogene, GTPase	Atividade GTPase	GC12M 025204

NF1	Neurofibromin 1	Parece funcionar como um regulador negativo da via de transdução do sinal RAS	GC17P 031094
NPM1	Nucleophosmin 1	Duplicação do centríolo, chaperonamento de proteínas e proliferação celular	GC05P 171387
NRAS	NRAS Proto-Oncogene, GTPase	Atividade GTPase, navega entre o complexo de golgi e a membrana plasmática	GC01M 114704
PHF6	PHD Finger Protein 6	Potencial papel na regulação da transcrição, que se localiza no nucléolo	GC0XP 134373
PTPN1 1	Protein Tyrosine Phosphatase Non-Receptor Type 11	Regulação do crescimento celular, diferenciação, mitose e transformação oncogênica.	GC21M 034787
RUNX1	Runt-Related Transcription Factor 1	Subunidade alfa da CBF; Envolvimento no desenvolvimento normal da hematopoiese	GC21M 034787
SETBP 1	SET Binding Protein 1	Liga o oncogene nuclear SET que está envolvido na replicação do DNA	GC18P 044680
SF3B1	Splicing Factor 3b Subunit 1	Liga o pré-mRNA a montante do local de ramificação do íntron de uma maneira independente da sequência e pode ancorar o snRNP U2 ao pré-mRNA	GC02M 197389
SRSF2	Serine And Arginine Rich Splicing Factor 2	Envolvido na exportação de mRNA do núcleo e na tradução necessária para splicing de mRNA	GC17M 076734

STAG2	Stromal Antigen 2	Subunidade do complexo de coesina que regula a separação das cromátides irmãs durante a divisão celular	GC0XP 123960
TET2	Tet Methylcytosine Dioxygenase 2	Conversão da metilcitosina para 5-hidroximetilcitosina Envolvida na mielopoiese	GC04P 105145
TP53	Tumor Protein P53	Parada do ciclo celular, apoptose, senescência, reparo do DNA, mudanças metabólicas	GC17M 007661
U2AF1	U2 Small Nuclear RNA Auxiliary Factor 1	Necessário para a ligação de U2 snRNP ao local de ramificação do pré-mRNA	GC21M 043092
WT1	WT1 Transcription Factor	Papel essencial no desenvolvimento do sistema urogenital	GC11M 032365

Fonte: Stelzer *et al.*, 2016.

2.2.2 Tratamentos

Sami *et al.* (2020) resumem as terapias mais utilizadas no tratamento da AML, sendo elas a quimioterapia convencional e as novas terapias de alvos específicos. A quimioterapia consiste na combinação da citarabina (ARA-C), um antimetabólito específico da fase S, com daunorrubicina ou idarrubicina, antraciclinas que interferem na replicação do DNA. Entre as novas terapias encontram-se os inibidores da BMI-1, NF- κ B, FLT3, BCL-2, AURKA/B e IDH1/2.

Os inibidores da FLT3 são subdivididos em fármacos de primeira geração, representados pelos inibidores de tirosina quinase (TKIs) Sorafenib, Sunitinib, Lestaurtinib, Tandutinib e Midostaurin, e de segunda geração, representados pelos TKIs específicos para FLT3 Quizartinib, Crenolanib e Gilteritinib. Já para a inibição seletiva das proteínas IDH1 e IDH2, os TKIs ivosidenib e enasidenib encontram-se disponíveis (WU, LI, ZHU, 2018; MYERS *et al.*, 2018).

Dentre os estudos ainda em desenvolvimento, encontram-se os fármacos inibidores da AURKA (MLN8054, ENMD-2076), AURKB (AZD1152, PHA-680632) e

BMI-1 (PTC-209) (YANG *et al.*, 2012; WANG *et al.*, 2010; LÖWENBERG *et al.*, 2011; SONCINI *et al.*, 2006; DARWISH *et al.*, 2016)

2.3 BIOLOGIA DE SISTEMAS

A biologia de sistemas é um novo campo da biologia que visa desenvolver uma compreensão de sistemas biológicos em nível de sistema (Kitano, 2000). Sua teoria data da década de 50, com a publicação do livro *General System Theory* por Ludwig von Bertalanffy, abordando conceitos como sistemas abertos, retroalimentação e homeostase com uma perspectiva teórica e matemática. Em 1968, Mihajlo Mesarovic, em seu artigo *Systems Theory and Biology*, acaba por definir biologia de sistemas como uma disciplina.

Grandes cientistas da época começaram a utilizar artifícios matemáticos em seus estudos, como Alan Hodgkin e Andrew Huxley nos estudos dos impulsos nervosos e contração muscular (1952), Alan Turing em *A base química da morfogênese* (1952) e Denis Noble com o modelo matemático da função cardíaca (1960), contribuindo para o uso de modelos computacionais.

Apesar da biologia de sistemas propor uma abordagem holística, ou seja, não apenas analisar as partes, mas as propriedades emergentes a partir das interações entre as partes, a quantidade de informação biológica disponível na época acabou inviabilizando o desenvolvimento da área, fazendo com que o termo entrasse em desuso. Após quatro décadas, com o desenvolvimento de metodologias de sequenciamento genético, melhora do poder computacional e de diversas áreas da biologia, entre elas a bioinformática e as ômicas, foi possível com que a disciplina fosse popularizada, em grande parte graças ao cientista Hiroaki Kitano.

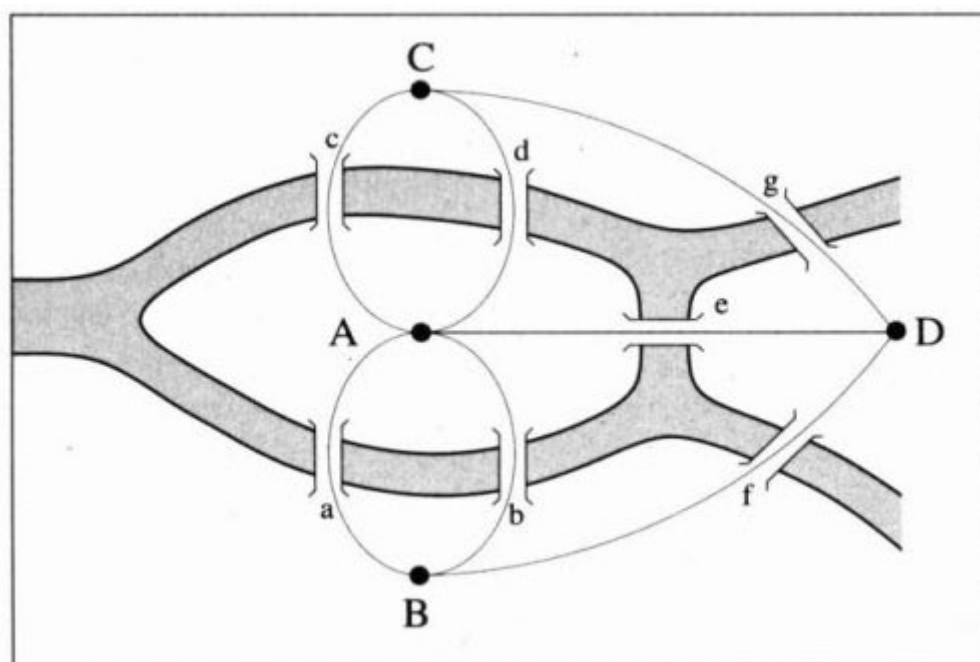
Barabási e Oltvai (2004) enfatizam a necessidade de entender os padrões de comunicação dos componentes celulares por meio do que eles chamam de redes biológicas, um conjunto de rotas e interações entre diferentes moléculas governadas por leis universais que ditam a organização e o comportamento celular.

2.3.1 Grafos

Um grafo consiste em um conjunto finito de vértices, um conjunto finito de arestas e uma regra que nos diz quais arestas unem quais pares de vértices (BIGGS, LLOYD, WILSON, 1986, p.9). Vértices, ou nós, são os elementos que

queremos representar, e as arestas são as relações entre esses elementos. A teoria dos grafos data do século 18, com a tentativa de resolução do problema das pontes de Kaliningrado pelo matemático e físico Leonhard Euler em seu artigo “Solutio problematis ad geometriam situs pertinentis” de 1741, como mostrado na figura x.

Figura 1. Representação do grafo apresentado por Leonhard Euler.



Fonte: Cardoso, 2017.

Para a biologia molecular, os elementos representados em grafos são constituídos de moléculas biológicas, como proteínas, ácidos nucleicos e seus substratos e as arestas são interações físicas ou funcionais, que determinados elementos possuem entre si. Estas interações podem ser direcionadas, como em rotas metabólicas, indicando uma precedência ou fluxo de informação, ou não direcionadas, indicando uma relação (BARABÁSI, OLTVAI, 2004).

A análise de grafos, ou redes biológicas para a biologia, perpassa por leis e parâmetros que nos auxiliam a entender a complexidade dos dados apresentados. Dentre eles encontram-se os conceitos de grau, menores caminhos, centralidades, coeficiente de agrupamento e módulos e as ideias de rede livre de escala e o efeito de pequeno mundo (ALBERT, BARABÁSI, 2002).

A característica mais elementar de um nó é seu grau (ou conectividade), que nos diz quantos links o nó tem para outros nós (BARABÁSI, OLTVAI, 2004, p.102). Ou seja, para um determinado nó de grau 3, em uma rede não direcionada, isso significa que ele possui conexão com outros 3 nós desta rede. Já os menores caminhos perpassam pela busca da menor distância necessária para se chegar de um nó A para um nó B (BARABÁSI, OLTVAI, 2004).

Para redes biológicas, um módulo é uma entidade discreta, cuja função é separável das de outros módulos (HARTWELL *et al.*, 1999, p.49), o que permite a análise separada de partes de uma rede, culminando em discussões mais robustas.

As centralidades são características que distinguem a importância de determinado nó para com a rede inteira. Dentre elas encontram-se a centralidade de grau, ou seja, quanto mais conexões um nó possui, maior sua influência, caracterizando-o como um Hub (LYRA, OLIVEIRA, 2011), centralidade de proximidade, dada pelo comprimento médio dos caminhos mais curtos, estando o vértice mais próximo de todos os outros vértices (LYRA, OLIVEIRA, 2011), e centralidade de intermediação, que relaciona a quantidade de vezes que um vértice encontra-se no menor caminho entre outros dois vértices.

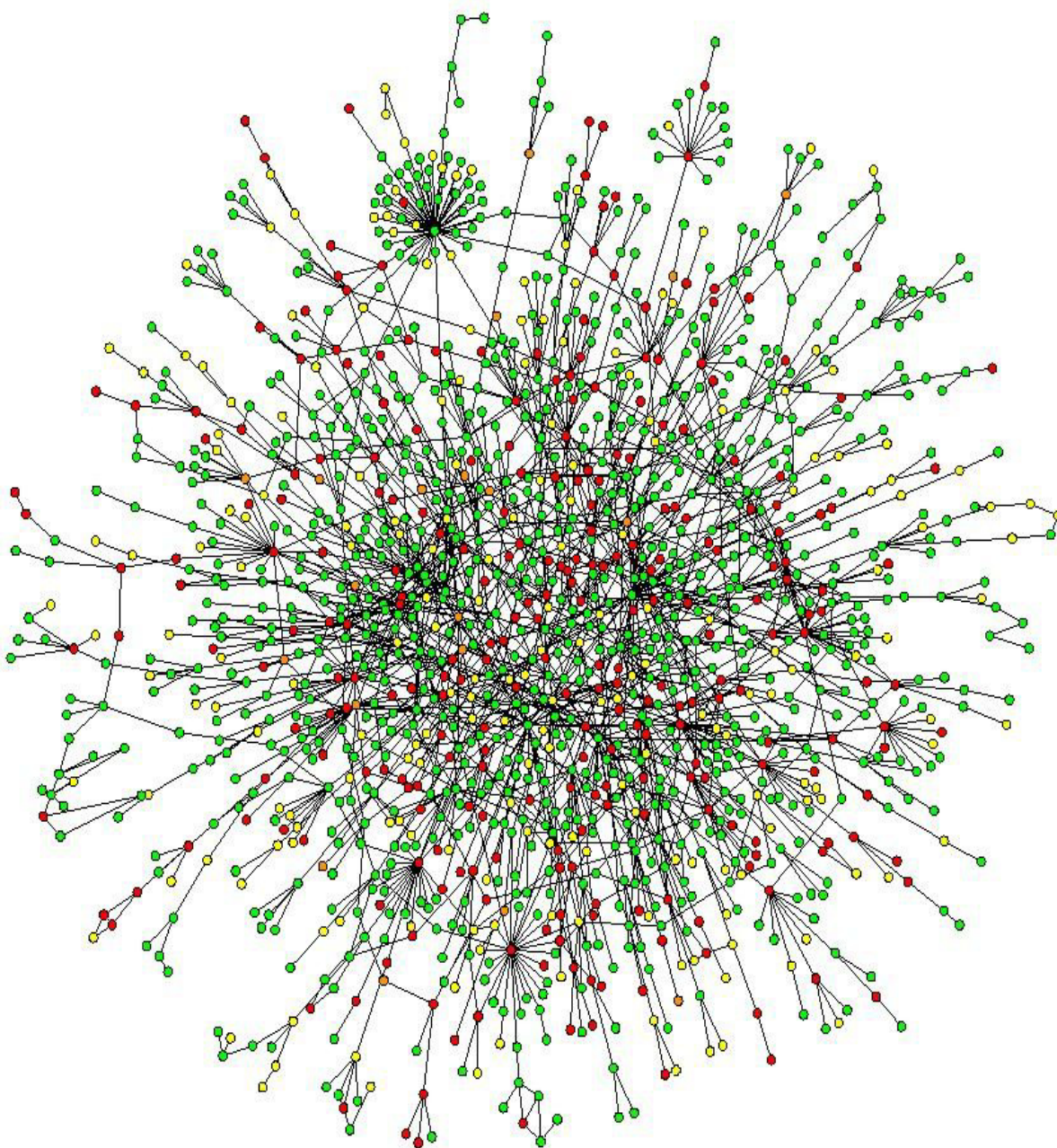
Redes livre de escala caracterizam-se por ter muitos nós com poucas conexões e poucos nós com muitas conexões, indo em contraponto a redes aleatórias, obedecendo a lei de potência, o que impossibilita a atribuição de um nó que caracteriza toda a rede (BARABÁSI, OLTVAI, 2004). Essa característica permite que redes biológicas tenham maior resistência à interferências ou falhas. Já a ideia de mundo pequeno indica que quaisquer dois nós podem ser conectados com um caminho de apenas alguns links (BARABÁSI, OLTVAI, 2004, p.106), o que na prática indica que uma pequena mudança em um ponto é rapidamente percebido pelo restante da rede.

2.3.2 Redes de Interação Proteína-Proteína (PPI)

Em redes de interação proteína-proteína (PPI), um sistema biológico é descrito em termos de proteínas, ou seja, os nós e suas relações (interações físicas/funcionais), ou seja, as linhas (VELLAI *et al.*, 2018, p.1). Por se tratarem de possíveis realidades celulares, redes PPI são ferramentas viáveis para entender as funções celulares, os mecanismos de doenças e o design/reposicionamento de

medicamentos (VELLA *et al.*, 2018, p.1). A partir dessas características podemos ter uma visão holística dos mais complexos mecanismos de interação e a relevância que determinadas proteínas possam ter para com a rede. A figura x nos mostra um exemplo de rede PPI de levedura.

Figura 2. Representação de rede de interações proteína-proteína de levedura



Fonte: Barabási e Oltvai, 2004.

Outra característica das redes PPI é a possibilidade de trabalhar com sub grafos dentro da rede principal a partir da criação de módulos funcionais e complexos proteicos. Os complexos de proteínas são grupos de proteínas que interagem umas com as outras, por isso são geralmente sub gráficos densos em redes PPI (LIU, WONG, CHUA, 2009, p.1981), já os módulos funcionais são compostos por proteínas que compartilham determinadas funções, mas não necessariamente interagem entre si ou estão próximas na rede (VELLA *et al.*, 2018).

3 METODOLOGIA

3.1 COLETA DE DADOS PRIMÁRIOS

Para a lista inicial de genes relacionados à Leucemia Mieloide Aguda, foi feita uma pesquisa no banco de dados de genes My Cancer Genome, que retornou 119 genes. Em seguida, foi feita uma rede de interação proteína-proteína com o STRING v.11 com a finalidade de observar a qualidade de confiança das interações para o escore 0.7 a partir de evidências de mineração de texto, dados experimentais e banco de dados.

Dos 119 genes presentes na rede apresentada, 9 genes não apresentaram interações com outros genes, mostrando-se necessário o aumento da rede para tentar incluí-los.

3.2 COLETA DE DADOS ADICIONAIS

Foram realizadas buscas no banco de dados UniProt usando como termos de busca “*acute myeloid leukemia*”, “*acute myelocytic leukemia*” e “*acute myelogenous leukemia*” para aumentar a rede com possíveis proteínas relacionadas a lista inicial a fim de integrar à rede proteínas que ficaram sem interações após a filtragem de escore de confiança. Foram obtidas 165 proteínas para a busca “*acute myeloid leukemia*”, 37 para “*acute myelogenous leukemia*” e 15 para “*acute myelocytic leukemia*”. Uma pesquisa de literatura científica com o intuito de encontrar genes relacionados à LMA que não estivessem presentes nas buscas e na lista inicial foi realizada em paralelo, obtendo-se um total de 49 genes.

Unindo todos os genes em uma lista consenso, foram obtidos 295 genes relacionados à LMA. Em seguida, foi feita uma nova rede PPI através do STRING v.11 seguindo os mesmos critérios da primeira. Ao final, 48 proteínas que não apresentaram interações com nenhum outro gene de acordo com o escore de confiança 0.7 foram removidas da rede, dentre elas 5 da lista primária, restando assim 249 proteínas na rede consenso.

3.3 RENDERIZAÇÃO DA REDE E ANÁLISE TOPOLÓGICA

Através da tabela de interações baixada do site do STRING foi utilizado o aplicativo Cytoscape v3.8.2 para a análise da rede consenso. Com a utilização do aplicativo do STRING presente no Cytoscape, foi possível recuperar as mesmas características visuais utilizadas no site.

Com a rede finalizada e modificada para o layout circular, as proteínas foram marcadas de acordo com a busca na qual elas foram encontradas e foi feita então a análise topológica utilizando a função “*Analisar rede*” do Cytoscape. A análise retornou os dados de grau, centralidades de intermediação e proximidade, estresse, comprimento médio dos menores caminhos e coeficiente de clusterização.

Com os dados da análise, foram separadas as trinta proteínas com maiores graus e as proteínas mais mutadas de acordo com o banco de dados my cancer genome a fim de descobrir quais estão presentes nas duas listas. A partir disso, foram mapeadas quais proteínas relacionadas a estas possuem maior grau para com a rede e interações com estas.

3.4 BUSCA NO DRUG GENE INTERACTION DATABASE

Foram feitas buscas no banco de dados The Drug Gene Interaction Database (DGIdb) para as 45 proteínas de interesse selecionadas na rede PPI-LMA com a finalidade de encontrar proteínas que não possuam medicamentos registrados.

3.5 COMPARATIVO COM LEUCEMIA LINFOBLÁSTICA AGUDA E LEUCEMIA LINFOCÍTICA CRÔNICA

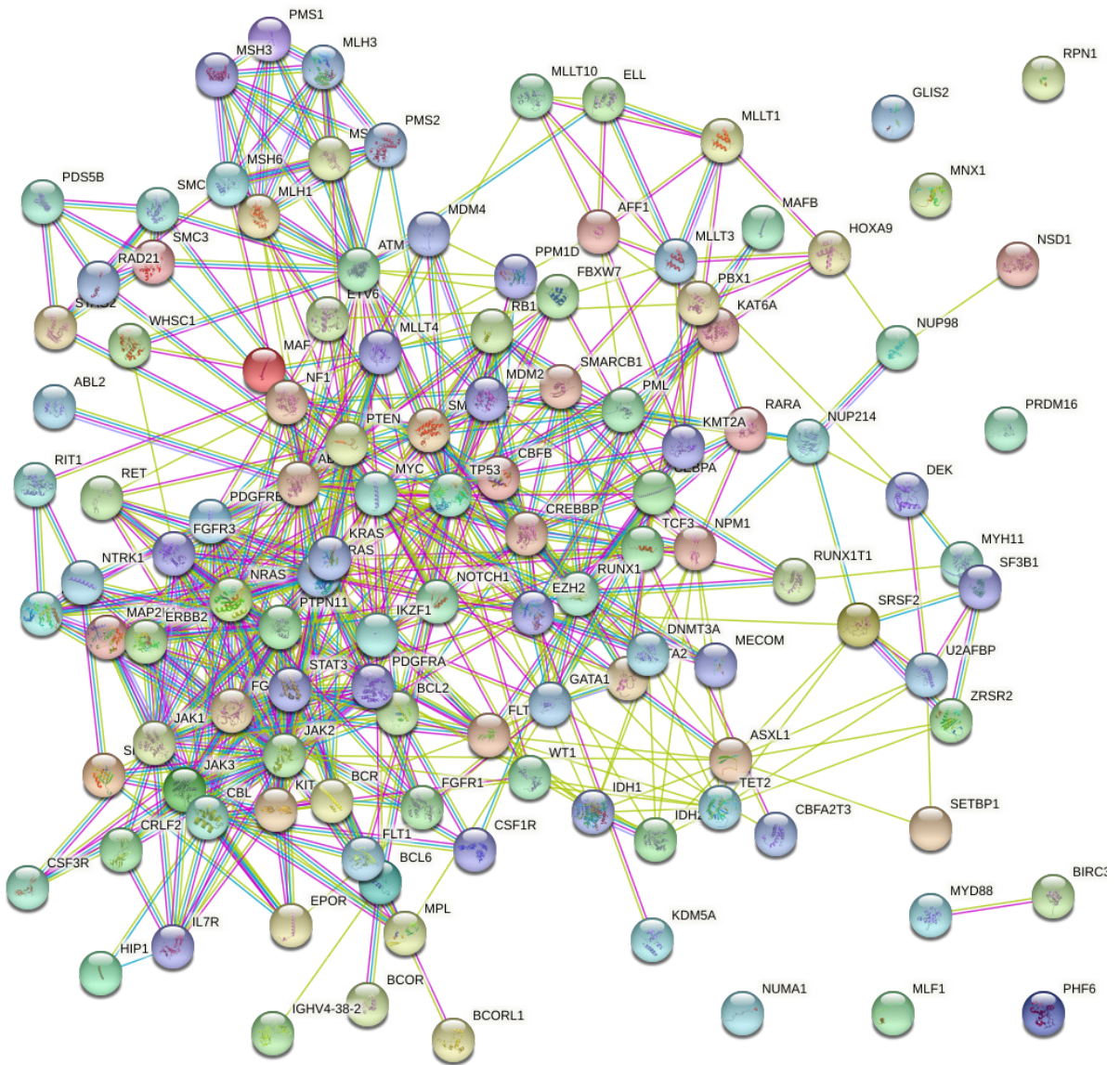
Foram realizados comparativos da rede de LMA com redes de LLA e LLC feitas a partir dos mesmos critérios para observar quais proteínas que não possuem medicamentos registrados no DGIdb são compartilhadas pelas doenças a fim de determinar a possibilidade do estudo dessas proteínas.

4 RESULTADOS E DISCUSSÃO

4.1 REDES PRIMÁRIA E CONSENSO

A rede primária feita pelo servidor on-line do banco de dados STRING retornou 119 proteínas, das quais 9 não possuem interações quando o escore de confiança de 0,7 é aplicado (Figura 3).

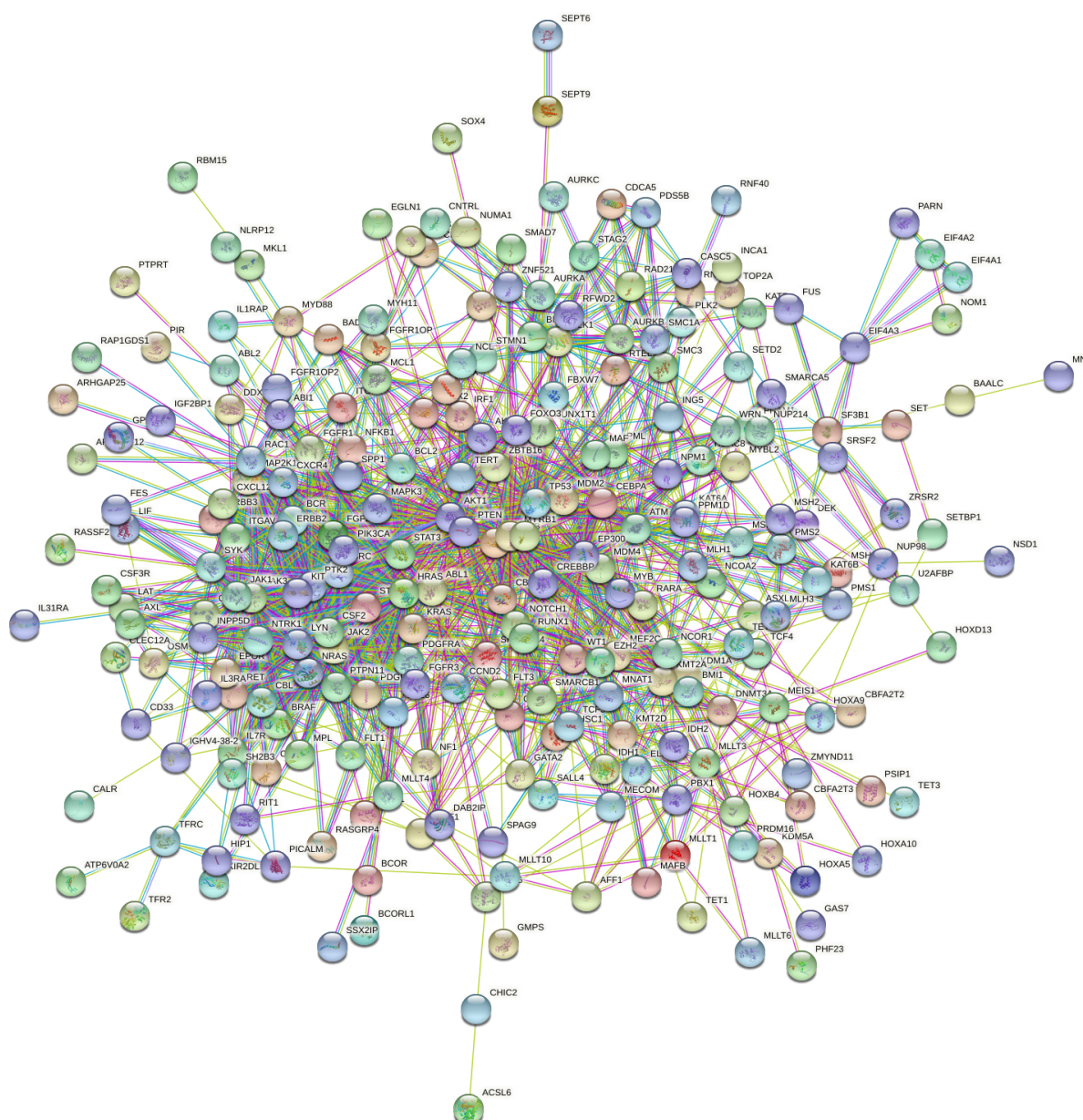
Figura 3. Rede de interação proteína-proteína primária feita a partir do banco de dados My Cancer Genome. As arestas indicam o tipo de informação que comprova a interação entre duas proteínas, amarela para evidência por mineração de texto, azul para evidência de banco de dados e rosa para dados experimentais.



Fonte: Szklarczyk *et al.*, 2019.

A rede consenso, feita com o intuito de incluir as 9 proteínas que ficaram sem interações na rede primária retornou 295 proteínas, e após a filtragem com escore 0,7, 48 destas proteínas foram removidas por não apresentarem interações, sobrando 249 (Figura 4). O número de proteínas descartadas na rede consenso possui um aumento considerável em relação a rede primária devido a falta de evidências que permitissem a elas terem interações na rede de acordo com o escore 0,7, necessário para aumentar a confiabilidade dos dados apresentados para futuras análises.

Figura 4. Rede de interação proteína-proteína consenso feita a partir dos bancos de dados my cancer genome e uniprot e pesquisa literária. As arestas indicam o tipo de informação que comprova a interação entre duas proteínas, sendo a amarela para evidência por mineração de texto, azul para evidência por banco de dados e rosa dados experimentais.

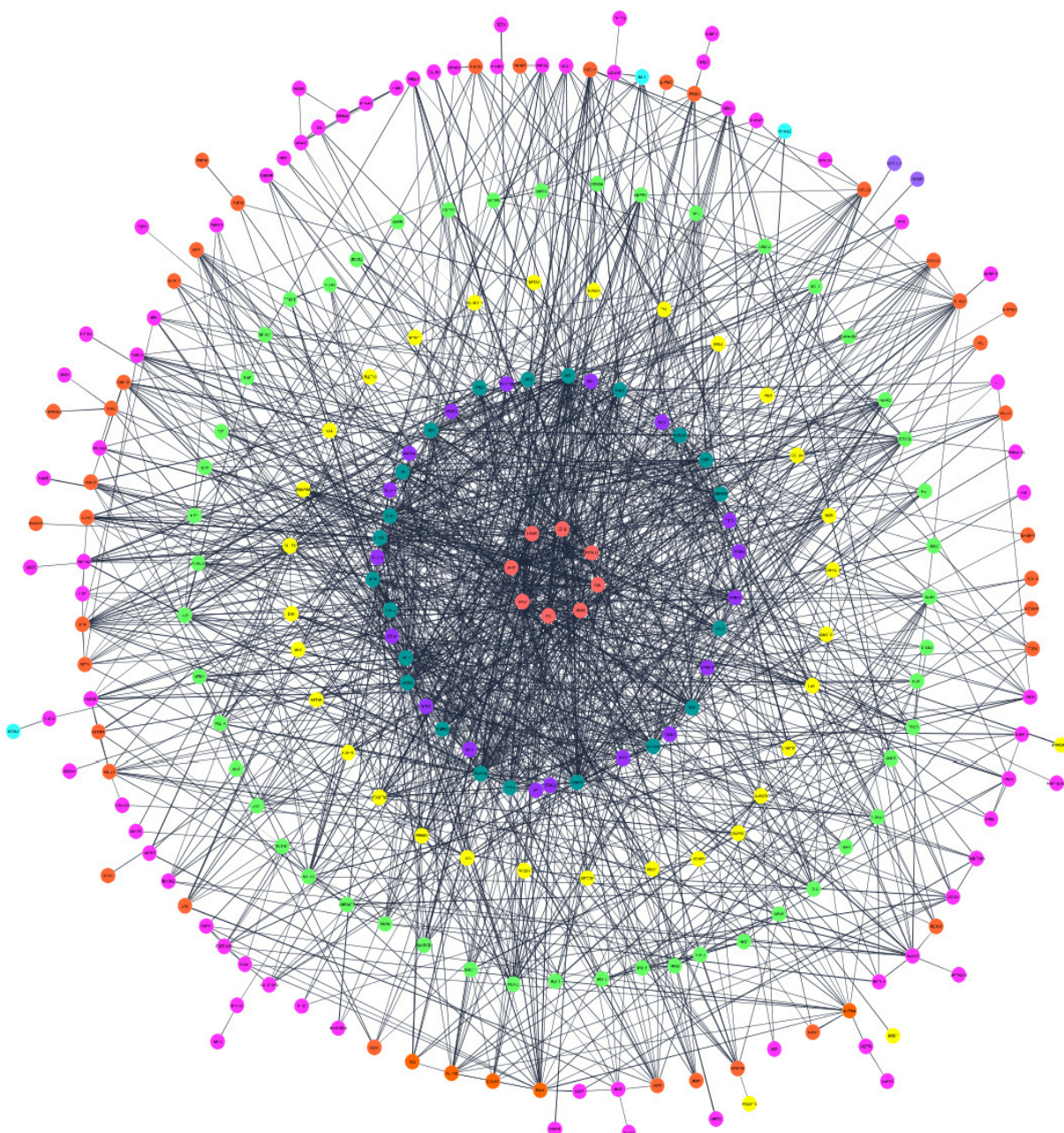


Fonte: Szklarczyk *et al.*, 2019.

4.2 RENDERIZAÇÃO DA REDE CONSENSO E ANÁLISE TOPOLÓGICA

A mudança para o layout circular (figura 5) foi feita para permitir melhor visualização das proteínas e distribuí-las de acordo com o grau apresentado e com a busca na qual foram obtidas.

Figura 5. Rede consenso com layout circular feito a partir do Cytoscape. Vermelho: Proteínas com maiores graus e mais mutadas; Roxo: Proteínas mais mutadas de acordo com o MCG; Turquesa: Proteínas com maiores graus; Verde: Proteínas do MCG; Rosa: Busca “acute+myeloid+leukemia” Azul: Busca “acute+myelogenous+leukemia” Laranja: Literatura ; Amarelo: Proteínas encontradas em mais de uma busca.



Fonte: Shannon *et al.*, 2003.

A tabela 1 contém dados referentes às proteínas mais mutadas de acordo com o banco de dados my cancer genome e a tabela 2 as proteínas com maiores graus, com as proteínas pertencentes às duas listas destacadas em vermelho. Já a tabela 3 contém as proteínas relacionadas às das tabelas 1 e 2 que possuem maiores graus para com a rede.

Tabela 1. Proteínas mais mutadas de acordo com o My Cancer Genome e proteínas com maiores graus. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.

Proteína	Grau	C.I	C.P	C.C
ASXL1	13	14.901.096.405.64 6.100	37.237.237.237.23 7.200	34.615.384.615.38 4.600
BCOR	1	0.0	29.488.703.923.90 0.100	0.0
BCORL1	1	0.0	29.488.703.923.90 0.100	0.0
CBL	26	26.009.255.196.45 6.400	38.390.092.879.25 6.900	3.384.615.384.615. 380
CEBPA	15	8.013.094.500.113. 100	3.821.263.482.280 .430	24.761.904.761.90 4.700
DNMT3A	13	11.691.215.319.40 6.600	3.887.147.335.423 .190	358.974.358.974.3 59
ETV6	5	16.820.808.573.88 5.900	34.206.896.551.72 4.100	0.1
EZH2	30	6.203.056.902.529. 950	4.635.514.018.691 .580	2.597.701.149.425. 280
FLT3	18	15.460.200.992.01 9.900	40.456.769.983.68 6.700	3.137.254.901.960. 780
GATA2	11	2.155.134.678.154. 730	3.803.680.981.595 .090	38.181.818.181.81 8.100
IDH1	9	9,79E+10	3.887.147.335.423 .190	6.666.666.666.666. 660
IDH2	8	6,51E+11	3.740.573.152.337 .850	7.142.857.142.857. 140
JAK2	46	42.434.227.912.73 6.500	4.592.592.592.592 .590	3.169.082.125.603. 860
KIT	27	1.580.960.740.456. 460	4.239.316.239.316 .230	3.561.253.561.253. 560
KMT2A	13	1.245.772.274.663. 780	3.641.703.377.386 .190	3.717.948.717.948. 710
KRAS	46	26.983.496.242.06 4.000	4.824.902.723.735 .400	3.497.584.541.062. 800

NF1	11	6,70E+11	3.875	5.818.181.818.181.810
NPM1	22	3.514.259.248.587.800	4.224.872.231.686.540	18.614.718.614.718.600
NRAS	37	1.745.246.915.439.050	4.373.897.707.231.040	4.039.039.039.039.030
PTPN11	37	14.881.002.084.915.400	4.328.097.731.239.090	44.594.594.594.594.500
RUNX1	23	11.540.841.412.123.300	4.161.073.825.503.350	2.648.221.343.873.510
SETBP1	3	7,75E+11	27.647.714.604.236.300	3.333.333.333.333.330
SF3B1	8	16.383.659.036.663.100	3.502.824.858.757.060	32.142.857.142.857.100
SRSF2	9	12.055.175.140.070.900	34.065.934.065.934.000	3.611.111.111.111.110
STAG2	11	1.426.280.967.948.630	36.795.252.225.519.200	8.181.818.181.818.180
TET2	12	763.541.256.431.637	37.181.409.295.352.300	3.787.878.787.878.780
TP53	79	1.641.564.930.514.610	5.299.145.299.145.290	1.460.564.751.703.990
U2AFBP	6	10.751.155.175.609.700	2.970.059.880.239.520	0.4
WT1	5	5,46E+11	3.780.487.804.878.040	0.4

Fonte: Shannon *et al.*, 2003.

Tabela 2. Proteínas da rede PPI-LMA com maiores graus. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.

Proteína	Grau	C.I	C.P	C.C
ABL1	26	29.120.616.379.725.700	4.381.625.441.696.110	22.153.846.153.846.100
AKT1	60	6.851.733.331.641.690	5.030.425.963.488.840	22.937.853.107.344.600
ATM	34	37.756.070.694.934.000	44.364.937.388.193.200	20.677.361.853.832.400
CBFB	28	12.127.085.145.629.200	42.176.870.748.299.300	25.396.825.396.825.300
CBL	26	26.009.255.196.456.400	38.390.092.879.256.900	3.384.615.384.615.380
CREBBP	37	20.431.673.331.955.900	45.173.041.894.353.300	2.747.747.747.747.740
CSF2	29	8.163.733.026.468.640	4.203.389.830.508.470	41.625.615.763.546.700

EP300	53	11.489.536.749.41 0.000	501.010.101.010.1 01	16.981.132.075.47 1.600
EZH2	30	6.203.056.902.529. 950	4.635.514.018.691 .580	2.597.701.149.425 .280
FGF2	37	12.138.306.712.42 8.100	43.508.771.929.82 4.500	34.684.684.684.68 4.600
HRAS	52	53.106.683.368.18 6.800	49.011.857.707.50 9.800	2.918.552.036.199 .090
JAK1	36	16.263.688.700.88 2.000	4.099.173.553.719 .000	37.142.857.142.85 7.100
JAK2	46	42.434.227.912.73 6.500	4.592.592.592.592 .590	3.169.082.125.603 .860
JAK3	29	4.293.311.319.295. 240	4.025.974.025.974 .020	4.458.128.078.817 .730
KIT	27	1.580.960.740.456. 460	4.239.316.239.316 .230	3.561.253.561.253 .560
KRAS	46	26.983.496.242.06 4.000	4.824.902.723.735 .400	3.497.584.541.062 .800
LYN	29	8.662.344.773.584. 240	4.189.189.189.189 .180	45.566.502.463.05 4.100
MAPK3	40	162.010.865.447.0 91	4.509.090.909.090 .900	3.294.871.794.871 .790
MYC	50	4.363.272.240.872. 500	4.815.533.980.582 .520	23.755.102.040.81 6.300
NOTCH1	26	13.623.735.352.62 8.800	4.452.423.698.384 .200	3.569.230.769.230 .760
NRAS	37	1.745.246.915.439. 050	4.373.897.707.231 .040	4.039.039.039.039 .030
PIK3CA	55	38.807.627.049.21 8.600	46.792.452.830.18 8.600	2.841.750.841.750 .840
PTEN	37	17.923.413.379.09 3.800	455.045.871.559.6 33	3.108.108.108.108 .100
PTK2	26	4.410.265.284.406. 580	42.105.263.157.89 4.700	5.169.230.769.230 .760
PTPN11	37	14.881.002.084.91 5.400	4.328.097.731.239 .090	44.594.594.594.59 4.500
RAC1	32	40.247.443.905.68 3.500	40.789.473.684.21 0.500	2.903.225.806.451 .610
SRC	44	3.246.347.516.687. 530	47.784.200.385.35 6.400	35.095.137.420.71 8.800
STAT3	58	6.824.599.945.667. 940		25.105.868.118.57 2.200
STAT5B	34	1.520.134.694.205. 910	4.525.547.445.255 .470	3.244.206.773.618 .530
TP53	79	1.641.564.930.514. 610	5.299.145.299.145 .290	1.460.564.751.703 .990

Fonte: Shannon *et al.*, 2003.

Tabela 3. Proteínas de interesse relacionadas às proteínas das tabelas x e x. C.I - Centralidade de Intermediação, C.P - Centralidade de Proximidade e C.C - Coeficiente de Clusterização.

Proteína	Grau	C.I	C.P	C.C
AKT2	16	3.163.403.180.741 .970	40.456.769.983.68 6.700	36.666.666.666.66 6.600
AURKB	17	10.955.924.623.41 2.700	39.178.515.007.89 8.800	3.897.058.823.529 .410
BCR	17	24.030.576.976.90 7.000	39.116.719.242.90 2.200	4.632.352.941.176 .470
BIRC5	16	18.663.189.919.08 5.000	4.032.520.325.203 .250	375
CXCL12	16	45.321.731.528.52 3.700	4.065.573.770.491 .800	6.833.333.333.333 .330
CXCR4	16	4.404.268.590.170 .740	40.589.198.036.00 6.500	6.166.666.666.666 .660
ERBB2	24	7.010.880.563.020 .490	4.203.389.830.508 .470	5.797.101.449.275 .360
FGFR3	16	34.196.030.184.86 2.200	3.993.558.776.167 .470	6.333.333.333.333 .330
GATA1	18	9.422.631.161.709 .890	4.012.944.983.818 .770	29.411.764.705.88 2.300
ITGAV	16	4.019.057.949.845 .700	36.849.925.705.79 4.900	35.833.333.333.33 3.300
MAP2K1	21	5.102.762.314.194 .950	4.012.944.983.818 .770	5.904.761.904.761 .900
MDM2	23	13.686.983.977.36 4.600	4.328.097.731.239 .090	3.675.889.328.063 .240
NCOR1	18	2.120.845.393.783 .100	40.589.198.036.00 6.500	27.450.980.392.15 6.800
NFKB1	24	2.222.433.855.533 .400	4.381.625.441.696 .110	391.304.347.826.0 87
PDGFRA	22	38.861.872.747.42 7.800	41.471.571.906.35 4.500	5.367.965.367.965 .360
PDGFRB	24	12.141.656.783.31 7.000	431.304.347.826.0 87	5.869.565.217.391 .300
PLK1	23	23.259.930.010.17 1.600	418.212.478.920.7 42	3.241.106.719.367 .580
PML	17	11.657.542.850.65 3.500	41.264.559.068.21 9.600	36.764.705.882.35 2.900
RARA	16	10.554.512.689.82 6.100	40.456.769.983.68 6.700	30.833.333.333.33 3.300
RB1	19	454.486.642.087.5 43	41.471.571.906.35 4.500	4.093.567.251.461 .980
SYK	24	53.211.061.721.52 3.300	389.937.106.918.2 39	42.391.304.347.82 6.000
TCF3	19	12.064.627.197.95 0.200	389.937.106.918.2 39	27.485.380.116.95 9.000

TERT	17	45.440.399.002.13 7.000	40.924.092.409.24 0.900	38.235.294.117.64 7.000
------	----	----------------------------	----------------------------	----------------------------

Fonte: Shannon *et al.*, 2003.

O grau das proteínas têm papel relevante para o propósito da pesquisa pois com ele podemos inferir quais proteínas são mais centrais na rede, tendo maior impacto caso ocorra falhas ou interferências, o que no caso do câncer seriam mutações nas proteínas ou alterações no processamento destas.

4.3 BUSCA NO DRUG GENE INTERACTION DATABASE

Das 45 proteínas selecionadas para a busca no DGIdb (ABL1, AKT1, AKT2, ATM, AURKB, BCR, BIRC5, CBF2, CREBBP, CSF2, CXCL12, CXCR4, EP300, ERBB2, FGF2, FGFR3, GATA1, HRAS, ITGAV, JAK1, JAK3, LYN, MAP2K1, MAPK3, MDM2, MYC, NCOR1, NFKB1, NOTCH1, PDGFRA, PDGFRB, PIK3CA, PLK1, PML, PTEN, PTK2, RAC1, RARA, RB1, SRC, STAT3, STAT5B, SYK, TCF3, TERT), apenas duas (NCOR1 e TCF3) não apresentaram medicamentos registrados no banco de dados, o que corrobora com a metodologia adotada para a busca por possíveis proteínas alvo na terapia anticâncer.

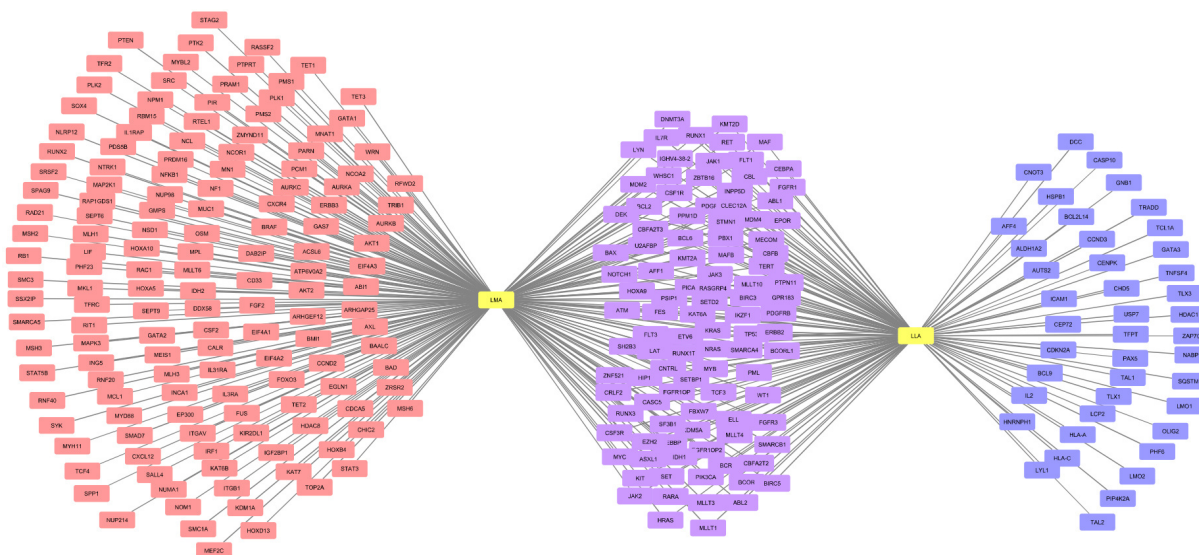
A proteína NCOR1 interage com 6 proteínas dentre as que apresentam maiores graus e são mais mutáveis, sendo elas TP53, EZH2, EP300, CEBPA, CREBBP e NOTCH1. Já a proteína TCF3 interage com 7, sendo elas GATA2, CBLB, EP300, ABL1, NOTCH1, CREBBP, RUNX2. Dessas proteínas, apenas 3, EP300, CREBBP e NOTCH1 são compartilhadas entre as duas.

A proteína TCF3 possui maiores valores com relação às centralidades de intermediação e proximidade, estando assim mais próxima de todas as outras proteínas da rede e presente em mais menores caminhos entre outras proteínas, apresentando assim uma maior influência sobre os acontecimentos celulares.

4.4 COMPARATIVO COM LEUCEMIA LINFOBLÁSTICA AGUDA E LEUCEMIA LINFOCÍTICA CRÔNICA

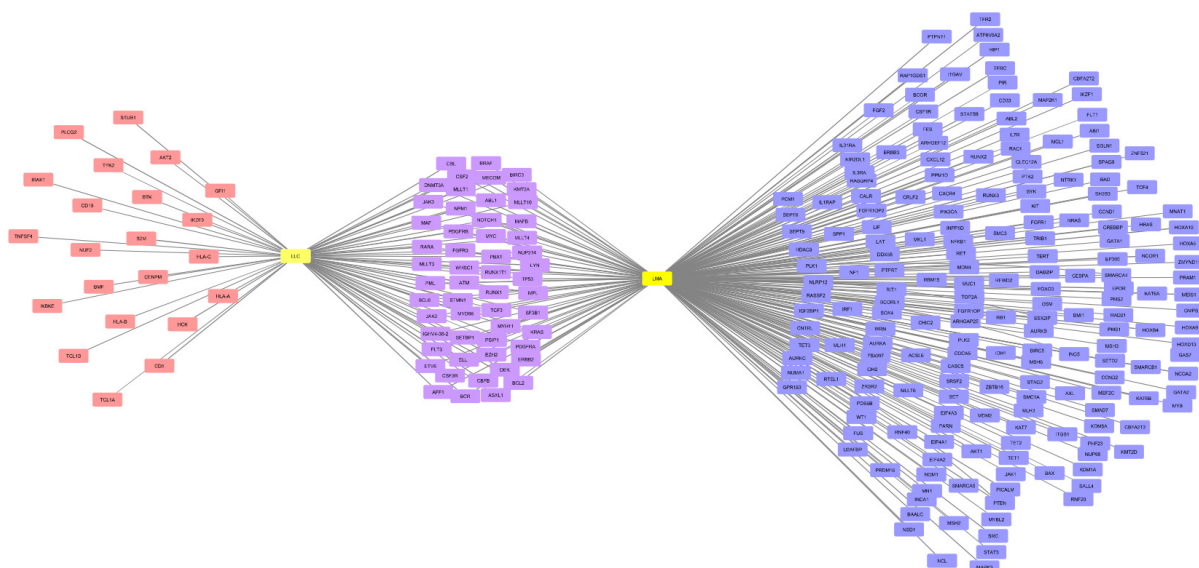
As figuras 6 e 7 permitem visualizar quais proteínas são compartilhadas pela leucemia mieloide aguda e pelas leucemias linfoides aguda e crônica. A LMA compartilha um total de 116 proteínas com a LLA e 55 proteínas com a LLC.

Figura 6. Comparativo das proteínas presentes nas leucemias mieloide aguda e linfóide aguda. Em vermelho estão destacadas as proteínas presentes apenas na rede da LMA, em azul as proteínas presentes apenas na rede da LLA e em rodo as proteínas compartilhadas entre as duas condições.



Fonte: Shannon *et al.*, 2003

Figura 7. Comparativo das proteínas presentes nas leucemias mieloide aguda e linfóide aguda. Em vermelho estão destacadas as proteínas presentes apenas na rede da LMA, em azul as proteínas presentes apenas na rede da LLA e em rodo as proteínas compartilhadas entre as duas condições.



Fonte: Shannon *et al.*, 2003

Com relação às proteínas que não possuem nenhum registro de medicamentos no DGIdb (NCOR1 e TCF3), apenas a proteína TCF3 é compartilhada entre a LMA e as outras duas doenças. Essa proteína já é bem

estudada em casos de LLA, estando presente em genes fusionados e associada a hipogamaglobulinemia em pacientes com essa doença (LEE *et al.*, 2020; BEN-ALI *et al.*, 2017). O compartilhamento dessa proteína entre as leucemias pode indicar um possível papel que ainda não foi descoberto e compreendido para as outras enfermidades.

5 CONCLUSÃO

Conclui-se que a construção de redes de interação proteína-proteína para a análise de dados moleculares é de relevância para a pesquisa científica, pois permite uma visualização holística da composição celular a ser estudada, favorecendo pesquisas oncológicas e farmacogenômicas, sendo importante para o avanço no combate ao câncer e beneficiando toda a sociedade.

O intuito do trabalho concentra-se na aplicação da teoria de sistemas para o estudo de doenças complexas, a fim de encontrar proteínas centrais relacionadas à leucemia mieloide aguda e corroborar dados presentes na literatura. Com os dados apresentados, foi possível atingir os objetivos propostos e também realizar comparativo com outras leucemias.

A rede construída possui um total de 249 proteínas, sendo 45 de interesse para o trabalho, onde 2 não apresentaram medicamentos registrados no DGIdb. A presença de proteínas sem medicamentos registrados demonstra a complexidade das doenças e como a pesquisa ainda pode avançar.

Este trabalho pode servir de referencial para futuras pesquisas relacionadas a esse tema, mostra-se importante a utilização de ferramentas de análise do Gene ontology e criação de programas computacionais que permitam visualizar a progressão das interações presentes neste trabalho.

REFERÊNCIAS

- ALBERT, Réka; BARABÁSI, Albert-László. Statistical mechanics of complex networks. **Reviews of modern physics**, v. 74, n. 1, p. 47, 2002.
- ASHBURNER, Michael *et al.* Gene ontology: tool for the unification of biology. **Nature genetics**, v. 25, n. 1, p. 25-29, 2000.
- APWEILER, Rolf *et al.* UniProt: the universal protein knowledgebase. **Nucleic acids research**, v. 32, n. suppl_1, p. D115-D119, 2004.
- BADER, Gary D.; HOGUE, Christopher WV. An automated method for finding molecular complexes in large protein interaction networks. **BMC bioinformatics**, v. 4, n. 1, p. 1-27, 2003.
- BARABASI, Albert-Laszlo; OLTVAI, Zoltan N. Network biology: understanding the cell's functional organization. **Nature reviews genetics**, v. 5, n. 2, p. 101-113, 2004.
- BAYLIN, Stephen B.; JONES, Peter A. A decade of exploring the cancer epigenome—biological and translational implications. **Nature Reviews Cancer**, v. 11, n. 10, p. 726-734, 2011.
- BEN-ALI, Meriem *et al.* Homozygous transcription factor 3 gene (TCF3) mutation is associated with severe hypogammaglobulinemia and B-cell acute lymphoblastic leukemia. **Journal of Allergy and Clinical Immunology**, v. 140, n. 4, p. 1191, 2017.
- BERG, Kristi L. *et al.* History of databases. **International Journal of Management & Information Systems (IJMIS)**, v. 17, n. 1, p. 29-36, 2013.
- BIGGS, Norman; LLOYD, E. Keith; WILSON, Robin J. **Graph Theory, 1736-1936**. Oxford University Press, 1986.
- BINDEA, Gabriela *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. **Bioinformatics**, v. 25, n. 8, p. 1091-1093, 2009.
- BIRD, Adrian. The essentials of DNA methylation. **Cell**, v. 70, p. 5-8, 1992.
- BIRD, Adrian. DNA methylation patterns and epigenetic memory. **Genes & development**, v. 16, n. 1, p. 6-21, 2002.

CARDOSO, Bruno N. **Grafos Eulerianos na Educação Básica**. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2017.

COFFELT, Seth B. *et al.* IL-17-producing $\gamma\delta$ T cells and neutrophils conspire to promote breast cancer metastasis. **Nature**, v. 522, n. 7556, p. 345-348, 2015.

COTTO, Kelsy C. *et al.* DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. **Nucleic acids research**, v. 46, n. D1, p. D1068-D1073, 2018.

CURRAN, Stephanie; MURRAY, Graeme Ian. Matrix metalloproteinases: molecular aspects of their roles in tumour invasion and metastasis. **European journal of cancer**, v. 36, n. 13, p. 1621-1630, 2000.

CHALLENGEN, Grant A. *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. **Nature genetics**, v. 44, n. 1, p. 23-31, 2012.

CHOU, W. C. *et al.* Acute myeloid leukemia bearing t (7; 11)(p15; p15) is a distinct cytogenetic entity with poor outcome and a distinct mutation profile: comparative analysis of 493 adult patients. **Leukemia**, v. 23, n. 7, p. 1303-1310, 2009.

CHU, Edward; SARTORELLI, A. C. Cancer chemotherapy. **Basic & clinical pharmacology**, v. 9, p. 898-930, 2004.

CROCE, Carlo M. Oncogenes and cancer. **New England journal of medicine**, v. 358, n. 5, p. 502-511, 2008.

DARWISH, Noureldien HE *et al.* Acute myeloid leukemia stem cell markers in prognosis and targeted therapy: potential impact of BMI-1, TIM-3 and CLL-1. **Oncotarget**, v. 7, n. 36, p. 57811, 2016.

DELANGE, Robert J.; SMITH, Emil L. Histones: structure and function. **Annual review of biochemistry**, v. 40, n. 1, p. 279-314, 1971.

DUDA, Dan G. *et al.* Malignant cells facilitate lung metastasis by bringing their own soil. **Proceedings of the National Academy of Sciences**, v. 107, n. 50, p. 21677-21682, 2010.

EULER, Leonhard. Solutio problematis ad geometriam situs pertinentis. **Commentarii academiae scientiarum Petropolitanae**, p. 128-140, 1741.

EGEBLAD, Mikala; WERB, Zena. New functions for the matrix metalloproteinases in cancer progression. **Nature reviews cancer**, v. 2, n. 3, p. 161-174, 2002.

EMILY, H.-YA Cheng *et al.* BCL-2, BCL-XL sequester BH3 domain-only molecules preventing BAX-and BAK-mediated mitochondrial apoptosis. **Molecular cell**, v. 8, n. 3, p. 705-711, 2001.

FANTL, Wendy J.; JOHNSON, Daniel E.; WILLIAMS, Lewis T. Signalling by receptor tyrosine kinases. **Annual review of biochemistry**, v. 62, n. 1, p. 453-481, 1993.

FARES, Jawad *et al.* Molecular principles of metastasis: a hallmark of cancer revisited. **Signal transduction and targeted therapy**, v. 5, n. 1, p. 1-17, 2020.

FARKONA, Sofia; DIAMANDIS, Eleftherios P.; BLASUTIG, Ivan M. Cancer immunotherapy: the beginning of the end of cancer?. **BMC medicine**, v. 14, n. 1, p. 1-18, 2016.

FOULDS, Leslie. The natural history of cancer. **Journal of chronic diseases**, v. 8, n. 1, p. 2-37, 1958.

FORSYTHE, Jo A. *et al.* Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. **Molecular and cellular biology**, v. 16, n. 9, p. 4604-4613, 1996.

FLORES, Mauricio *et al.* P4 medicine: how systems medicine will transform the healthcare sector and society. **Personalized medicine**, v. 10, n. 6, p. 565-576, 2013.

GIANCOTTI, Filippo G. Mechanisms governing metastatic dormancy and reactivation. **Cell**, v. 155, n. 4, p. 750-764, 2013.

GREEN, Douglas R.; KROEMER, Guido. Cytoplasmic functions of the tumour suppressor p53. **Nature**, v. 458, n. 7242, p. 1127-1130, 2009.

GREEN, Douglas R.; REED, John C. Mitochondria and apoptosis. **science**, p. 1309-1312, 1998.

GRIMWADE, David *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom

Medical Research Council trials. **Blood, The Journal of the American Society of Hematology**, v. 116, n. 3, p. 354-365, 2010.

GRIVENNIKOV, Sergei I.; GRETEN, Florian R.; KARIN, Michael. Immunity, inflammation, and cancer. **Cell**, v. 140, n. 6, p. 883-899, 2010.

GRIVENNIKOV, Sergei I.; KARIN, Michael. Inflammation and oncogenesis: a vicious connection. **Current opinion in genetics & development**, v. 20, n. 1, p. 65-71, 2010.

HANAHAN, Douglas; WEINBERG, Robert A. Hallmarks of cancer: the next generation. **cell**, v. 144, n. 5, p. 646-674, 2011.

HARTWELL, Leland H. *et al.* From molecular to modular cell biology. **Nature**, v. 402, n. 6761, p. C47-C52, 1999.

HOFKIRCHNER, Wolfgang; SCHAFRANEK, Matthias. General system theory. In: **Philosophy of complex systems**. North-Holland, 2011. p. 177-194.

HOOD, Leroy; FLORES, Mauricio. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. **New biotechnology**, v. 29, n. 6, p. 613-624, 2012.

IKUSHIMA, Hiroaki; MIYAZONO, Kohei. TGF β signalling: a complex web in cancer progression. **Nature reviews cancer**, v. 10, n. 6, p. 415-424, 2010.

Instituto Nacional do Câncer (INCA). Estimativa de câncer no brasil. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2020-incidencia-de-cancer-no-brasil.pdf>. Acesso em: 29 de ago. 2020

JAIN, Neha *et al.* The My Cancer Genome clinical trial data model and trial curation workflow. **Journal of the American Medical Informatics Association**, v. 27, n. 7, p. 1057-1066, 2020.

JENUWEIN, Thomas; ALLIS, C. David. Translating the histone code. **Science**, v. 293, n. 5532, p. 1074-1080, 2001.

KAPLAN, Rosandra N. *et al.* VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. **Nature**, v. 438, n. 7069, p. 820-827, 2005.

KATSUNO, Yoko; LAMOUILLE, Samy; DERYNCK, Rik. TGF- β signaling and epithelial–mesenchymal transition in cancer progression. **Current opinion in oncology**, v. 25, n. 1, p. 76-84, 2013.

KATZUNG, Bertram G.; TREVOR, Anthony J. **Farmacologia Básica e Clínica-13**. McGraw Hill Brasil, 2017.

KESHET, Ilana; LIEMAN-HURWITZ, Judy; CEDAR, Howard. DNA methylation affects the formation of active chromatin. **Cell**, v. 44, n. 4, p. 535-543, 1986.

KITANO, Hiroaki. Perspectives on systems biology. **New Generation Computing**, v. 18, n. 3, p. 199-216, 2000.

KHUON, Satya *et al.* Myosin light chain kinase mediates transcellular intravasation of breast cancer cells through the underlying endothelial cells: a three-dimensional FRET study. **Journal of cell science**, v. 123, n. 3, p. 431-440, 2010.

KVINLAUG, Brynn T. *et al.* Common and overlapping oncogenic pathways contribute to the evolution of acute myeloid leukemias. **Cancer research**, v. 71, n. 12, p. 4117-4129, 2011.

LÄUBLI, Heinz; BORSIG, Lubor. Selectins promote tumor metastasis. In: **Seminars in cancer biology**. Academic Press, 2010. p. 169-177.

LAMOUILLE, Samy; XU, Jian; DERYNCK, Rik. Molecular mechanisms of epithelial–mesenchymal transition. **Nature reviews Molecular cell biology**, v. 15, n. 3, p. 178-196, 2014.

LEE, Shawn H. R. *et al.* Genome-wide association study of susceptibility loci for TCF3-PBX1 acute lymphoblastic leukemia in children. **JNCI: Journal of the National Cancer Institute**, v. 113, n. 7, p. 933-937, 2021.

LIAO, Debbie; JOHNSON, Randall S. Hypoxia: a key regulator of angiogenesis in cancer. **Cancer and Metastasis Reviews**, v. 26, n. 2, p. 281-290, 2007.

LIU, Guimei; WONG, Limsoon; CHUA, Hon Nian. Complex discovery from weighted PPI networks. **Bioinformatics**, v. 25, n. 15, p. 1891-1897, 2009.

LÖWENBERG, Bob *et al.* Phase 1/2 study to assess the safety, efficacy, and pharmacokinetics of barasertib (AZD1152) in patients with advanced acute myeloid

leukemia. **Blood, The Journal of the American Society of Hematology**, v. 118, n. 23, p. 6030-6036, 2011.

LUGER, Karolin *et al.* Crystal structure of the nucleosome core particle at 2.8 Å resolution. **Nature**, v. 389, n. 6648, p. 251-260, 1997.

LYRA, Taíse Ferraz; OLIVEIRA, Carla Silva. Um estudo sobre confiabilidade de redes e medidas de centralidade em uma rede de co-autoria. **Pesquisa Operacional para o Desenvolvimento**, v. 3, n. 2, p. 160-172, 2011.

MESAROVIĆ, Mihajlo D. Systems theory and biology—view of a theoretician. In: **Systems theory and biology**. Springer, Berlin, Heidelberg, 1968. p. 59-87.

MILES, Fayth L. *et al.* Stepping out of the flow: capillary extravasation in cancer metastasis. **Clinical & experimental metastasis**, v. 25, n. 4, p. 305-324, 2008.

MOSS, Laurie A. Shuman; JENSEN-TAUBMAN, Sandra; STETLER-STEVENSON, William G. Matrix metalloproteinases: changing roles in tumor progression and metastasis. **The American journal of pathology**, v. 181, n. 6, p. 1895-1899, 2012.

MYERS, Rebecca A. *et al.* Enasidenib: an oral IDH2 inhibitor for the treatment of acute myeloid leukemia. **Journal of the advanced practitioner in oncology**, v. 9, n. 4, p. 435, 2018.

NIH, Types of cancer treatment. Disponível em: <https://www.cancer.gov/about-cancer/treatment/types>. Acesso em: 17 de ago. 2020.

NORTH, Sophie; MOENNER, Michel; BIKFALVI, Andreas. Recent developments in the regulation of the angiogenic switch by cellular stress factors in tumors. **Cancer letters**, v. 218, n. 1, p. 1-14, 2005.

PEINADO, Héctor *et al.* Pre-metastatic niches: organ-specific homes for metastases. **Nature Reviews Cancer**, v. 17, n. 5, p. 302-317, 2017.

PORPORATO, Paolo E. *et al.* A mitochondrial switch promotes tumor metastasis. **Cell reports**, v. 8, n. 3, p. 754-766, 2014.

RAZA, Azra; GALILI, Naomi. The genetic basis of phenotypic heterogeneity in myelodysplastic syndromes. **Nature Reviews Cancer**, v. 12, n. 12, p. 849-859, 2012.

REYMOND, Nicolas; D'AGUA, Barbara Borda; RIDLEY, Anne J. Crossing the endothelial barrier during metastasis. **Nature Reviews Cancer**, v. 13, n. 12, p. 858-870, 2013.

RODEN, Dan M. *et al.* Pharmacogenomics: challenges and opportunities. **Annals of internal medicine**, v. 145, n. 10, p. 749-757, 2006.

ROSES, Allen D. Pharmacogenetics and the practice of medicine. **Nature**, v. 405, n. 6788, p. 857-865, 2000.

RUAN, Wen-Jing; LAI, Mao-De. Autocrine stimulation in colorectal carcinoma (CRC). **Medical Oncology**, v. 21, n. 1, p. 1-7, 2004.

SAULTZ, Jennifer N.; GARZON, Ramiro. Acute myeloid leukemia: a concise review. **Journal of clinical medicine**, v. 5, n. 3, p. 33, 2016.

SAMI, Shaheedul A. *et al.* Current and future molecular targets for acute myeloid leukemia therapy. **Current treatment options in oncology**, v. 21, n. 1, p. 1-16, 2020.

SIEGEL, Rebecca. L; MILLER, Kimberly. D; JEMAL, A. Cancer statistics, 2020. **CA: a cancer journal for clinicians**. v. 70, n. 1, p. 7–30, 2020.

SONCINI, Chiara *et al.* PHA-680632, a novel Aurora kinase inhibitor with potent antitumoral activity. **Clinical Cancer Research**, v. 12, n. 13, p. 4080-4089, 2006.

SULLIVAN, Lucas B.; CHANDEL, Navdeep S. Mitochondrial reactive oxygen species and cancer. **Cancer & metabolism**, v. 2, n. 1, p. 1-12, 2014.

SHANNON, Paul *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, v. 13, n. 11, p. 2498-2504, 2003.

SHARMA, Shikhar; KELLY, Theresa K.; JONES, Peter A. Epigenetics in cancer. **Carcinogenesis**, v. 31, n. 1, p. 27-36, 2010.

SHIH, Alan H. *et al.* The role of mutations in epigenetic regulators in myeloid malignancies. **Nature Reviews Cancer**, v. 12, n. 9, p. 599-612, 2012.

SHLUSH, Liran I. *et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. **Nature**, v. 506, n. 7488, p. 328-333, 2014.

SPARMANN, Anke; BAR-SAGI, Dafna. Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. **Cancer cell**, v. 6, n. 5, p. 447-458, 2004.

STRACHAN, Tom *et al.* Human molecular genetics. **CHROMOSOME RESEARCH**, v. 4, p. 475-475, 1996.

STELZER, Gil *et al.* The GeneCards suite: from gene data mining to disease genome sequence analyses. **Current protocols in bioinformatics**, v. 54, n. 1, p. 1.30. 1-1.30. 33, 2016

SZCZERBA, Barbara Maria *et al.* Neutrophils escort circulating tumour cells to enable cell cycle progression. **Nature**, v. 566, n. 7745, p. 553-557, 2019.

SZKLARCZYK, Damian *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. **Nucleic acids research**, v. 47, n. D1, p. D607-D613, 2019.

TAKAI, Daiya; JONES, Peter A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. **Proceedings of the national academy of sciences**, v. 99, n. 6, p. 3740-3745, 2002.

TEFFERI, A. *et al.* IDH mutations in primary myelofibrosis predict leukemic transformation and shortened survival: clinical evidence for leukemogenic collaboration with JAK2 V617F. **Leukemia**, v. 26, n. 3, p. 475-480, 2012.

TIWARI, Neha *et al.* EMT as the ultimate survival mechanism of cancer cells. In: **Seminars in cancer biology**. Academic Press, 2012. p. 194-207.

The Gene Ontology resource: enriching a GOld mine. **Nucleic Acids Research**, v. 49, n. D1, p. D325-D334, 2021.

VARDIMAN, James W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. **Blood, The Journal of the American Society of Hematology**, v. 114, n. 5, p. 937-951, 2009.

VELLA, Danila *et al.* MTGO: PPI network analysis via topological and functional module identification. **Scientific reports**, v. 8, n. 1, p. 1-13, 2018.

WANG, Xiaojing *et al.* Preclinical activity of a novel multiple tyrosine kinase and aurora kinase inhibitor, ENMD-2076, against multiple myeloma. **British journal of haematology**, v. 150, n. 3, p. 313-325, 2010.

WEINBERG, Robert A. Tumor suppressor genes. **Science**, v. 254, n. 5035, p. 1138-1146, 1991.

WEST, Douglas Brent *et al.* **Introduction to graph theory**. Upper Saddle River: Prentice hall, 2001.

WU, Mei; LI, Chuntuan; ZHU, Xiongpeng. FLT3 inhibitors in acute myeloid leukemia. **Journal of hematology & oncology**, v. 11, n. 1, p. 1-11, 2018.

WYCKOFF, Jeffrey B. *et al.* Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. **Cancer research**, v. 67, n. 6, p. 2649-2656, 2007.

YANG, Ying *et al.* Molecular dynamics and free energy studies on Aurora kinase A and its mutant bound with MLN8054: insight into molecular mechanism of subtype selectivity. **Molecular BioSystems**, v. 8, n. 11, p. 3049-3060, 2012.

ZILFOU, Jack T.; LOWE, Scott W. Tumor suppressive functions of p53. **Cold Spring Harbor perspectives in biology**, v. 1, n. 5, p. a001883, 2009.