UNIVERSIDADE FEDERAL DE PELOTAS

Centro de Desenvolvimento Tecnológico (CDTec)

Curso de Graduação em Biotecnologia

Trabalho de Conclusão de Curso

O câncer e a bioinformática: Ferramentas utilizadas na descoberta de novos fármacos para o tratamento.

João Carlos Rodrigues Junior

Pelotas, 2021

João Carlos Rodrigues Junior

O câncer e a bioinformática: Ferramentas utilizadas na descoberta de novos fármacos para o tratamento.

Trabalho acadêmico apresentado ao Curso de Bacharelado em Biotecnologia da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Bacharel em Biotecnologia.

Orientadora: Profa. Dra. Fabiana Kömmling Seixas

Pelotas, 2021.

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

R696c Rodrigues Junior, João Carlos
O câncer e a bioinformática : ferramentas utilizadas na descoberta de novos fármacos para o tratamento / João Carlos Rodrigues Junior ; Fabiana Kömmling Seixas, orientadora. — Pelotas, 2021.
105 f. : il.

Trabalho de Conclusão de Curso (Bacharelado em Biotecnologia) — Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2021.

1. Oncologia. 2. Docking molecular. 3. *In silico*. 4. Dinâmica molecular. 5. Banco de dados biológico. I. Seixas, Fabiana Kömmling, orient. II. Título.

CDD : 616.992

Elaborada por Ubirajara Buddin Cruz CRB: 10/901

João Carlos Rodrigues Junior

O câncer e a bioinformática: Ferramentas utilizadas na descoberta de novos fármacos para o tratamento.

Trabalho acadêmico apresentado ao Curso de Bacharelado em Biotecnologia da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Bacharel em Biotecnologia.

Data da Defesa: 18/06/2021

Banca examinadora:

_____

Profa. Dra. Fabiana Kömmling Seixas (Orientadora)

Doutora em Biotecnologia pela Universidade Federal de Pelotas

_____

Prof. Dr. Frederico Schmitt Kremer

Doutor em Biotecnologia pela Universidade Federal de Pelotas

_____

Dra. Gabriela Klein Couto

Doutora em Bioquímica e Bioprospecção pela Universidade Federal de Pelotas

_____

Dra. Martha Lucia Ruiz Benitez

Doutora em Biotecnologia pela Universidade Federal de Pelotas

Dedico este trabalho à minha família, minha namorada, amigos, professores, colegas de laboratório e a todos que de alguma forma se tornaram parte desta história. Muito obrigado pela companhia durante estes anos de graduação e estágio!

Agradecimentos

Primeiramente agradeço o grande apoio de meus pais, João e Simone, pois sem seus esforços incondicionais nada disso teria sido possível.

Agradeço a minha namorada Darling e família, que sempre me incentivaram e ajudaram sempre que possível, me incentivando a olhar para frente e sempre ser uma pessoa melhor.

A meus orientadores Fabiana Seixas e Tiago Collares, que sempre auxiliariam na minha carreira acadêmica e pessoal durante todos esses anos.

A meus colegas de laboratório que sempre compartilharam seus conhecimentos, motivações e ideias, sendo sempre professores dentro e fora do ambiente de trabalho.

A meus amigos e colegas que sempre estiveram presentes sendo na sala de aula ou fora dela, compartilhando momentos e lembranças que irei carregar para o resto da vida.

A todos os professores e servidores do curso de Biotecnologia, pelo seu empenho em além de trazer conhecimento, fazer do curso o nosso lar.

E a Deus por me conceder saúde e todas as ferramentas necessárias para chegar até este momento.

Muito obrigado!

“O que sabemos é uma gota; o que ignoramos é um oceano”.

Isaac Newton

Resumo

RODRIGUES JUNIOR, João Carlos. **O câncer e a bioinformática: Ferramentas utilizadas na descoberta de novos fármacos para o tratamento.** 105f. 2021. Trabalho de Conclusão de Curso – Curso de bacharelado em Biotecnologia. Universidade Federal de Pelotas, Pelotas.

O câncer é um grupo de doenças que estão relacionadas com o crescimento descontrolado e desorientado de células, e que possuem a capacidade de se espalhar para outros tecidos circundantes através da metástase. O câncer é a segunda maior causa de morte no mundo, em 2020 estimou-se 19,3 milhões de novos casos levando a óbito por volta de 10 milhões de pessoas em todo o mundo. O principal fator de falha no seu tratamento é a resistência a multidrogas, esse problema afeta diretamente os tratamentos com quimioterapia principalmente de pacientes com câncer de mama, ovário, pulmão ou outros tipos de tumores sólidos ou líquidos. A descoberta de novas drogas para seu tratamento é de suma importância, a questão é que os métodos utilizados tradicionalmente como triagem de drogas *in vitro* ou experimentação em animais *in vivo* para a descoberta de novas drogas é caro e laborioso. Devido ao avanço da tecnologia e das ciências ômicas hoje pode se utilizar de ferramentas computacionais para analisar e testar drogas melhorando sua eficiência, descobrindo o alvo do medicamento de forma mais rápida, caracterizando os efeitos colaterais e prevendo se pode ocorrer resistência a droga utilizada. O objetivo do presente trabalho foi elaborar uma revisão da literatura visando mostrar os conceitos básicos de câncer assim como mostrar de forma breve algumas ferramentas da bioinformática que podem ajudar o pesquisador no descobrimento de novos medicamentos para o tratamento de câncer. Para a realização desde trabalho foi realizado uma pesquisa em forma de revisão de literatura com intuito de coletar dados e informações cientificas advindas de artigos científicos originais, de revisões e sites governamentais tanto brasileiros quanto estrangeiros.

Palavras-chave: oncologia, *in silico*, banco de dados biológico, docking molecular, dinâmica molecular, ADME.

Abstract

RODRIGUES JUNIOR, João Carlos. **Cancer and bioinformatics: Tools used in the discovery of new drugs treatment**. 105f. 2021.
Term paper – Bachelor of Biotechnology. Federal University of Pelotas, Pelotas.

Cancer is a group of diseases related to the uncontrolled and disoriented growth of cells and that can spread to other surrounding tissues through metastasis. Cancer is the second leading cause of death in the world, in 2020 an estimated 19.3 million new cases leading to death for around 10 million people. The main factor of failure in its treatment is multidrug resistance, this problem directly affects chemotherapy treatments mainly for patients with breast, ovarian, lung, or other types of solid or liquid tumors. The discovery of new drugs for their treatment is of high importance, the problem is that the methods traditionally used as *in vitro* drug screening or *in vivo* animal testing for the discovery of new drugs are expensive and laborious. Thanks to the advancement of technology and the omic sciences today, computer tools can be used to analyze and test drugs, improving their efficiency, discovering the drug's target more quickly, characterizing side effects, and predicting its resistance to the used drugs. The present work aimed to elaborate a literature review aiming to show the basic concepts of cancer and show several tools of bioinformatics that can help the researcher discover new drugs for the treatment of cancer. To carry out this work, a survey was carried out in the form of a literature review in order to collect data and scientific information from original scientific articles, reviews and government websites, both Brazilian and foreign.

Keywords: cancer, bioinformatics, biological database, molecular docking, molecular dynamics, ADME.

LISTA DE FIGURAS

Figura 1. Os 10 hallmarks do câncer 17

Figura 2. Gráfico de Incidência do câncer no Mundo no ano de 2020 26

Figura 3. Gráfico de número de óbitos de câncer no mundo no ano de 2020. 28

Figura 4. Gráfico de Incidência do câncer no Brasil no ano de 2020 30

Figura 5. Gráfico de número de óbitos de câncer no Brasil no ano de 2020 32

Figura 6. Linha do tempo da descoberta de novos fármacos 45

Figura 7. Esquema de docking molecular 70

Figura 8. Demonstração de uma simulação de macromoléculas na água em uma caixa no formato dodecaedro 73

Figura 9. Representação da interface do software PyMOL 76

Figura 10. Representação da interface gráfico do software QSAR Toolbox 77

Figura 11. Exemplo de fluxograma da utilização de ferramentas de bioinformática na pesquisa por novos fármacos 80

LISTA DE TABELAS

Tabela 1. Incidência de casos de câncer no mundo em 2020. 25

Tabela 2. Número de mortes por câncer no ano de 2020..... 27

Tabela 3. Incidência de casos de câncer no Brasil em 2020. 29

Tabela 4. Número de mortes por câncer no Brasil durante o ano de 2020..... 31

Tabela 5. Links de acesso aos bancos de dados. 56

Tabela 6. Programas gratuitos ou com licença acadêmica de alinhamento de sequências. 61

Tabela 7. Programas gratuitos ou com licença acadêmica de montagem de sequências. 62

Tabela 8. Softwares para anotação de genomas estruturais. 64

Tabela 9. Softwares para anotação de genomas funcionais. 65

Tabela 10. Softwares para predição de estrutura de proteínas. 66

Tabela 11. Softwares para predição da função de proteínas..... 67

Tabela 12. Sites e programas gratuitos ou com licença acadêmica de docking molecular. 70

Tabela 13. Softwares gratuitos ou com licença acadêmica de dinâmica molecular e simulações computacionais. 73

Tabela 14. Sites e programas gratuitos ou com licença acadêmica de análise de ADME. 75

LISTA DE UNIDADES, ABREVIAÇÕES E SIGLAS

3D	Tridimensionais
5-FU	5-fluorouracil
6-MP	6-mercaptopurina
ADCD	Citotoxicidade Mediada por Células Dependentes de Anticorpos
ADME	Absorção, Distribuição, Metabolização e Excreção
APC	Células Apresentadoras de Antígenos
AS	Alinhamento de Sequências
ASM	Alinhamento de Sequência Múltipla
ASP	Alinhamento de Sequência de Pares
BCG	Bacillus Calmette– Guerin
BCL-2	Proteína 2 do Linfoma de Células B
BLAST	*Basic Local Alignment Search Tool*
BLAT	*BLAST-Like Alignment Tool*
BMRB	*BioMagResBank*
bp	Pares de Base
CCLE	*Cancer Cell Line Encyclopedia*
CD	Células Dendríticas
CDK	Cinases Dependentes de Ciclinas
cDNA	DNA Complementar
COSMIC	*Catalogue of Somatic Mutations in Cancer*
dbSNP	*Single Nucleotide Polymorphism Database*
DDBJ	*DNA Data Bank of Japan – National Institute of Genetics*
DEVH	Doença Enxerto Versus Hospedeiro
dLN	Linfonodos Drenantes
DM	Dinâmica Molecular
DNA	Ácido Desoxirribonucleico
EGF	Fator de Crescimento Epidérmico
EMBL	*European Bioinformatics Institute*
ERO	Espécie Reativa de Oxigênio
GAD	*The Genetic Association Database*

GDSC	*The Genomics of Drug Sensitivity in Cancer*
GEO	*Gene Expression Omnibus*
GO	*Gene Ontology*
GWAS	*Genome-Wide Association*
HER2	*Human Epidermal Growth Factor Receptor 2*
HGVS	*Human Genome Variation Society*
HMMs	Modelos Ocultos de Markov
IFN-α	Interferon-Alfa
IL-2	Interleucina-2
INSD	*International Nucleotide Sequence Database*
IUPAC	*International Union of Pure and Applied Chemistry*
kb	Kilobases
KEGG	*Kyoto Encyclopedia of Genes and Genomes*
LDL	Lipoproteínas de Baixa Densidade
mAbs	Anticorpos Monoclonais
MHC	Complexo Principal de Histocompatibilidade
MMP	Metaloproteinases da Matriz
mRNA	Ácido Ribonucleico Mensageiro
mtDNA	DNA Mitocondrial
NCBI	*National Center for Biotechnology Information*
NCI	*National Cancer Institute*
NGS	*Next Generation Sequencing*
NMR	Ressonância Magnética Nuclear
OECD QSAR	*The Organisation for Economic Cooperation and Development Quantitative Structure- Activity*
OMS	Organização Mundial da Saúde
PARP	Poli ADP-Ribose Polimerase
PDB	*Protein Data Bank*
PDBe	*Protein Data Bank Europe*
PDBj	*Protein Data Bank Japan*
PDGF	Fator de Crescimento Derivado de Plaquetas
RB	Retinoblastoma
RCSB PDB	RCSB *Protein Data Bank*
RefSeq	*Reference Sequence*

RNA	Ácido Ribonucleico
RO5	Regra de Cinco de Lipinski
siRNA	*Small Interfering* RNA
SNP	*Simple Nucleotide Polymorphism*
TAA	Antígenos Associados a Tumores
TCGA	*The Cancer Genome Atlas*
TDD	*Therapeutic Target Database*
TGF-β	Fator Transformador de Crescimento Beta
TMO	Transplante de Medula Óssea
TPD	Terapia Fotodinâmica
VEGF	Fator de crescimento endotelial vascular
XIAP	Proteína Inibidora da Apoptose Ligada ao X
WGS	*Whole Genome Shotgun*
WHO	*World Health Organization*
wwPDB	*Worldwide Protein Data Bank*

SUMÁRIO

1. INTRODUÇÃO 13
2. OBJETIVOS 14
2.1 Objetivo geral 14
2.2 Objetivos específicos 14
3 METODOLOGIA 15
4 REVISÃO DA LITERATURA 16
4.1 O Câncer 16
4.1.1 Hallmarks do Câncer 16
4.1.1.1 Autossuficiência de sinais estimuladores 17
4.1.1.2 Insensibilidade aos fatores de supressão 18
4.1.1.3 Invasão de tecidos e metástase 18
4.1.1.4 Multiplicação ilimitada 19
4.1.1.5 Angiogênese 20
4.1.1.6 Resistência a morte celular programada 20
4.1.1.7 Instabilidade genômica e mutação 21
4.1.1.8 Inflamação tumoral 22
4.1.1.9 Desregulação da energia celular 23
4.1.1.10 Evasão do sistema imune 23
4.1.2 Epidemiologia do Câncer 24
4.1.3 Terapias para o câncer 33
4.1.3.1 Cirurgia 33
4.1.3.2 Quimioterapia 33
4.1.3.3 Radioterapia 35
4.1.3.4 Transplante de medula óssea 36
4.1.3.5 Terapia Alvo Molecular 36
4.1.3.5.1 Pequenas Moléculas 38

4.1.3.5.2 Vacinas terapêuticas contra o câncer 38
4.1.3.5.3 Terapia gênica 39
4.1.3.5.4 Anticorpos Monoclonais 39
4.1.3.6 Imunoterapia 40
4.1.3.7 Terapia fotodinâmica 41
4.1.4 Importância da descoberta de novos compostos para o tratamento 43
4.2 A Bioinformática 46
4.3 Ferramentas de bioinformática utilizadas na descoberta de novas drogas. 46
4.3.1 Bancos de dados biológicos 48
4.3.1.1 Cancer Cell Line Encyclopedia (CCLE) 48
4.3.1.2 Catalogue of Somatic Mutations in Cancer (COSMIC) 49
4.3.1.3 ChEMBL 49
4.3.1.4 ClinVar 49
4.3.1.5 Single Nucleotide Polymorphism Database (dbSNP) 50
4.3.1.6 DisGeNET 50
4.3.1.7 DrugBank 50
4.3.1.8 Ensembl 51
4.3.1.9 Expression Atlas 51
4.3.1.10 Gene Expression Omnibus (GEO) 51
4.3.1.11 Gene Ontology (GO) 52
4.3.1.12 InterPro 52
4.3.1.13 Kyoto Encyclopedia of Genes and Genomes (KEGG) 52
4.3.1.14 Pfam 53
4.3.1.15 PROSITE 53
4.3.1.16 Protein Data Bank (PDB) 54
4.3.1.17 Reference Sequence (RefSeq) 54
4.3.1.18 The Cancer Genome Atlas (TCGA) 54

4.3.1.19 The Genetic Association Database (GAD) 55
4.3.1.20 The Genomics of Drug Sensitivity in Cancer (GDSC) 55
4.3.1.21 Therapeutic Target Database (TDD) 55
4.3.1.22 UniProt 56
4.3.2 Alinhamento de sequências 60
4.3.3 Montagem de sequências 61
4.3.4 Anotação de genoma 63
4.3.5 Predição de estruturas e funções proteicas 65
4.3.6 Docking Molecular 68
4.3.7 Dinâmica Molecular 71
4.3.8 Absorção, Distribuição, Metabolização e Excreção (ADME) 74
4.3.9 PyMOL 76
4.3.10 OECD QSAR Toolbox 77
4.4 Discussão 78
5 CONCLUSÃO 81
6 REFERÊNCIAS BIBLIOGRÁFICAS 82

1. INTRODUÇÃO

O câncer é um grupo de doenças que estão relacionadas com o crescimento descontrolado e desorientado de células, e que possuem a capacidade de se espalhar para outros tecidos circundantes através da metástase. É considerado uma doença genética, ou seja, acontece quando ocorre alterações nos genes responsáveis por controlar como nossas células funcionam, principalmente o crescimento e divisão celular, também é uma doença multifatorial (NCI, 2015).

O câncer é a segunda maior causa de morte no mundo, em 2020 estimou-se 19,3 milhões de novos casos levando a óbito por volta de 10 milhões de pessoas, a cada seis mortes ocorridas no mundo uma foi de câncer no ano de 2020 (SUNG et al., 2021; WHO, 2021).

O principal fator de falha no tratamento contra o câncer é a resistência a multidrogas, esse problema afeta diretamente os tratamentos com quimioterapia principalmente de pacientes com câncer de mama, ovário, pulmão ou outros tipos de tumores sólidos ou líquidos. A quimioterapia elimina células que são sensíveis às drogas utilizadas no tratamento, mas deixa para trás grandes quantidades de células com resistência ao tratamento, o que aumenta a preocupação caso o tumor cresça novamente pois agora suas células serão resistentes a este tratamento (CMR, 2000).

A descoberta de novas drogas e de suma importância para a terapia contra o câncer, a questão é que os métodos utilizados tradicionalmente como triagem de drogas *in vitro* ou experimentação em animais *in vivo* para a descoberta de novas drogas é caro e laborioso. Com a utilização da bioinformática para a análise de dados de alto rendimento, como dados genômicos, epigenéticos, arquitetura de genoma, transcriptômica, proteômica e dados de perfil ribossômico tiveram uma boa contribuição na descoberta e na realocação de drogas. O desenvolvimento de modelagem por homologia e simulação de estruturas proteicas por exemplo, tiveram grande relevância quando utilizaram-se de grandes banco de dados como os de pequenas moléculas e metabólitos para se obter experimentos mais realísticos de docking molecular na triagem virtual (XIA, 2017). Assim hoje pode se utilizar ferramentas computacionais para analisar e testar drogas melhorando sua eficiência, descobrindo o alvo do medicamento de forma mais rápida, caracterizando os efeitos colaterais e prevendo se pode ocorrer resistência a droga utilizada.

2. OBJETIVOS

2.1 Objetivo geral

O presente trabalho teve o intuito de elaborar uma revisão da literatura visando mostrar os conceitos básicos de câncer como suas características, tratamentos e a importância do estudo desta doença. Assim como mostrar ferramentas e abordagens da bioinformática que podem ajudar no desenvolvimento de pesquisas e na descoberta de novos tratamentos para o câncer.

2.2 Objetivos específicos

I) Descrever sobre os *hallmarks* do câncer e seu impacto no tratamento oncológico;
II) Mostrar a tendência de casos e o número da mortalidade de pessoas que sofrem com esta doença;
III) Sumarizar os principais tipos de tratamento do câncer;
IV) Ressaltar a importância da descoberta de novos medicamentos para o tratamento de câncer;
V) Apresentar as principais ferramentas de bioinformática que podem ser utilizados na descoberta de novos compostos terapêuticos para o câncer;

3 METODOLOGIA

Para a elaboração do presente trabalho foi realizado uma pesquisa em forma de revisão de literatura com intuito de coletar dados e informações científicas de artigos originais, revisão, *sites* governamentais e especializados. Os principais critérios de escolhas de artigos foram a data de publicação, sempre tentando obter a mais recente possível, já a escolha de *sites* se baseou em *sites* governamentais ou portais de noticiais consagrados no ramo da medicina. Para este estudo, foram utilizados banco de dados como *ScienceDirect e PubMed*, bem como *sites* oficias dos *softwares* comentados, *sites* nacionais e internacionais de referência na área da saúde e câncer como *National Cancer Institute* (NCI), *World Health Organization* (WHO), Instituto Nacional do Câncer (INCA), para este trabalho se utilizou cerca de 220 artigos científicos, onde após a leitura e o compreendimento do mesmo os mais relevantes foram utilizados. As principais palavras-chave empregadas foram: *cancer, hallmarks, cancer epidemology, cancer therapy, drug discovery, bioinformatics, biological databank, sequence assembly, genome annotation, sequence alignment, molecular docking, molecular dynamics, adme, pymol, geneness, qsar, qsar toolbox*.

4 REVISÃO DA LITERATURA

4.1 O Câncer

O câncer, segundo o *National Cancer Institute* (NCI), é um grupo de doenças que estão relacionadas com o crescimento descontrolado e desorientado de células, e que possuem a capacidade de se espalhar para outros tecidos circundantes através da metástase. O câncer é considerado uma doença genética, ou seja, acontece quando ocorrem alterações nos genes responsáveis por controlar o funcionamento das células, principalmente o crescimento e divisão celular. Também é uma doença multifatorial, ou seja, pode ser causada por diversos fatores, como por exemplo fatores de risco ambientais, fatores reprodutivos e hormonais entre outros. Esta doença, por ser genética, pode ser herdada, assim como pode ocorrer pelas exposições ambientais que a pessoa passa ao longo da vida, isso inclui substâncias e produtos químicos como fumaça do tabaco e raios ultravioleta emitidos pelo sol (NCI, 2015).

4.1.1 Hallmarks do Câncer

Os *hallmarks* do câncer (marcas registradas do câncer) apontam que existem dez capacidades biológicas diferentes que a célula adquire para que aconteça o desenvolvimento, progressão, e a disseminação dessa doença. Essas capacidades biológicas incluem a autossuficiência de sinais estimuladores, insensibilidade aos fatores de supressão de crescimento, invasão de tecidos e metástases, multiplicação ilimitada, capacidade de angiogênese, resistência a morte celular programada, instabilidade genômica e mutação, inflamação tumoral, desregulação da energia celular e evasão do sistema imune. (BROERTJES, 2015; HANAHAN; WEINBERG, 2011).

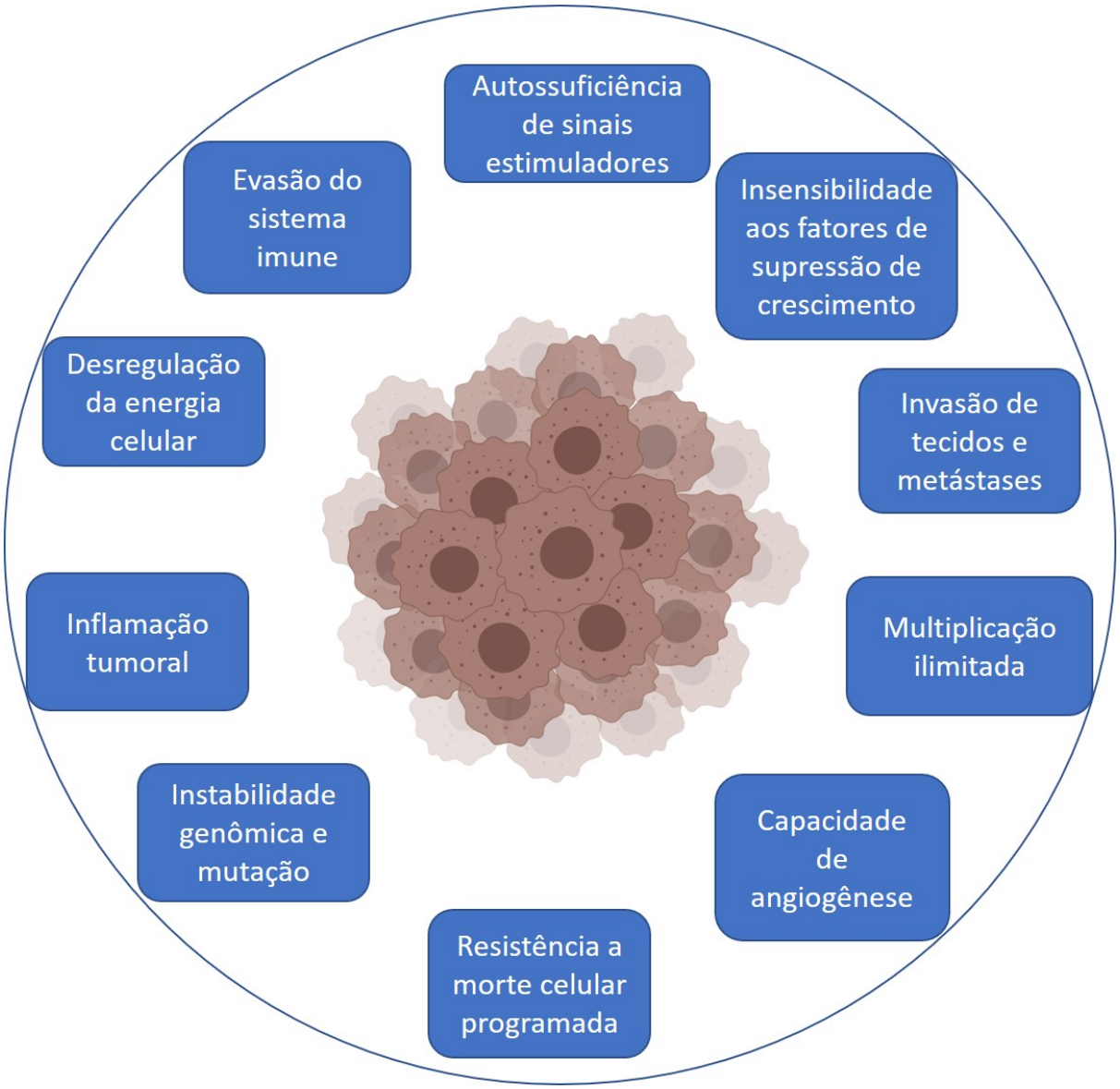

Figura 1. Os 10 *hallmarks* do câncer. **Fonte:** Adaptado de GARCÍA-CABALLERO; MEDINA; QUESADA, 2015

4.1.1.1 Autossuficiência de sinais estimuladores

A autossuficiência de sinais estimuladores é sustentada devido a capacidade de produção de ligantes dos fatores de crescimento, dessa forma a célula começa a expressar receptores cognatos fazendo com que ocorra uma proliferação celular de forma autócrina. Existem vários fatores de crescimento no qual podem ser expressos, por exemplo o Fator de Crescimento Epidérmico (EGF), Fator de crescimento endotelial vascular (VEGF), Fator de Crescimento Derivado de Plaquetas (PDGF) e entre outros.

Uma forma alternativa desse marcador acontecer se dá quando as células cancerosas começam a enviar sinais com intuito de estimular as células saudáveis dentro do estroma associado ao suporte do tumor, essas células estimuladas acabam fornecendo fatores de crescimento para as células cancerosas (BHOWMICK; NEILSON; MOSES, 2008; CHENG *et al.*, 2008; WITSCH; SELA; YARDEN, 2010).

4.1.1.2 Insensibilidade aos fatores de supressão

A insensibilidade aos fatores de supressão acontece devido a capacidade da célula cancerosa de contornar sistemas que regulam negativamente a proliferação celular. A maior parte desses sistemas dependem da transcrição de genes supressores de tumor que codificam para proteínas supressoras tumorais, como a proteína retinoblastoma (RB) e o fator de transcrição p53, que atuam no controle da proliferação, senescência e apoptose celular (HANAHAN; WEINBERG, 2011). A proteína RB possui diversos sinais tanto de fontes extracelulares quanto intracelulares, dessa forma tem o poder de decidir se a célula vai ou não prosseguir para o ciclo de divisão celular. As células cancerosas que possuem defeitos na via RB não conseguem distinguir células sadias de cancerosas fazendo com que o seu crescimento seja persistente (BURKHART; SAGE, 2008; DESHPANDE; SICINSKI; HINDS, 2005). O fator de transcrição p53 é induzido quando ocorre algum tipo de dano no DNA (ácido desoxirribonucleico), hipóxia ou ativação de oncogene, e possui a capacidade de regular o processo de expressão gênica no qual pode parar o ciclo celular ou induzir a morte celular programada caso alguma anormalidade no DNA da célula seja encontrada, dessa forma caso a via desse fator de transcrição estiver com algum problema ele não conseguirá mais induzir a apoptose em células defeituosas (ARNOLD J. LEVINE, 1997; SHERR; MCCORMICK, 2002).

4.1.1.3 Invasão de tecidos e metástase

A invasão de tecidos e metástase são processos nos quais as células cancerosas se disseminam a partir de uma lesão primária aos órgãos distais, sendo essa a principal causa da alta letalidade do câncer. Essa disseminação de células cancerosas do tumor primário para os outros órgãos do corpo é bem complexa, e envolve uma variedade de

mecanismos celulares, como por exemplo, a esquiva e modulação do microambiente tecidual, desenvolvimento de resistência ao tratamento terapêutico, evasão da vigilância imunológica e a inibição de processos antitumorigênicos (SUHAIL *et al.*, 2019). O câncer utiliza duas vias principais para sua disseminação, a via linfática a qual leva à invasão dos linfonodos que drenam os órgãos por onde evolui o tumor, e a via do sangue que leva à invasão de órgãos não tão próximos do tumor primário, como fígado, cérebro, ossos e pulmão. De todas as vias de sinalização envolvidas no processo de metástase, pode-se dizer que a via da integrina, do fator transformador de crescimento beta (TGF-β), das quimiocinas, e do receptor de dependência, estão intimamente relacionados com o direcionamento terapêutico, pois pode utilizar-se de anticorpos terapêuticos ou pequenas moléculas que inibam as cinases envolvidas nessas vias de sinalização (ROBERT, 2013).

4.1.1.4 Multiplicação ilimitada

A multiplicação ilimitada é uma capacidade bem marcante em células cancerosas pois a maioria das linhagens celulares normais do corpo são capazes de passar apenas por um número limitado de ciclos sucessivos de crescimento e divisão celular, isso graças a duas barreiras que estão intimamente associadas à proliferação. A primeira barreira é a senescência, isto é, a entrada tradicionalmente irreversível a um estado não proliferativo, mas viável da célula; a segunda barreira é o estado de crise, o qual envolve a morte celular. Quando as células são propagadas em vários ciclos repetitivos na cultura de tecidos, primeiramente as células entram no estado de senescência e para que consigam contornar essa barreira acontece uma fase de crise, a qual grande parte das células morrem; em raras ocasiões, as células emergem de uma população em crise com potencial replicativo ilimitado, essa transição é chamada de imortalização (HANAHAN; WEINBERG, 2011).

Evidências demonstram que os telômeros que protegem as extremidades dos cromossomos estão envolvidos na capacidade de proliferação ilimitada. Essas estruturas são compostas de várias repetições de hexanucleotídeos em tandem, que são reduzidos progressivamente em células normais quando propagadas em culturas, e assim perdem a capacidade de proteger as extremidades do DNA cromossômico de possíveis fusões de ponta a ponta. As fusões geram cromossomos dicêntricos instáveis, que resulta num

embaralhamento do cariótipo que ameaça a viabilidade celular. Dessa forma o tamanho do DNA telomérico de uma célula é o responsável por quantos ciclos essa célula pode passar antes que os telômeros sejam completamente perdidos e percam sua função protetora, desencadeando sua entrada em crise (SOBINOFF; PICKETT, 2017).

A senescência e a crise/apoptose são defesas cruciais anticâncer, pois impedem o crescimento de clones de células pré neoplásicas e neoplásicas, a maioria das neoplasias exaure sua cota replicações e é interrompida por uma dessas barreiras, já as células que avançam para a fase de tumores possuem a capacidade de manter seu DNA telomérico em comprimentos suficientes para evitar o desencadeamento de senescência ou apoptose (HANAHAN; WEINBERG, 2011).

4.1.1.5 Angiogênese

A capacidade de angiogênese é um processo biológico no qual o tumor induz o crescimento de novos vasos capilares sanguíneos a partir da vasculatura pré-existente assim fornecendo oxigênio e outros tipos de nutrientes aos tecidos. Esse processo é complexo e envolve inúmeras interações entre vários componentes biológicos, como vários tipos celulares, fatores angiogênicos solúveis e componentes da matriz extracelular (LI *et al.*, 2018a).

O processo de angiogênese acontece baseando-se em quatros processos distintos, começando pela ação de enzimas proteolíticas como as metaloproteinases da matriz (MMP), que degradam as glicoproteínas da membrana basal junto de componentes da matriz extracelular que envolvem os vasos sanguíneos. Após ocorre a ativação e migração de células endoteliais, seguido pela proliferação dessas células, por último as células endoteliais formam estruturas tubulares que passam a se transformar em vasos capitares no qual acabam desenvolvendo membranas basais, assim nutrindo e auxiliando a proliferação do câncer (FAN; JAGGER; BICKNELL, 1995; LI *et al.*, 2018a).

4.1.1.6 Resistência a morte celular programada

A resistência a morte celular programada acontece devido as células cancerosas estarem em constante estresse, como o estresse oncogênico, a instabilidade genômica e a hipóxia celular, como resposta a esses estímulos apoptóticos internos a via intrínseca

do apoptose deve ser ativada, porém, essas células podem evitar a resposta celular desativando as vias apoptóticas, uma dessas formas acontece através da expressão da Proteína Inibidora da Apoptose Ligada ao X (XIAP) no qual previne a apoptose celular principalmente pelo bloqueio da ativação e maturação da caspase-3 / -7 / -9, que são os principais iniciadores e efetores da apoptose (FERNAND; MANABU, 2013; TU; COSTA, 2020).

Modelos genéticos de camundongos já demonstraram que a inativação genética da proteína BH3 ou de uma caspase como a caspase-2 pode levar a resistência a certos estímulos pró-apoptóticos e além disso também pode acelerar a tumorigênese em estudos utilizando ratos como modelos animais (HÜBNER *et al.*, 2008; PARSONS *et al.*, 2013).

Também foi constatado que a expressão forçada de proteínas da família BCL-2 anti apoptóticas pode aumentar o desenvolvimento do tecido tumoral quando induzido por um oncogene como o MYC, uma vez que a expressão de uma proteína anti-apoptótica da família BCL-2 isolada não consegue resultar na formação de um tumor, como no caso de um oncogene sozinho (BRUNELLE *et al.*, 2009; ZHOU *et al.*, 1998).

Levando isso em consideração, sabe-se que a inibição da apoptose tem um papel crítico na sobrevivência das células cancerosas e no desenvolvimento de tumores, devido a sua capacidade de modular as vias apoptóticas de forma transcricional, translacional e pós-traducional. Assim, graças a essa capacidade as células cancerosas conseguem escapar da apoptose diminuindo ou aumentando a expressão de genes anti ou pró apoptóticos; também podem inibir a apoptose através da estabilização, desestabilização, ou alterando funções por meio de modificações pós traducionais de proteínas anti ou pro apoptóticas (FERNAND; MANABU, 2013).

4.1.1.7 Instabilidade genômica e mutação

Um dos fatores para o desenvolvimento de tumores malignos é a sua instabilidade genômica, que é a capacidade de facilitar com que o genoma celular sofra mutações devido a alguma falha nos mecanismos de reparo do DNA. Graças a ela além de aumentar a taxa de mutações também é responsável pelo desenvolvimento de outros *hallmarks*, aumentando a progressão do tumor (BROERTJES, 2015).

A instabilidade genômica do câncer colabora com o fenômeno da heterogeneidade genética intratumoral, pois fornece diversidade genética para a seleção gênica e ainda possibilita uma extensa diversidade fenotípica que é frequentemente observada entre pacientes. Antigamente a instabilidade genômica era associada a um mal prognóstico, no entanto estudos evidenciaram que tumores sólidos de origem epitelial, possuem níveis extremos de instabilidade genômica, onde mais de 75% do genoma está sujeito a alterações somáticas no número de cópias, e por isso acabam tendo um prognóstico potencialmente melhor em comparação com níveis intermediários abaixo deste limite (ANDOR; MALEY; JI, 2017).

4.1.1.8 Inflamação tumoral

A inflamação é um processo que envolve ativação, recrutamento e ação de células de imunidade, tanto inata quanto adaptativa, e é destacada por seu papel essencial na defesa do hospedeiro contra qualquer tipo de patógenos; também é importante para a reparação, regeneração e remodelação de tecidos (GRETEN; GRIVENNIKOV, 2019; MEDZHITOV, 2008).

A inflamação tumoral predispõe o desenvolvimento do câncer e auxilia todos os estágios da tumorigênese. As células cancerosas, assim como as células inflamatórias circundantes e estromais, se envolvem em interações recíprocas bem administradas para que seja formado um microambiente tumoral inflamatório (GRETEN; GRIVENNIKOV, 2019).

A inflamação crônica desempenha um papel fundamental na tumorigênese, atuando em diferentes funções acompanhadas do avanço do tumor. Gradualmente tem se ilustrado que o microambiente tumoral é construído não apenas por células tumorais, mas também por células do estroma, assim como células inflamatórias, fibroblastos e células endoteliais (MA *et al.*, 2020). Algumas células inflamatórias como macrófagos associados a tumor, auxiliam a hiperplasia de células tumorais pois secretam fatores de crescimento, ajudam na angiogênese por meio da secreção de fator de crescimento endotelial vascular , promovem metástase tumoral por secreção de metaloproteinases de matriz e suprimem a imunidade adaptativa pela secreção de várias citocinas imunossupressoras (MA *et al.*, 2020).

4.1.1.9 Desregulação da energia celular

As células tumorais, diferentemente das células normais, possuem uma atividade glicolítica aumentada e respiração mitocondrial reduzida, mesmo na presença de oxigênio. A energia celular desregulada possui mecanismos subjacentes que estão associados à disfunção mitocondrial causada por mutações do DNA mitocondrial (mtDNA), defeitos em enzimas mitocondriais ou supressores de tumores/oncogenes alterados. Esses que, utilizando da reprogramação metabólica das células cancerosas, desencadeiam mudanças promotoras de tumor mediadas por espécies reativas de oxigênio e prótons de cálcio (Ca^{2+}), carbonos ou outros metabólitos de pequenas moléculas liberadas pela mitocôndria, fazendo com que ocorra a desregulação da energia celular (HSU; TSENG; LEE, 2016).

O mtDNA dos humanos é uma molécula de DNA circular de fita dupla de 16,6kb (kilobases) o qual codifica 13 polipeptídios dos complexos de enzimas respiratórias, 22 RNAs (ácido ribonucleico) de transferência e 2 RNAs ribossômicos necessários para a síntese de proteínas mitocondriais. Como o mtDNA é peça fundamental na manutenção de mitocôndrias funcionalmente competentes, espera-se que o acúmulo de mutações ou números reduzidos de cópias do mtDNA afete a produção de energia e aumente a geração de espécies reativas de oxigênio (ERO) e a sobrevivência celular, esses processos podem estar envolvidos no câncer, doenças mitocôndrias ou envelhecimento (HSU; TSENG; LEE, 2016).

4.1.1.10 Evasão do sistema imune

A evasão do sistema imunológico do câncer é um grande obstáculo na elaboração de estratégias terapêuticas anticâncer de sucesso. Mesmo se obtendo um enorme progresso no entendimento de como o câncer evita a imunidade destrutiva, as maneiras de neutralizar ou evitar que o tumor fuja não acompanharam o ritmo, pois existem vários fatores que contribuem para a persistência do tumor, isso independente de um hospedeiro com sistema imune normal ou não. O principal aspecto pelo qual os tumores escapam da vigilância do sistema imune é a imunoedição, pois esta faz com que tumores fiquem latentes em pacientes por anos durante o “equilíbrio” e “senescência” antes de ressurgir. Além disso, os tumores exploram outros diversos processos imunológicos,

como apresentação de antígenos, modificação da produção de mediadores imunossupressores, tolerância, desvio imunológico e direcionamento da função das células T regulatórias ou de suas secreções (VINAY *et al.*, 2015).

A imunoedição normalmente ocorre em três fases sequenciais: eliminação, equilíbrio e escape. A fase de eliminação é um conceito moderno da antiga imunovigilância do câncer, no qual destaca que a imunidade inata e a adaptativa trabalham em conjunto para detectar e destruir células transformadas bem antes de se tornarem clinicamente visíveis. Porém, em alguns casos, as variantes das células tumorais não são completamente eliminadas, podendo entrar em fase de equilíbrio no qual o sistema imunológico controla o crescimento dessas células; nesta fase, a imunidade adaptativa delimita o crescimento das células tumorais clinicamente indetectáveis e edita a imunogenicidade das células tumorais. Por último, o estado de dormência funcional das células tumorais pode ser finalizada, levando à progressão das células para a fase de escape, essa na qual tumores editados para imunogenicidade reduzida começam a crescer progressivamente de maneira irrestrita, estabelecendo um microambiente tumoral imunosupressor e o tumor torna-se clinicamente aparente (VESELY; SCHREIBER, 2013).

4.1.2 Epidemiologia do Câncer

A epidemiologia é uma ciência que tem evoluído constantemente devido as mudanças da sociedade e ao surgimento de novas doenças. Essa evolução fez com que a epidemiologia se tornasse uma ferramenta extremamente relevante para compreender doenças e eventos ocorridos na saúde (FRÉROT *et al.*, 2018). Pode ser definida como a ciência que descreve e explica padrões de doenças nas populações, e utiliza deste conhecimento para prevenir, controlar doenças e melhorar a saúde (BROPHAL, 2016).

O câncer é a segunda maior causa de mortes no mundo. Em escala global, no ano de 2020, a cada seis óbitos ocorridos, um foi de câncer; no mesmo ano estimou-se 19,3 milhões de novos casos levando a óbito por volta de 10 milhões de pessoas (SUNG *et al.*, 2021; WHO, 2021). A grande taxa de mortalidade do câncer nos países que estão em desenvolvimento deve aumentar as tendências como envelhecimento da população e mudanças socioeconômicas. Além do impacto na saúde pública, o câncer também impõe custos econômicos para os acometidos e para a sociedade, como a perda de

produtividade dessas pessoas que faz com que a sociedade perca a contribuição desses indivíduos para a economia de mercado devido a morte prematura por câncer (PEARCE *et al.*, 2018).

As tabelas a seguir dispõem de informações fornecidas pela Organização Mundial da Saúde (OMS; tradução livre de *World Health Organization*, WHO) as quais mostraram a incidência (Tabela 1), o número de mortes (Tabela 2) e a porcentagem em relação a todos os tipos de cânceres no ano de 2020, as tabelas estão dispostas levando em conta os dez tipos de cânceres mais incidentes e com maior número de mortes no mundo, desconsiderando o sexo ou a idade; a porcentagem está relacionada ao total de casos de cânceres sem distinção de tipo.

Tabela 1. Incidência de casos de câncer no mundo em 2020.

Tipo de Câncer	Incidência (número de casos)	Incidência em %
Mama	2.264.419	11,7%
Pulmão	2.206.771	11,4%
Colorretal	1.931.590	10,0%
Próstata	1.414.259	7,3%
Estomago	1.089.103	5,6%
Fígado	905.677	4,7%
Cérvix Uterina	604.127	3,1%
Esôfago	604.100	3,1%
Tireoide	586.202	3,0%
Bexiga	573.278	3,0%
Outros tipos de cânceres	7.116.263	36,9%

Fonte: Elaborado pelo autor com dados da OMS, 2021.

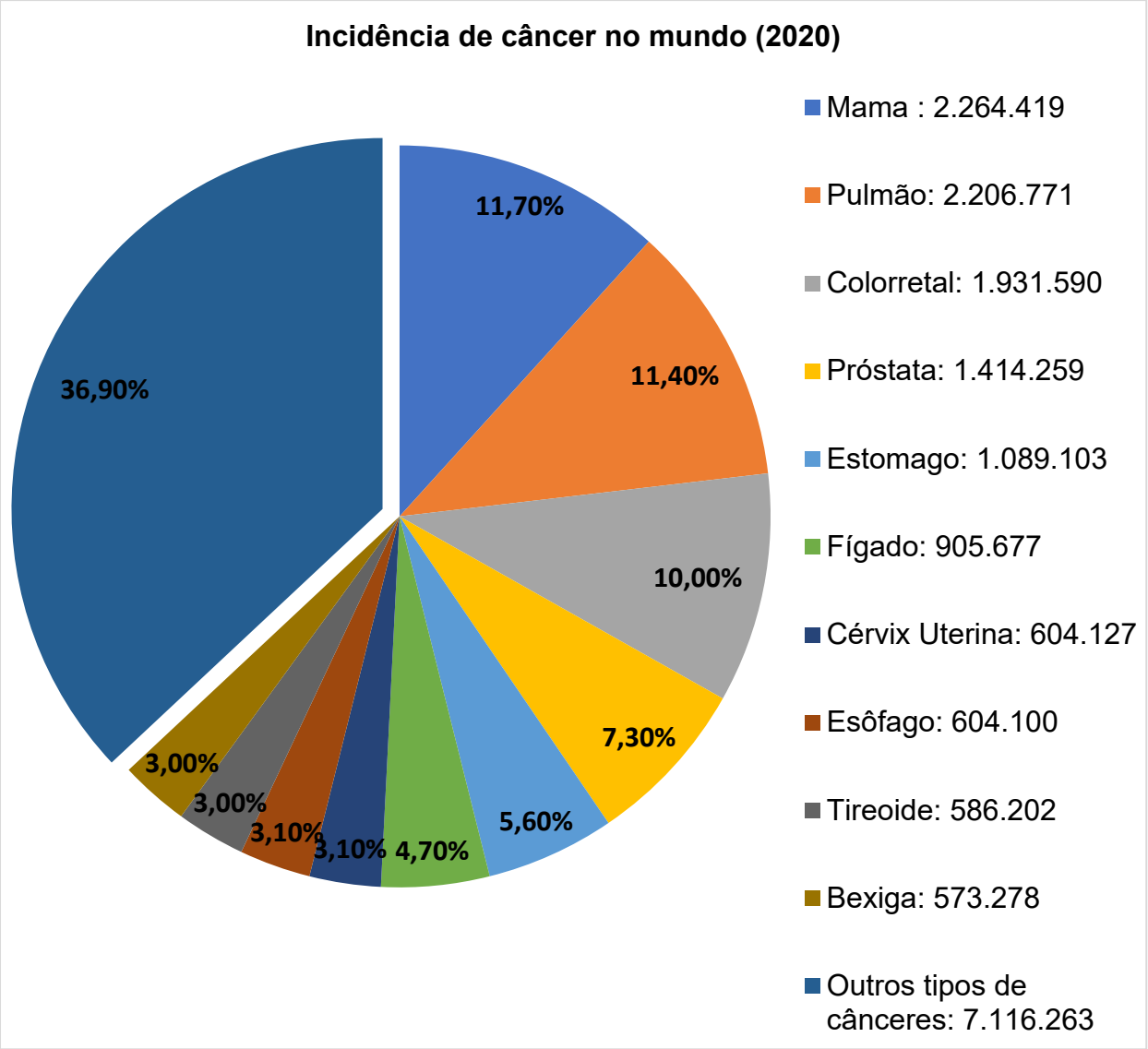

Figura 2. Gráfico de Incidência do câncer no Mundo no ano de 2020. **Fonte:** Elaborado pelo autor com dados da OMS, 2021.

Tabela 2. Número de mortes por câncer no ano de 2020.

Tipo de Câncer	Número de mortes	Número de mortes em %
Pulmão	1.796.144	18,0%
Colorretal	935.173	9,4%
Fígado	830.180	8,3%
Estômago	768.793	7,7%
Mama	684.996	6,9%
Esôfago	544.076	5,5%
Pâncreas	466.003	4,7%
Próstata	375.304	3,8%
Cérvix Uterina	341.831	3,4%
Leucemia	311.594	3,1%
Outros tipos de cânceres	2.904.039	29,2%

Fonte: Elaborado pelo autor com dados da OMS, 2021.

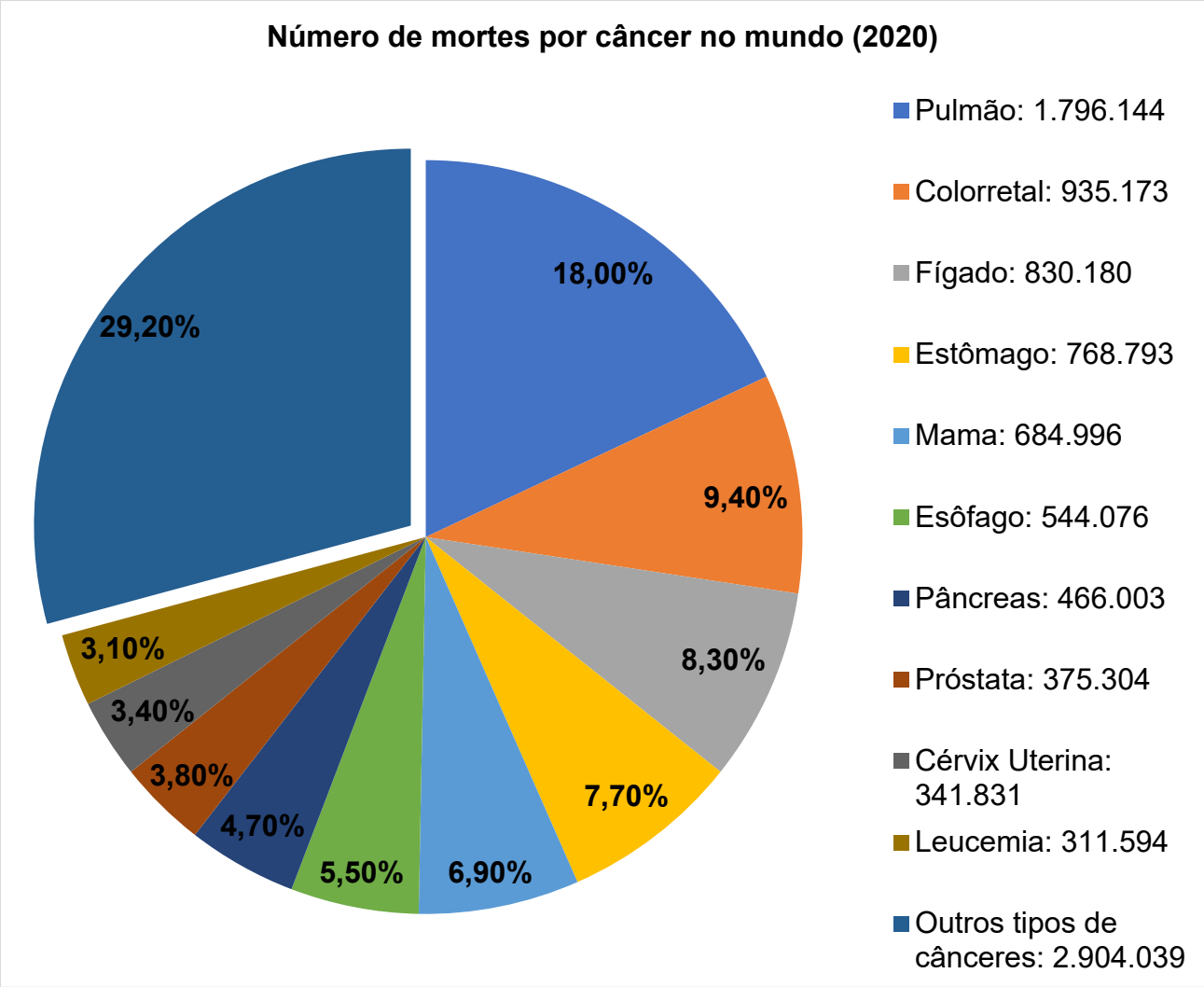

Figura 3. Gráfico de número de óbitos de câncer no mundo no ano de 2020. **Fonte:** Elaborado pelo autor com dados da OMS, 2021

Ao analisar as tabelas pode-se afirmar que os 10 principais tipos de cânceres são responsáveis por 60% dos casos recém diagnosticados e também responsáveis por 70% das mortes de câncer. O câncer de mama feminino é o câncer com maior número de diagnósticos cerca de 11,7% do total de casos, em seguida aparece o câncer de pulmão com 11,4%, colorretal (10%), próstata (7,3%) e estômago (5,6%). Em relação ao número de causa de mortes, em primeiro lugar está o câncer de pulmão, responsável por 18% das mortes, seguido pelo câncer de colorretal com 9,4%, fígado com 8,3%, estômago com 7,7% e mama com 6,9% (SUNG *et al.*, 2021).

No Brasil segundo a OMS estimou-se 600 mil novos casos durante o ano de 2020, com cerca de 260 mil casos levados a óbito, o que equivale aproximadamente 43% dos

casos. Abaixo segue duas tabelas mostrando os dez tipos de cânceres mais incidentes (Tabela 3) e com maior mortalidade (Tabela 4) no Brasil, desconsiderando sexo ou idade, a porcentagem está relacionada ao total de casos de canceres sem distinção de tipo.

Tabela 3. Incidência de casos de câncer no Brasil em 2020.

Tipo de câncer	Incidência (número de casos)	Incidência em %
Próstata	97.278	16,4%
Mama	88.492	14,9%
Colorretal	55.102	9,3%
Pulmão	40.409	6,8%
Tireoide	30.607	5,2%
Estômago	20.139	3,4%
Cérvix uterina	17.743	3,0%
Bexiga	15.854	2,7%
Pâncreas	13.307	2,2%
Fígado	12.674	2,1%
Outros tipos de cânceres	200.607	33,9%

Fonte: Elaborado pelo autor com dados da OMS, 2021.

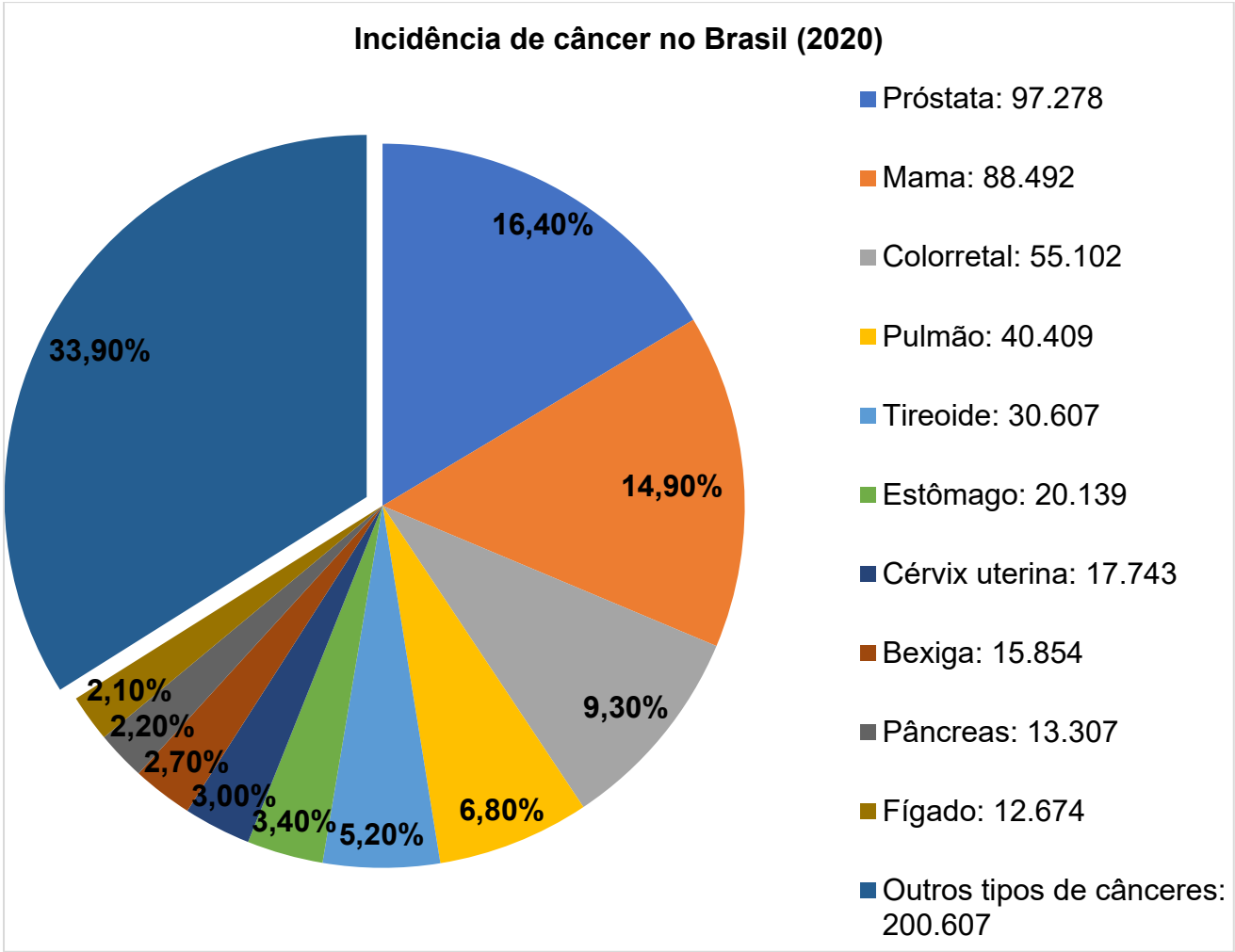

Figura 4. Gráfico de Incidência do câncer no Brasil no ano de 2020. **Fonte:** Elaborada pelo autor com dados da OMS ,2021.

Tabela 4. Número de mortes por câncer no Brasil durante o ano de 2020.

Tipo de câncer	Número de mortes	Número de mortes em %
Pulmão	35.160	13,5%
Colorretal	26.170	10,1%
Mama	20.725	8,0%
Próstata	18.345	7,1%
Estômago	15.783	6,1%
Pâncreas	12.911	5,0%
Fígado	12.139	4,7%
Bexiga	10.920	4,2%
Esôfago	9.786	3,8%
Cérvix uterina	9.168	3,5%
Outros tipos de cânceres	88.842	34,2%

Fonte: Elaborado pelo autor com dados da OMS, 2021

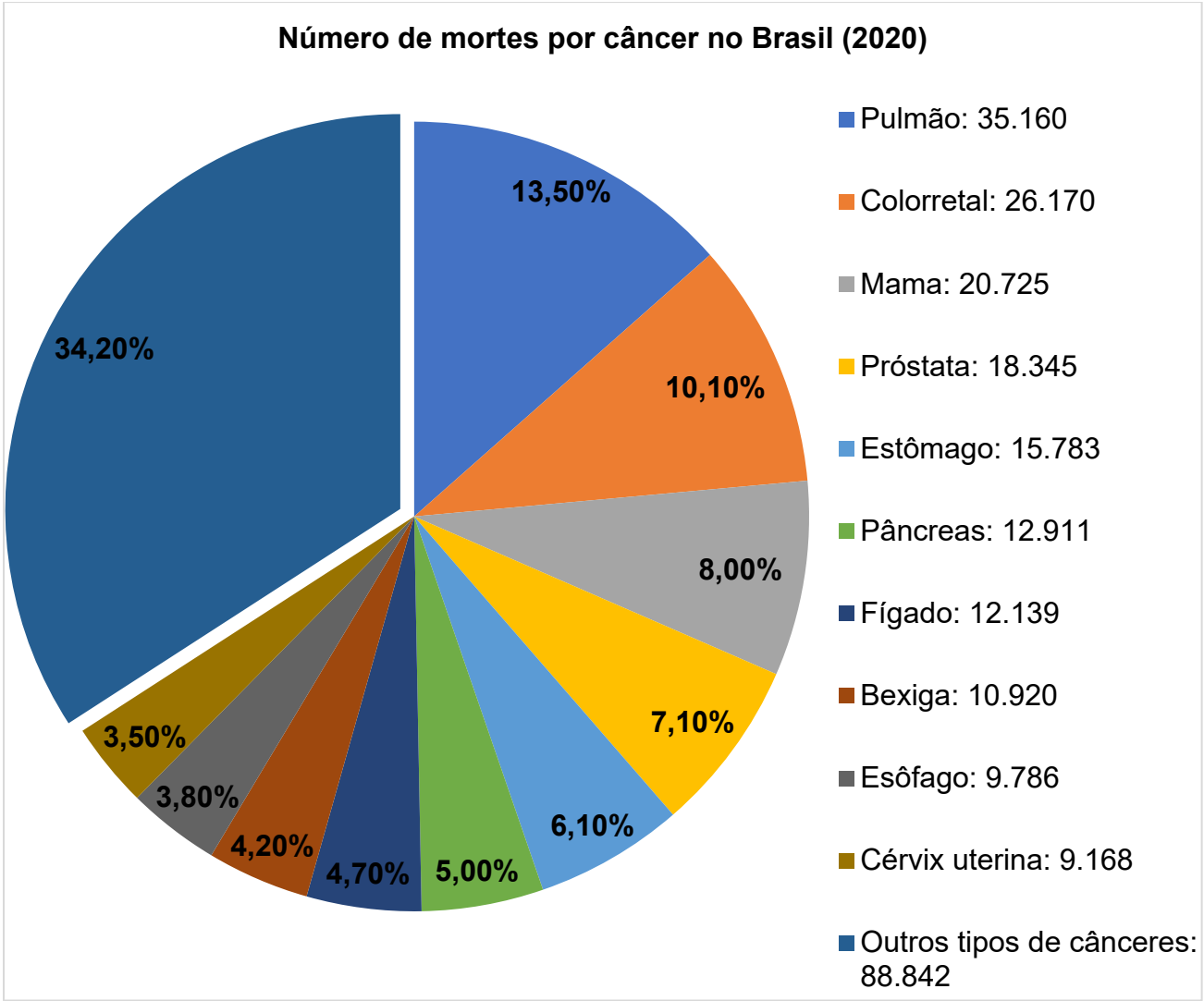

Figura 5. Gráfico de número de óbitos de câncer no Brasil no ano de 2020. **Fonte:** Elaborado pelo autor com dados da OMS, 2021.

Segundo as Tabelas 3 e 4, é possível afirmar que os dez cânceres mais diagnosticados no Brasil somam cerca de 67% de todos os casos, e juntos possuem em torno de 66% do número de mortes, em relação a todos os tipos de cânceres. No Brasil o tipo de câncer que apresenta o maior número de diagnóstico é o de próstata com 16,4% de todos os casos, seguido pelo câncer de mama com 14,5%, colorretal (9,3%), pulmão (6,8%), tireoide (5,2%) e outros. Em quesito ao número de mortes, em primeiro lugar está o câncer de pulmão com 13,5% de todas as mortes por câncer, seguido pelo câncer colorretal com 10,1%, mama (8,0%), próstata (7,1%), estômago (6,1%) e outros (OMS, 2021).

4.1.3 Terapias para o câncer

Com a crescente evolução das áreas biomédicas, como por exemplo a biologia molecular e tumoral, criou-se novos paradigmas no tratamento do câncer, pois anteriormente o câncer era classificado e tratado de acordo com o local ou órgão de origem e suas características histomorfológicas, hoje em dia o câncer possui vários tratamentos diferentes dependendo do tipo de câncer e do grau no qual se encontra (ZUGAZAGOITIA *et al.*, 2016).

4.1.3.1 Cirurgia

A cirurgia oncológica consiste em remover o tumor com uma margem circundada de tecido saudável, com o intuito de remover totalmente o tumor e minimizar o risco do retorno do mesmo. Essa prática é datada de milhares de anos, antes mesmo da invenção da anestesia e da antissepsia; no século passado somente os pacientes em casos graves eram submetidos a cirurgia, devido a baixíssima taxa de cura e altíssima taxa de mortalidade. Com o avanço das tecnologias e a descoberta da anestesia, da assepsia e de diagnósticos por imagem como a radiografia e a tomografia computadorizada no qual permitem o delineamento do local do tumor e a extensão no qual ele ocupa, a cirurgia se tornou uma intervenção médica aceitável e amplamente utilizada (WYLD; AUDISIO; POSTON, 2015).

4.1.3.2 Quimioterapia

A quimioterapia é um tratamento oncológico no qual se utiliza um ou mais fármacos com intuito de combater o câncer, prolongar a vida ou reduzir os sintomas do paciente. O uso da quimioterapia acontece a mais de 50 anos devido seu grande sucesso no tratamento de tumores testiculares, alguns tipos de leucemias, tumores de mama de tecido epitelial comum, colón e pulmão. (DEVITA; CHU, 2008; JOHNSTONE; RUEFLI; LOWE, 2002).

As drogas utilizadas para a quimioterapia foram desenvolvidas com base de estudos que mostram que as células que se dividem rapidamente são particularmente mais sensíveis ao dano em sua síntese de DNA e no seu fuso mitótico, pois a grande

maioria das células normais são quiescentes, em repouso na fase G1/G0 do ciclo celular, assim faz com que as drogas que danificam seletivamente as células em divisão nas fases S ou M tenham uma tendência a preferir atacar as células tumorais (BAGNYUKOVA *et al.*, 2010).

A utilização de drogas platinóides como Cisplatina, Carboplatina, Dedaplatina e Oxaliplatina em conjunto de agentes alquilantes como Altretamina, Bendamustina, Busulfan, Carmustina, Clorambucil, Ciclofosfamida Dacarbazina, Ifosfamida, Lomustine, paralisa as bifurcações de replicação de DNA durante a fase S e bloqueia a transcrição por RNA polimerases (BAGNYUKOVA *et al.*, 2010; CANCER.ORG, 2019; ROSSI; DI MAIO, 2016).

Também pode ser utilizado antimetabólitos como Azacitidina, 5-fluorouracil (5-FU), 6-mercaptopurina (6-MP), Capecitabina (Xeloda), Cladribina, Clofarabina, Floxuridina, Gencitabina (Gemzar), Hidroxiureia, Pentostatina, e outros. Os antimetabólitos agem no DNA e no RNA, servindo como um substituto para os blocos de construções normais de RNA e DNA, fazendo com que os ácidos nucleicos não consigam se replicar, assim impedindo a replicação celular. Outra classe de drogas são os antibióticos antitumorais, estes que alteram o DNA dentro das células cancerosas para sua multiplicação, exemplo deles são: Daunorrubicina, Doxorrubicina (Adriamicina), Doxorrubicina lipossomal, Epirrubicina, Idarrubicina e Valrubicina. Também pode se utilizar de inibidores das enzimas topoisomerase I (Irinotecan e Topotecan) e II (Etoposide (VP-16), Mitoxantrone, Teniposide) e inibidores mitóticos como: Cabazitaxel, Docetaxel, Nab-paclitaxel, Paclitaxel, Vinblastine, Vincristine e Vinorelbine (CANCER.ORG, 2019).

Utilizando essas drogas e causando esses tipos de danos no DNA pode ativar pontos de verificação para eliminação de lesões fornecendo tempo para o DNA ser reparado, com isso a exposição prolongada das células durante a quimioterapia pode levar as células que se dividem rapidamente à apoptose reduzindo a massa tumoral, mas também ocorre um efeito colateral as células normais de divisão frequente como as células epiteliais intestinais e células precursoras da medula óssea (BAGNYUKOVA *et al.*, 2010; JOHNSTONE; RUEFLI; LOWE, 2002).

4.1.3.3 Radioterapia

A radioterapia é uma terapia que utiliza da radiação para tratar tumores tanto malignos quanto benignos, pode ser utilizada em conjunto da quimioterapia, cirurgia ou ambos. No caso de tumores primários ou depósitos metastáticos a radioterapia pode ser utilizada de forma paliativa com intuito de reduzir a dor ou efeitos colaterais do tumor (BREUR, 2017). A radioterapia continua sendo de grande importância no tratamento contra o câncer, pois é um tratamento de modalidade única e altamente rentável, contabilizando apenas 5% do custo total de um tratamento contra câncer, aproximadamente 70% dos pacientes com câncer receberão radioterapia como tratamento, estima-se que a radioterapia contribui cerca de 40% no tratamento curativo (BASKAR *et al.*, 2012; WHO, 2021).

O tratamento inicialmente consiste em um planejamento do local onde será bombardeado pela radiação, para isso se utiliza de tomografia computadorizada em conjunto de imagens de ressonância magnética, assim se terá profundo estudo do local mais apropriado para o tratamento (BREUR, 2017).

A radiação utilizada nessa terapia é a radiação ionizante pois ela forma partículas carregadas eletricamente (íons) e deposita energia nas células por qual passa, essa energia pode eliminar as células cancerosas ou causar anomalias genéticas resultando também na morte do câncer pois o dano no material genético ira bloquear a capacidade do mesmo de se dividir e proliferar (BASKAR *et al.*, 2012).

Esse tratamento causa danos tanto às células cancerosas quanto às células normais, sendo assim se maximiza a dose de radiação nas células anormais enquanto se reduz a exposição de radiação em células normais no qual estão próximas as células cancerosas. Com essa redução as células normais podem se reparar rapidamente e retornar a sua função normal, diferentemente das célula cancerosas no qual normalmente não são eficientes no reparo do dano causado pela radiação, fazendo assim com que este tratamento seja eficiente (BASKAR *et al.*, 2012; BEGG; STEWART; VENS, 2011).

4.1.3.4 Transplante de medula óssea

O transplante de medula óssea (TMO) é um tratamento usado no combate de doenças hemolíticas, tanto malignas quanto benignas como no caso de leucemias e linfomas. É caracterizado pela remoção das células troncos de um doador, no qual pode ser o próprio paciente ou outro doador compatível (transplante autólogo e alógeno) e a colocação dessas células novamente no paciente após um período de condicionamento, células progenitoras infundidas na corrente sanguínea são implantadas na medula óssea fazendo com que ocorra uma reconstituição hematopoiética (DE ARAÚJO LIMA *et al.*, 2012).

Esse tratamento tem obtido grandes chances de sucesso e crescido de forma exponencial, devido ao conhecimento que se tem sobre o sistema de histocompatibilidade humana e ao extremo cuidado durante a seleção de doadores de medula óssea, pois pode ocorrer complicações após transplantes alogênicos como a doença enxerto versus hospedeiro (DEVH) (TREISTER *et al.*, 2010). A DEVH é uma complicação que acontece mediante a ativação de células T em resposta às moléculas do complexo principal de histocompatibilidade (MHC), ou seja, as células T doadas reconhecem moléculas do tecido do hospedeiro como se fossem invasoras, e desencadeiam reações autoimunes. A incidência dessa doença pode variar de 6 a 80% de acordo com a idade do paciente, o tipo de fonte da célula tronco, o tipo de doador, a manipulação do enxerto e o uso de infusões de linfócitos após o transplante (DE ARAÚJO LIMA *et al.*, 2012).

4.1.3.5 Terapia Alvo Molecular

As terapias alvo, ou terapias biológicas veem se tornando uma das melhores escolhas durante o tratamento de câncer, pois ela atua especificamente bloqueando moléculas essenciais para o crescimento celular e tumorigênese. Devido a essa especificidade, essas novas terapias possuem grandes expectativas de terem uma melhor eficácia e poucos efeitos colaterais quando comparado com outros tratamentos (SEEBACHER *et al.*, 2019).

Para se ter sucesso no desenvolvimento dessa terapia se deve identificar os alvos ideais, uma das ocorrências mais comum no câncer é a alteração do perfil genético no

qual leva a mutações ou mudanças em proteínas e receptores que promovem a sobrevivência e a proliferação celular. Usando dessas alterações genéticas específicas é possível distinguir células normais de células cancerosas e pode ser usado como alvos moleculares para uma terapia alvo (LEE; TAN; OON, 2018).

Conhecendo a fisiologia e as características dos alvos moleculares fica possível identificar estratégias moleculares com grande potencial de inibição do crescimento tumoral e sua progressão. Os marcadores do câncer podem ser escolhidos utilizando de sequenciamento genômico o que permite a comparação da expressão de genes e proteínas de células normais e células cancerosas afim de identificar as mudanças em suas expressões e revelar sua heterogeneidade genética, no qual é de suma importância para o desenvolvimento de novas drogas (LEE; TAN; OON, 2018; VOGELSTEIN *et al.*, 2013).

Os agentes utilizados na terapia alvo molecular devem exibir diferentes funções e características, devem atuar em antígenos da superfície celular, fatores de crescimento, receptores ou vias de transdução de sinal no qual regulam a progressão do ciclo celular, morte celular, angiogênese e metástase (SAIJO, 2010). Os medicamentos utilizadas na terapia alvo bloqueiam sinais no qual favorecem o crescimento celular do câncer interferindo na regulação do ciclo celular ou induzindo a morte celular das células cancerosas (PADMA, 2015). Esses agentes podem ser classificados em pequenas moléculas, vacinas imunoterapêuticas para câncer, terapia gênica e anticorpos monoclonais.

As terapias alvos são consideradas como um grande marco no tratamento do câncer devido ao seu caráter de entregar drogas com alta especificidade, eficiência e com uma menor toxicidade comparado com a quimioterapia que tem o problema de seus agentes citotóxicos não afetarem somente as células cancerosas (DE PALMA; HANAHAN, 2012). Contudo, também existem também limitações, pois devido ao fato de que a terapia alvo molecular trabalha inibindo biomarcadores específicos para a progressão do câncer, o tratamento só será efetivo em pacientes com tumores que expressam esses biomarcadores (LEE; TAN; OON, 2018).

4.1.3.5.1 Pequenas Moléculas

As pequenas moléculas são definidas como compostos com baixo peso molecular relativo (<900 Da) os quais conseguem penetrar nas células do alvo molecular e localizar por exemplo uma proteína específica dentro da célula (JOO; VISINTIN; MOR, 2013), normalmente essas pequenas moléculas focam na inativação de cinases e na interrupção de vias sinalizadoras que estão desreguladas por causa da carcinogênese. Outras formas de utilizá-las é como alvo de proteosomos, cinases dependentes de ciclinas (CDKs), e inibidores da poli ADP-ribose polimerase (PARP), assim ativando gatilhos do *checkpoint* do ciclo celular, levando à apoptose e ao reparo do DNA (LEE; TAN; OON, 2018; LHEUREUX *et al.*, 2017).

4.1.3.5.2 Vacinas terapêuticas contra o câncer

As vacinas terapêuticas para câncer focam em uma resposta antitumoral mediada pelo sistema imune. São divididas em duas principais classes, a primeira é o tipo paciente específico, no qual a vacina é gerada a partir das células tumorais do paciente; o segundo tipo são as vacinas do tipo paciente não específico, elaboradas para induzir uma resposta imunológica generalizada a qual talvez tenha efeito antitumoral sobre o paciente, um exemplo disso é o interferon-alfa (IFN-α) e a interleucina-2 (IL-2), que induzem uma resposta imune geral, mas que em pouca parte dos pacientes é antitumoral. Essas vacinas têm como alvo antígenos associados a tumores específicos (TAA), que ocorre por meio da estimulação de células T ligadas ao peptídeo do complexo MHC induzindo uma resposta imune antitumoral (MARIJO BILUSIC; RAVI A. MADAN, 2012).

A quantidade de expressão de TAA no câncer é relacionada ao tamanho da especificidade do tumor em comparação ao tecido normal e ao grau de “tolerância” inerente dada ao TAA, esses fatores são críticos para a eficácia dessas vacinas para o tratamento do câncer (LEE; TAN; OON, 2018; SCHLOM *et al.*, 2014).

4.1.3.5.3 Terapia gênica

A terapia gênica tem o propósito de alterar ou inserir novos materiais genéticos dentro de seu alvo estabelecido, como células cancerosas. O intuito desse tratamento é destruir ou inibir o crescimento tumoral, e sua vantagem é não causar nenhum tipo de dano a células e tecidos saudáveis que são circundantes ao alvo (AYLLÓN BARBELLIDO *et al.*, 2008).

Como dito anteriormente, a terapia gênica pode ser efetuada através da substituição do gene de supressão tumoral mutado pela sua versão normal fazendo com que restaure a sua função. A outra forma é realizada através da introdução de materiais genéticos com intuito de inibir a expressão gênica do oncogene, como por exemplo a utilização de siRNA (*small interfering* RNA) ou oligonucleotídeos *antisense*, estimulando a resposta imunológica e inibindo o processo de angiogênese do tumor, fazendo com que as células desse tumor possam ser tratadas mais facilmente (LEE; TAN; OON, 2018).

As metodologias utilizadas dentro da terapia gênica incluem o uso de viroterapia oncolítica, RNA de interferência, aptameros, modulação do micro ambiente, terapia com oligonucleotídeos, intervenções genômicas específicas, edição gênica e transferência de genes (SADELAIN, 2016).

4.1.3.5.4 Anticorpos Monoclonais

Os anticorpos monoclonais (mAbs) são desenvolvidos para alvos localizados na parte externa da célula devido a seu tamanho (JOO; VISINTIN; MOR, 2013).

Especialmente proteínas extracelulares e inibidores de crescimento tumoral, no qual irá interromper as interações entre receptores e ligantes, podendo ter ações pelos mecanismos diretos e indiretos após a ligação nas células cancerosas (WEINER; SURANA; WANG, 2010). O mecanismo de forma indireta atua utilizando a resposta de defesa do corpo, como o recrutamento de células efetoras ou fagócitos após a estimulação causada pela ligação do anticorpo monoclonal com os antígenos específicos da célula cancerosa (FOLTZ; KAROW; WASSERMAN, 2013). Na forma direta os anticorpos monoclonais se ligam ao antígeno, receptor celular ou proteína de membrana

da célula cancerígena e a partir dessa ligação consegue exercer seu efeito diretamente no seu alvo especifico, induzindo a morte celular (VAN DE DONK *et al.*, 2016).

Um exemplo disso é o Trastuzumabe que funciona de forma direta inibindo a proteína *Human Epidermal Growth Factor Receptor 2* (HER2) pela ativação de sua internalização, degradação ou dimerização, também pode atuar de forma indireta ao recrutar a citotoxicidade mediada por células dependentes de anticorpos (ADCD) mediado pelos CD-16 em células de cânceres de mama, esse mecanismo reduz a divisão celular do câncer, induz parada de ciclo celular e a internalização do receptor (LEE; TAN; OON, 2018; LUQUE-CABAL *et al.*, 2016; VU; CLARET, 2012). Outros mAbs que funcionam através das ADCD são: rituximab, cetuximab e pertuzumab (KIMIZ-GEBOLOGLU; GULCE-IZ; BIRAY-AVCI, 2018).

Outra forma de se utilizar os anticorpos monoclonais é conjugando-os com drogas, chamados de conjugado da anticorpo-droga (ADC), que compreendem em utilizar um anticorpo monoclonal com alvo específico e agregar a ele por meio de ligações covalentes moléculas pequenas de drogas com função anticâncer, o ADC irá internalizar nas células cancerosas alvo através da endocitose mediada pelo receptor, a seguir acontece a liberação das citotoxinas que levam a morte apoptótica das células cancerosas (LEE; TAN; OON, 2018; PETERS; BROWN, 2015).

4.1.3.6 Imunoterapia

O sistema imunológico tem uma série complexa de mecanismos para detectar e erradicar as células cancerosas, mas essas mesmas vias que protegem contra o câncer também podem ser as responsáveis por promover a seleção das células tumorais que quais estão preparadas para enganar o sistema imunológico. As imunoterapias contra o câncer manipulam o sistema imunológico a fim de reativar a resposta imune antitumoral e assim superar as vias que levam ao escape das células cancerosas. Um exemplo do emprego dessa terapia é a utilização de citocinas para afetar a função das células imunológicas, por exemplo, altas doses de IL-2 e IFN α-2b podem levar a múltiplos efeitos benéficos e têm sido usados no tratamento de melanoma avançado e carcinoma de células renais (KENNEDY; SALAMA, 2020).

A eliminação de tumores por *clusters* de células de diferenciação tumor específico 8 T (CD8 T) passam por vários processos, sendo cada um deles necessário, e regulados dependendo da necessidade de força para que não se tenha efeitos colaterais à tecidos saudáveis. Primeiramente, os neoantígenos tumorais devem estar presentes no tumor e devem ter sidos apresentados as células apresentadoras de antígenos (APC), assim como ás células dendríticas (CD) onde seus antígenos possam ser apresentados as células imunes nos linfonodos, as células T tumor especifico devem “reconhecer” esses antígenos através do complexo peptídico do MHC e devem receber os sinais estimulatórios corretos e ainda pedir uma “ajuda” (por exemplo a ligação do CD40) antes de ele ser ativado e começar a proliferar. As células T tumor específicas ativadas então saem dos linfonodos drenantes (dLN) e circulam ao redor antes de entrar no tumor, após entrarem essas células T encontram uma série de mecanismos imunossupressores que precisam superar, e possivelmente irão necessitar de uma re-estimulação pelas células APCs dentro do tumor. Uma vez que essas células T ativadas tenham passado por essas etapas, as mesmas atacam o tumor em um processo que envolve o reconhecimento do antígeno expresso pelo tumor e a liberação de moléculas potentes, como por exemplo as perforinas e as granzimas, que irão eliminar o tumor; por fim esse processo pode gerar células de memória no qual podem ser importante no controle de micro metástases (STEVEN; FISHER; ROBINSON, 2016).

A imunoterapia pode ser utilizada com agentes biológicos, segundo Benitez e colaboradores (2019) pode se utilizar de *Mycobacterium bovis* Bacillus Calmette– Guerin (BCG), como agente imunoterapêutico para tratar melanoma e carcinoma urotelial superficial. O BCG quando utilizado no tratamento de melanoma cutâneo primário é comumente administrado sozinho, ou com uma vacina de células tumorais autólogas, ou com algum tipo de droga terapêutica, são administradas por injeção intralesional para diferentes estágios do melanoma, cominando na redução da lesão de tumores locais e regionais e obtendo melhora na sobrevida do paciente (BENITEZ *et al.*, 2019).

4.1.3.7 Terapia fotodinâmica

A terapia fotodinâmica (TPD) é uma nova forma de terapia seletiva e não-invasiva a qual tem mostrado resultados promissores para o tratamento de câncer. Pode ser

utilizada de forma local ou sistêmica, através de compostos fotossensibilizantes que são acumulados intensamente nos tecidos danificados (KWIATKOWSKI *et al.*, 2018).

A reação da TPD consiste em três componentes: droga foto-sensitiva (sensibilizador), luz no comprimento de onda adequado e oxigênio. O processo começa pela excitação da droga foto-sensitiva pela luz, que passa do estado singleto, após a iluminação a droga foto-sensitiva sai do estado singleto de baixa energia e vai para o estado de alta energia conhecido como estado singleto excitado, no estado excitado o fotossensibilizador pode sofrer cruzamento intersistema, o qual faz com que o estado singleto de vida curta seja convertido em tripleto excitado, que possui uma vida relativamente longa. Nesse estado os elétrons possuem um spin paralelo em sua orbital no qual podem reagir em três mecanismos diferentes (DÜZGÜNEŞ *et al.*, 2018).

No mecanismo I o agente fotossensível excitado interage com a molécula intracelular, transferindo um próton ou um elétron, formando assim um radical. A molécula que recebeu a carga reage com oxigênio e forma uma espécie reativa de oxigênio (ERO), moléculas de água também podem ser afetadas pelo fotossensibilizador e caso isso ocorra vão formar radicais hidroxilas e o radical ânion superóxido. (DÜZGÜNEŞ *et al.*, 2018). No mecanismo II a energia do fotossensibilizador no estado tripleto é transferida diretamente a o oxigênio molecular, que está em seu estado enérgico básico; isso faz com que várias partículas de oxigênio sejam excitadas, formando o oxigênio singleto, o qual é caracterizado pela sua grande propriedade oxidante (KWIATKOWSKI *et al.*, 2018). Por último o mecanismo III acontece quando se tem uma interação direta entre o fotossensibilizante excitado e as células cancerosas alvo do tratamento sem a formação de espécies reativas de oxigênio (DÜZGÜNEŞ *et al.*, 2018).

O processo de morte do câncer pode acontecer de forma direta pela via da apoptose, assim como também pode acontecer por formas não apoptóticas como por exemplo necrose e autofagia. Também é possível ocorrer de formas indiretas, quando ocorre dano na vasculatura que fornece nutrientes e oxigênio para o câncer. Esses mecanismos são induzidos por vias de sinalizações distintas, isso vai depender do agente fotossensibilizante, protocolo, localização e dosagem do fotossensibilizante, genótipo das células no qual serão submetidas ao tratamento, nível de oxigenação e entre outros fatores. Um exemplo disso é quando se utiliza um fotossensibilizador localizado na mitocôndria, é mais fácil ele desenvolver uma resposta apoptótica assim

como os que estão localizados na membrana plasmática e lisossomos estimularem necrose (CHILAKAMARTHI; GIRIBABU, 2017).

A terapia fotodinâmica possui uma boa seletividade devido a capacidade dos fotossensibilizadores se acumularem em altas concentrações em células cancerosas comparado às células saudáveis, isso acontece devido a preferência dos fotossensibilizadores de se ligarem com lipoproteínas de baixa densidade (LDL) (DÜZGÜNEŞ *et al.*, 2018). A função da LDL é fornecer colesterol aos tecidos para a criação de membranas durante a divisão celular, como as células cancerosas estão em frequente divisão celular e precisam de uma alta quantidade da LDL, a mesma atua como um transportador para os fotossensibilizantes fazendo com que cheguem em maior quantidade nas células cancerosas (CRUZ *et al.*, 2013).

Trabalhos recentes como o de Couto e colaboradores (2020) identificaram moléculas de porfirinas com platina como uma possível molécula promissora para se utilizar no tratamento contra melanoma metastático, seus testes *in vitro* apontaram inibição da viabilidade das células do melanoma metastático quando expostas a luz e não alteraram a viabilidade celular no escuro, além de não demonstrar citotoxicidade em células não tumorais.

4.1.4 Importância da descoberta de novos compostos para o tratamento

O principal fator de falha no tratamento contra o câncer é a resistência a multidrogas, esse problema afeta diretamente os tratamentos com quimioterapia principalmente de pacientes com câncer de mama, ovário, pulmão ou outros tipos de tumores sólidos ou líquidos. A quimioterapia elimina células que são sensíveis a droga utilizada no tratamento, mas também resulta em grandes quantidades de células com resistência ao tratamento, o que aumenta a preocupação caso o tumor cresça novamente pois agora as células serão resistente a este tratamento (CMR, 2000).

A resistência a os medicamentos utilizados na quimioterapia pode acontecer de maneira intrínseca ou de maneira adquirida. A resistência intrínseca indica que antes de receber o tratamento quimioterápico já preexistia fatores mediadores de resistência na maior parte das células tumorais, fazendo assim com que o tratamento seja ineficaz (HOLOHAN *et al.*, 2013). A resistência adquirida a os medicamentos pode se desenvolver ao decorrer do tratamento de tumores inicialmente sensíveis no qual se

tornaram resistentes devido a mutações que surgem durante o tratamento, também como várias outras respostas adaptativas (HOLOHAN *et al.*, 2013; LONGLEY; JOHNSTON, 2005). A resistência a medicamentos contra o câncer é um sistema complexo no qual recebe influência de fatores como a inativação, alteração do alvo e efluxo do medicamento, reparo de danos no DNA, inibição da morte celular, matriz extra celular, heterogeneidade celular inerente, efeitos epigenéticos ou quaisquer combinações desses mecanismos (HOUSMAN *et al.*, 2014).

Todavia o desafio atual é obter experiência com a utilização de medicamentos tradicionais e com as recentes terapias alvos moleculares (HOLOHAN *et al.*, 2013) e sua utilização de forma combinada para poder obter a melhor opção de tratamento visando prevenir o desenvolvimento de resistência a medicamentos e ser mais eficaz do que o uso em isolado. Com isso vem se empregando a utilização de técnicas de alto rendimento em conjunto de abordagens bioinformáticas e biologia de sistemas com intuito de analisar amostras clínicas e permitindo a identificação de assinaturas moleculares e genótipos que predizem à resposta a certos medicamentos, além disso tais abordagens podem identificar novos alvos terapêuticos para superar ou contornar a resistência a esses medicamentos (HOUSMAN *et al.*, 2014).

Para se iniciar o desenvolvimento de um medicamento baseado em uma pequena molécula ocorre um processo exaustivo que inclui pesquisa básica, identificação e validação de alvos, geração e otimização de *leads*, testes pré-clínicos, testes clínicos em humanos e aprovação regulatória pelos órgãos responsáveis. O processo de desenvolvimento de medicamentos começa com a identificação de um novo alvo (proteína, DNA, RNA, metabólito, entre outros), seguido de sua validação subsequente para confirmar um efeito terapêutico. Isso inclui o desenvolvimento e otimização do ensaio podendo ser bioquímico, baseado em células, citotoxicidade, e entre outros, em que uma metodologia objetiva é derivada para capturar a interação pretendida entre uma biblioteca de compostos e o alvo específico. Os "acertos" confirmados de telas de alto rendimento são então organizados por tipo químico para identificar "ligações" que podem posteriormente serem refinados por química medicinal e por relações estrutura-atividade (SAR), para otimizar propriedades físico-químicas e farmacológicas para potência aprimorada e seletividade. A geração e otimização preliminares de *leads* são um processo *in vitro*, em que os leads selecionados são então progredidos por meio de uma série de ensaios complexos para avaliar as características bifarmacológicas eficazes,

propriedades toxicológicas específicas da classe do composto e absorção, distribuição, metabolismo e excreção favoráveis (MATTHEWS; HANISON; NIRMALAN, 2016).

O processo de ensaio clínico em si pode começar somente após a aprovação dos órgãos regulamentadores através de um pedido de novo medicamento investigacional que detalha os resultados pré-clínicos, o modo de ação do medicamento proposto, efeitos colaterais potenciais e informações de fabricação. Por fim os ensaios clínicos em fases (I, II e III) são supervisionados por uma equipe de pesquisa clínica em estreita comunicação com os órgãos regulamentadores, com a Fase I focando na segurança do medicamento, a Fase II na efetividade e a Fase III na confirmação dos resultados em uma coorte populacional maior. Um resultado bem-sucedido nos testes clínicos levará ao envio de um Pedido de Novo Medicamento para análise e aprovação pelos órgãos reguladores (MATTHEWS; HANISON; NIRMALAN, 2016).

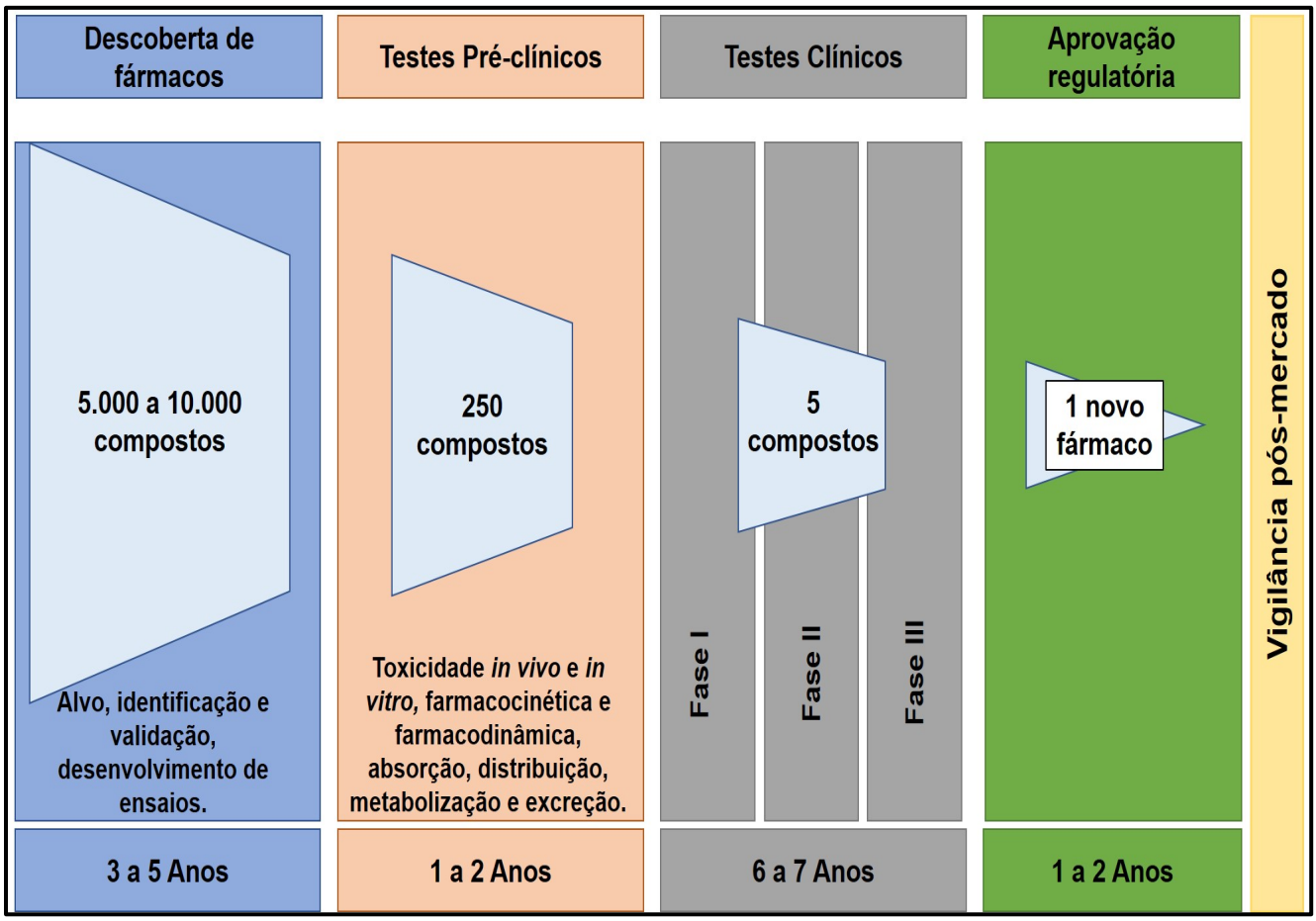

Figura 6. Linha do tempo da descoberta de novos fármacos. A processo atual pode levar em torno de 15 anos. É estimado que de cada 5.000 a 10.000 compostos somente um poderá a se tornar um novo fármaco aprovado. **Fonte:** Adaptado de MATTHEWS; HANISON; NIRMALAN, 2016.

4.2 A Bioinformática

A bioinformática teve sua origem nos anos de 1960 cerca de uma década antes do sequenciamento de DNA se tornar viável (HAGEN, 2000). Essa área pode ser definida como o emprego de ferramentas computacionais no estudo de problemas e questões biológicas, incluindo também as aplicações na saúde humana e planejamento de novos fármacos (VERLI, 2014).

Ocorreram alguns momentos de grande impacto na ciência e isso contribuiu com o desenvolvimento da bioinformática, um desses momentos foi a publicação da estrutura do DNA por Watson e Crick no ano de 1953 além da melhor elucidação da bioquímica e da estrutura proteica graças aos estudos de Pauling, Coren e Ramachandran na década de 1960 (DINIZ; CANDURI, 2017; VERLI, 2014).

Outro fator de grande influência foi os estudos de Margaret O. Dayhoff, também conhecida como a mãe da bioinformática, foi pioneira na sistematização do conhecimento da estrutura tridimensional das proteínas, fazendo com que computadores fossem capazes de determinar a sequência de peptídeos, programas de reconhecimento e exibição de estruturas para serem utilizados em cristalografia de raios-x e métodos computacionais com capacidade de comparação de sequências de proteínas, assim sendo capaz de inferir conexões evolutivas entre reinos (DINIZ; CANDURI, 2017; HAGEN, 2000; VERLI, 2014).

4.3 Ferramentas de bioinformática utilizadas na descoberta de novas drogas.

A descoberta de novas drogas é de suma importância para a terapia contra o câncer como foi previamente descrito. Os métodos tradicionalmente utilizados, como triagem de drogas *in vitro* ou experimentação em animais *in vivo* para a descoberta de novas drogas são de extrema importância, mas são processos caros e laboriosos. Graças ao avanço da tecnologia e o surgimento das ciências ômicas hoje pode se utilizar ferramentas computacionais para analisar e testar drogas melhorando sua eficiência, descobrindo o alvo do medicamento de forma mais rápida, caracterizando os efeitos colaterais e prevendo se pode ocorrer resistência a droga.

Um dos maiores problemas das terapias anticânceres é a heterogeneidade do paciente, que a anos vem mostrando que pacientes com o mesmo tipo de câncer podem ter respostas totalmente diferente quando tratados com o mesmo medicamento (RICE *et al.*, 2010). Com a utilização da bioinformática para a análise de dados de alto rendimento, como dados genômicos, epigenéticos, arquitetura de genoma, transcriptômica, proteômica e dados de perfil ribossômico tiveram uma boa contribuição na descoberta e na realocação de drogas. O desenvolvimento de modelagem por homologia e simulação de estruturas proteicas por exemplo, tiveram grande relevância quando utilizaram-se de grandes banco de dados como os de pequenas moléculas e metabólitos para se obter experimentos mais realísticos de *docking* molecular na triagem virtual (XIA, 2017).

Um ponto fundamental na descoberta de novas drogas é a seleção dos melhores alvos no qual a droga deverá atuar eficientemente, assim pode-se priorizar candidatos como possíveis drogas terapêuticas (PATEL *et al.*, 2013). Um medicamento desejável é aquele que reduz os sintomas do paciente sem causar efeitos secundários fortes e que apresente baixa chance de resistência ao medicamento e baixo efeito deletério ao meio ambiente por exemplo, reativação por bactérias após uso humano (BOXALL *et al.*, 2012; XIA, 2017).

Inicialmente a bioinformática contribuiu com a descoberta de drogas através da identificação das sequências homólogas entre o gene onc do vírus simian do sarcoma símio, v-sis e um fator de crescimento derivado de plaquetas (PDGF) através de um alinhamento simples de sequências (DOOLITTLE *et al.*, 1983; WATERFIELD *et al.*, 1983), essa descoberta resultou na utilização do PDGF como droga para o tratamento anticâncer (ÖSTMAN; HELDIN, 2007; XIA, 2017).

Na *internet* pode ser encontrados *sites* com grande compilação de programas, banco de dados e servidores *online* que estão diretamente relacionados com a descoberta de novas drogas, um desses sites é o Click2Drug (https://click2drug.org/). Também existem alguns bancos de dados que fornecem informações diretas de oncogenes ou conjuntos de genes supressores de tumor de diferentes tipos de câncer, como os bancos de dados OMIM (https://www.omim.org/), GAD (https://geneticassociationdb.nih.gov/), GWAS Catalog (https://www.ebi.ac.uk/gwas/) e DisGeNET (https://www.disgenet.org/). Ademais existem outros *sites* que disponibilizam dados brutos que precisam ser analisados por bioinformática para identificar genes relacionados ao câncer, como TCGA (https://www.cancer.gov/about-

nci/organization/ccg/research/structural-genomics/tcga) e o GEO (https://www.ncbi.nlm.nih.gov/geo/). No que tange a simulação computacional, existem técnicas como a de *docking* molecular e dinâmica molecular onde utilizam de recursos tridimensionais para testar a funcionalidade de moléculas, esses sistemas possuem *software*s gratuitos disponíveis, como é o exemplo do Autodock4, Autodock Vina, Gromacs, LAMMPS, MDynaMix e entre outros.

4.3.1 Bancos de dados biológicos

Os bancos de dados biológicos são bibliotecas de informações relacionadas a ciências da vida, consideradas uma ferramenta de extrema importância para a ciência, pois com o passar dos anos foram geradas milhares de informações relacionadas a proteínas, genes, fármacos e entre outros. Assim surgiu a necessidade de armazenar, organizar, categorizar e fornecer anotações biológicas desses conteúdos com intuito de otimizar trabalhos e facilitar a vida dos pesquisadores.

Com o avanço nas técnicas de sequenciamento gênico por exemplo, os bancos de dados de sequencias de nucleotídeos foram altamente incrementados de conteúdos e atualmente servem como base em diversas áreas de pesquisa, entre os principais banco de dados de sequências de nucleotídeos estão o GenBank (*National Center for Biotechnology Information* – NCBI), EMBL (*European Bioinformatics Institute*) e DDBJ (*DNA Data Bank of Japan – National Institute of Genetics*), o conjunto desses três banco de dados formam o *International Nucleotide Sequence Database* (INSD), esse que além de ser o centro das informações biológicas sobre qualquer tipo de organismo, também troca informações e são utilizados como fonte para outros bancos de dados especializados (BENSON *et al.*, 2009; NEOPROSPECTA, 2018).

4.3.1.1 Cancer Cell Line Encyclopedia (CCLE)

O projeto *Cancer Cell Line Encyclopedia* (CCLE) tem por objetivo fornecer dados ômicos sistemáticos de experimentos farmacológicos que utilizam linhagens de células cancerosas, ajudando na identificação de assinaturas moleculares de drogas, possui mais de 1500 linhagens celulares de câncer humano sequenciadas e já determinou mais de 2317 genes mutantes. Possui perfil de expressão genica, no qual foi realizado usando

54.000 sondas gênicas em 1037 linhagens celulares. Além disso traçou o perfil de 24 compostos em 504 linhagens celulares com intuito de associar as sensibilidades as drogas com perfis ômicos (BARRETINA *et al.*, 2012; LI *et al.*, 2019).

4.3.1.2 Catalogue of Somatic Mutations in Cancer (COSMIC)

O *Catalogue of Somatic Mutations in Cancer* (COSMIC) é um banco de dados curado por especialistas no qual abrange uma vasta variedade de mecanismos de mutação somática que causam câncer humano. O COSMIC possui detalhes sobre milhões de mutações em variados tipos de cânceres, sua curadoria manual fornece detalhes aprofundados sobre as distribuições e efeitos da mutação, enquanto a curadoria semiautomática de genomas do câncer fornece amplas anotações somáticas para a descoberta de alvos e identificação de padrões e assinaturas (FORBES *et al.*, 2009a, 2009b, 2011).

4.3.1.3 ChEMBL

O ChEMBL é um banco de dados de larga escala que contém informações relacionadas a bioatividade de compostos, como ligações, funções, absorção *in vivo*, distribuição, metabolização, excreção e toxicidade. Os dados do ChEMBL são manualmente curados da literatura publicada, juntamente com dados extraídos de outros banco de dados, é padronizado para que possa ser utilizado na biologia química e na descoberta de novas drogas (BENTO *et al.*, 2014).

4.3.1.4 ClinVar

O ClinVar é um arquivo público de relatórios das relações entre variações humanas e fenótipos, é hospedado pelo NCBI. Seu potencial está intimamente relacionado a integração de mais de dez tipos diferentes de bancos de dados como, dbSNP, dbVar, Pubmed, OMIM e outros bancos de dados com informações sobre variações genéticas e fenótipo clínico, assim formando um padrão confiável de variação genética e correlação clínica (LANDRUM *et al.*, 2014; LI *et al.*, 2019).

4.3.1.5 Single Nucleotide Polymorphism Database (dbSNP)

O *Single Nucleotide Polymorphism Database* (dbSNP) é um banco de dados para variações genéticas curtas de uma variedade de organismos, ele cataloga variações de nucleotídeo único, pequenas inserções e deleções de nucleotídeos, repetições curtas em tandem e microssatélites. A busca nesse banco de dados é bem completa e os detalhes do registro de variação correspondida são exibidos como o relatório de *cluster* do polimorfismo de nucleotídeo único (SNP, do inglês *Simple Nucleotide Polymorphism*) de referência que contém o resumo do alelo, informações de mapeamento na nomenclatura da *Human Genome Variation Society* (HGVS), visão centrada no gene, tabela de mapa com coordenadas cromossômicas e visão de variação (BARRETT *et al.*, 2015; CHEN; HUANG; WU, 2017).

4.3.1.6 DisGeNET

O DisGeNET é uma plataforma de busca no qual se tem uma das maiores coleções disponíveis publicamente de genes e variantes associadas a doenças humanas. O DisGeNET integra dados de repositórios curados por profissionais qualificados, catálogos de estudo de associação genômica ampla (GWAS, do inglês *genome-wide association*), modelos animais e literatura cientifica (LI *et al.*, 2019; PIÑERO *et al.*, 2015, 2017).

4.3.1.7 DrugBank

O *DrugBank* é um banco de dados ou enciclopédia de livre acesso que contém informações sobre drogas e seus respectivos alvos. Possui amplo escopo, referências abrangentes e descrições de dados excepcionalmente detalhadas como os alvos das drogas, sequência, estrutura e via metabólica. Seus extensos dados sobre medicamentos e alvos de medicamentos permitiram a descoberta e o reaproveitamento de uma série de medicamentos existentes para tratar doenças raras e recentemente identificadas, também é fortemente utilizado pela indústria farmacêutica, químicos medicinais, farmacêuticos, médicos, pesquisadores, estudantes e público em geral (LAW *et al.*, 2014; LI *et al.*, 2019; WISHART *et al.*, 2006, 2008).

4.3.1.8 Ensembl

O *Ensembl* é um banco de dados de anotações de genoma que processa esses dados genômicos em grande escala para cordados e organismos modelos, resume esses dados e disponibiliza maneiras de acessar essas informações. Suas anotações descrevem localização de genes e transcrições, evolução da sequência gênica, evolução do genoma, sequência e variantes estruturais e elementos regulatórios (CUNNINGHAM *et al.*, 2015). Além disso o Ensembl fornece um conjunto de ferramentas de *software* de processamento de dados, como Preditor de efeito variante, BLAST (*Basic Local Alignment Search Tool*) / BLAT (*BLAST-Like Alignment Tool*), conversor de montagem, conversor de histórico de ID e entre outros (CHEN; HUANG; WU, 2017).

4.3.1.9 Expression Atlas

O *Expression Atlas* é um banco de dados que fornece padrões de expressão de genes, proteínas e *splicing* alternativo em diferentes tipos de células, partes de organismos, condições biológicas e experimentais. Os dados de alta qualidade importados para este banco de dados, são curados, anotados e processados usando métodos de análise padronizados para detectar os padrões de expressão sob as condições experimentais originais. O *Expression Atlas* possui dois componentes importantes, o *Baseline Atlas* e o *Differential Atlas*. O *Baseline Atlas* trata de genes e seu padrão de expressão em condições “normais” utilizando apenas de dados obtidos de RNA-Seq. Já o *Differential Atlas* trata de genes que são suprarregulados e infrarregulados em condições biológicas (PETRYSZAK *et al.*, 2014).

4.3.1.10 Gene Expression Omnibus (GEO)

O *Gene Expression Omnibus* (GEO) é um repositório de dados genômicos funcionais que possui tanto conteúdo baseados em arranjos quanto baseados em sequencias. O GEO possui um grande número de experimentos e dados de expressão gênica curados, seus dados são amplamente explorados em estudos de reposicionamento de drogas, comumente é utilizado para procurar por relações entre drogas e doenças comparando sistematicamente os perfis de expressão de genes de

drogas e de genes de doenças (EDGAR; DOMRACHEV; LASH, 2002; HU; AGARWAL, 2009; KIM *et al.*, 2019).

4.3.1.11 Gene Ontology (GO)

O *Gene Ontology* (GO) tem o intuito de através da bioinformática criar uma representação computacional consistente das funções dos genes a nível celular, molecular e tecidual de todos os organismos. Fornece um vocabulário de termos controlados (ontologias) para descrever produtos gênicos em termos de processos biológicos, componentes celulares e funções moleculares associadas. O GO permite uma consulta uniforme em conjunto de uma associação de muitos bancos de dados biológicos, no *site* é possível pesquisar anotações para produtos genéticos e metadados em várias espécies e realizar análises de enriquecimento utilizando os termos GO (BLAKE *et al.*, 2015).

4.3.1.12 InterPro

O InterPro é um recurso integrado de “assinaturas” ou modelos preditivos que representam domínios de proteínas, famílias, regiões, repetições e locais dos principais bancos de dados de assinaturas de proteínas, como CATH-Gene3D, HAMAP, PANTHER, Pfam, PIRSF, PRINTS, ProDom, PROSITE, SMART, SUPERFAMILY e TIGRFAMs. Todas as entradas no banco de dados é anotada com um nome abstrato descritivo, após é feito um cruzamento de referências nas fontes de dados originais (CHEN; HUANG; WU, 2017; MITCHELL *et al.*, 2015).

4.3.1.13 Kyoto Encyclopedia of Genes and Genomes (KEGG)

O *Kyoto Encyclopedia of Genes and Genomes* (KEGG) é uma base de dados utilizada para análise sistemática das funções do gene, ligando informações genômicas com informações funcionais. O KEGG possui três bancos de dados disponíveis, são eles: GENES, PATHWAY e LIGAND, cada um possuindo uma classificação especifica de dados. O GENES armazena toda a informação genômica, possui uma coleção de catálogos de genes para todos os genomas completamente sequenciados e também de

alguns genomas parciais que possuem anotações de funções atualizadas. O PATHWAY armazena todas as informações funcionais de ordem superior, possuindo também representações gráficas de processos celulares, como metabolismo, transporte de membrana, transdução de sinal e ciclo celular. O terceiro banco de dados é o LIGAND, este no qual armazena informações sobre compostos químicos, enzimas e reações enzimáticas (KANEHISA, 2000).

4.3.1.14 Pfam

O Pfam é um banco de dados de famílias de proteínas representados por alinhamentos de múltiplas sequências e por modelos ocultos de Markov (HMMs). O Pfam classifica suas entradas como Família (proteínas com regiões relacionadas), Domínio (unidade estrutural da proteína), Repetição (múltiplas unidades estruturais de proteínas de tamanho pequeno) e Motivos (pequena unidade estrutural de proteína fora dos domínios globais). As entradas Pfam que são relacionadas são agrupadas em grupos usando como base sua sequência, estrutura ou similaridade de perfil HMM (CHEN; HUANG; WU, 2017; FINN *et al.*, 2014).

4.3.1.15 PROSITE

PROSITE é um banco de dados de identificação e anotação de regiões conservadas em sequências de proteínas, essas regiões são identificadas utilizando dois tipos de assinaturas, por perfis gerais (matrizes de peso) que descrevem as famílias das proteínas e seus domínios modulares e por padrões (expressões regulares), que descrevem motivos de sequências curtas que na maioria das vezes corresponde a funcionalidade ou a resíduos estruturais importantes. O PROSITE fornece uma extensa documentação, como informações de nomenclatura, função, características de sequência, ponteiros para estrutura 3D, arquiteturas de proteínas nas quais a assinatura é encontrada, sua distribuição taxonômica e referências bibliográficas (SIGRIST *et al.*, 2013).

4.3.1.16 Protein Data Bank (PDB)

O *Protein Data Bank* (PDB) foi criado em 1971 e é o único repositório global de estruturas tridimensionais (3D) que foram experimentalmente determinadas de macroméculas biológicas e seus complexos. É administrado pela organização *Worldwide Protein Data Bank* (wwPDB), que inclui o *RCSB Protein Data Bank* (RCSB PDB), o *Protein Data Bank Japan* (PDBj), o *Protein Data Bank Europe* (PDBe) e o *BioMagResBank* (BMRB). A junção de todos esses quatro servidores operam um sistema de *software* global unificado que impõe padrões aos dados de acordo com a comunidade e oferece suporte a deposição, biocuração e validação de dados (BURLEY *et al.*, 2017).

4.3.1.17 Reference Sequence (RefSeq)

O *Reference Sequence* (RefSeq) é um banco de dados de nucleotídeos e sequencias de proteínas com uma grande anotação bibliográfica, criado pelo NCBI. Fornece sequências curadas e não redundantes de regiões genômicas, transcritos e proteínas para diversos organismos taxonômicos, como Archaeas, Bactérias, Vírus e Eucariotos. As sequências do RefSeq contemplam regiões de codificação, domínios conservados, variações, anotações aprimoradas como publicações, nomes, símbolos, identidades de gene e referências cruzadas de banco de dados (BARRETT *et al.*, 2015; CHEN; HUANG; WU, 2017).

4.3.1.18 The Cancer Genome Atlas (TCGA)

O *The Cancer Genome Atlas* (TCGA) é um banco de dados que tem por objetivo utilizar dados multiômicos de larga escala e traçar o perfil de pacientes com câncer para descobrir suas alterações moleculares com intuito de melhorar o diagnóstico, tratamento e prognóstico. Possui dados genômicos, transcriptômicos, proteômicos e de metilação do DNA de milhares de amostras de tecidos de pacientes. O conjunto desses dados fornece uma alta abrangência de alterações genéticas e regulação de transcrição em variados tipos de câncer, além disso, o TCGA fornece informações clínicas dos

pacientes, como detalhes do tratamento e tempo de sobrevida, fazendo com que seja possível a estratificação do paciente, a análise da resposta ao medicamento e a identificação do marcador de prognóstico (BRENNAN *et al.*, 2014; LI *et al.*, 2019; TCGA RESEARCH NETWORK, 2008).

4.3.1.19 The Genetic Association Database (GAD)

O *Genetic Association Database* (GAD) é um banco de dados de estudos de associação genética de doenças e distúrbios genéticos humanos, seu objetivo é acompanhar o usuário para identificar algum polimorfismo que seja clinicamente relevante. Fornece um repositório online amplo de dados de parâmetros clínicos e moleculares de mais de 5000 estudos de associação genética humana, além disso permite a análise sistemática do complexo comum de doenças genéticas humanas (BECKER *et al.*, 2004; LI *et al.*, 2019).

4.3.1.20 The Genomics of Drug Sensitivity in Cancer (GDSC)

O *Genomics of Drug Sensitivity in Cancer* (GDSC) é um dos maiores bancos de dados públicos disponíveis, nele é possível obter informações sobre a resposta de várias drogas em relação a sensibilidade das células cancerosas e marcadores moleculares. O GDSC na data de sua criação, contava com dados sobre a sensibilidade de mais de 75.000 experimentos, descrevendo a resposta a 138 medicamentos anticâncer em quase 700 tipos de linhagens celulares. Para identificar marcadores moleculares que possuem resposta a droga, utilizam se do agrupamento de dados de sensibilidade a drogas testadas em linhagens celulares com grandes conjuntos de dados genômicos obtidos do banco de dados chamado de Catálogo de Mutações Somáticas no Câncer, incluindo informações sobre mutações somáticas em genes de câncer, amplificação e deleção de genes, tipo de tecido e dados de transcrição (YANG *et al.*, 2013).

4.3.1.21 Therapeutic Target Database (TDD)

O *Therapeutic Target Database* (TTD) é um banco de dados que contém informações sobre proteínas conhecidas que possuem alguma capacidade terapêutica,

também possui informações sobre alvos de ácidos nucleicos, doença alvo, informações sobre vias de sinalização e drogas no qual podem ser diretamente enviadas para cada um desses alvos. Também possuem *links* para outros bancos de dados que possam ser relevantes para pesquisas, como banco de dados com estruturas 3D, propriedades de ligação do ligante, nomenclatura enzimática e estrutura do medicamento, classe terapêutica e status de desenvolvimento clínico (CHEN; JI; CHEN, 2002; LI *et al.*, 2019, 2018b).

4.3.1.22 UniProt

O UniProt possui uma extensa e importante coleção de sequências de proteínas e suas anotações, que vem crescendo continuamente com o passar dos anos, o *site* possui seções onde é caracterizado seus dados. A seção UniProtKB/ Swiss-Prot possui dados curados e revisados manualmente, já a seção UniProtKB / TrEMBL é onde todas as outras sequências são armazenadas, nessa seção as informações não são revisadas, mas possuem um sistema de complemento que gera anotações para as sequências automaticamente. O UniProt também possui cruzamento de referências entre 150 bancos de dados e atua como um hub central para organizar as informações geradas sobre proteínas (BATEMAN *et al.*, 2015).

Tabela 5. Links de acesso aos bancos de dados.

Banco de dados	Função	*Site*
Cancer Cell Line Encyclopedia (CCLE)	Arquivos genômicos.	https://portals.broadinstitute.org/ccle
Catalogue of Somatic Mutations in Cancer (COSMIC)	Mutação somática em câncer de humanos.	https://cancer.sanger.ac.uk/cosmic
ChEMBL	Banco de dados de drogas bioativas	https://www.ebi.ac.uk/chembl/

Clinicaltrials	Banco de dados de estudos clínicos.	https://www.clinicaltrials.gov/
ClinVar	Relações entre variações humanas e fenótipo	https://www.ncbi.nlm.nih.gov/clinvar/
dbSNP	Banco de dados de variações genéticas curtas.	https://www.ncbi.nlm.nih.gov/snp/
DisGeNET	Coleções de genes e variantes associadas a doenças humanas.	https://www.disgenet.org/
DrugBank	Banco de dados de medicamentos e alvos de medicamentos.	https://go.drugbank.com/
Ensembl	Banco de dados de anotação do genoma eucariótico	https://www.ensembl.org/index.html
Expression Atlas	Genes, proteínas e padrões de expressão de variantes de splice.	https://www.ebi.ac.uk/gxa/home
Gene Expression Omnibus (GEO)	Dados genômicos funcionais	https://www.ncbi.nlm.nih.gov/geo/
Gene Ontology (GO)	Ontologia Genética	http://geneontology.org/
Genomics of Drug Sensitivity in Cancer (GDSC)	Banco de dados de compostos anticâncer.	https://www.cancerrxgene.org/
GWAS Catalog	Estudos de associação do genoma.	https://www.ebi.ac.uk/gwas/
InterPro	Recurso integrado de famílias de proteínas,	https://www.ebi.ac.uk/interpro/

	domínios e locais funcionais.	
Kyoto Encyclopedia of Genes and Genomes	Genes associados ao câncer.	https://www.genome.jp/kegg/
MetaCyc	Vias metabólicas elucidadas e enzimas.	https://metacyc.org/
Mouse Genome Informatics	Dados genômicos, genéticos e biológicos de rato de laboratório.	http://www.informatics.jax.org/
Naturally Occurring Plant-based Anti-cancer Compound-Activity-Target database (NPACT)	Banco de dados com curadoria de compostos naturais derivados de plantas que exibem atividade anticancerígena.	https://www.npact.org/
NCI60	Grupo de 60 linhagens celulares de câncer humano usadas para a triagem de compostos.	https://discover.nci.nih.gov/cellminer/
neXtProt	Banco de dados de proteínas humanas.	https://www.nextprot.org/
Online Mendelian Inheritance in Man (OMIM)	Relação entre genes e fenótipos genéticos em relação a distúrbios.	https://www.omim.org/
PeptideAtlas	Bancos de dados proteômicos.	http://www.peptideatlas.org/
Pfam	Banco de dados de famílias de proteínas Pfam.	http://pfam.xfam.org/

PharmGKB	Farmacogenômica	https://www.pharmgkb.org/
PROSITE	Banco de dados de domínios de proteínas, famílias e locais funcionais, bem como padrões e perfis associados.	https://prosite.expasy.org/
Protein Data Bank	Proteínas em 3D	https://www.rcsb.org/
RefSeq	Banco de dados de sequência de referência do NCBI	https://www.ncbi.nlm.nih.gov/refseq/
The Cancer Genome Atlas (TCGA)	Dados genômicos (por exemplo, Exoma, SNP, Metilação, mRNA, miRNA, clínico) de amostras de tecido de paciente.	https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga
The Connectivity Map (Cmap)	Perfis de expressão gênica de drogas em linhagens celulares primárias.	https://portals.broadinstitute.org/cmap/
The Genetic Association Database (GAD)	Estudos de associação genética humana de doenças e distúrbios.	https://geneticassociationdb.nih.gov/
The Genomics of Drug Sensitivity in Cancer (GDSC)	Informações sobre sensibilidade a drogas em células cancerosas	https://www.cancerrxgene.org/
The Orthologous Matrix	Método e banco de dados que infere relações evolutivas entre proteomas completos.	https://omabrowser.org/oma/home/

Therapeutic Target Database	Banco de dados contendo informações sobre proteínas terapêuticas e alvos de ácido nucleico.	http://db.idrblab.net/ttd/
UniProt	Proteínas e anotações	https://www.uniprot.org/

Fonte: Adaptada de CHEN; HUANG; WU, 2017; LI et al., 2019.

4.3.2 Alinhamento de sequências

O alinhamento de sequências (AS) normalmente é a primeira etapa realizada na bioinformática quando se pretende entender a filogenia molecular de uma sequência desconhecida. O processo é realizado alinhando a sequência desconhecida com uma ou mais sequências conhecidas de um banco de dados, assim será possível prever as partes em comuns, pois seus resíduos desempenham um papel funcional e estrutural e tendem a ser preservados pela seleção natural durante o curso evolucionário da espécie (THOMPSON *et al.*, 2011).

Para se ter um alinhamento ideal, deve-se organizar duas ou mais sequências de forma que o número máximo de resíduos idênticos ou semelhantes sejam correspondidos. As sequências utilizadas para o alinhamento são de nucleotídeos (DNA ou RNAs) ou de aminoácidos (proteínas), este processo de rearranjo pode ocasionar em um ou mais espaços ou lacunas (*gap*) no alinhamento. Quando ocorre um *gap*, é sinal de que houve uma possível perda ou ganho de resíduo, assim através do alinhamento de sequências pode se observar eventos evolutivos como inserção ou deleção (*indel*), translocações e inversões (CHOWDHURY; GARAI, 2017).

O alinhamento é uma etapa essencial após o sequenciamento, uma vez que sequenciar um fragmento longo utilizando sequenciadores de nova geração (NGS, do inglês *Next Generation Sequencing*) geralmente podem produzir muitas sequências curtas (<~200 bases). Por isso deve se alinhar essas leituras com uma sequência de referência para assim poder encontrar a parte correspondente de cada leitura sequenciada dentro da sequência de referência. O alinhamento também é utilizado para a anotação gênica, pois pode se analisar leituras curtas de RNA-seq derivadas de ácido ribonucleico mensageiro (mRNA) e mapeá-las no genoma de referência. Existem dois

tipos de alinhamentos de sequências, o alinhamento de sequência de pares (ASP) e o alinhamento de sequência múltipla (ASM), o ASP utiliza duas sequências por vez, enquanto o ASM utiliza várias sequências relacionadas, assim apresentando vantagem em relação ao ASP, pois consegue considerar vários membros de uma família de sequências, assim gerando mais informações biológicas (CHOWDHURY; GARAI, 2017; KUMAR; FILIPSKI, 2007).

Tabela 6. Programas gratuitos ou com licença acadêmica de alinhamento de sequências.

Software	Tipo de alinhamento	*Site*
AlignMe	Alinhamento em pares	http://www.bioinfo.mpg.de/AlignMe/
Clustal Omega	Alinhamento múltiplo	https://www.ebi.ac.uk/Tools/msa/clustalo/
MAFFT	Alinhamento múltiplo	https://mafft.cbrc.jp/alignment/software/
MEGA	Alinhamento múltiplo	https://www.megasoftware.net/
MUSCLE	Alinhamento múltiplo	https://www.ebi.ac.uk/Tools/msa/muscle/
Stretcher	Alinhamento em pares	https://www.ebi.ac.uk/Tools/psa/emboss_stretcher/
T-Coffee	Alinhamento múltiplo	http://tcoffee.crg.cat/

4.3.3 Montagem de sequências

A montagem de sequências do genoma é um processo computacional utilizado para decifrar a composição da sequência do DNA que está dentro de uma célula de um organismo. Utiliza-se de várias pequenas sequências chamadas de *reads* que são derivadas de diferentes porções do DNA alvo. A montagem do genoma é realizada logo após o seu sequenciamento, assim garantindo a reconstrução do genoma a partir das leituras geradas pelo sequenciador. Cada nucleotídeo de uma fita pode ser representado por uma letra do alfabeto (a para adenina, c para citosina, g para guanina e t para timina) no qual são sempre expressas na direção 5’ para 3’, o comprimento do gênoma é medido pelo número de seus pares de bases (bp). O comprimento do gênoma pode variar dependendo do organismo, por exemplo, micróbios podem ter alguns milhões de pares de bases, já organismos eucarióticos podem passar de bilhões de pares de base (ARMSTRONG *et al.*, 2019; KALYANARAMAN *et al.*, 2011).

Uma estratégia bem aceita é a de *Whole Genome Shotgun* (WGS) no qual múltiplas cópias de um genoma longo são fragmentadas aleatoriamente em várias frações de comprimento sequenciável e as leituras correspondentes sequenciadas individualmente utilizando qualquer tecnologia padrão. Porém, a informação ordenada entre as leituras de sequências é perdida durante o sequenciamento de forma total ou parcial. Portanto, a informação primária que os programas montadores de gênoma devem utilizar é a sobreposição de sequências de ponta a ponta das sequências preservadas entre as leituras que foram sequenciadas de regiões sobrepostas ao longo do gênoma alvo. Para aumentar a chance de sobreposição, o genoma alvo é normalmente sequenciado de forma redundante (KALYANARAMAN *et al.*, 2011).

Tabela 7. Programas gratuitos ou com licença acadêmica de montagem de sequências.

Software	Função	*Site*
3D-DNA	Criar contigs usando leituras Hi-C.	https://github.com/aidenlab/3d-dna
A5	Automatizar o pipeline de montagem do genoma.	https://sourceforge.net/projects/ngopt/files/
ALE-Assembly	Avaliar a precisão de montagens sem a necessidade de um genoma de referência.	https://github.com/sc932/ALE
ELOPER	Pré-processa leituras curtas emparelhadas para um melhor desempenho durante a montagem.	https://sourceforge.net/projects/eloper/files/
GRAbB	Montar loci genômicos específicos usando regiões iscas para encontrar leituras correspondentes	https://github.com/b-brankovics/grabb
HINGE	Montagem de genoma *de novo* que aborda o desafio de usar leituras longas sujeitas a erros.	https://github.com/HingeAssembler/HINGE
Kermit	Usar mapas de ligação para orientar a montagem do genoma.	https://github.com/rikuu/kermit
PAGIT	Melhorar a qualidade de genomas já montados	https://www.sanger.ac.uk/tool/pagit/
RAMPART	Montagem *de novo* implementada como um sistema de gerenciamento de fluxo de trabalho	https://github.com/TGAC/RAMPART

	que identifica automaticamente montadores adequados de acordo com os dados sequenciados dos usuários.	
SGA (String Graph Assembler)	Montagem de genoma *de novo* e é capaz de usar leituras curtas para montagem sem a necessidade do gráfico de Bruijn	https://github.com/jts/sga
Velvet	Montagem *de novo* baseada em gráficos de Bruijn e é adequada para dados de leitura curta com alta cobertura.	https://github.com/dzerbino/velvet
VGA	Reunir genomas virais individuais a partir de uma amostra que consiste em diversas populações de vírus.	http://genetics.cs.ucla.edu/vga/

Fonte: Bioinformatics Home (https://bioinformaticshome.com/tools/wga/wga.html)

4.3.4 Anotação de genoma

A anotação de genomas é um processo de identificação de elementos genômicos e suas funções. Inicialmente este termo era utilizado para estruturas de genes (mRNAs) em genomas, mas atualmente ele tem sido empregado em um amplo sentido para quaisquer elementos genômicos. As anotações veem demonstrando uma variedade de elementos funcionais encontrados em genomas humanos, funções bioquímicas de segmentos genômicos e complexidade inesperada de transcritos como isoformas variáveis de mRNA e RNAs de codificação não proteica. (ABUGESSAISA; KASUKAWA; KAWAJI, 2017).

A ideia de anotar genomas de forma automática foi visada desde que os primeiros genomas completos foram lançados em meados da década de 1990 (LETOVSKY, 1998). Normalmente o processo de anotação é dividido em duas categorias: previsão *ab initio,* ou previsão computacional da estrutura exon-intron utilizando modelos estatísticos e abordagens em alinhamento de sequência, que mapeiam qualquer marcador de sequência expressa, DNA complementar (cDNA) ou sequências de proteínas em uma sequência montada com intuito de descobrir seus transcritos (AKEN *et al.*, 2016; ARMSTRONG *et al.*, 2019).

A anotação do genoma além de ser essencial para a interpretação das sequências do genoma também é uma área de pesquisa ativa no campo da genômica, ou seja, a plena utilização das anotações é essencial tanto para pesquisas abrangentes de “lógica” e biologia de sistemas, quanto para outras áreas, como biologia celular, biologia do desenvolvimento e pesquisa translacional, podendo melhorar o processo de seleção de alvos moleculares como genes-chaves, alvos de medicamentos, biomarcadores, e muito mais. Também pode auxiliar na abordagem de sistemas subjacentes, como vias de sinalização e patogênese (ABUGESSAISA *et al.*, 2014; ABUGESSAISA; KASUKAWA; KAWAJI, 2017).

A anotação do genoma pode se dividir em anotações estruturais e anotações funcionais. As anotações estruturais têm o objetivo de encontrar as características do DNA, por exemplo, íntrons, exons, promotores e transposons. A anotação funcional é a associação de informações biológicas nas sequências de gene ou proteínas que foram identificadas por anotação estrutural, ou seja, identifica a função biológica da sequência (EJIGU; JUNG, 2020).

Tabela 8. *Software*s para anotação de genomas estruturais.

Software	Função	*Site*
MAKER2	Aplicativo de gerenciamento de dados e anotação de genoma paralelizado e multithread.	http://www.yandell-lab.org/software/maker.html
BRAKER1	Pipeline de anotação de genoma baseado em RNA-seq não supervisionado totalmente automatizado, altamente preciso para genomas eucarióticos.	http://bioinf.uni-greifswald.de/bioinf/braker/
Comparative Annotation Toolkit (CAT)	Kit de ferramentas de *software* totalmente aberto para anotação do inicio ao fim.	https://github.com/ComparativeGenomicsToolkit/Comparative-Annotation-Toolkit
NCBI Eukaryotic Annotation Pipeline	Pipeline automatizada para eucariotos, no qual genes codificadores e não codificantes, transcritos e proteínas em genomas podem ser anotados.	https://www.ncbi.nlm.nih.gov/genome/annotation_euk/process/

Fonte: EJIGU; JUNG, 2020

Tabela 9. *Software*s para anotação de genomas funcionais.

Software	Função	*Site*
MANATEE	Ferramenta usada para realizar anotações funcionais manuais.	http://manatee.sourceforge.net/

Fonte: EJIGU; JUNG, 2020

4.3.5 Predição de estruturas e funções proteicas

Com a chegada da genômica estrutural, funcional e comparativa, inúmeras sequências de proteínas puderam ser determinadas em uma velocidade altíssima, fazendo com que seus estudos experimentais não conseguissem acompanhar a quantidade de informação gerada, fazendo com que a única forma de anotar essas funções provisórias fosse por meio da análise automática de sequências (GALPERIN *et al.*, 2015).

Métodos experimentais os quais conseguem determinar a estrutura da proteína, como a cristalografia de raios-X ou a ressonância magnética nuclear (NMR) são procedimentos caros e demorados, assim a utilização de métodos computacionais de predição de estrutura de proteína se tornam uma excelente opção para reduzir a lacuna entre o número de sequências de proteínas existentes e o número de estruturas (ÁLVAREZ-MACHANCOSES; FERNÁNDEZ-MARTÍNEZ; KLOCZKOWSKI, 2020; ZHANG, 2008).

Também existe um problema em relação a função dessas proteínas, pois mais de um terço das estruturas proteicas possuem função desconhecida, assim essas proteínas só atingirão seu respectivo valor quando for descoberto os papeis biológicos no que desempenham, devido a grande quantidade de sequências e proteínas não é viável de forma experimental a caracterização dessas estruturas (CRUZ *et al.*, 2017).

Com isso, tem-se utilizado da biologia computacional onde se tem ferramentas que podem fornecer informações sobre a função das proteínas com base em sua sequência, estrutura, história evolutiva e sua associação com outras proteínas (CRUZ *et al.*, 2017; PUNTA; OFRAN, 2008).

Tabela 10. *Software*s para predição de estrutura de proteínas.

Software	Função	*Site*
CPHmodels	Estrutura 3D por modelagem por homologia	http://www.cbs.dtu.dk/services/CPHmodels/
COILS	Regiões enroladas dentro da proteína	https://embnet.vital-it.ch/software/COILS_form.html
FUGUE	Estrutura 3D por *threading*	https://mizuguchilab.org/fugue/
GlobPlot	Região desordenada	http://globplot.embl.de/
HHPred	Estrutura 3D por modelagem por homologia	https://toolkit.tuebingen.mpg.de/tools/hhpred
HMMTOP	Domínio transmembrana	http://www.enzim.hu/hmmtop/
I-TASSER	Estrutura 3D por *threading*	https://zhanglab.dcmb.med.umich.edu/I-TASSER/
MODELLER	Estrutura 3D por modelagem por homologia	https://salilab.org/modeller/
Phyre2	Estrutura 3D por *threading*	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index
PONDR	Região desordenada	http://www.pondr.com/
PORTER	Estrutura secundária	http://distillf.ucd.ie/porterpaleale/quickhelp.html
PredictProtein	Estrutura secundária e outras	https://predictprotein.org/
PSIPRED	Estrutura secundária	http://bioinf.cs.ucl.ac.uk/psipred/
Robetta	Estrutura 3D por *ab initio*	https://robetta.bakerlab.org/
SABLE	Estrutura secundária e acessibilidade de solventes	https://sable.cchmc.org/
SPARKS	Estrutura 3D por *threading*	https://sparks-lab.org/server/sparks-x/
SWISS-MODEL	Estrutura 3D por modelagem por homologia	https://swissmodel.expasy.org/
TMHMM	Domínio transmembrana	http://www.cbs.dtu.dk/services/TMHMM/

Tabela 11. *Software*s para predição da função de proteínas.

Software	Função	*Site*
BLAST	Busca por homologia	https://blast.ncbi.nlm.nih.gov/Blast.cgi (selecione protein-protein blast)
ELM	Busca pelo motivo funcional em eucariotos	http://elm.eu.org/
FASTA	Busca por homologia	https://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml
Pfam	Identificação de família	http://pfam.xfam.org/
PFP	Predição de função através da mineração de resultados do PSI-BLAST	https://kiharalab.org/web/pfp.php
PROSITE	Busca pelo motivo funcional	https://prosite.expasy.org/
PSI-BLAST	Busca por homologia	https://blast.ncbi.nlm.nih.gov/Blast.cgi (selecione psi- e phi-blast)
PSROT	Predição da localização subcelular	http://www.psort.org/
SMART	Busca pelo motivo conservado	http://smart.embl-heidelberg.de/
STRING	Predição funcional por comparação genômica	https://string-db.org/

4.3.6 Docking Molecular

O *docking* molecular (ancoramento molecular) é uma ferramenta computacional que tenta prever a ligação não covalente de macromoléculas ou de uma macromolécula (receptor) e uma pequena molécula (ligante) com eficiência, mantendo sempre o objetivo de prever as conformações ligadas e a afinidade de ligação. A predição da ligação de pequenas moléculas a uma proteína é de grande importância pratica, pois é usada para rastrear bibliotecas virtuais de moléculas semelhantes a drogas para obter pistas sobre o desenvolvimento de novas drogas. O docking também pode ser utilizado para prever a conformação ligada de ligantes conhecidos, quando as estruturas experimentais não estão disponíveis (O. TROTT, A. J. OLSON, 2010).

Foi desenvolvido na década de 1970 e desde então tem se mostrado como uma ferramenta importante no auxílio do estudo de como os compostos químicos interagem com seus alvos moleculares, e também para a descoberta e desenvolvimento de medicamentos. Além disso, com o passar dos anos, houve um aumento nos estudos relatando o uso de *docking* molecular para identificar determinantes estruturais necessários para a ligação eficiente do ligante-receptor e o desenvolvimento de métodos de *docking* molecular mais precisos (PINZI; RASTELLI, 2019).

Baseando-se em seu funcionamento, os programas de *docking* molecular geralmente utilizam uma função de pontuação, que pode ser vista como uma tentativa de aproximar os potencias químicos padrão do sistema. Quando os termos utilizados baseados na física são utilizados, como interações de Van der Walls e energias de Coulomb, eles precisam ser significativamente ponderados empiricamente, em parte, para que possa explicar essa diferença entre energias e energias livres (O. TROTT, A. J. OLSON, 2010).

A realização do *docking* molecular envolve a previsão da conformação e orientação do ligante (ou posicionamento) dentro de um sítio de ligação direcionado. Normalmente, ao utilizar o *docking* molecular, se tem como objetivos a modelagem estrutural precisa e a previsão correta da atividade, o processo inicia-se com a aplicação de algoritmos de encaixe que avaliam a pose de pequenas moléculas nos sítios de interesse de uma macromolécula. É um processo desafiador, pois mesmo moléculas orgânicas simples podem conter muitos graus de liberdade conformacionais, a amostragem desses graus de liberdade deve ser realizada com precisão suficiente para

identificar a conformação que melhor corresponde à estrutura do receptor e também precisa ser rápida o suficiente para permitir a avaliação de muitos compostos em somente uma execução do teste (KITCHEN *et al.*, 2004).

Os algoritmos utilizados no processo de docking molecular são complementados por funções de pontuação as quais são projetadas para prever a atividade biológica por meio da avaliação das interações entre os compostos e os alvos potenciais, as primeiras funções de pontuação avaliadas combinam os ajustes com base em cálculos de forma aproximada e complementariedades eletrostáticas. As funções de pontuação relativamente simples são utilizadas principalmente em estágios iniciais das simulações. Para testes mais avançados são utilizados esquemas de pontuação mais complexos com o tratamento mais detalhado sobre interações eletrostáticas e forças de Van Der Walls, e o acréscimo de pelo menos alguns efeitos de solvatação ou entrópico (KITCHEN *et al.*, 2004).

Contudo, o *docking* molecular possui limitações como uma amostragem restrita das conformações do ligante e do receptor na previsão da pose, o uso de funções de pontuação aproximadas, estas que podem fornecer resultados que não se correlacionam com as afinidades de ligação experimentais. Também é importante salientar que para a aplicação do *docking* molecular no projeto de fármacos é limitada a alvos biológicos que possuem estruturas cristalinas, para contornar isso foram adotadas técnicas de construção de modelos por homologia derivados de modelos estruturais com sequências altamente homólogas (PINZI; RASTELLI, 2019).

Atualmente o *docking* molecular vem sendo empregado em diversas áreas e trabalhos científicos, um exemplo disso foi no estudo de Couto e colaboradores (2020), no qual utilizou a técnica para avaliar a força de interação de porfirinas de platina com receptores de LDL com intuito de utilizá-los de forma a carrear essa moléculas ao alvo, no tratamento fotodinâmico para o melanoma.

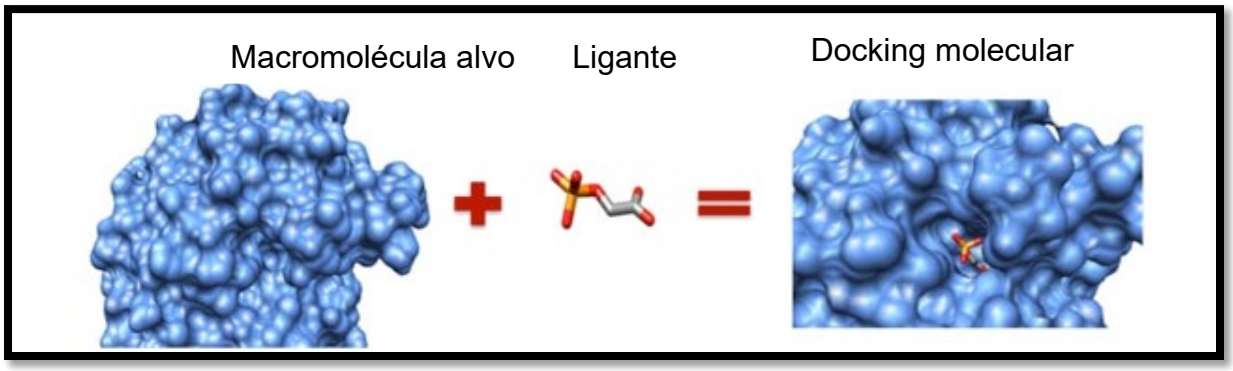

Figura 7. Esquema de docking molecular. **Fonte:** Adaptado de Creative Proteomics (https://www.creative-proteomics.com/services/molecular-docking-service.htm)

Tabela 12. *Site*s e programas gratuitos ou com licença acadêmica de docking molecular.

Software/*Web server*	*Site*
1-Click Docking	https://mcule.com/apps/1-click-docking/
Autodock/ Autodock Vina	http://autodock.scripps.edu/
BetaDock	http://voronoi.hanyang.ac.kr/software.htm
BSP-SLIM	https://zhanglab.ccmb.med.umich.edu/BSP-SLIM/
CABS-dock	http://biocomp.chem.uw.edu.pl/CABSdock
Dock Blaster	https://blaster.docking.org/
DockThor	https://dockthor.lncc.br/v2/
Fitted	http://fitted.ca/
FlexPepDock	http://flexpepdock.furmanlab.cs.huji.ac.il/
FLIPDock	http://flipdock.scripps.edu/
GalaxyPepDock	http://galaxy.seoklab.org/index.html
GEMDOCK	http://gemdock.life.nctu.edu.tw/dock/
HADDOCK 2.4	https://wenmr.science.uu.nl/haddock2.4/
LeDock	http://www.lephar.com/software.htm
LightDock	https://lightdock.org/
MEDock	https://bio.tools/medock
MEdusaDock	https://dokhlab.med.psu.edu/medusadock/#/
MEGADOCK 4.0	https://www.bi.cs.titech.ac.jp/megadock/
OED Docking	https://www.eyesopen.com/oedocking
SwissDock	http://www.swissdock.ch/
UCSF DOCK	http://dock.compbio.ucsf.edu/

4.3.7 Dinâmica Molecular

Os átomos que formam uma biomolécula estão em constante movimento, fazendo com que tanto a sua função molecular quanto suas interações intermoleculares dependem da dinâmica no qual estão envolvidas. Dito isto, é muito mais interessante observarmos essas biomoléculas em ação e não apenas um momento estático das mesmas, além da capacidade de perturbá-las em nível atômico e ver suas devidas respostas ao estímulo. Atualmente, perturbar átomos individuais ou até mesmo observa-los de maneira desejada é difícil, uma alternativa é utilizar de métodos de simulações computacionais de nível atômico como a dinâmica molecular (DM), para observar as biomoléculas de interesse (HOLLINGSWORTH; DROR, 2018).

Segundo a *International Union of Pure and Applied Chemistry* (IUPAC), a "dinâmica molecular é um procedimento de simulação que consiste na computação do movimento dos átomos em uma molécula ou de átomos individuais ou moléculas em sólidos, líquidos e gases, de acordo com as leis de movimento de Newton", ou seja, a dinâmica molecular é um processo que analisa a variação do comportamento molecular em função do tempo. Os comportamentos moleculares de uma molécula são suas propriedades, ou seja, sua conformação de alças, orientação de cadeias laterais, conteúdo de estrutura secundária e a energia de interação entre diferentes moléculas (fármaco e receptor, proteína e DNA, enzima e substrato e proteína e proteína). O fato de que as propriedades das moléculas se alteram com a função do tempo indica que de fato não são estáticas, mas se modificam em soluções biológicas, assim fazendo com que as moléculas se aproximem muito de métodos experimentais como a RMN a qual consegue gerar as médias temporais de uma molécula durante seu experimento. Com isso ao final um ensaio de dinâmica molecular buscamos estas propriedades médias e representativas de comportamentos biológicos medidos experimentalmente (VERLI, 2014).

Em resumo, a ideia de uma simulação de DM é receber as posições de todos os átomos em um sistema biomolecular (como uma proteína cercada por água e por uma bicamada lipídica) e calcular a força exercida em cada átomo por todos os outros átomos. Também pode-se utilizar das leis de Newton para prever a posição espacial de cada átomo em função do tempo, ou seja, se avança no tempo calculando repetidamente as forças em cada átomo e em seguida usando essas forças para atualizar a posição e a

velocidade de cada átomo, por fim se resulta em uma trajetória, um filme tridimensional que descreve a configuração em nível atômico do sistema em cada ponto durante o intervalo de tempo simulado (HOLLINGSWORTH; DROR, 2018).

As forças utilizadas durante uma simulação de DM são calculadas usando um modelo chamado de campo de força da mecânica molecular, este que é adequado aos resultados dos cálculos da mecânica quântica e casualmente a alguns outros experimentos. Na simulação um campo de força típico agrega termos que capturam interações eletrostáticas (Coulombic) entre os átomos, estes termos que podem ser do tipo mola, no qual irão modelar o comprimento preferido de cada ligação covalente e termos que irão capturar vários outros tipos de interações interatômicas (HOLLINGSWORTH; DROR, 2018).

Ao comparar simulações com uma variedade de dados experimentais, tudo indica que os campos de força utilizados na DM melhoraram absurdamente na última década , mas mesmo assim permanecem imperfeitos e isto deve ser levado em conta na hora de analisar os resultados (LINDORFF-LARSEN *et al.*, 2012).

Também é importante salientar que durante uma simulação DM clássica, nenhuma ligação covalente deve se formar ou se romper, nesses casos utilizam-se simulações de mecânica quântica / mecânica molecular, nas quais uma pequena parte do sistema é modelada usando cálculos de mecânica quântica e o restante por simulação DM, para estudar as reações que envolvam mudanças nas ligações covalentes ou são impulsionadas pelo absorção de luz (HOLLINGSWORTH; DROR, 2018; SENN; THIEL, 2009).

A DM pode ser empregada em diversas pesquisas, principalmente na busca de novos medicamentos contra o câncer, assim como também pode ser utilizada na busca ou melhoramento de diagnósticos para o câncer. Um exemplo disso foi feito por Jeevanandam e colaboradores (2020), que utilizou da DM para otimizar os marcadores apta (biomarcação mediada por aptâmero) e melhorar o seu direcionamento a receptores específicos de câncer e de superfície tumoral com intuito de desenvolver novos e aprimorados métodos de diagnóstico e tratamento para o câncer.

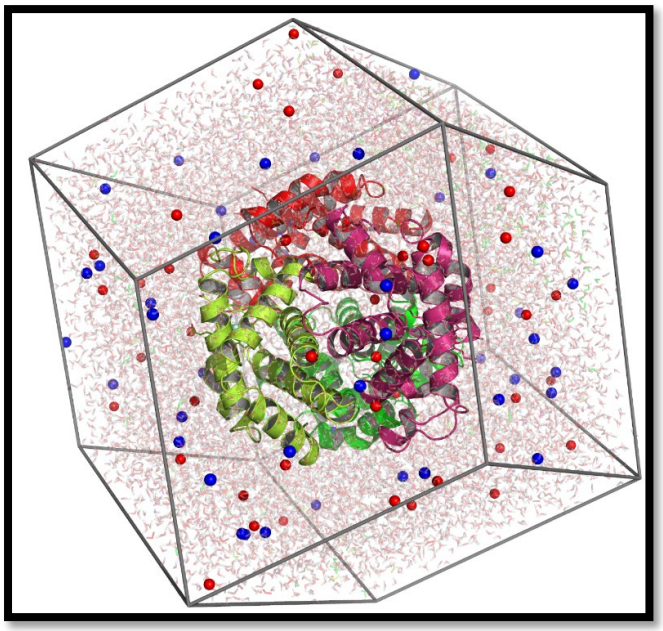

Figura 8. Demonstração de uma simulação de macromoléculas na água em uma caixa no formato dodecaedro. **Fonte:** HUB; KUBITZKI; DE GROOT, 2010

Tabela 13. *Softwares* gratuitos ou com licença acadêmica de dinâmica molecular e simulações computacionais.

Software	*Site*
Ascalaph Designer	http://www.biomolecular-modeling.com/Products.html
Avogadro	http://avogadro.cc/
CP2K	https://www.cp2k.org/
Desmond	https://www.deshawresearch.com/resources_desmond.html
Fold.it	http://fold.it/portal/
FoldX	http://foldxsuite.crg.eu/
GROMACS	http://www.gromacs.org/
LAMMPS	https://lammps.sandia.gov/
MDynaMix	http://www.fos.su.se/~sasha/mdynamix/
Orac	http://www1.chim.unifi.it/orac/
NAMD + VMD	http://www.ks.uiuc.edu/Research/vmd/
NWChem	https://nwchemgit.github.io/
Q	https://www.icm.uu.se/cbbi/aqvist-lab/q/
TINKER	https://dasher.wustl.edu/tinker/
UCSF	http://www.cgl.ucsf.edu/chimera/

4.3.8 Absorção, Distribuição, Metabolização e Excreção (ADME)

A busca e o desenvolvimento de novos medicamentos é um empreendimento complexo, oneroso e de alto custo, pois demanda a seleção de doenças, identificação e validação de alvos, otimização de dados, testes pré-clínicos e clínicos. Com o desenvolvimento de novas ferramentas *in silico* o número de drogas aprovadas por órgãos regulamentadores vem aumentando significativamente. Porém, ainda existem várias moléculas promissoras que que podem se tornar fármacos, mas que não conseguem atingir esse objetivo devido à falta de eficácia e segurança (GUAN *et al.*, 2019).

Para um composto provocar algum tipo de efeito biológico (seja terapêutico ou tóxico) depende da atividade intrínseca do composto e do seu potencial de atingir o local de ação em concentrações suficientes pelo período de tempo necessário, com isso a probabilidade de um composto atingir qualquer local depende da exposição externa e interna. A exposição externa é a possibilidade de o composto atingir um organismo, ou seja, a concentração da molécula presente no mesmo ambiente do organismo, já a exposição interna é relacionada com a absorção, distribuição do composto aos locais de ação dentro de um organismo, metabolização e excreção (ADME) do mesmo. Esses fatores irão determinar o perfil de concentração e tempo no qual um determinado composto se mantenha no local desejado ou no corpo como um todo. Dessa forma, os fatores ADME são de grande importância na determinação da resposta *in vivo* a um poderoso agente terapêutico ou a um tóxico, pois eles podem ter efeito modular significativo na atividade geral (PRZYBYLAK *et al.*, 2018).

Para uma molécula ser ideal ela precisa estar em conformidade com as diretrizes de propriedade físico-química da Regra de Cinco de Lipinski (RO5), o ADME utiliza esta regra no qual prevê a semelhança de um composto químico com uma determinada atividade biológica projetada para a via de administração oral. De acordo com o RO5, um composto semelhante a um medicamento deve ter um peso molecular (MW) de <500 g / mol, um valor log p de <5 representando sua hidrofobicidade, possuir >5 doadores de ligação de hidrogênio (HBDs) e >10 sítios aceitadores de ligação de hidrogênio (HBA). Futuramente foram acrescentadas mais duas regras novas; área de superfície polar (PSA) de £ 140 A˚ e <10 ligações rotativas (Rot B), essas novas regras estão correlacionadas com a permeabilidade e flexibilidade do fármaco, respectivamente. Se

o composto atende aos cinco princípios, ele terá propriedades farmacocinéticas aprimoradas e maior biodisponibilidade no processo metabólico do organismo, assim podem ser aprovados (CHEN *et al.*, 2020).

Sendo assim o ADME se torna de conhecimento crítico no desenvolvimento de novos medicamentos e na avaliação do risco representado por compostos, como é o caso de pesticidas, poluentes ambientais, aditivos alimentares e produtos químicos industriais, esses no qual um organismo pode ser exposto (HOU; WANG, 2008; PRZYBYLAK *et al.*, 2018).

Tabela 14. *Site*s e programas gratuitos ou com licença acadêmica de análise de ADME.

Software/*Web service*	*Site*
admetSAR2	http://lmmd.ecust.edu.cn/admetsar2/
CAESAR/VEGA	http://www.caesar-project.eu/index.php?page=links
Chem Prop (OSIRIS EDITION)	https://www.ufz.de/index.php?en=34593
EPI Suite™	https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface
Lazar	https://lazar.in-silico.ch/predict
Mollispiration	https://www.molinspiration.com/
NERDD	https://nerdd.univie.ac.at/
OpenTox platform: ToxPredict	https://toxbank.net/about/opentox
OSIRIS property explorer	https://www.organic-chemistry.org/prog/peo/
ProTox-II	https://tox-new.charite.de/protox_II/index.php?site=compound_input
SMARTCyp	https://smartcyp.sund.ku.dk/mol_to_som
SwissADME	http://www.swissadme.ch/
T.E.S.T.	https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test
ToxTree	http://toxtree.sourceforge.net/download.html#Toxtree_3.1.0
VCCLAB	http://www.vcclab.org/

Fonte: Adaptado de HU et al., 2019

4.3.9 PyMOL

O PyMOL é um sistema gráfico molecular que ao longo dos anos vem evoluindo constantemente, hoje não é somente considerado um poderoso e excepcional visualizador de moléculas e recursos 3D como também é uma plataforma que agrega vários programas e aplicativas que tornam as propriedades de visualizações do PyMOL mais versáteis. Utilizando de uma arquitetura multicamadas em conjunto com a poderosa linguagem de programação Python, o PyMOL é relativamente fácil de se expandir e personalizar sem precisar recompilar o seu código-fonte. Suas extensões podem utilizar a interface do assistente ou a interface do *plug-in*, tornando assim seu uso descomplicado. Um exemplo de extensão (*plugin*) é o APBS este que é uma interface para o popular programa *adaptive Poisson-Boltzmann solver*, este no qual fornece fácil acesso a cálculos eletrostáticos e a visualização de superfícies de energia potencial e densidades de carga em superfícies de proteínas (SEELIGER; DE GROOT, 2010). O PyMOL está disponível para *download* no seu *site* oficial (https://pymol.org/2/).

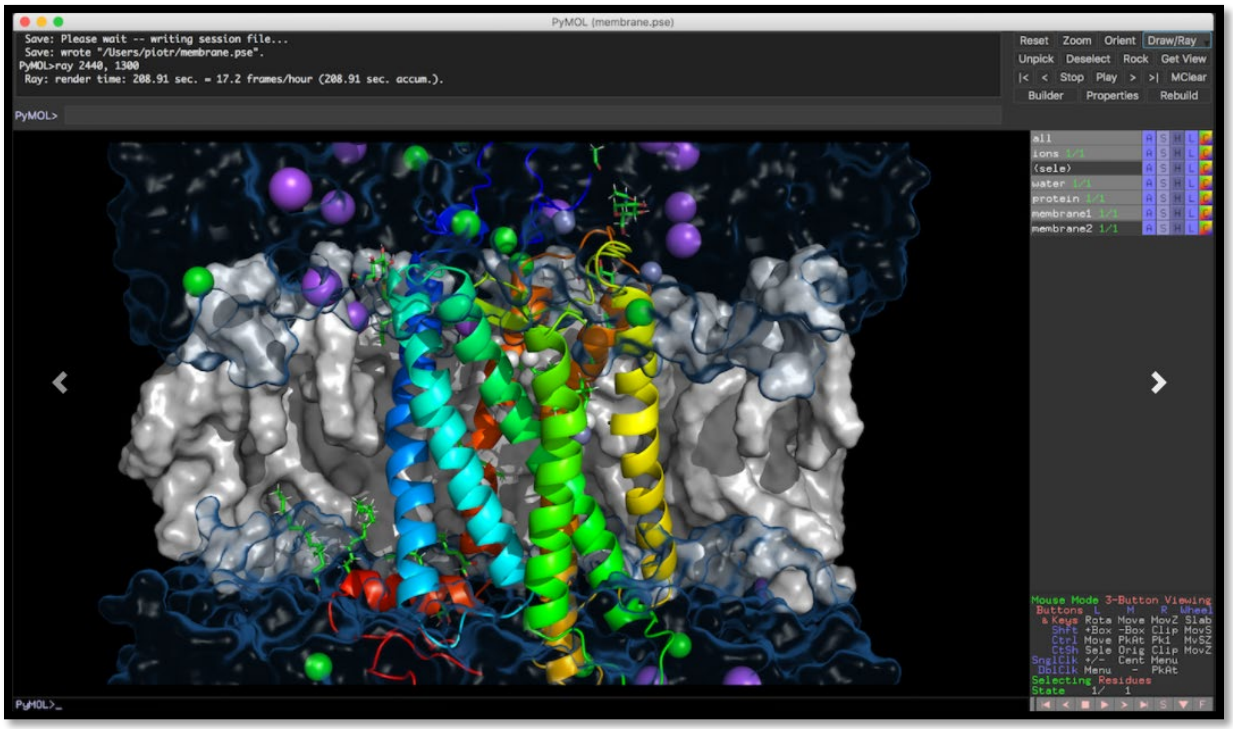

Figura 9. Representação da interface do *software* PyMOL. **Fonte:** *Site* oficial do PyMOL (https://pymol.org/2/#screenshots).

4.3.10 OECD QSAR Toolbox

O OECD QSAR Toolbox (*The Organisation for Economic Cooperation and Development Quantitative Structure- Activity*) é um programa criado para fazer previsões de maneira objetiva da toxicidade de compostos por métodos qualitativos e quantitativos de estrutura-atividade (SCHULTZ *et al.*, 2018). Segundo o *site* oficial do programa o seu objetivo é apoiar a avaliação de perigos de produtos químicos, bem como aumentar o conhecimento sobre substâncias químicas de uma forma econômica, promovendo o uso de métodos de avaliação alternativos aos animais e minimiza testes desnecessários em animais sem reduzir a segurança da saúde humana e do meio ambiente.

O programa consiste de ferramentas que fornece informações sobre produtos químicos, conta com um sistema de busca bem elaborado, principalmente por arquivos padronizados que estão associados a dados químicos e de toxicidade, assim garantindo que análogos estruturas adequados possam ser identificados (SCHULTZ *et al.*, 2018). O OECD QSAR Toolbox pode ser baixado diretamente do seu *site* oficial (https://qsartoolbox.org/).

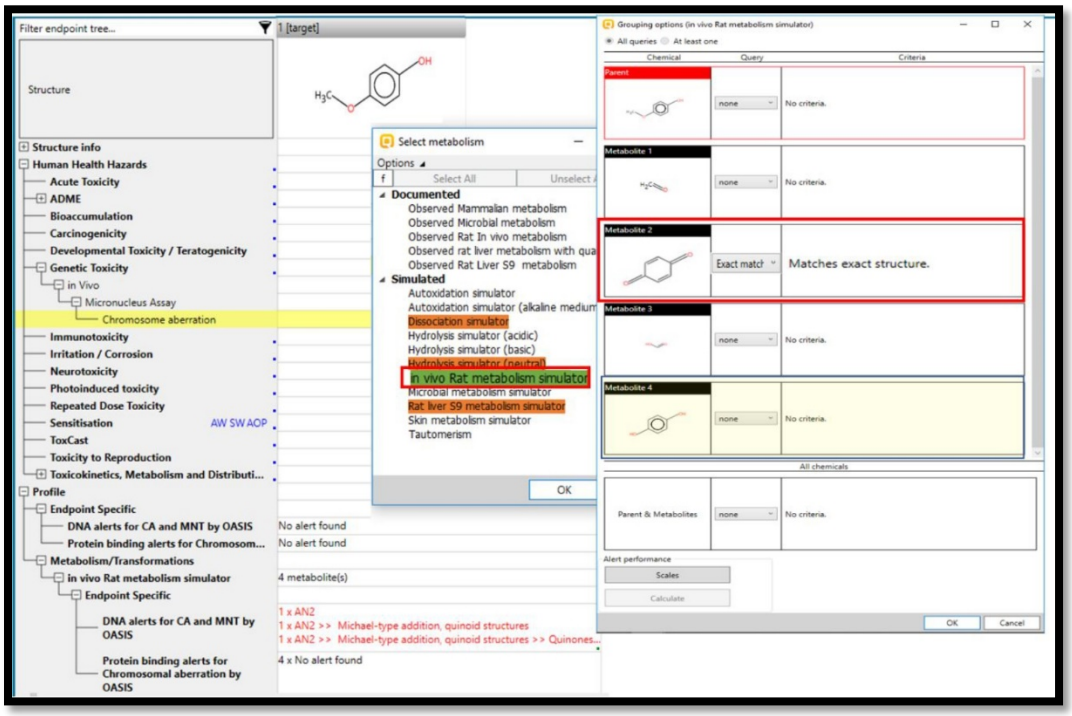

Figura 10. Representação da interface gráfico do *software* QSAR Toolbox. **Fonte:** *Site* oficial (https://qsartoolbox.org/).

4.4 Discussão

O câncer segundo a OMS é a segunda doença com maior taxa de mortalidade no mundo e com o tempo tende a aumentar seus números, mesmo existindo inúmeros tipos de tratamentos como cirurgia, quimioterapia, radioterapia, terapia alvo molecular e entre outras, o câncer continua com seus números altos. Isso ocorre, pois, tratamentos como terapia alvo molecular são caros e grande parte da população mundial não possui condições financeiras para adquirir esse tipo de tratamento, com isso partem para as terapias mais convencionais como cirurgias, quimioterapia e radioterapia.

Porém a quimioterapia, mesmo sendo eficiente em grande parte dos casos, não é absoluta como tratamento, pois existem casos de compostos quimioterápicos que não atingem o efeito esperado durante o tratamento, seja por incompatibilidade genética ou por conta da resistência do câncer. A resistência a quimioterápicos é preocupante em casos de reincidência de tumores, pois como grande parte do mesmo já foi eliminado devido ao tratamento anterior, as partes que ficam ainda têm grandes chances de se tornarem resistente ao tratamento anterior. Com isso é de extrema importância que existam variações de compostos quimioterápicos com intuito de diversificar os tratamentos oncológicos.

Dessa forma a bioinformática é uma ferramenta poderosa que pode ajudar na pesquisa e desenvolvimento de novos medicamentos para se utilizar na quimioterapia ou em outros tratamentos, pois devido ao avanço da tecnologia hoje é possível realizar testes *in silico* com grande taxa de assertividade, assim sendo capaz de reduzir em muito o tempo e os gastos de busca e desenvolvimento de uma nova molécula.

Hoje existem muitas maneiras nas quais a bioinformática pode ajudar na pesquisa de novas moléculas, como a utilização de banco de dados biológicos e de ferramentas como o QSAR, onde é possível fazer a análise de moléculas que já foram testadas anteriormente e que possuem resultados já bem especificados, e a partir desta desenvolver uma nova molécula que seja mais eficiente que as anteriores. Para testar essas novas moléculas desenhadas pode se utilizar de testes de ADME, que realiza a análise de farmacocinética dessas moléculas, o que é de grande importância para se ter segurança em sua utilização.

Após estes testes, outra ferramenta interessante no qual pode se utilizar é o docking molecular, que resulta em uma predição da farmacodinâmica das moléculas

testadas. O docking pode mostrar se as moléculas possuem afinidade por algum receptor (normalmente proteínas) de interesse, caso um alvo molecular já bem elucidado tenha sido definido. Se o alvo não possuir uma estrutura tridimensional cristalográfica, é possível utilizar da bioinformática e tentar predizer a estrutura de interesse utilizando de softwares de modelagem como o SWISS-MODEL, obtendo assim, uma versão tridimensional do alvo molecular (receptor) assim podendo prosseguir com os testes de farmacodinâmica. O docking pode ser utilizado como mais um filtro de moléculas, no qual é possível selecionar as que possuem maior afinidade pelo receptor de interesse, e após a utilização do docking, pode se utilizar dinâmica molecular para filtrar melhor ainda as moléculas. A dinâmica irá simular um ambiente próximo ao biológico, e assim poderá analisar se as moléculas ligadas através do docking realmente possuem ligações estáveis com o receptor.

O processo discutido acima é uma forma, dentre muitas outras possíveis, no qual a bioinformática pode ser utilizada na pesquisa de novos medicamentos, isso pode ser tanto para o câncer quando para outras doenças. Assim é possível utilizar essas plataformas e reduzir o tempo e o gasto de materiais de testes in vitro ou in vivo. Mesmo com o avanço da tecnologia e todas as ferramentas disponíveis, os testes in vitro e in silico ainda são de grande importância para a pesquisa, pois é o que pode simular da melhor maneira possível um sistema biológico. Em suma, a bioinformática tem como função filtrar melhor as moléculas no qual futuramente podem ser analisadas in vitro e in silico e seguir com seus processos de desenvolvimento.

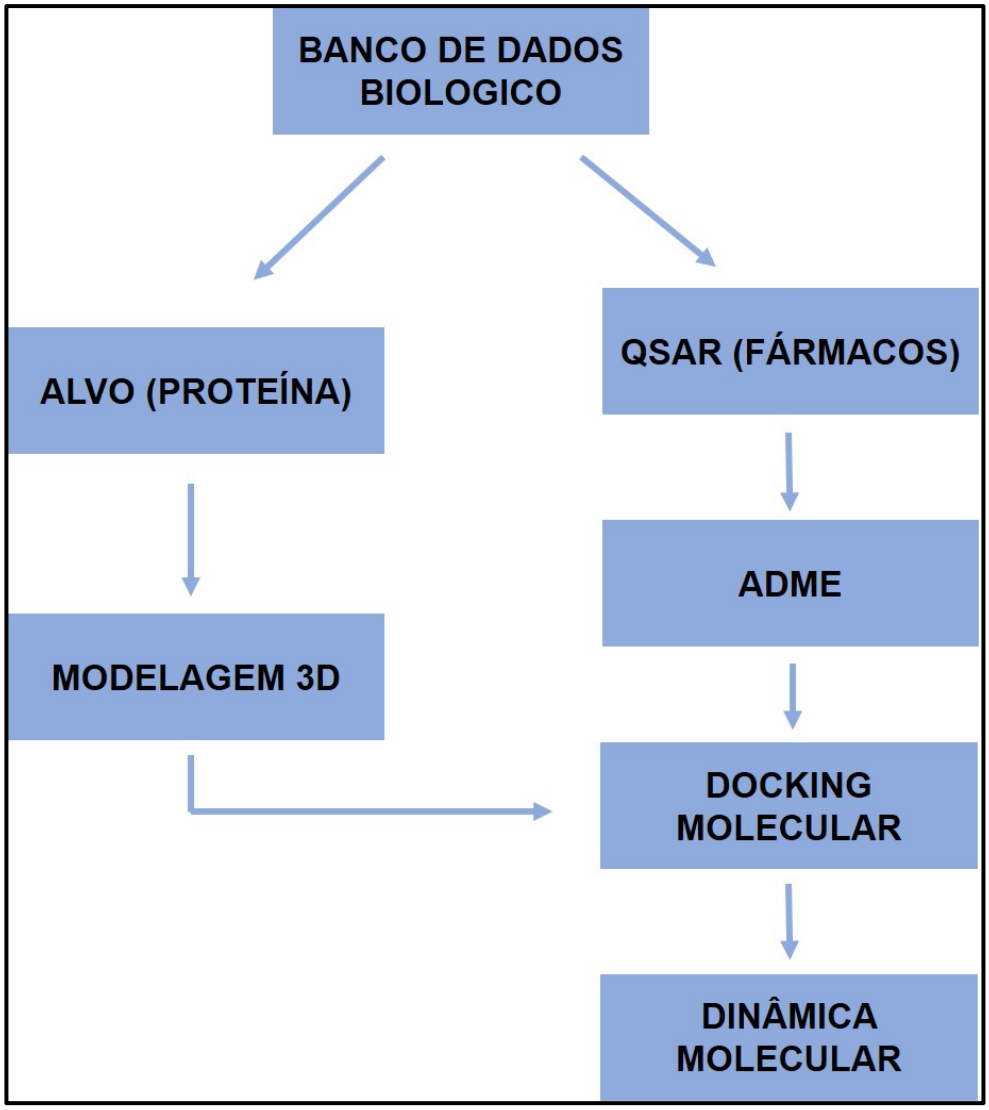

Figura 11. Exemplo de fluxograma da utilização de ferramentas de bioinformática na pesquisa por novos fármacos. **Fonte:** Elaborado pelo autor.

5 CONCLUSÃO

Como destacado no presente trabalho, o câncer é uma das doenças com maiores taxas de incidência e mortalidade existente. O seu tratamento ao longo das décadas vem ficando cada vez melhor e apresentando menos efeitos colaterais, mas mesmo assim, ainda não é o suficiente devido a sua grande capacidade de sofrer mutações, o que pode causar resistência aos tratamentos.

Com o grande avanço da tecnologia as ferramentas da área de bioinformática estão cada vez mais próximas de cientistas e pesquisadores, auxiliando-os em pesquisas, tornando as mesmas mais rápidas, baratas e confiáveis. Ademais, o uso da bioinformática proporciona uma redução significativa do número de animais modelos e da quantidade de reagentes poupados de testes que podem ser feitos por computadores.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ABUGESSAISA, Imad *et al.* Accelerating Translational Research by Clinically Driven Development of an Informatics Platform–A Case Study. **PLoS ONE**, *[S. l.]*, v. 9, n. 9, p. e104382, 2014. Disponível em: https://doi.org/10.1371/journal.pone.0104382

ABUGESSAISA, Imad; KASUKAWA, Takeya; KAWAJI, Hideya. Genome Annotation. *In*: *[S. l.: s. n.]*. v. 1525p. 107–121. *E-book*. Disponível em: https://doi.org/10.1007/978-1-4939-6622-6_5

AKEN, Bronwen L. *et al.* The Ensembl gene annotation system. **Database**, *[S. l.]*, v. 2016, p. baw093, 2016. Disponível em: https://doi.org/10.1093/database/baw093

ÁLVAREZ-MACHANCOSES, Óscar; FERNÁNDEZ-MARTÍNEZ, Juan Luis; KLOCZKOWSKI, Andrzej. Prediction of protein tertiary structure via regularized template classification techniques. **Molecules**, *[S. l.]*, v. 25, n. 11, 2020. Disponível em: https://doi.org/10.3390/molecules25112467

ANDOR, Noemi; MALEY, Carlo C.; JI, Hanlee P. Genomic instability in cancer: Teetering on the limit of tolerance. **Cancer Research**, *[S. l.]*, v. 77, n. 9, p. 2179–2185, 2017. Disponível em: https://doi.org/10.1158/0008-5472.CAN-16-1553

ARMSTRONG, Joel *et al.* Whole-Genome Alignment and Comparative Annotation. **Annual Review of Animal Biosciences**, *[S. l.]*, v. 7, n. 1, p. 41–64, 2019. Disponível em: https://doi.org/10.1146/annurev-animal-020518-115005

ARNOLD J. LEVINE. p53, the Cellular Gatekeeper for Growth and Division. **European Journal of Cancer Part A**, *[S. l.]*, v. 88, n. 5, p. 323–331, 1997. Disponível em: https://doi.org/10.1016/S0959-8049(97)00064-6

AYLLÓN BARBELLIDO, Sonia *et al.* Gene therapy in the management of oral cancer: Review of the literature. **Medicina Oral, Patologia Oral y Cirugia Bucal**, *[S. l.]*, v. 13, n. 1, p. 15–21, 2008.

BAGNYUKOVA, Tetyana *et al.* Chemotherapy and signaling: How can targeted therapies supercharge cytotoxic agents? **Cancer Biology and Therapy**, *[S. l.]*, v. 10, n. 9, p. 839–853, 2010. Disponível em: https://doi.org/10.4161/cbt.10.9.13738

BARRETINA, Jordi *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. **Nature**, *[S. l.]*, v. 483, n. 7391, p. 603–607, 2012. Disponível em: https://doi.org/10.1038/nature11003

BARRETT, Tanya *et al.* Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D6–D17, 2015. Disponível em: https://doi.org/10.1093/nar/gku1130

BASKAR, Rajamanickam *et al.* Cancer and radiation therapy: Current advances and future directions. **International Journal of Medical Sciences**, *[S. l.]*, v. 9, n. 3, p. 193–199, 2012. Disponível em: https://doi.org/10.7150/ijms.3635

BATEMAN, Alex *et al.* UniProt: A hub for protein information. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D204–D212, 2015. Disponível em: https://doi.org/10.1093/nar/gku989

BECKER, Kevin G. *et al.* The Genetic Association Database. **Nature Genetics**, *[S. l.]*, v. 36, n. 5, p. 431–432, 2004. Disponível em: https://doi.org/10.1038/ng0504-431

BEGG, Adrian C.; STEWART, Fiona A.; VENS, Conchita. Strategies to improve radiotherapy with targeted drugs. **Nature Reviews Cancer**, *[S. l.]*, v. 11, n. 4, p. 239–253, 2011. Disponível em: https://doi.org/10.1038/nrc3007

BENITEZ, Martha Lucia Ruiz *et al.* Mycobacterium bovis BCG in metastatic melanoma therapy. **Applied Microbiology and Biotechnology**, *[S. l.]*, v. 103, n. 19, p. 7903–7916, 2019. Disponível em: https://doi.org/10.1007/s00253-019-10057-0

BENSON, Dennis A. *et al.* GenBank. **Nucleic Acids Research**, *[S. l.]*, v. 37, n. SUPPL. 1, p. 26–31, 2009. Disponível em: https://doi.org/10.1093/nar/gkn723

BENTO, A. Patrícia *et al.* The ChEMBL bioactivity database: An update. **Nucleic Acids Research**, *[S. l.]*, v. 42, n. D1, p. 1083–1090, 2014. Disponível em: https://doi.org/10.1093/nar/gkt1031

BHOWMICK, Neil A.; NEILSON, Eric G.; MOSES, Harol L. Emerging Role of Stromal Fibroblasts in Epithelial Cancer. **Current Signal Transduction Therapy**, *[S. l.]*, v. 1, n. 3, p. 273–283, 2008. Disponível em: https://doi.org/10.2174/157436206778226969

BLAKE, J. A. *et al.* Gene ontology consortium: Going forward. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D1049–D1056, 2015. Disponível em: https://doi.org/10.1093/nar/gku1179

BOXALL, Alistair B. A. *et al.* Pharmaceuticals and personal care products in the environment: What are the big questions? **Environmental Health Perspectives**, *[S. l.]*, v. 120, n. 9, p. 1221–1229, 2012. Disponível em: https://doi.org/10.1289/ehp.1104477

BRENNAN, Cameron W. *et al.* The Somatic Genomic Landscape of Glioblastoma. **Cell**, *[S. l.]*, v. 157, n. 3, p. 753, 2014. Disponível em: https://doi.org/10.1016/j.cell.2014.04.004

BREUR, K. Recent developments in radiotherapy. **Archivum Chirurgicum Neerlandicum**, *[S. l.]*, v. 25, n. 4, p. 341–350, 2017. Disponível em: https://doi.org/10.1056/nejmra1608986

BROERTJES, Jorrit. The Ten Hallmarks of Cancer in Cutaneous Malignant Melanoma. **UNAV-JMS (University of Navarra-Journal for Medical Students)**, *[S. l.]*, v. 01, n. April, p. 6–12, 2015. Disponível em: https://doi.org/10.15581/005.1.6-12

BROPHAL, Raj S. **Concepts of Epidemiology: Integrating the ideas, theories, principles, and methods of epidemiology (3 ed.)**. *[s. l.]*, 2016. Disponível em: https://oxfordmedicine.com/view/10.1093/med/9780198739685.001.0001/med-9780198739685.

BRUNELLE, Joslyn K. *et al.* MCL-1-dependent leukemia cells are more sensitive to

chemotherapy than BCL-2-dependent counterparts. **Journal of Cell Biology**, *[S. l.]*, v. 187, n. 3, p. 429–442, 2009. Disponível em: https://doi.org/10.1083/jcb.200904049

BURKHART, Deborah L.; SAGE, Julien. Cellular mechanisms of tumour suppression by the retinoblastoma gene. **Nature Reviews Cancer**, *[S. l.]*, v. 8, n. 9, p. 671–682, 2008. Disponível em: https://doi.org/10.1038/nrc2399

BURLEY, Stephen K. *et al.* Protein Data Bank (PDB): The single global macromolecular structure archive. **Methods in Molecular Biology**, *[S. l.]*, v. 1607, p. 627–641, 2017. Disponível em: https://doi.org/10.1007/978-1-4939-7000-1_26

CANCER.ORG. **How Chemotherapy Drugs Work**. *[s. l.]*, 2019. Disponível em: https://www.cancer.org/treatment/treatments-and-side-effects/treatment-types/chemotherapy/how-chemotherapy-drugs-work.html#:~:text=Alkylating agents keep the cell,%2C multiple myeloma%2C and sarcoma.

Cancer multidrug resistance. **Bio/Technology**, *[S. l.]*, v. 18, n. 10s, p. IT18–IT20, 2000. Disponível em: https://doi.org/10.1038/80051

CHEN, Chuming; HUANG, Hongzhan; WU, Cathy H. **Protein bioinformatics databases and resources**. *[S. l.: s. n.]*. v. 1558*E-book*. Disponível em: https://doi.org/10.1007/978-1-4939-6783-4_1

CHEN, X.; JI, Z. L.; CHEN, Y. Z. TTD: Therapeutic Target Database. **Nucleic Acids Research**, *[S. l.]*, v. 30, n. 1, p. 412–415, 2002. Disponível em: https://doi.org/10.1093/nar/30.1.412

CHEN, Xiaoxia *et al.* Analysis of the Physicochemical Properties of Acaricides Based on Lipinski's Rule of Five. **Journal of Computational Biology**, *[S. l.]*, v. 27, n. 9, p. 1397–1406, 2020. Disponível em: https://doi.org/10.1089/cmb.2019.0323

CHENG, Nikki *et al.* Transforming growth factor-β signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering

and invasion. **Molecular Cancer Research**, *[S. l.]*, v. 6, n. 10, p. 1521–1533, 2008. Disponível em: https://doi.org/10.1158/1541-7786.MCR-07-2203

CHILAKAMARTHI, Ushasri; GIRIBABU, Lingamallu. Photodynamic Therapy: Past, Present and Future. **Chemical Record**, *[S. l.]*, v. 17, n. 8, p. 775–802, 2017. Disponível em: https://doi.org/10.1002/tcr.201600121

CHOWDHURY, Biswanath; GARAI, Gautam. A review on multiple sequence alignment from the perspective of genetic algorithm. **Genomics**, *[S. l.]*, v. 109, n. 5–6, p. 419–431, 2017. Disponível em: https://doi.org/10.1016/j.ygeno.2017.06.007

COUTO, Gabriela Klein *et al.* Tetra-cationic platinum(II) porphyrins like a candidate photosensitizers to bind, selective and drug delivery for metastatic melanoma. **Journal of Photochemistry and Photobiology B: Biology**, *[S. l.]*, v. 202, n. November 2019, p. 111725, 2020. Disponível em: https://doi.org/10.1016/j.jphotobiol.2019.111725

CRUZ, Leonardo Magalhães *et al.* Protein Function Prediction. *In*: *[S. l.: s. n.]*. v. 1654p. 55–75. *E-book*. Disponível em: https://doi.org/10.1007/978-1-4939-7231-9_5

CRUZ, Pedro M. R. *et al.* The role of cholesterol metabolism and cholesterol transport in carcinogenesis: A review of scientific findings, relevant to future cancer therapeutics. **Frontiers in Pharmacology**, *[S. l.]*, v. 4 SEP, n. September, p. 1–7, 2013. Disponível em: https://doi.org/10.3389/fphar.2013.00119

CUNNINGHAM, Fiona *et al.* Ensembl 2015. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D662–D669, 2015. Disponível em: https://doi.org/10.1093/nar/gku1010

DE ARAÚJO LIMA, Emeline Das Neves *et al.* Bone marrow transplantation: Graft versus host disease and oral changes. **Revista Odonto Ciencia**, *[S. l.]*, v. 27, n. 1, p. 10–15, 2012. Disponível em: https://doi.org/10.1590/s1980-65232012000100002

DE PALMA, Michele; HANAHAN, Douglas. The biology of personalized cancer medicine: Facing individual complexities underlying hallmark capabilities. **Molecular Oncology**, *[S.*

l.], v. 6, n. 2, p. 111–127, 2012. Disponível em: https://doi.org/10.1016/j.molonc.2012.01.011

DESHPANDE, Amit; SICINSKI, Peter; HINDS, Philip W. Cyclins and cdks in development and cancer: A perspective. **Oncogene**, *[S. l.]*, v. 24, n. 17, p. 2909–2915, 2005. Disponível em: https://doi.org/10.1038/sj.onc.1208618

DEVITA, Vincent T.; CHU, Edward. A history of cancer chemotherapy. **Cancer Research**, *[S. l.]*, v. 68, n. 21, p. 8643–8653, 2008. Disponível em: https://doi.org/10.1158/0008-5472.CAN-07-6611

DINIZ, W. J. S.; CANDURI, F. Bioinformatics: An overview and its applications. **Genetics and Molecular Research**, *[S. l.]*, v. 16, n. 1, p. 1–21, 2017. Disponível em: https://doi.org/10.4238/gmr16019645

DOOLITTLE, Russell F. *et al.* Simian sarcoma virus onc gene, v-sis, is derived from the gene (or genes) encoding a platelet-derived growth factor. **Science**, *[S. l.]*, v. 221, n. 4607, p. 275–277, 1983. Disponível em: https://doi.org/10.1126/science.6304883

DÜZGÜNEŞ, Nejat *et al.* Photodynamic therapy of cancer with liposomal photosensitizers. **Therapeutic Delivery**, *[S. l.]*, v. 9, n. 11, p. 823–832, 2018. Disponível em: https://doi.org/10.4155/tde-2018-0050

EDGAR, Ron; DOMRACHEV, Michael; LASH, Alex E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. **Nucleic Acids Research**, *[S. l.]*, v. 30, n. 1, p. 207–210, 2002. Disponível em: https://doi.org/10.1093/nar/30.1.207

EJIGU, Girum Fitihamlak; JUNG, Jaehee. Review on the computational genome annotation of sequences obtained by next-generation sequencing. **Biology**, *[S. l.]*, v. 9, n. 9, p. 1–27, 2020. Disponível em: https://doi.org/10.3390/biology9090295

FAN, Tai Ping D.; JAGGER, Rhys; BICKNELL, Roy. Controlling the vasculature: Angiogenesis, anti-angiogenesis and vascular targeting of gene therapy. **Trends in**

Pharmacological Sciences, *[S. l.]*, v. 16, n. 2, p. 57–66, 1995. Disponível em: https://doi.org/10.1016/S0165-6147(00)88979-8

FERNAND, Kaleigh; MANABU, Kurokawa. Evading apoptosis in cancer. **Trends Cell Biol**, *[S. l.]*, v. 23, n. 1, p. 620–633, 2013. Disponível em: https://doi.org/10.1016/j.tcb.2013.07.006.Evading

FINN, Robert D. *et al.* Pfam: The protein families database. **Nucleic Acids Research**, *[S. l.]*, v. 42, n. D1, p. 222–230, 2014. Disponível em: https://doi.org/10.1093/nar/gkt1223

FOLTZ, Ian N.; KAROW, Margaret; WASSERMAN, Scott M. Evolution and emergence of therapeutic monoclonal antibodies what cardiologists need to know. **Circulation**, *[S. l.]*, v. 127, n. 22, p. 2222–2230, 2013. Disponível em: https://doi.org/10.1161/CIRCULATIONAHA.113.002033

FORBES, S. a *et al.* **Europe PMC Funders Group The Catalogue of Somatic Mutations in Cancer (COSMIC)**. *[S. l.: s. n.]*. *E-book*. Disponível em: https://doi.org/10.1002/0471142905.hg1011s57

FORBES, Simon A. *et al.* COSMIC (the Catalogue of Somatic Mutations In Cancer): A resource to investigate acquired mutations in human cancer. **Nucleic Acids Research**, *[S. l.]*, v. 38, n. SUPPL.1, p. 652–657, 2009 b. Disponível em: https://doi.org/10.1093/nar/gkp995

FORBES, Simon A. *et al.* COSMIC: Mining complete cancer genomes in the catalogue of somatic mutations in cancer. **Nucleic Acids Research**, *[S. l.]*, v. 39, n. SUPPL. 1, p. 945–950, 2011. Disponível em: https://doi.org/10.1093/nar/gkq929

FRÉROT, Mathilde *et al.* What is epidemiology? Changing definitions of epidemiology 1978-2017. **PLoS ONE**, *[S. l.]*, v. 13, n. 12, 2018. Disponível em: https://doi.org/10.1371/journal.pone.0208442

GALPERIN, Michael Y. *et al.* Expanded Microbial genome coverage and improved protein

family annotation in the COG database. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D261–D269, 2015. Disponível em: https://doi.org/10.1093/nar/gku1223

GARCÍA-CABALLERO, Melissa; MEDINA, Miguel Ángel; QUESADA, Ana R. Toluquinol, A Marine Fungus Metabolite, Inhibits Some of the Hallmarks of Cancer. *In*: **Handbook of Anticancer Drugs from Marine Origin**. Cham: Springer International Publishing, 2015. p. 269–299. *E-book*. Disponível em: https://doi.org/10.1007/978-3-319-07145-9_14

GRETEN, Florian R.; GRIVENNIKOV, Sergei I. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. **Immunity**, *[S. l.]*, v. 51, n. 1, p. 27–41, 2019. Disponível em: https://doi.org/10.1016/j.immuni.2019.06.025

GUAN, Longfei *et al.* ADMET-score-a comprehensive scoring function for evaluation of chemical drug-likeness. **MedChemComm**, *[S. l.]*, v. 10, n. 1, p. 148–157, 2019. Disponível em: https://doi.org/10.1039/C8MD00472B

HAGEN, Joel B. The origins of bioinformatics. **Nature Reviews Genetics**, *[S. l.]*, v. 1, n. 3, p. 231–236, 2000. Disponível em: https://doi.org/10.1038/35042090

HANAHAN, Douglas; WEINBERG, Robert A. Review Hallmarks of Cancer : The Next Generation. **Cell**, *[S. l.]*, v. 144, n. 5, p. 646–674, 2011. Disponível em: https://doi.org/10.1016/j.cell.2011.02.013

HOLLINGSWORTH, Scott A.; DROR, Ron O. Molecular Dynamics Simulation for All. **Neuron**, *[S. l.]*, v. 99, n. 6, p. 1129–1143, 2018. Disponível em: https://doi.org/10.1016/j.neuron.2018.08.011

HOLOHAN, Caitriona *et al.* Cancer drug resistance: An evolving paradigm. **Nature Reviews Cancer**, *[S. l.]*, v. 13, n. 10, p. 714–726, 2013. Disponível em: https://doi.org/10.1038/nrc3599

HOU, Tingjun; WANG, Junmei. Structure – ADME relationship : still a long way to go ? *[S. l.]*, p. 759–770, 2008.

HOUSMAN, Genevieve *et al.* Drug resistance in cancer: An overview. **Cancers**, *[S. l.]*, v. 6, n. 3, p. 1769–1792, 2014. Disponível em: https://doi.org/10.3390/cancers6031769

HSU, Chia Chi; TSENG, Ling Ming; LEE, Hsin Chen. Role of mitochondrial dysfunction in cancer progression. **Experimental Biology and Medicine**, *[S. l.]*, v. 241, n. 12, p. 1281–1295, 2016. Disponível em: https://doi.org/10.1177/1535370216641787

HU, Guanghui; AGARWAL, Pankaj. Human disease-drug network based on genomic expression profiles. **PLoS ONE**, *[S. l.]*, v. 4, n. 8, 2009. Disponível em: https://doi.org/10.1371/journal.pone.0006536

HU, Yanqiu *et al.* Pharmacophore modeling, multiple docking, and molecular dynamics studies on Wee1 kinase inhibitors. **Journal of Biomolecular Structure and Dynamics**, *[S. l.]*, v. 37, n. 10, p. 2703–2715, 2019. Disponível em: https://doi.org/10.1080/07391102.2018.1495576

HUB, Jochen S.; KUBITZKI, Marcus B.; DE GROOT, Bert L. Spontaneous Quaternary and Tertiary T-R Transitions of Human Hemoglobin in Molecular Dynamics Simulation. **PLoS Computational Biology**, *[S. l.]*, v. 6, n. 5, p. e1000774, 2010. Disponível em: https://doi.org/10.1371/journal.pcbi.1000774

HÜBNER, Anette *et al.* Multi-site Phosphorylation Regulates Bim Stability and Apoptotic Activity. **Mol Cell**, *[S. l.]*, v. 30, p. 415–425, 2008.

JEEVANANDAM, Jaison *et al.* Advancing Aptamers as Molecular Probes for Cancer Theranostic Applications—The Role of Molecular Dynamics Simulation. **Biotechnology Journal**, *[S. l.]*, v. 15, n. 3, p. 1–65, 2020. Disponível em: https://doi.org/10.1002/biot.201900368

JOHNSTONE, Ricky W.; RUEFLI, Astrid A.; LOWE, Scott W. Apoptosis: A link between cancer genetics and chemotherapy. **Cell**, *[S. l.]*, v. 108, n. 2, p. 153–164, 2002. Disponível em: https://doi.org/10.1016/S0092-8674(02)00625-6

JOO, Won Duk; VISINTIN, Irene; MOR, Gil. Targeted cancer therapy - Are the days of systemic chemotherapy numbered? **Maturitas**, *[S. l.]*, v. 76, n. 4, p. 308–314, 2013. Disponível em: https://doi.org/10.1016/j.maturitas.2013.09.008

KALYANARAMAN, Ananth *et al.* Genome Assembly. *In*: **Encyclopedia of Parallel Computing**. Boston, MA: Springer US, 2011. p. 755–768. *E-book*. Disponível em: https://doi.org/10.1007/978-0-387-09766-4_402

KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, *[S. l.]*, v. 28, n. 1, p. 27–30, 2000. Disponível em: https://doi.org/10.1093/nar/28.1.27

KENNEDY, Lucy Boyce; SALAMA, April K. S. A review of cancer immunotherapy toxicity. **CA: A Cancer Journal for Clinicians**, *[S. l.]*, v. 70, n. 2, p. 86–104, 2020. Disponível em: https://doi.org/10.3322/caac.21596

KIM, In Wha *et al.* Computational Drug Repositioning for Gastric Cancer using Reversal Gene Expression Profiles. **Scientific Reports**, *[S. l.]*, v. 9, n. 1, p. 1–10, 2019. Disponível em: https://doi.org/10.1038/s41598-019-39228-9

KIMIZ-GEBOLOGLU, Ilgin; GULCE-IZ, Sultan; BIRAY-AVCI, Cigir. Monoclonal antibodies in cancer immunotherapy. **Molecular Biology Reports**, *[S. l.]*, v. 45, n. 6, p. 2935–2940, 2018. Disponível em: https://doi.org/10.1007/s11033-018-4427-x

KITCHEN, Douglas B. *et al.* Docking and scoring in virtual screening for drug discovery: Methods and applications. **Nature Reviews Drug Discovery**, *[S. l.]*, v. 3, n. 11, p. 935–949, 2004. Disponível em: https://doi.org/10.1038/nrd1549

KUMAR, S.; FILIPSKI, A. Multiple sequence alignment: In pursuit of homologous DNA positions. **Genome Research**, *[S. l.]*, v. 17, n. 2, p. 127–135, 2007. Disponível em: https://doi.org/10.1101/gr.5232407

KWIATKOWSKI, Stanisław *et al.* Photodynamic therapy – mechanisms, photosensitizers

and combinations. **Biomedicine and Pharmacotherapy**, *[S. l.]*, v. 106, n. July, p. 1098–1107, 2018. Disponível em: https://doi.org/10.1016/j.biopha.2018.07.049

LANDRUM, Melissa J. *et al.* ClinVar: Public archive of relationships among sequence variation and human phenotype. **Nucleic Acids Research**, *[S. l.]*, v. 42, n. D1, p. 980–985, 2014. Disponível em: https://doi.org/10.1093/nar/gkt1113

LAW, Vivian *et al.* DrugBank 4.0: Shedding new light on drug metabolism. **Nucleic Acids Research**, *[S. l.]*, v. 42, n. D1, p. 1091–1097, 2014. Disponível em: https://doi.org/10.1093/nar/gkt1068

LEE, Yeuan Ting; TAN, Yi Jer; OON, Chern Ein. Molecular targeted therapy: Treating cancer with specificity. **European Journal of Pharmacology**, *[S. l.]*, v. 834, n. January, p. 188–196, 2018. Disponível em: https://doi.org/10.1016/j.ejphar.2018.07.034

LETOVSKY, S. GDB: the Human Genome Database. **Nucleic Acids Research**, *[S. l.]*, v. 26, n. 1, p. 94–99, 1998. Disponível em: https://doi.org/10.1093/nar/26.1.94

LHEUREUX, S. *et al.* Molecularly targeted therapies in cancer: a guide for the nuclear medicine physician. **European Journal of Nuclear Medicine and Molecular Imaging**, *[S. l.]*, v. 44, p. 41–54, 2017. Disponível em: https://doi.org/10.1007/s00259-017-3695-3

LI, Kening *et al.* Bioinformatics Approaches for Anti-cancer Drug Discovery. **Current Drug Targets**, *[S. l.]*, v. 21, n. 1, p. 3–17, 2019. Disponível em: https://doi.org/10.2174/1389450120666190923162203

LI, Tinglu *et al.* Tumor angiogenesis and anti-angiogenic gene therapy for cancer (Review). **Oncology Letters**, *[S. l.]*, v. 16, n. 1, p. 687–702, 2018 a. Disponível em: https://doi.org/10.3892/ol.2018.8733

LI, Ying Hong *et al.* Therapeutic target database update 2018: Enriched resource for facilitating bench-to-clinic research of targeted therapeutics. **Nucleic Acids Research**, *[S. l.]*, v. 46, n. D1, p. D1121–D1127, 2018 b. Disponível em:

https://doi.org/10.1093/nar/gkx1076

LINDORFF-LARSEN, Kresten *et al.* Systematic Validation of Protein Force Fields against Experimental Data. **PLoS ONE**, *[S. l.]*, v. 7, n. 2, p. e32131, 2012. Disponível em: https://doi.org/10.1371/journal.pone.0032131

LONGLEY, D. B.; JOHNSTON, P. G. Molecular mechanisms of drug resistance. **Journal of Pathology**, *[S. l.]*, v. 205, n. 2, p. 275–292, 2005. Disponível em: https://doi.org/10.1002/path.1706

LUQUE-CABAL, María *et al.* Mechanisms behind the resistance to trastuzumab in HER2-amplified breast cancer and strategies to overcome It. **Clinical Medicine Insights: Oncology**, *[S. l.]*, v. 10, p. 21–30, 2016. Disponível em: https://doi.org/10.4137/CMO.S34537

MA, Sheng *et al.* Neutralizing tumor-promoting inflammation with polypeptide-dexamethasone conjugate for microenvironment modulation and colorectal cancer therapy. **Biomaterials**, *[S. l.]*, v. 232, p. 119676, 2020. Disponível em: https://doi.org/10.1016/j.biomaterials.2019.119676

MARIJO BILUSIC, MD; RAVI A. MADAN, MD. Therapeutic Cancer Vaccines: The Latest Advancement in Targeted Therapy. *[S. l.]*, v. 23, n. 1, p. 1–7, 2012. Disponível em: https://doi.org/10.1097/MJT.0b013e3182068cdb.

MATTHEWS, Holly; HANISON, James; NIRMALAN, Niroshini. “Omics”-informed drug and biomarker discovery: Opportunities, challenges and future perspectives. **Proteomes**, *[S. l.]*, v. 4, n. 3, p. 1–12, 2016. Disponível em: https://doi.org/10.3390/proteomes4030028

MEDZHITOV, Ruslan. Origin and physiological roles of inflammation. **Nature**, *[S. l.]*, v. 454, n. 7203, p. 428–435, 2008. Disponível em: https://doi.org/10.1038/nature07201

MITCHELL, Alex *et al.* The InterPro protein families database: The classification resource after 15 years. **Nucleic Acids Research**, *[S. l.]*, v. 43, n. D1, p. D213–D221, 2015.

Disponível em: https://doi.org/10.1093/nar/gku1243

NATIONAL CANCER INSTITUTE. **What is Cancer**. *[s. l.]*, 2015. Disponível em: https://www.cancer.gov/about-cancer/understanding/what-is-cancer.

NEOPROSPECTA. **BANCOS DE DADOS BIOLÓGICOS**. *[s. l.]*, 2018. Disponível em: https://blog.neoprospecta.com/bancos-de-dados-biologicos/.

O. TROTT, A. J. OLSON. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. **Journal of computational chemistry**, *[S. l.]*, v. 31, p. 455–461, 2010. Disponível em: https://doi.org/10.1002/jcc

ÖSTMAN, Arne; HELDIN, Carl Henrik. PDGF Receptors as Targets in Tumor Treatment. **Advances in Cancer Research**, *[S. l.]*, v. 97, n. 06, p. 247–274, 2007. Disponível em: https://doi.org/10.1016/S0065-230X(06)97011-0

PADMA, Viswanadha Vijaya. An overview of targeted cancer therapy. **BioMedicine (Netherlands)**, *[S. l.]*, v. 5, n. 4, p. 1–6, 2015. Disponível em: https://doi.org/10.7603/s40681-015-0019-4

PARSONS, M. J. *et al.* Genetic deletion of caspase-2 accelerates MMTV/c-neu-driven mammary carcinogenesis in mice. **Cell Death and Differentiation**, *[S. l.]*, v. 20, n. 9, p. 1174–1182, 2013. Disponível em: https://doi.org/10.1038/cdd.2013.38

PATEL, Mishal N. *et al.* Objective assessment of cancer genes for drug discovery. **Nature Reviews Drug Discovery**, *[S. l.]*, v. 12, n. 1, p. 35–50, 2013. Disponível em: https://doi.org/10.1038/nrd3913

PEARCE, Alison *et al.* Productivity losses due to premature mortality from cancer in Brazil, Russia, India, China, and South Africa (BRICS): A population-based comparison. **Cancer Epidemiology**, *[S. l.]*, v. 53, n. December 2017, p. 27–34, 2018. Disponível em: https://doi.org/10.1016/j.canep.2017.12.013

PETERS, Christina; BROWN, Stuart. Antibody-drug conjugates as novel anti-cancer chemotherapeutics. **Bioscience Reports**, *[S. l.]*, v. 35, n. 4, 2015. Disponível em: https://doi.org/10.1042/BSR20150089

PETRYSZAK, Robert *et al.* Expression Atlas update - A database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. **Nucleic Acids Research**, *[S. l.]*, v. 42, n. D1, p. 926–932, 2014. Disponível em: https://doi.org/10.1093/nar/gkt1270

PIÑERO, Janet *et al.* DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes. **Database**, *[S. l.]*, v. 2015, p. 1–17, 2015. Disponível em: https://doi.org/10.1093/database/bav028

PIÑERO, Janet *et al.* DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants. **Nucleic Acids Research**, *[S. l.]*, v. 45, n. D1, p. D833–D839, 2017. Disponível em: https://doi.org/10.1093/nar/gkw943

PINZI, Luca; RASTELLI, Giulio. Molecular docking: Shifting paradigms in drug discovery. **International Journal of Molecular Sciences**, *[S. l.]*, v. 20, n. 18, 2019. Disponível em: https://doi.org/10.3390/ijms20184331

PRZYBYLAK, K. R. *et al.* Characterisation of data resources for in silico modelling: benchmark datasets for ADME properties. **Expert Opinion on Drug Metabolism and Toxicology**, *[S. l.]*, v. 14, n. 2, p. 169–181, 2018. Disponível em: https://doi.org/10.1080/17425255.2017.1316449

PUNTA, Marco; OFRAN, Yanay. The rough guide to in silico function prediction, or how to use sequence and structure information to predict protein function. **PLoS Computational Biology**, *[S. l.]*, v. 4, n. 10, 2008. Disponível em: https://doi.org/10.1371/journal.pcbi.1000160

RICE, Shara D. *et al.* Analysis of chemotherapeutic response heterogeneity and drug

clustering based on mechanism of action using an in vitro assay. **Anticancer Research**, *[S. l.]*, v. 30, n. 7, p. 2805–2811, 2010.

ROBERT, Jacques. Cancer Metastase. *[S. l.]*, v. 100, p. 333–342, 2013.

ROSSI, Antonio; DI MAIO, Massimo. Platinum-based chemotherapy in advanced non-small-cell lung cancer: optimal number of treatment cycles. **Expert Review of Anticancer Therapy**, *[S. l.]*, v. 16, n. 6, p. 653–660, 2016. Disponível em: https://doi.org/10.1586/14737140.2016.1170596

SADELAIN, Michel. A new induction to the gene and cell therapy hall of fame: Genome editing. **Molecular Therapy**, *[S. l.]*, v. 24, n. 3, p. 407–408, 2016. Disponível em: https://doi.org/10.1038/mt.2016.25

SAIJO, Nagahiro. Progress in cancer chemotherapy with special stress on molecular-targeted therapy. **Japanese Journal of Clinical Oncology**, *[S. l.]*, v. 40, n. 9, p. 855–862, 2010. Disponível em: https://doi.org/10.1093/jjco/hyq035

SCHLOM, Jeffrey *et al.* **Therapeutic cancer vaccines**. 1. ed. *[S. l.]*: Elsevier Inc., 2014. v. 121*E-book*. Disponível em: https://doi.org/10.1016/B978-0-12-800249-0.00002-0

SCHULTZ, Terry W. *et al.* The OECD QSAR toolbox starts its second decade. **Methods in Molecular Biology**, *[S. l.]*, v. 1800, p. 55–77, 2018. Disponível em: https://doi.org/10.1007/978-1-4939-7899-1_2

SEEBACHER, N. A. *et al.* **Clinical development of targeted and immune based anti-cancer therapies**. *[S. l.]*: Journal of Experimental & Clinical Cancer Research, 2019. v. 38*E-book*. Disponível em: https://doi.org/10.1186/s13046-019-1094-2

SEELIGER, Daniel; DE GROOT, Bert L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. **Journal of Computer-Aided Molecular Design**, *[S. l.]*, v. 24, n. 5, p. 417–422, 2010. Disponível em: https://doi.org/10.1007/s10822-010-9352-6

SENN, Hans Martin; THIEL, Walter. QM/MM Methods for Biomolecular Systems.

Angewandte Chemie International Edition, *[S. l.]*, v. 48, n. 7, p. 1198–1229, 2009. Disponível em: https://doi.org/10.1002/anie.200802019

SHERR, Charles J.; MCCORMICK, Frank. The RB and p53 pathways in cancer. **Cancer Cell**, *[S. l.]*, v. 2, n. 2, p. 103–112, 2002. Disponível em: https://doi.org/10.1016/S1535-6108(02)00102-2

SIGRIST, Christian J. A. *et al.* New and continuing developments at PROSITE. **Nucleic Acids Research**, *[S. l.]*, v. 41, n. D1, p. 344–347, 2013. Disponível em: https://doi.org/10.1093/nar/gks1067

SOBINOFF, Alexander P.; PICKETT, Hilda A. Alternative Lengthening of Telomeres: DNA Repair Pathways Converge. **Trends in Genetics**, *[S. l.]*, v. 33, n. 12, p. 921–932, 2017. Disponível em: https://doi.org/10.1016/j.tig.2017.09.003

STEVEN, Antonius; FISHER, Scott A.; ROBINSON, Bruce W. Immunotherapy for lung cancer. **Respirology**, *[S. l.]*, v. 21, n. 5, p. 821–833, 2016. Disponível em: https://doi.org/10.1111/resp.12789

SUHAIL, Yasir *et al.* Systems Biology of Cancer Metastasis. **Cell Systems**, *[S. l.]*, v. 9, n. 2, p. 109–127, 2019. Disponível em: https://doi.org/10.1016/j.cels.2019.07.003

SUNG, Hyuna *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, *[S. l.]*, v. 0, n. 0, p. 1–41, 2021. Disponível em: https://doi.org/10.3322/caac.21660

THE CANCER GENOME ATLAS (TCGA) RESEARCH NETWORK. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. **Nature**, *[S. l.]*, v. 455, n. 7216, p. 1061–1068, 2008. Disponível em: https://doi.org/10.1038/nature07385

THOMPSON, Julie D. *et al.* A Comprehensive Benchmark Study of Multiple Sequence

Alignment Methods: Current Challenges and Future Perspectives. **PLoS ONE**, *[S. l.]*, v. 6, n. 3, p. e18093, 2011. Disponível em: https://doi.org/10.1371/journal.pone.0018093

TREISTER, Nathaniel S. *et al.* Oral Chronic Graft-versus-Host Disease Scoring Using the NIH Consensus Criteria. **Biology of Blood and Marrow Transplantation**, *[S. l.]*, v. 16, n. 1, p. 108–114, 2010. Disponível em: https://doi.org/10.1016/j.bbmt.2009.09.010

TU, Huailu; COSTA, Max. XIAP's Profile in Human Cancer. **Biomolecules**, *[S. l.]*, v. 10, n. 11, p. 1493, 2020. Disponível em: https://doi.org/10.3390/biom10111493

VAN DE DONK, Niels W. C. J. *et al.* Clinical efficacy and management of monoclonal antibodies targeting CD38 and SLAMF7 in multiple myeloma. **Blood**, *[S. l.]*, v. 127, n. 6, p. 681–695, 2016. Disponível em: https://doi.org/10.1182/blood-2015-10-646810

VERLI, Hugo. **Bioinformática da Biologia à Flexibilidade Molecular**. *[S. l.: s. n.]. E-book*.

VESELY, Matthew D.; SCHREIBER, Robert D. Cancer immunoediting: Antigens, mechanisms, and implications to cancer immunotherapy. **Annals of the New York Academy of Sciences**, *[S. l.]*, v. 1284, n. 1, p. 1–5, 2013. Disponível em: https://doi.org/10.1111/nyas.12105

VINAY, Dass S. *et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. **Seminars in Cancer Biology**, *[S. l.]*, v. 35, p. S185–S198, 2015. Disponível em: https://doi.org/10.1016/j.semcancer.2015.03.004

VOGELSTEIN, Bert *et al.* Cancer genome landscapes. **Science**, *[S. l.]*, v. 340, n. 6127, p. 1546–1558, 2013. Disponível em: https://doi.org/10.1126/science.1235122

VU, Thuy; CLARET, Francois X. Trastuzumab: Updated mechanisms of action and resistance in breast cancer. **Frontiers in Oncology**, *[S. l.]*, v. 2 JUN, n. June, p. 1–6, 2012. Disponível em: https://doi.org/10.3389/fonc.2012.00062

WATERFIELD, Michael D. *et al.* Platelet-derived growth factor is structurally related to

the putative transforming protein p28sis of simian sarcoma virus. **Nature**, *[S. l.]*, v. 304, n. 5921, p. 35–39, 1983. Disponível em: https://doi.org/10.1038/304035a0

WEINER, Louis M.; SURANA, Rishi; WANG, Shangzi. Monoclonal antibodies: Versatile platforms for cancer immunotherapy. **Nature Reviews Immunology**, *[S. l.]*, v. 10, n. 5, p. 317–327, 2010. Disponível em: https://doi.org/10.1038/nri2744

WISHART, David S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. **Nucleic acids research**, *[S. l.]*, v. 34, n. Database issue, p. 668–672, 2006. Disponível em: https://doi.org/10.1093/nar/gkj067

WISHART, David S. *et al.* DrugBank: A knowledgebase for drugs, drug actions and drug targets. **Nucleic Acids Research**, *[S. l.]*, v. 36, n. SUPPL. 1, p. 901–906, 2008. Disponível em: https://doi.org/10.1093/nar/gkm958

WITSCH, Esther; SELA, Michael; YARDEN, Yosef. Roles for Growth Factors in Cancer Progression. **Physiology**, *[S. l.]*, v. 25, n. 2, p. 85–101, 2010. Disponível em: https://doi.org/10.1152/physiol.00045.2009

WORLD HEALTH ORGANIZATION. **Cancer**. *[s. l.]*, 2021a. Disponível em: https://www.who.int/news-room/fact-sheets/detail/cancer#:~:text=Globally%2C about 1 in 6,- and middle-income countries.&text=The total annual economic cost,data in 2019 (6).

WORLD HEALTH ORGANIZATION. **Cancer Today**. *[s. l.]*, 2021b. Disponível em: https://doi.org/10.17226/18700.

WYLD, Lynda; AUDISIO, Riccardo A.; POSTON, Graeme J. The evolution of cancer surgery and future perspectives. **Nature Reviews Clinical Oncology**, *[S. l.]*, v. 12, n. 2, p. 115–124, 2015. Disponível em: https://doi.org/10.1038/nrclinonc.2014.191

XIA, Xuhua. Bioinformatics and Drug Discovery. **Current Topics in Medicinal Chemistry**, *[S. l.]*, v. 17, n. 15, p. 1709–1726, 2017. Disponível em:

https://doi.org/10.2174/1568026617666161116143440

YANG, Wanjuan *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. **Nucleic Acids Research**, *[S. l.]*, v. 41, n. D1, p. 955–961, 2013. Disponível em: https://doi.org/10.1093/nar/gks1111

ZHANG, Yang. Progress and challenges in protein structure prediction. **Current Opinion in Structural Biology**, *[S. l.]*, v. 18, n. 3, p. 342–348, 2008. Disponível em: https://doi.org/10.1016/j.sbi.2008.02.004

ZHOU, Ping *et al.* Mcl-1 in transgenic mice promotes survival in a spectrum of hematopoietic cell types and immortalization in the myeloid lineage. **Blood**, *[S. l.]*, v. 92, n. 9, p. 3226–3239, 1998. Disponível em: https://doi.org/10.1182/blood.v92.9.3226

ZUGAZAGOITIA, Jon *et al.* Current Challenges in Cancer Treatment. **Clinical Therapeutics**, *[S. l.]*, v. 38, n. 7, p. 1551–1566, 2016. Disponível em: https://doi.org/10.1016/j.clinthera.2016.03.026