IN SILICO TARGET DETERMINATION AND IDENTIFICATION OF NOVEL AGENTS AGAINST CHEMORESISTANT ACUTE LYMPHOBLASTIC LEUKEMIA

BAŞAK ÖZAY

Thesis for the Master's Program in Bioengineering

Graduate School

Izmir University of Economics

Izmir

2024

IN SILICO TARGET DETERMINATION AND IDENTIFICATION OF NOVEL AGENTS AGAINST CHEMORESISTANT ACUTE LYMPHOBLASTIC LEUKEMIA

BAŞAK ÖZAY

THESIS ADVISOR: ASST. PROF. DR. YAĞMUR KİRAZ DURMAZ

A Master's Thesis

Submitted to

the Graduate School of Izmir University of Economics

the Department of Bioengineering

Izmir

2024

ETHICAL DECLARATION

I hereby declare that I am the sole author of this thesis and that I have conducted my work in accordance with academic rules and ethical behaviour at every stage from the planning of the thesis to its defence. I confirm that I have cited all ideas, information and findings that are not specific to my study, as required by the code of ethical behaviour, and that all statements not cited are my own.

Name, Surname: Başak Özay

Date: 25.01.2024

Signature:

ABSTRACT

Determination of Differentially Expressed Genes and Possible Inhibitors Against Chemoresistant Subtypes of Acute Lymphoblastic Leukemia Through in Silico Methods

Özay, Başak

Master's Program in Bioengineering

Advisor: Asst. Prof. Dr. Yağmur Kiraz Durmaz

January, 2024

Acute lymphoblastic leukemia is a malignancy of lymphocyte origin. Although it has high survival rates among children, relapse and drug resistance are still obstacles in treatment. As such, identifying novel genes associated with chemoresistance against first-line treatment drugs and determining possible inhibitors remains crucial. As such, this study utilized a GEO dataset, GSE635, containing gene expression data of asparaginase, prednisolone, daunorubicin and vincristine resistant and sensitive ALL patients, GSE22529 with expression data of eleven healthy subjects, and GSE19143 for validation. Using RMA normalization and LIMMA, the files were analyzed for differentially expressed genes, pathway, and protein-protein interactions. Two proteins selected from this analysis went through molecular docking. Molecular dynamics simulations were done on GROMACS for further validation. 1294 upregulated differentially expressed genes, 25 hub genes, and 12 common genes were identified in resistant ALL types. Among KEGG pathways, PI3K-Akt and pathways in cancer were

significantly enriched. 3556 small molecules were screened against two proteins and following ADMET analysis, three possible inhibitor candidates emerged. MD analysis against one of the proteins corroborated the findings, and while results were similar across all three, they pointed to Eltrombopag having better potential. Additional cytotoxic analyses on Ph+ ALL cell line SUP-B15 and Ph- ALL cell line Jurkat demonstrated similar effects, and analysis on HUVEC cells revealed the drugs had significantly less anti-proliferative effects on healthy cells. This study reveals potential common target genes for chemoresistant ALL as well as other novel genes for diagnostic screening, and proposes three potential inhibitors through drug repurposing.

Keywords: Acute lymphoblastic leukemia, drug resistance, differentially expressed genes, in silico screening, drug repurposing, molecular dynamics.

ÖZET

Akut Lenfoblastik Löseminin Kemorezistan Alt Tiplerine Karşı Diferansiyel Eksprese Edilen Genlerin ve Olası İnhibitörlerin in Silico Yöntemleriyle Belirlenmesi

Özay, Başak

Biyomühendislik Yüksek Lisans Programı

Tez Danışmanı: Dr. Öğretim Üyesi Yağmur Kiraz Durmaz

Ocak, 2024

Akut lenfoblastik lösemi, lenfosit kökenli bir malignitedir. Hayatta kalma oranı yüksek olmasına rağmen nüks ve ilaca direnç tedavide engel teşkil etmektedir. Bu nedenle, kemoterapi ilaçlarına karşı direnç ile ilişkili yeni genlerin tanımlanması ve bu direncin üstesinden gelebilecek olası inhibitörlerin belirlenmesi hayati önem taşımaktadır. Bu çalışma, asparaginaz, prednizolon, daunorubisin ve vinkristin dirençli ve hassas ALL hastalarının gen ekspresyon verilerini içeren GEO veri seti GSE635'in yanı sıra, onbir sağlıklı bireyin ekspresyon verilerini içeren GSE22529'u ve doğrulama seti olarak GSE19143'ü kullanmıştır. RMA normalizasyonu ve LIMMA kullanılarak dosyalar diferansiyel olarak eksprese edilen genler açısından analiz edilmiştir ve bu analizden seçilen iki protein, moleküler yerleştirme işleminden geçirildikten sonra bulunan olası inhibitörlerin GROMACS üzerinde moleküler dinamik simülasyonları yapılmıştır. Bunun sonucunda 1294 tane ekspresyonu anlamlı derecede artmış gen ve 25 merkez gen bulunmakla birlikte, 12 gen dört dirençli tipte ortak çıkmıştır. KEGG yolakları

arasında PI3K-Akt ve kanserdeki yolaklarda önemli ölçüde zenginleşmiş gen olduğu görülmüştür. 3556 küçük molekülün iki proteine karşı taranması ve düşük bağlanma enerjili moleküllerin ADMET analizi ile incelenmesinin ardından üç inhibitör adayı ortaya çıkmıştır. Proteinlerden birine karşı MD analizi bağlanma bulgularını doğrulamak için kullanmış ve Eltrombopag'ın daha iyi bir inhibitör olma potansiyeli olduğunu göstermiştir. Ek olarak, Ph+ ALL hücresi SUP-B15 ve Ph- ALL hücresi Jurkat üzerin yapılan sitotoksik analizler benzer etkiler göstermiş ve ayrıca HUVEC hücreleri üzerinde yapılan analizler, ilaçların sağlıklı hücreler üzerinde önemli ölçüde daha az anti-proliferatif etkiye sahip olduğunu ortaya çıkarmıştır. Bu çalışma, kemorezistan ALL için potansiyel ortak hedef genleri ortaya çıkarmakta ve ilacın yeniden kullanılması yoluyla üç potansiyel inhibitör önermektedir.

Anahtar Kelimeler: Akut lenfoblastik lösemi, anlamlı gen ifadesi, ilaç direnci, in siliko analiz, ilaç yeniden konumlandırılması, moleküler dinamik.

ACKNOWLEDGEMENT

For everything she has done throughout my Master's education, I would like to first thank my advisor, Asst. Prof. Dr. Yağmur Kiraz Durmaz. Not only has she helped me gain important technical skills that will be invaluable to my future education and career, she has had patience with me when I was stuck, and always had feedback to improve my work. I feel privileged to be among her first graduate students.

I would also like to extend my thanks to my jury members Asst. Prof. Atakan Ekiz, for his insightful feedback, as well as Asst. Prof. Gizem Ayna Duran, for helping me navigate challenges and providing aid as I was working on my thesis.

I would also like to thank my lab members and friends for all their support and feedback during the past two years, especially my best friend Yağmur Tükel who helped me overcome many challenges both in and out of the lab.

Lastly, I would like to thank my parents, Neval Özay and Özcan Özay, for all their patience and encouragement as I worked on my thesis. Their curiosity and belief in what I do has been priceless.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1. Acute Lymphoblastic Leukemia

Throughout the world, cancer is among the leading causes of death and a significant issue that any person of all ages can face. The United States National Center of Health Statistics expects near 2 million new cancer cases and around 600.000 cancer deaths to arise in the US (Siegel et al., 2023). Similarly, trend-based predictions using data from the International Agency for the Research on Cancer (IARC), anticipates global cancer incidence to double by 2070 (Soerjomataram and Bray, 2021). As such, identifying novel genes and mechanisms related to cancer progression and discovering new treatments is necessary to lower the global burden. Amongst cancer types, leukemias have the highest incidence rate in children and adolescents. According to GLOBOCAN statistics, in 2020 there were 474,519 new leukemia cases worldwide, of which 7,023 were seen in Türkiye (Sung et al., 2021). There are four main types of leukemia: acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), and chronic myeloid leukemia (CML) (Pejovic and Schwartz, 2002). While ALL is a rarer type of cancer in general, seen in 1.3 per 100,000 people (Pejovic and Schwartz, 2002), with a 25% incidence rate it is the most common cancer for those under 15 years old (Kakaje et al., 2020) and notably, constitutes 97% of childhood leukemias (Carroll and Bhatla, 2016).

ALL is a hematological malignancy that emerges from the aberrant and excessive proliferation of immature cells of lymphoid lineage, and can spread to extramedullary sites, blood, and bone marrow (Malard and Mohty, 2020). It can mainly be categorized into two types by lineage: T-lymphoblastic leukemia and B-lymphoblastic leukemia (Iacobucci and Mullighan, 2017), both of which can be further classified into various subgroups. The exact cause of ALL is unknown still, as most patients are healthy individuals before the outset of the disease (Terwilliger and Abdul-Hay, 2017). Nonetheless, assorted studies show ALL can be related to genetic predisposition and environmental factors. Disorders such as ataxia-telangiectasia and Down syndrome have been associated with increased rates of T-ALL and B-ALL, respectively (Inaba and Mullighan, 2020). Additionally, Bloom syndrome and neurofibromatosis type I, as well as Li-Fraumeni syndrome, DNA repair syndromes

such as Nijmegen breakage, and constitutional mismatch repair deficiency syndrome are related to a higher risk of disease occurrence (Inaba and Mullighan, 2020; Onciu, 2009). Moreover, some viral infections, such as human immunodeficiency virus (HIV) (Bacci et al., 2013) and human T-cell leukemia virus-1 (HTLV-1) (Vadillo et al., 2018), along with environmental factors such as pesticide exposure and ionizing radiation (Malard and Mohty, 2020; Onciu, 2009) have been associated with the development of ALL.

1.1.1. Classification of ALL

ALL can be classified in various ways: cytochemical features, morphological characteristics, immunological traits, and cytogenic and molecular characteristics, which was developed by the World Health Organization (WHO) (Carroll and Bhatla, 2016). According to the French-American-British (FAB) classification based on light microscopy results, ALL can be separated into types L1, L2, and L3. In the most common FAB1 L1 subtype, ALL cells have sizes ranging between small and medium, have condensed nuclear chromatin, a sparse amount of cytoplasm, and barely perceptible or no nucleoli. If cells are larger, with thinly distributed chromatin, a middling amount of pale basophilic cytoplasm, and notable nucleoli, they are classified into FAB L2. In the rarest FAB L3 subtype, the cells are large, with roughly clustered nuclear chromatin, irregularly notable nucleoli, and a hefty amount of basophilic and sometimes vacuolated cytoplasm (Onciu, 2009). Later on, following the need for a classification model that focused on the functionality of cellular differentiation and maturation, the morphology-immunology-cytogenetics (MIC) (First MIC Cooperative Study Group, 1986) classification was developed. According to MIC, B-ALL has four major immunologic subtypes: early B-precursor ALL, common ALL, pre-B-ALL, and mature B-cell ALL, whereas T-ALL is categorized into two as early T-precursor ALL and T-ALL (Harrison and Johansson, 2015). A more comprehensive classification was created by the WHO and updated in 2016, which classified ALL into three main categories: T-cell lymphoblastic leukemia/lymphomas, B-cell lymphoblastic leukemia/lymphoma, and B-cell lymphoblastic leukemia/lymphoma with recurrent genetic abnormalities (Terwilliger and Abdul-Hay, 2017), which will be explained in further detail in the following text.

T-lymphoblastic ALL

T-ALL is seen in 15% of childhood ALL and 25% of adult ALL patients. It is a more aggressive, high-risk risk, and considerably heterogeneous type with a more unfavorable prognosis (Harrison and Johansson, 2015). In recent years, early T-cell progenitor ALL (ETP-ALL) has been discovered as a more aggressive, infiltrative type, taking the name from its similarities to the transcriptional and phenotypical profiles of early T-cell progenitors (Vadillo et al., 2018), which are cells that have recently moved from the bone marrow to the thymus with a high level of multilineage pluripotency (Jain et al., 2016). ETP-ALL is determined with a specific immunophenotype (CD1a negative, CD8 negative, CD5 negative, and myeloid or stem cell marker positive)(Morita et al., 2021). T-ALL itself emerges from the abnormal growth and transformation of T-cell precursors. Based on the intrathymic differentiation stages, The European Group for the Immunologic Classification of Leukemia separates T-ALL into pro-T (cCD31, sCD32, CD1a2, CD21, CD52, CD71, CD342), pre-T/immature (cCD31,sCD32, CD1a2, CD21, CD51, CD71, CD342), cortical-T (cCD31, sCD31 /2, CD1a1, CD21, CD51, CD71, CD342), and mature-T (cCD31, sCD31, CD1a2, CD21, CD51, CD71, CD342) (Szczepański et al., 2003). However, the prognostic value of the differentiation stages changes according to studies, and the modern chemotherapy and minimal residual diseases (MRD) based risk stratification has led to further loss of prognostic value for this classification (Raetz and Teachey, 2016).

Although T-ALL is a greatly diverse disease, its lesions can be separated into two types: those that have mutations and deletions that alter the cell cycle or signaling, and those with chromosomal translocations related to specific gene expression (Raetz and Teachey, 2016). The malignant transformation of T-cells involves various alterations of cell growth, propagation, and differentiation pathways of thymocyte development, the most notable being the constitutive activation of Notch signaling (Van Vlierberghe and Ferrando, 2012). In over 70% of T-ALL cases, the p16/INK4A and p14/ARF suppressor genes are deleted (Aifantis et al., 2008), which, along with Notch activity, establishes the root of T-ALL pathogenesis and oncogenic programming (Van Vlierberghe and Ferrando, 2012). One of the hallmarks of cancer, the deregulation of cell cycle is an issue in T-ALL that is also related to the loss of

p16/INK4A and p14/ARF, which arrest cells at the G1 phase by the inactivation of cyclin D-CDK4 and CKD6 complexes and inhibit MDM2, a negative P53 regulator, due to cellular stress related cell cycle arrest and apoptosis, respectively (Kamijo et al., 1998). Additionally, 15% of T-ALL cases have a loss of function of retinoblastoma 1 (Van Vlierberghe et al., 2013), 12% have deletions on the gene that encodes for p27/KIP1, a cyclin E-CDK2 and cyclin D-CDK4 complex inhibitor (Remke et al., 2009), and 3% have t(12;14)(p13;q11) and t(7;12)(q34;p13) translocations that are related to increased cyclin D2 (CCND2) expression, which is normally downregulated during the differentiation of T-cells (Clappier et al., 2006). Asides from these, abnormal expressions of transcription factors, such as oncogenes bHLH, LMO, and HOX family genes, and MYC, as well as tumor suppressors RUNX1, ETV6, and WT1 are prevalent in T-ALL (Belver and Ferrando, 2016).

B-lymphoblastic ALL

Much more common when compared to T-ALL, B-ALL is seen in around 75% of adult and 85% of childhood ALL cases (Cobaleda and Sánchez-García, 2009). Diagnosed in 26% of children between ages 0-14, B-ALL is among the most common childhood cancers (Malouf and Ottersbach, 2018). Typical clinical symptoms of B-ALL include bone marrow failure and cytopenia, sometimes with leukocytosis (Loghavi et al., 2015). It's a disease with a variety of genetic subtypes, depending on chromosomal alterations or rearrangements. According to the International Consensus Classification, there are over 20 B-ALL subtypes (Figure 1) (Duffield et al., 2023). Identification of the subtypes present in a given patient is a very important part of diagnosis due to this abundance of genetic aberrations.

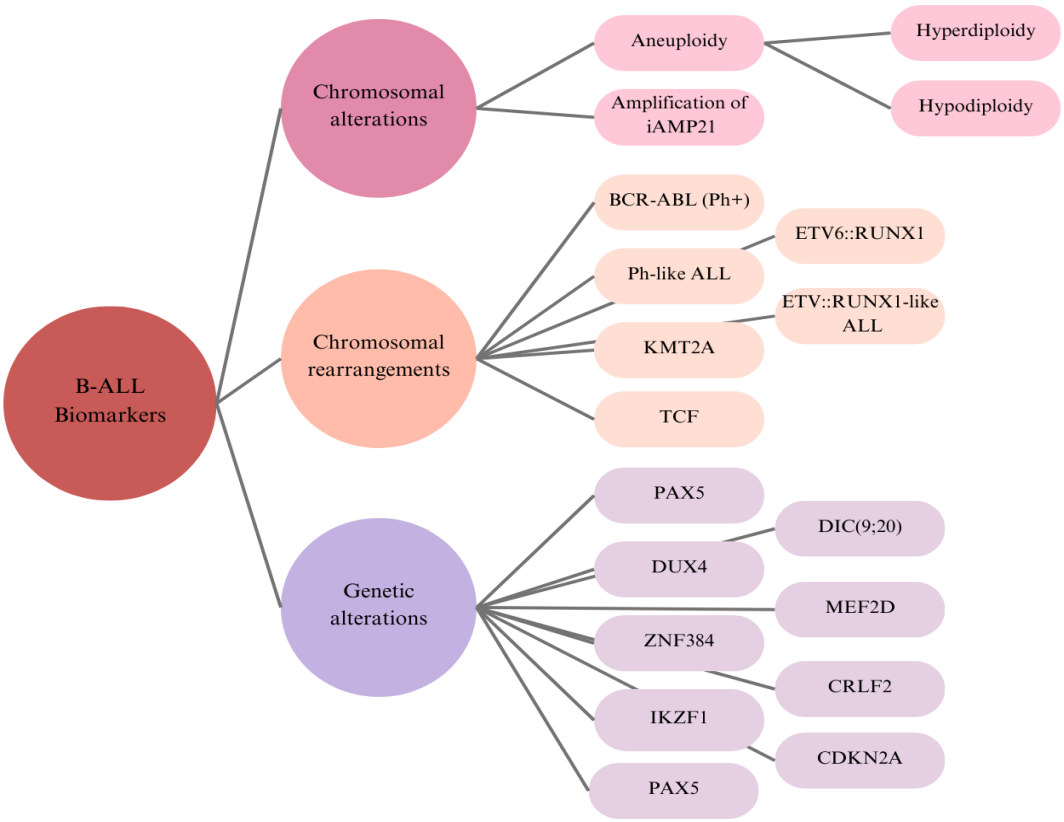

Figure 1. Biomarkers of B-ALL separated based on chromosomal alterations, chromosomal rearrangements and genetic alterations. Adapted from (Lejman et al., 2022).

B-ALL subtypes generally depend on the first genetic hit, commonly either aneuploidy or chromosomal translocations such as ETV6-RUNX1, with additional mutations and rearrangements that follow (Malouf and Ottersbach, 2018), predominantly involving genes such as tyrosine kinases, cytokine receptors, and transcription factors (Iacobucci and Mullighan, 2017). Aneuploidy in ALL presents as either hypodiploidy or hyperdiploidy (Lejman et al., 2022). The most common B-ALL genetic aberration, seen in over 35% of childhood cases, hyperdiploidy can be separated into several categories depending on the the presence of extra copies of chromosome 21, as hyperdiploid (47 to 58 chromosomes), hypotriploid (59 to 68 chromosomes) and hypertriploid (70 to 80 chromosomes) (Haas and Borkhardt, 2022). Seen in around 30% of childhood B-ALL cases, the ETV6-RUNX1 translocation t(12;21)(q13;q22) is related with a more favorable prognosis, with higher 5-year survival rates and better treatment response (Wang et al., 2018). Another one of the

common B-ALL subtypes contains the BCR-ABL fusion gene, also referred to as Philadelphia positive or Ph+ B-ALL, that is created through the translocation of chromosomes 9 and 22, with t(9;22)(q34:q11.2) on chromosome 22 being the most commonly seen abnormality among that subtype (Wieduwilt, 2022). BCR-ABL fusion is seen in up to 30% of adult patients, and 5% of children (Loghavi et al., 2015), and is known for its worse prognosis: lower 5-year survival rates or overall survival (Moorman et al., 2007), and a higher change of chemoresistance to traditional cytotoxic drugs (Wieduwilt, 2022). The secondary mutations that follow the aforementioned subtype determining ones can be acquired during disease progression and therapy. Typically, these mutations are seen in lymphoid transcription factors such as PAX5, IKZF1, ETV6, tumor suppressors such as TP53 and RB1, lymphoid signaling regulators such as CD200 and BTLA and, RAS pathway proteins such as KRAS and NRAS, and chromatin modifiers such as CREBBP and SETD2 (Roberts and Mullighan, 2020) .

Other Lymphoblastic Abnormalities

In addition to T- and B-ALL, there are two smaller subtypes of ALL: Mixed Phenotype Acute Leukemia (MPAL) and natural killer cell ALL (NK-ALL). Seen in less than 4% of patients (Vardiman et al., 2009), MPAL refers to leukemias of unclear cell lineage, that can have T and B-cell populations as well as monocytic or myeloid cells, one type of blast cells with B- or T-cell antigens with myeloid antigens, and/or monocytic antigens (Chiaretti et al., 2014). NK-ALL is an even rarer subtype, seen in around 3% of adult patients (Vardiman et al., 2009), presenting with CD56, one of the NK cell differentiation markers, and early T-cell antigens such as CD2, CD5, CD7 (Chiaretti et al., 2014).

1.1.2. Treatment of ALL

The chemotherapy treatment of ALL takes around two to three years and consists of three stages: remission induction, consolidation, and maintenance, as well as intermittent central nervous system (CNS) prophylaxis (Terwilliger and Abdul-Hay, 2017). The induction stage uses a glucocorticoid (either prednisone or dexamethasone), vincristine, and asparaginase, in some cases anthracycline is given as well, and takes around four to six weeks to induce complete remission in nearly 98% of pediatric ALL patients (Inaba and Mullighan, 2020). In the consolidation phase, patients are administered cyclophosphamide, cytarabine, and mercaptopurine. It was found that, for B-ALL patients, if MRD is negative when both induction and consolidation phases end, the 5-year event-free survival (EFS) is 92.3%. However, if MRD is positive and less than 10^{-3} or over 10^{-3} at the end of consolidation, the EFS falls down to 77.6% and 50.1%, respectively (Conter et al., 2010). After the consolidation stage, drugs similar to those in the induction and consolidation stages are administered to both standard- and high-risk patients as reinduction therapy (Inaba and Mullighan, 2020), a crucial part of ALL treatment that lowers the chance of relapse in standard-risk patients (Schrappe et al., 2018). This is followed by the maintenance stage, which lasts at least one year, with daily mercaptopurine and weekly methotrexate administration, which can be supplemented with vincristine and steroid (Inaba and Mullighan, 2020).

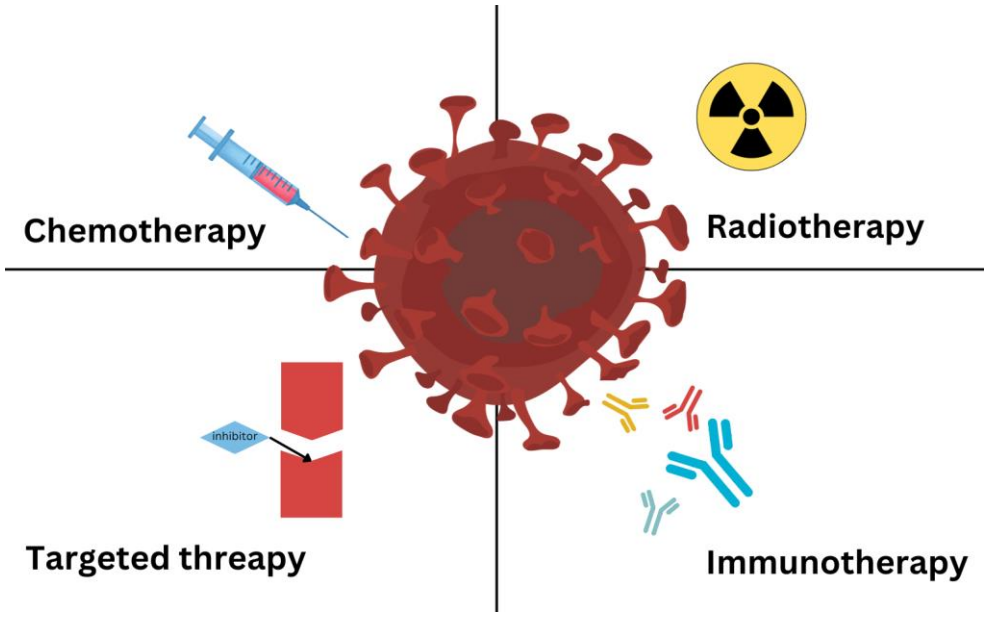

Figure 2. Types of ALL treatments, adapted from (Malczewska et al., 2022).

Besides chemotherapy, radiotherapy, targeted therapy such as proteasome and JAK inhibitors, and immunotherapy are alternative treatment methods for ALL (Figure 2) (Terwilliger and Abdul-Hay, 2017). Currently, three types of immunotherapy drugs are in use for pediatric ALL: CAR-T cells, bispecific antibodies such as blinatumomab, and antibody-drug conjugates such as inotuzumab (Sotillo et al., 2015). Among targeted therapy drugs, bortezomib, a proteasome inhibitor, wasn't found effective by itself in relapsed ALL patients (Cortes et al., 2004), although a phase 2 study with children with relapsed pre-B-ALL shows its combined use with vincristine, doxorubicin, dexamethasone and pegylated asparaginase results improved response in 80% patients (Messinger et al., 2012). Recent clinical study results show that while the addition of bortezomib to the induction and reinduction phases was linked to improved 3-year-EFS in standard- and intermediate-risk T-ALL patients, it was linked to worse outcomes for high-risk patients (Teachey et al., 2020). Other proteasome inhibitors, such as ixazomib and carfilzomib are currently being studied for ALL (Inaba and Pui, 2021). Similarly, other targeted therapy drugs, such as tyrosine kinase inhibitors are used in combination with standard chemotherapy in order to increase effectiveness. ABL1 inhibitors such as imatinib, nilotinib, dasatinib and ponatinib are used in the treatment of BCR/ABL+ ALL and ABL1-class fusions in BCR/ABL like ALL and T-ALL (Slayton et al., 2020; Tasian et al., 2017). Nonetheless, at present, targeted therapy can be used in only around 10% of pediatric ALL patients (Inaba and Pui, 2021). As such, gaining a further understanding of ALL etiology and pathology, and identifying target genes is necessary for developing new inhibitory drugs.

This thesis is mainly related to the chemotherapy drugs asparaginase, daunorubicin, prednisolone, and vincristine, and as such further information on these four will be provided in the following text.

Asparaginase

Asparagine is a nonessential amino acid necessary for the synthesis of DNA, RNA, and proteins (Kawedia and Rytting, 2014). While healthy cells can synthesize it using asparagine synthetase, ALL cells are unable to make this amino acid due to hypermethylation on the promoter of asparagine synthetase, rendering asparagine an

essential amino acid that they must get from the serum (Ren et al., 2004; Worton et al., 1991). As such, the possibility of using asparaginase, an enzyme required to hydrolyze asparagine to aspartic acid in the serum, therefore stopping cancer cells from using it, emerges with using asparaginase in humans back in the early 1960s, followed by larger clinical trials in the 1970s, that illuminated the enzyme's benefit for ALL patients (Egler et al., 2016).

Daunorubicin

Daunorubicin is a type of anthracycline antibiotic that was first used as a treatment method for AML (Gewirtz, 1999), but has later begun to be used in various cancers such as small lung cancer and breast cancer, and is currently still a part of first-line ALL treatment (Shandilya et al., 2021). Its anti-neoplastic effects work through inhibiting DNA and RNA transcription of cancer cells, although its specific pathways have not reached a consensus (Samosir et al., 2021). Clinical trials have reported that in combination with prednisone and vincristine, daunorubicin increases the complete remission of pediatric ALL (Willemze et al., 1975). Due to recent studies that link Daunorubicin with cardiotoxicity, several studies are underway to determine analogues in its place (Shandilya et al., 2021).

Prednisolone

Prednisolone is a glucocorticoid and the active form of prednisone, and its first clinical trial for ALL treatment was done in 1971 by the Cancer and Leukemia Group B (CALGB) with 493 children (Inaba and Pui, 2010). The mechanisms of actions of glucocorticoids is dependent on binding the glucocorticoid receptor to stop growth and cause apoptosis (Pufall, 2015). They work predominantly through the AP-1 and NfkB pathways (Inaba and Pui, 2010).

Vincristine

In use against cancer since the 1960s, vincristine is still among the most common and effective drugs in childhood cancers (Moore and Pinkerton, 2009). Its mechanism of action is based on its obstruction of microtubule formation, and intracellular transport, eventually leading to apoptosis (Gidding et al., 1999).

1.1.3. Drug Resistance in B-cell AL

Drug resistance emerges as a combination of various different pathways and mechanisms, and effects different drugs. Cancer cells in general are able to develop drug resistance throughout or after treatment, through more mutations or metabolic changes, or have more mutations related to resistance at diagnosis (Housman et al., 2014). In ALL, although the survival rates in high income countries are over 90% for pediatric patients, the remaining are deemed incurable, and the relapse rate is around 20% (Jędraszek et al., 2022). Resistance in ALL arises through different ways, at diagnosis patients may already have minor relapse initiating subclones, or may develop mutations in genes such as PRPS1 and NT5C2 during therapy, or have increased mutations after relapse, such as CREBBP and SETD2, which are known to have increased mutations at B-ALL relapse (Inaba and Mullighan, 2020). CELSR2 decrease has been linked to increased BCL2 expression and glucocorticoid resistance in ALL cells, and PI3K-Akt and mTOR pathways are known to be associated with chemoresistance in ALL. Furthermore, epigenetic modifications, such as DNA methylation and histone modifications have been found to be related to drug resistance and relapse in ALL as well (Inaba and Pui, 2021). Several studies also point to hypoxia as a cause of resistance (Petit et al., 2016). Another driver of drug resistance in leukemia is the loss of IKZF1, which disrupts cell adhesion, metabolic pathways, and the target gene regulators of glucocorticoid receptors, as well as PI3K and mTOR, leading to glucocorticoid resistance (Aberuyi et al., 2019). The RAS pathway has also been linked to drug resistance in various cancers including ALL. Resistance to targeted therapy and immunotherapy is also an issue, such as loss of CD19, which causes resistance against both CAR-T and blinatumomab (Inaba and Pui, 2021).

1.2. Computational Approaches in Cancer

Researching and gaining new insights against cancer, its pathology, the mechanisms and pathways it uses to escape cell death, and to develop drug resistance, is important to discover ways to prevent, diagnose, and cure this complicated disease. To this end, computational and bioinformatics approaches have been becoming more and more popular.

1.2.1. Microarray Analysis

Microarrays can be used for the analysis of proteins, mRNA, DNA or other biological material (Wu et al., 2008). In this thesis, the microarrays chosen for analysis contain gene expression profiles of a given cell or patient data, what is essentially a snapshot of all transcriptional activity, which enables a straightforward analysis of a substantial amount of genes in concert with each other, instead of the more conventional single genes approach or a small group of gene studies (Slonim and Yanai, 2009). This global approach has expedited the determination of new disease subtypes, novel genes, and mechanisms related to the disease progression or drug resistance. Gene expression microarray technology has evolved in recent years to concurrently measure all mRNA transcripts (transcriptome) (Masuda and Yamada, 2015), with recent microarrays containing over a hundred thousand DNA probes. The technology behind this depends on base-pair hybridization: the mRNA is extracted from tissue or cell lines, complementary RNA is produced, labeled with fluorescent dyes, and hybridized to a DNA sequence, and the color intensities of probes is determined with a laser scanner utilizing a software specifically designed for microarrays (Tao et al., 2017). The R programming language is designed for statistical analysis, and through Bioconductor (https://www.bioconductor.org/), an open-source software, packages can be used to analyze microarray data.

One of the types of microarray analysis is classification studies. Previous research shows that microarrays can be used to identify genes that increase survival in early-stage lung carcinoma (Beer et al., 2002) and to determine patient subgroups of AML based on molecular signatures (Bullinger et al., 2004; Valk et al., 2004). More recently, microarray analysis has been used to analyze biomarkers related to the

prognosis and treatment response of cervical cancer (Lin et al., 2019) and identify microRNA biomarkers for stage two colorectal cancer (Gungormez et al., 2019). In recent years, microarray analysis of differentially expressed genes (DEGs) has started to be done either separately or with integrated RNA-Seq analyses of various cancers and diseases (Chen et al., 2017; Nisar et al., 2021; Wolff et al., 2018). Although researchers can develop the microarrays they will use for DEG analysis themselves, existing microarray studies can be accessed for analysis as well. Established in 2000, Gene Expression Omnibus (Edgar et al., 2002), GEO for short, is a free, international archive that contains data from various types of experiments, such as high-throughput sequencing, or microarrays, of gene expression, genome occupancy, single nucleotide polymorphism or protein profiles (Clough and Barrett, 2016). Since its establishment, researchers and institutions across the world have uploaded raw and processed data that can be accessed by anyone wishing to utilize it for a different research topic, abiding by journal or grant rules that ask for their data to be uploaded to a publicly available database (Barrett et al., 2005). The data uploaded to GEO is stored in three types: Platform, Series, and Sample. Platform depicts which elements, such as antibodies, oligonucleotide probes, cDNAs, etc., are assayed at the particular experiment. Sample provides an abundance measurement of each element and indicates the Platform, and Series compiles the affiliated samples of an experiment (Edgar et al., 2002). According to GEO's statistics, at the time of this thesis, the archive holds 212.544 public Series and 6.800.397 Samples. The various data in GEO come in different file formats depending on the experiment type. Some of these files can only be accessed using the R program, and Bioconductor packages such as "affy" (Gautier et al., 2004) and "oligo" (Carvalho and Irizarry, 2010). After accessing and normalizing the microarray data, Linear Models for Microarray Data, or the "LIMMA" package (Smyth, 2005) is seen as the golden standard for the analysis of microarrays (Caiazzo et al., 2011; Ritchie et al., 2015). LIMMA utilizes a combination of several statistical methods to perform large-scale expression studies (Zhang et al., 2009). In LIMMA analysis, a matrix is made containing expression values in rows, and probes in columns and a linear model is fit to every row. Its most characteristic method is using linear models to analyze the data as a whole, therefore making it easier to determine correlations that can be present for several reasons, such as repeated measures (Ritchie et al., 2015). LIMMA uses statistical methods that can enable information borrowing via empirical Bayesian methods (Casella, 1992) to tolerate data

quality differences, facilitating variance modeling for the adjustment of biological and technical heterogeneity and pre-processing to lessen the noise (Ritchie et al., 2015). At the end of LIMMA analysis of microarrays, a list of DEGs, both upregulated and downregulated is generated.

1.2.1.1 Gene Ontology and Enrichment Analysis

After the popularization of genome-wide studies, enrichment analysis has been developed to determine whether the set of DEGs discovered from microarray analysis is involved in specific pathways or functional groups, which have been characterized by gene ontology (GO): particularly molecular function (MF), cellular component (CC) and biological process (BP) (Ashburner et al., 2000). MF is predominantly related to molecular actions, such as catalytic behavior, and focuses on the actions themselves, whereas although similar, BP is centered more on assemblies of molecules, such as signal transduction. And lastly, CC describes the locations of genetic products, such as the nucleus (Gupta et al., 2021). Enrichment analysis, also referred to as pathway analysis, while similar to GO analysis, delivers a profound grasp of the mechanisms of diseases (Wang et al., 2007), calculating disease phenotype associations of pathways, and describing gene interactions in a reliable way that has made it valuable for researchers (Curtis et al., 2005). Several methods are used for pathway analysis: namely, functional class scoring (FCS), overrepresentation analysis (ORA), and pathway topology-based analysis (PTB) (Gupta et al., 2021). In the last two decades, various tools have been generated for GO and pathway enrichment analysis, such as DAVID (Huang et al., 2007), Enrichr (Kuleshov et al., 2016), ShinyGO (Ge et al., 2020) and FunCluster (Henegar et al., 2006). Out of these, due to its user-friendly interface, and ease of use, ShinyGo was selected as the annotation and pathway analysis tool.

1.2.1.2 Protein-Protein Interaction Networks

Upon receiving an annotated gene list, the next step in microarray analysis is to develop a PPI network. This analysis is important for determining the function of a

set of proteins and drug ability of small molecules (Rao et al., 2014). PPIs show an extensive reaction network of a given set of proteins depending on an existing database and can also be referred to as the interactome (Garland et al., 2013). They are related to almost all processes, especially in higher organisms, and determining PPIs are crucial in understanding cancer and identifying target genes. To this end, several databases have been made available, including String (Szklarczyk et al., 2023), MatrixDB (Chautard et al., 2011), MPDI (Goll et al., 2008). This thesis uses String V11 to develop a PPI network, due to its ease of use, and its comprehensive list of databases (Szklarczyk et al., 2019). The network generated by String can be exported into Cytoscape, which can be used to visualize the network and determine hub genes using plug-ins (Shannon et al., 2003). One plug-in, Cytohubba (Chin et al., 2014), can be used to find hub genes, which are genes that are believed to be capable of regulating networks, with some proposing that they are responsible for the entire network's functionality (Farber and Mesner, 2016).

1.2.2. Druggability and Pocket Analysis

Cavities within or on the surface of proteins are termed binding pockets, and the substrate binds to what is generally the biggest pocked, termed the active site (Liang et al., 1998). If a molecule binds to a different pocket than the active site and causes conformational changes, it can regulate a number of cell processes such as gene regulation and signaling (Lu et al., 2014). This can be referred to as allosteric regulation, and is the most direct method of protein function regulation (Du et al., 2016). Thus, in recent years, it has become an attractive target for drug discovery (Nussinov and Tsai, 2014). Some of the benefits of allosteric drugs include enhanced selectivity and increased efficiency of competitive inhibitors (W. Huang et al., 2017). For instance, GNF2, an allosteric inhibitor of BCR-ABL, was found to have additive inhibitory effects when combined with imatinib in CML (Zhang et al., 2010). To date, several allosteric modulators have been approved by the FDA, such as cinacalcet and plerixafor (Müller et al., 2012). With this popularity, several protein structure prediction softwares were developed to identify binding pockets and analyze them based on druggability which is the affinity for of target to therapeutic agents, and it is key for the early steps of drug discovery, as near 60% of projects have failed in the

past for being undruggable (Cheng et al., 2007). The druggability of binding pockets is contingent on several factors; hydrophobicity, the composition of aminoacid residues, electrostatistics, and overall geometry (Stank et al., 2016). In recent years, several bioinformatics tools have been developed for the assessment of druggable pockets. In this study, DogSiteScorer (Volkamer et al., 2012), a web-tool found in ProteinPlus, was selected for druggability analysis. This tool detects binding pockets based on protein atom coordinates, and outputs volume, surface, depth, and drug scores of the determined pockets based on these calculations (Volkamer et al., 2012). The steps of druggability analysis can be seen in Figure 3.

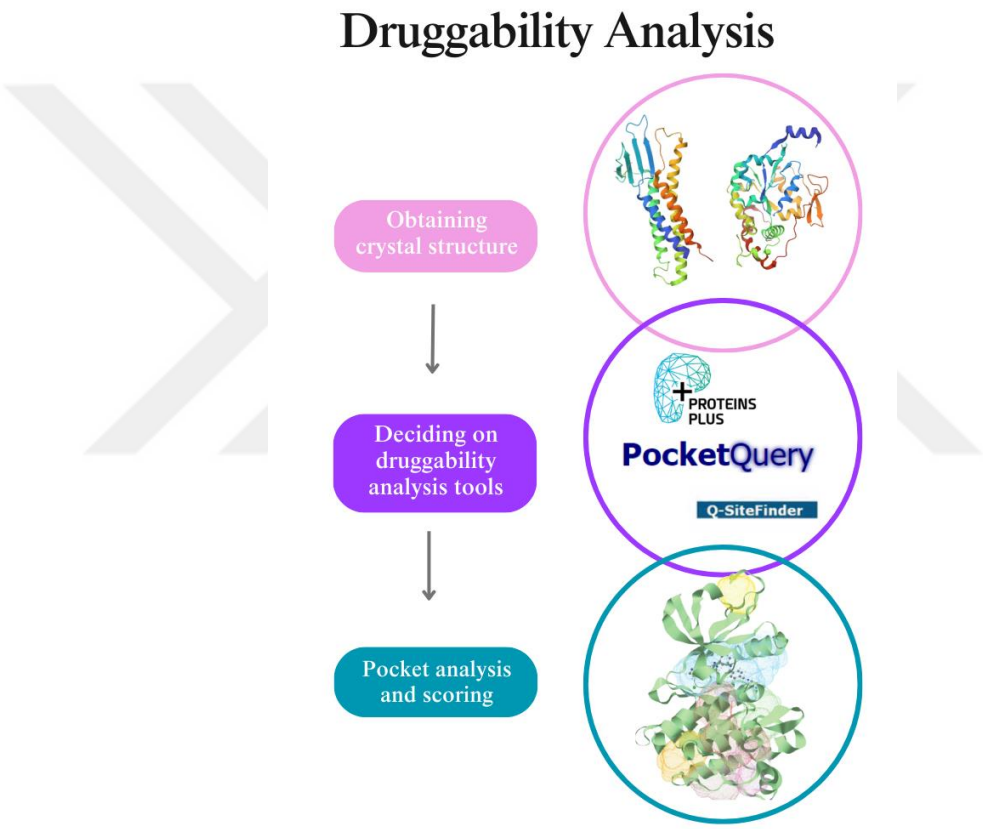

Figure 3. Workflow of druggability analysis steps and tools.

1.2.3. In Silico Screening Methods

Conventional drug discovery is a process that is long and high-cost. On average, it takes fifteen years for a single drug to be developed, approved, and marketed, costing around 2.6 billion dollars (Mohs and Greig, 2017). As nine out of

ten drugs fail clinical trials (Sun et al., 2022), reducing the time and cost of the drug discovery process is key. To this end, computer aided drug discovery (CADD) methods have become increasingly popular. In fact, a 2015 report asserts that CADD methods had been used for half of twenty new molecules in phase I clinical trials at the time (Rognan, 2017). Moreover, CADD methods have been used to effectively establish novel drugs to the market for various diseases, such as HIV1 inhibitors: saquinavir, ritonavir, and atazanavir; antibiotics: norfloxacin; and anti-cancer drugs: raltitrexed (Shaker et al., 2021). One of the CADD methods utilized in the drug discovery process is in silico screening. In this method, molecules from large drug libraries are virtually screened against macromolecules to determine pharmacological activity (Lin et al., 2020). This is also referred to as drug repurposing: investigating the different therapeutic potentials of pre-existing drugs and small molecules. There are two types of in silico screening methods: structure-based and ligand-based.

1.2.3.1. Ligand-based Pharmacophore Screening

In ligand-based drug screening, the structure of the protein is unknown. Instead, in this method novel drug molecules are determined based on chemical, physical, and structural characteristics of existing drugs, using similarity search, pharmacophore modeling, and quantitative structure-activity relationships (QSAR) modeling (Shaker et al., 2021). This method has some limitations to be considered: firstly, ligand stereochemistry is disregarded although ligand recognition generally depends on stereospecificity, and secondly, chemical similarity and biological similarity are not always interdependent, and structurally similar compounds may have significantly different potencies (Rognan, 2017).

1.2.3.2. Structure-based Virtual Screening

Unlike ligand-based screening, structure-based virtual screening works by calculating the binding affinity of ligands upon binding the crystal structure of the target protein through 3D characteristics, such as bruiedness, polarity, volume, hydrophobicity and curvature, of the ligand-bound pockets (Rognan, 2017).

Saquinavir and amprenavir are among the first FDA-approved drugs that were designed using structure-based virtual screening methods (Shaker et al., 2021). For this method, the ligands can be accessed from a compound library, such as DrugBank (Wishart et al., 2008), ZINC15 (Sterling and Irwin, 2015), ChEMBL (Gaulton et al., 2012), or ChemBridge. The target structure may be accessed from the publicly available archive, Protein Data Bank, or if it is not available, may be generated using homology modeling tools. Once it is downloaded, it often goes through preprocessing to get ready for molecular docking, which can be performed using programs such as PyMol or UCSF Chimera (Pettersen et al., 2004). Over sixty types of commercial and academic tools and softwares have been generated to perform the molecular docking simulations (Pagadala et al., 2017); such as AutoDock (Morris et al., 2009), AutoDock Vina (Trott and Olson, 2010), and UCSF Dock (Allen et al., 2015). Docking simulations work through two processes: a search algorithm, and a scoring function (Li and Shah, 2017). For each molecular docking program, the search algorithm that works to analyze the ligand and the binding pocket of the macromolecule is different, varying from shape-based algorithms, incremental construction approaches, and systematic search methods to Monte Carlo Simulations (Pagadala et al., 2017). The scoring function assesses all the possible binding poses and gives a score based on force-field, empirical and knowledge-based algorithms (Li and Shah, 2017). In AutoDock Vina, one of the most commonly used molecular docking programs, as well as the majority of programs, the macromolecule is treated as rigid while the ligand is flexible (Agu et al., 2023).

In this thesis, the molecular docking program was selected as PyRx, a publicly available *in silico* screening software. It combines various softwares including OpenBabel, to convert ligands to the .pdbqt format necessary for the docking process, and AutoDock Vina for molecular docking (Dallakyan and Olson, 2015).

1.2.4. Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) Analysis

After a list of molecules with suitable binding affinities are selected from molecular docking, the next important step is the analysis of absorption, distribution,

metabolism, excretion and toxicity (ADMET) parameters. One of the biggest reasons for novel drugs to fail in clinical trials is because of inadequate bioavailability as a result of high toxicity, and unsuitable pharmacodynamic and pharmacokinetic properties (Kar and Leszczynski, 2020). As such, analyzing these properties is necessary to determine the most ideal of the possible drug candidates. For this purpose, there are various easily accessible web-tools available. In this thesis, two such tools were used: SwissADME (Daina et al., 2017) and AdmetSAR (Yang et al., 2019). Both tools work based on structural analysis, and canonical SMILES notation of each molecule is used as the input. SwissADME analyzes physicochemical properties, lipophilicity, solubility, pharmacokinetics, and druglikeness of a given molecule. Using OpenBabel to measure physicochemical properties, SwissADME employs various models from previous research to calculate lipophilicity and solubility (Daina et al., 2017). In this thesis, SwissADME was used primarily for the drug-likeness results, meaning the possibility of a molecule being an oral drug. The determination of this quality depends on the physicochemical and structural analysis of molecules that are considered drug candidates. SwissADME uses five filters for this parameter: the Lipinski, Muegge, Veber, Egan, and Ghose filters (Daina et al., 2017). Among these filters, Lipinski (Lipinski et al., 2001) and Muegge (Muegge et al., 2001) are among the criterion of pharmacokinetics for drug-likeness (Benet et al., 2016), and the other filters were adapted to SwissADME for further validation and comparison of this feature. According to Lipinski's rule of five, an orally active molecule should have hydrogen bond acceptors ≤ 10, hydrogen bond donors ≤ 5, molecular weight less than 500 Da, and LogP (The logarithm of Octanol-water partition coefficient) ≤ 5 (Lipinski et al., 2001). In the Muegge filter, the numbers of hydrogen bond acceptors and donors are the same as in Lipinski but additionally, it states that the molecular weight must be between 200 and 600 Da, XLogP -2 to 5, topological polar surface area (TPSA) less than 150, and rotatable bonds less than 15 for a molecule to successfully be a drug molecule (Muegge et al., 2001). In this thesis, in addition to Lipinski and Muegge filters, the Ghose filter was used as a consensus filter and states that a drug molecule should have LogP between -0.4 and 5.6, molecular weight between 160 and 480, the total number of atoms between 20 and 70, and molar refractivity between 40 and 130 (Ghose et al., 1999). Another score of interest from this tool was the Abbot Bioavailability Score (Martin, 2005), which calculates the possibility of a compound having 10% oral bioavailability in rats at the minimum or determinable $CaCO^2$

permeability, depending on the violations of Lipinski filters, TPSA, and total charge of the molecules (Daina et al., 2017). Moreover, AdmetSar is used for the further analysis of pharmacokinetics and drug-likeness, as well as toxicity. Giving out biological and chemical information and ADMET predictions, the webtool contains 210.000 experimental data of 96000 molecules and 27 computational models (Yang et al., 2019). Similarly to SwissADME, LogP, TPSA, molecular weight, and hydrogen bond donors and acceptors are calculated with OpenBabel. Using 22 qualitative classification models and five quantitative regression models, SwissADME calculates blood-brain barrier penetration, $CaCO^2$ permeability, human intestinal absorption, rat acute toxicity, and many other factors (Cheng et al., 2012). Thus enabling researchers to gain a better understanding of ADMET properties, saving time and resources in the drug discovery process.

1.3. Molecular Dynamics

First used to analyze proteins in the 1970s (McCammon et al., 1977), molecular dynamics (MD) simulations are used to investigate the kinetics of a given system of particles (Karplus and Petsko, 1990). Since its origin, MD has become an indispensable tool for the analysis of macromolecular structures and protein-ligand or protein-protein interaction analysis. At the present, MD simulations are used for three main reasons: to gain an understanding of a structure's natural dynamics in various timeframes, to measure fluid properties and free energy changes of mechanisms such as ligand binding, and to analyze conformational changes of a protein or complexes in docking simulations (Hansson et al., 2002). In the last five to ten years, MD simulations have been used extensively in areas such as structural analysis and drug design. Especially in neuroscience and drug development for the nervous system (Dawe et al., 2016; Manglik et al., 2016; McCorvy et al., 2018), as well as drug discovery in various cancers (Morshed et al., 2023; Opo et al., 2021; Yoda et al., 2018; Zhang et al., 2017), MD has been a popular tool to validate protein-ligand binding. Based on the molecular configurations of the atoms of a system, such as a protein in water and the atomic interactions, the force of all atoms on each atom can be measured, and Newton's laws of motion can be utilized to infer the location of each atom in space as a function of time to calculate a trajectory which characterizes the system at all time

steps of a selected interval (Hollingsworth and Dror, 2018). To apply the theory to practice, several MD simulation softwares are available such as AMBER (Case et al., 2005), CHARMM (Brooks et al., 1983), GROMACS (Berendsen et al., 1995), and LAMMPS (Plimpton, 1995).

In this thesis, based on easy access, speed of simulation and user support, the MD software was selected as GROMACS 2021.2, which is among the most popular MD simulation softwares (Hospital, A et al., 2015). The next important step after selecting a software is to determine which force field to use. In broad terms, force fields are equations that define how a system's energy depends on its own particle coordinates (González, 2011). In MD simulations, the force field is reliant on quantum mechanical calculations as well as experimental data (Hollingsworth and Dror, 2018). Several force fields are available for MD simulations, typically AMBER (Lindorff-Larsen et al., 2010), CHARMM (J. Huang et al., 2017) or OPLS (Harder et al., 2016), each with several different versions best utilized for different types of analyses. For the analysis in this study, the selected force field was CHARMM27 (Bjelkmar et al., 2010). After the force field is decided, a solvent must be selected that fits the protein complex and the force field (Hollingsworth and Dror, 2018). TIP3P water molecules are typically utilized when CHARMM force fields are selected (Bjelkmar et al., 2010). Once the simulation is performed and trajectory data is obtained, essential parameters of structural analysis, such as root mean square deviation (RMSD), radius of gyration (Rg), root mean square fluctuation (RMSF), and the formation of hydrogen bonds can be inspected through GROMACS.

RMSD analysis is among the most basic methods of structure analysis. The RMSD of the structure is calculated through computing the mean displacement of each particle throughout the trajectory (Grossfield and Zuckerman, 2009). Through this, the conformational change the protein goes through can be determined based on the RMSD levels. RMSF calculates the fluctuations of each residue on the protein backbone based on their displacement from their mean position at all instances of the simulation (Saxena et al., 2009). As such, through this value the flexibility of the polypeptide backbone and how much of a conformational change occurs upon ligand binding can be extrapolated. Rg values assess the root mean square distance of each atom within the protein from the axis of rotation, usually at the center of mass of the

protein (Sneha and George Priya Doss, 2016). Through this value, it is possible to gain insight into the compactness of the protein, and again judge the conformational changes upon ligand binding depending on its results and stability based on its overall fluctuations throughout the simulation. Finally, hydrogen-bonds are calculated to determine the stability of the structure, and have been found to be related to better binding affinity and drug efficacy (Sneha and George Priya Doss, 2016).

In brief, MD simulations are a valuable tool to determine the stability and conformational changes of the protein upon ligand binding, through calculating several parameters, in order to further validate the results of molecular docking analysis.

1.4. Aim of the Study

Although ALL has a high survival rate for children at 90%, 20% of those are known to relapse. Additionally, the remaining 10% are deemed incurable once they don't respond to treatment (Jędraszek et al., 2022). With the current chemotherapy approaches reaching their maximum effectiveness, what is left for traditional ALL treatment is to determine novel genes related to chemoresistance, gain further understanding of drug resistance mechanisms and pathways, and utilize drug repurposing to discover inhibitors, or to drive inspiration to develop new inhibitors from these small molecules.

As such, this thesis aims to identify DEGs that are upregulated in Asp, Dnr, Pred, and Vcr resistant ALL patients, perform GO and PPI analysis to investigate pathways related to these genes, determine DEGs that are common in all four resistant subtypes, and consequently discover potential inhibitors to overcome drug resistance in ALL through docking methods and drug repurposing.

CHAPTER 2: METHODS

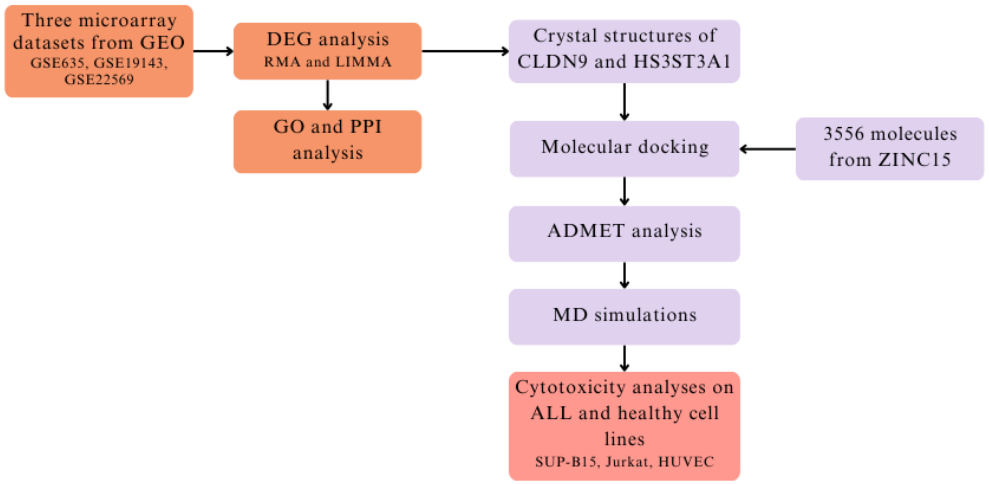

Figure 4. Graphical representation of the workflow of the study.

This study employs a bioinformatics approach to finding DEGs, and identifies potential inhibitors through in silico screening and molecular docking, using MD simulations for the validation of the binding of the potential drugs (Figure 4). In the following sections, a detailed explanation will be given for each step.

2.1. Datasets

Gene Expression Omnibus (GEO) is a public archive that contains data of various high-throughput gene expression and genomic hybridization research in different formats, one of which being microarrays (Edgar et al., 2002). The database enables free distribution and access to high-quality data for researchers across the world. Each unique object in GEO has an accession number which has a different prefix if the record is a GEO Platform (GPL), Sample (GSM), or Series (GSE) (Barrett et al., 2005). Any data within GEO is available for public use and most datasets make raw patient data available for all.

In this thesis, three datasets were selected for analysis after a thorough search of the GEO database. The raw microarray files were obtained in the .CEL file format, which contain probe intensities and are created directly by the array scanner software (Klaus and Reisenauer, 2016).

2.1.1. Healthy B-cell Data

Healthy B-cell data was acquired from a different dataset as there was none in the dataset of interest used in this study and no other resistant ALL dataset with healthy data was available. The gene expression data of eleven healthy individuals was accessed from the GSE22529 dataset (Gutierrez et al., 2010).

2.1.2. ALL Datasets

Acute lymphoblastic leukemia datasets were investigated for those containing sufficient patient data and were related to drug resistance. GSE635 (Holleman et al., 2004) was chosen as it contained expression data of patients (a total of 173) resistant and sensitive to common chemotherapy drugs, asparaginase, daunorubicin, vincristine and prednisolone. Containing prednisolone sensitive data, GSE19143 (Stam et al., 2010) was selected for validation purposes.

Table 1. List of GEO datasets utilized in the study and their characteristics.

Datasets	**Disease**	**Types**	**Sample Size**	**Platform**	**Reference**
GSE22529	Healthy B-cells	Control	11	GLP96	Gutierrez et al., 2010
GSE635	ALL	Drug resistance	173	GLP96	Holleman et al., 2004
GSE19143	ALL	Drug resistance	52	GPL96	Stam et al., 2010

2.2. Microarray Data Analysis

As mentioned, the raw data from the aforementioned datasets was downloaded in the CEL file format from GEO. To access the raw data files and in all the consequent steps, RStudio was the main program and Bioconductor packages "oligo"(Carvalho and Irizarry, 2010), "LIMMA" (Smyth, 2005), "EnhancedVolcano"(Blighe, 2018) and "VennDiagram" (Chen and Boutros, 2011) were utilized for accession, analysis and visualization.

2.2.1. Normalization of the Raw Microarray Data

To determine the differentially expressed genes between healthy and resistant data, the resistant samples from GSE635 were separated into four categories depending on which drug they are resistant to. Each subgroup was normalized separately alongside the control group. Additionally, to eliminate the genes that increase due to ALL but are unrelated to the development of drug resistance, the same normalization process was performed on drug sensitive data, and on GSE19143 for further validation.

Robust Multi-array Average (RMA) normalization was elected as the method for this process (Bolstad et al., 2003). The RMA normalization function is found within the oligo package, which was downloaded to RStudio using Bioconductor. With this function, background correction, quantile normalization, and summarization of the list of probes can be performed for the DEG analysis to be achieved. A confirmation of normalization was done by visualizing the expression levels of probes prior to and after normalization (Figure 5, Section 3.1.1).

2.2.2. Differential Gene Expression (DEG) Analysis

To determine the DEGs, the method used was based on the empirical Bayesian algorithm, and this process was done utilizing LIMMA and its built-in functions. The Benjamini-Hochberg method was used to control the false discovery rate (FDR) and after that adjustment on the p value, the adj.P value was selected as 0.05 (Benjamini and Hochberg, 1995). Then, the results of the resistant data of each group were

compared to the sensitive data of their counterparts, and those that were upregulated in both groups were eliminated. For example, the upregulated probes of the asparaginase resistant group were compared with those of the asparaginase sensitive group, and any that were in both were eliminated from the list of DEGs. A similar process was done using the pred sensitive data of GSE19143 for validating the pred group.

In the end, DEGs were selected as probes that were upregulated only in the resistant subtypes of GSE635. Volcano plots of the four subtypes were drawn using the "EnhancedVolcano" package (Blighe, 2018) for visualization of the variation of gene expression between the healthy and resistant subjects (Figure 6, Section 3.1.2). Heatmaps depicting 100 upregulated and 20 downregulated genes of each resistant type against their sensitive counterparts were also generated using the Heatmapper web-tool (Babicki et al., 2016). Additionally, a venn diagram of each DEG list from each resistant group was drawn using the "VennDiagram" package (Chen and Boutros, 2011) in order to determine and visualize genes that were found in common in all four types (Figure 8, Section 3.1.2).

2.2.3. Gene Ontology and Enrichment Analysis

After the upregulated DEGs were determined, the list of probes obtained from R was put into ShinyGO V0.77 (http://bioinformatics.sdstate.edu/go/), a web-based tool containing an exhaustive annotation and pathway database. It utilizes the Ensembl database for annotation and a compilation of various other sources, such as KEGG and STRING for pathway analysis, determining gene enrichment concerning different pathways using hierarchical clustering, and analyzing chromosomal distributions, GC content, and length of genes (Ge et al., 2020). In this analysis, the FDR corrected p-value cutoff was selected as 0.05, and after this process the probes that contained no gene information were removed from the list.

2.2.4. Protein-Protein Interaction Analysis

In this step, STRING V12.0 (https://string-db.org/) was used to investigate the interactions between the upregulated DEGs (Szklarczyk et al., 2019). The minimum required interaction score was selected as 0.4. Then, after downloading the network, CytoScape (Shannon et al., 2003) and CytoHubba (Chin et al., 2014) were used to visualize the protein-protein interaction (PPI) network and hub genes. Two sets of 10 hub genes were determined based on betweenness and degree.

2.3. In silico Screening for Drug Repurposing

The two proteins used in this step were determined previously, from amongst the DEGs found in common in all resistant subtypes. The following in silico screening and molecular docking analyses were done utilizing PyRx V.0.8, a virtual screening software (Dallakyan and Olson, 2015) and the ZINC15 drug library (Sterling and Irwin, 2015).

2.3.1. Crystal Structure and Drug Library Acquisition

The drug library for the screening was downloaded from the ZINC15 database. The selective tags were “named”, “for sale”, and “in trial”, in order to only screen drugs that would be easily available for sale and that had been or are currently being investigated. As a result, 3556 small molecules, including those that were approved by the FDA, were downloaded in the mol2 format.

The crystal structures for the proteins CLDN9 and HS3ST3A1 were accessed from the Protein Data Bank (PDB) (Berman et al., 2000). For CLDN9, the crystal structure 6OV2 (Vecchio and Stroud, 2019) and for HS3ST3A1, the crystal structure 1T8U (Moon et al., 2004) were downloaded in the .pdb format.

2.3.2. Preparation of the Crystal Structures and the Drug Library

The drug library was uploaded to PyRX as a list. The energies of the small molecules were then minimized using the program according to its default parameters. Finally, before the screening could take place, the drugs were converted to .pdbqt file format using Open Babel, which is built into PyRX to convert chemical files into formats suitable for the docking process (Dallakyan and Olson, 2015).

Before the crystal structures could be used for docking, they were preprocessed using USCF Chimera (Pettersen et al., 2004). In this step, all water molecules and any extra solvents were removed from the environment using UCSF Chimera's DockPrep tool, and then prepared for docking. Afterwards, both of the preprocessed structures were validated using two online tools: ProSA (Wiederstein and Sippl, 2007) and PROCHECK (Laskowski et al., 1996). Additionally, DogSiteScorer (Volkamer et al., 2012) was used to determine binding pockets and assess their druggability for both of the structures.

2.3.3. Molecular Docking Screen

Once the drug library and the crystal structures were prepared for docking, the location of the grid box, where the drug molecules will bind, was determined by calculating the center of mass of the selected binding pocket. For CLDN9, the grid box was calculated with the coordinates X: 4.53, Y: -15.33, Z: -18.72 and with dimensions 26 Å x 26 Å x 26 Å, for HS3ST3A1 the coordinates were X: 35.12, Y: 32.53, Z: 35.21 with dimensions 20 Å x 20 Å x 20 Å. Autodock Vina, built-in to PyRX, was utilized for the molecular docking step. After the docking was completed for each of the 3556 small molecules against both of the crystal structures, the binding energy threshold was determined as -8.0 kcal/mol for CLDN9 and -9.0 kcal/mol for HS3ST3A1, as the latter demonstrated better binding this was necessary to limit the number of molecules obtained in this step. Additionally, only molecules that satisfied both thresholds with an RMSD value of 0 were selected for the next step.

2.3.4. ADMET Analysis

Here, ADMET analysis was performed using the online tools SwissADME (Daina et al., 2017) and AdmetSAR (Yang et al., 2019). Smiles notations of the molecules that were selected with docking were accessed from the ZINC database and uploaded first to SwissADME and afterwards to AdmetSAR for further analysis. In the first step, the Lipinski (Lipinski et al., 2001), Muegge (Muegge et al., 2001), and Ghose (Ghose et al., 1999) filters were the selective criteria and the molecules that had < 2 violations for each were selected for the next step of analysis. The molecules that passed these filters were evaluated based on human intestinal absorption, human oral bioavailability, and acute oral toxicity (< 2.5). Subsequently, the three molecules that satisfied these criteria were put into molecular dynamics simulations to validate their binding with HS3ST3A1.

2.4. Molecular Dynamics Simulations

Upon determining three small molecules (Flunarizine, Eltrombopag, and Talniflumate) that satisfied the binding affinity thresholds for both proteins and had favorable ADMET scores according to both SwissADME and AdmetSAR, MD simulations were performed on all three against HS3ST3A1. The main goal of MD analysis is to validate the findings of molecular docking to enable a more thorough and sophisticated assessment. In this study, MD simulations were done utilizing the GROMACS 2021.2 software (Abraham et al., 2015). For this protein, CHARMM27 was chosen as the force field (Bjelkmar et al., 2010). SwissParam was used to build ligand topology and conditions (Zoete et al., 2011), where the conditions of each atom were added into the topology of the system and packed into complex topology files. TIP3P water molecules were used to solvate the protein-ligand complexes inside a cubic space (Price and Brooks, 2004). The particle-mesh Ewald (PME) method was selected to maintain long-range electrostatic interactions, and Cl- and Na+ ions were used instead of solvent molecules in order for the system to be an electrically neutral simulated system (Fadrná et al., 2005). Before the simulation started, the system went through 50.000 steps of energy minimization. Subsequently, the simulations were done with NPT, constant number of particles, system pressure and temperature, and NVT,

constant number of particles, system volume respectively, and temperature, for 100 nanoseconds, with a constant 300 K temperature and 1 bar pressure (Chen et al., 2023). Every 10-picosecond, the trajectory data of the system was taken, which was later analyzed for essential structural assessment values; RMSD, RMSF, Rg and H-bond formation using GROMACS and XQuartz (https://www.xquartz.org/).

2.5. Cell Culture and Subculture

2.5.1. SUP-B15 Cell Line

SUP-B15 cells are a cell line of Ph+ ALL cells. They were maintained in RPMI medium with 20% fetal bovine serum (FBS) and 1% penicillin/streptomycin and incubated at 37°C and %5 CO_2. The cells were subcultured every three days, with 400 g centrifuge speed for five minutes and diluted to ⅓.

2.5.2. Jurkat Cell Line

The Jurkat cell line consists of T-ALL cells. This line was maintained in RPMI medium with 10% FBS and 1% penicillin/streptomycin and incubated at 37°C and %5 CO_2. The cells were subcultured every two days, with 900 RPM for five minutes, and diluted to ⅓.

2.5.3. HUVEC Cell Line

HUVEC cell line is a healthy cell line consisting of human umbilical vein endothelial cells. These cells were maintained in DMEM medium, with 10% FBS and 1% penicillin/streptomycin and incubated at 37°C and %5 CO_2. They were subcultured every two days, using trypsin and incubated for five minutes, then centrifuged at 500 g for five minutes and diluted to ½.

2.6. Analysis of The Cytotoxic Effects of The Three Selected Drugs on Cancer Cell Lines

2.6.1. Cytotoxicity Analysis on the SUP-B15 Cell Line

The anti-proliferative effects of Eltrombopag, Talniflumate and Flunarizine were determined on SUP-B15 cell lines via MTT assays. 3 x 10^4 cells were seeded in each well in 96-well plates for each drug in triplicates and incubated for 72 h. Then 15 ul of MTT (5 mg / ml) was added to each well and after a 4-hour waiting period, 100 ul of DMSO was added and the plates were incubated for 30 minutes. The optical density (OD) was measured at 570 nm.

2.6.2. Cytotoxicity Analysis on the Jurkat Cell Line

The anti-proliferative effects of Eltrombopag, Talniflumate and Flunarizine were determined on Jurkat cell lines via MTT assays. 1 x 10^4 cells were seeded in each well in 96-well plates for each drug in triplicates and incubated for 48 h. Then 10 ul of MTT (5 mg / ml) was added to each well and after a 4-hour waiting period, 100 ul of DMSO was added and the plates were incubated for 30 minutes. The optical density (OD) was measured at 570 nm.

2.7. Analysis of the Selected Drugs on the HUVEC Cell Line

The IC50 values obtained from SUP-B15 cells for each drug was administered to HUVEC cells and their cytotoxic effects were determined by Trypan Blue assays. 5 x 10^4 cells were seeded in a 6-well plate and incubated for 24 h until they were attached to the wells, Consequently, the media of each well was changed, and IC50 doses for each drug was administered. 48 h proceeding this, each well was counted using Trypan Blue.

CHAPTER 3: RESULTS

3.1. DEG Analysis

3.1.1. Normalization of the Raw Data

As mentioned previously in Section 2.2.1, the “oligo” package of Bioconductor was used to access the raw data of GSE635, GSE19143 and GSE22529 datasets. Normalization of the gene expression data was done based on the RMA method, again utilizing a function of the “oligo” package. For each of the resistant groups, healthy and patient data were normalized together. In the GSE635 dataset, the asparaginase resistant group had the expression data 40 B-ALL patients, daunorubicin had 23 patients, prednisolone had 20 patients and vincristine had 29 patients.

To validate the normalization process in this step, a box plot visualization of the expression data from before and after RMA normalization was employed.

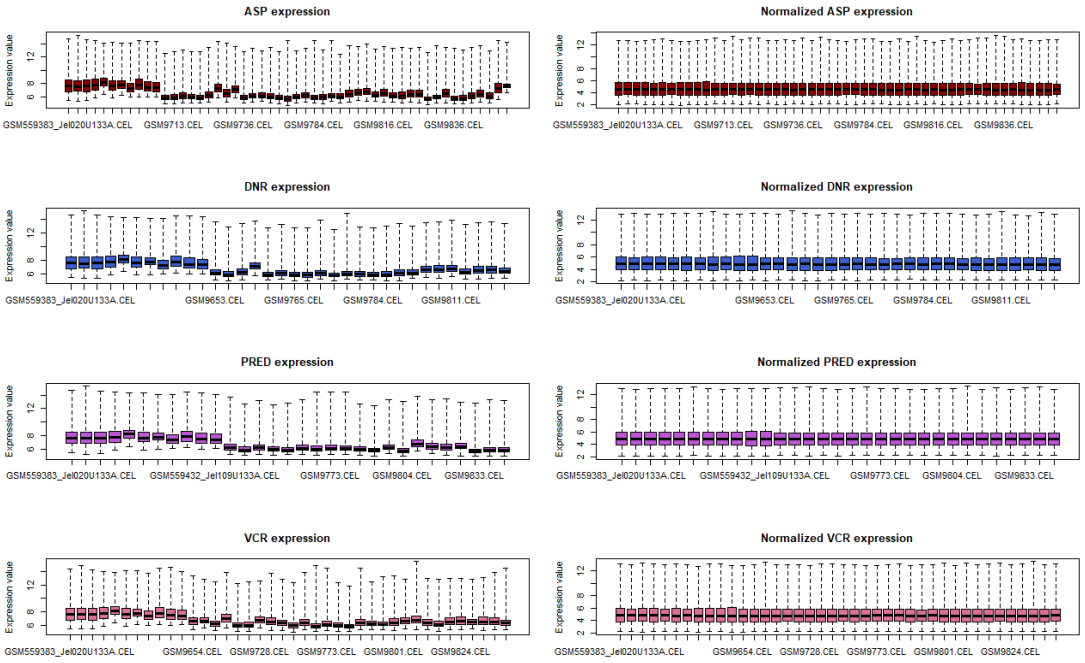

Figure 5. Gene expression levels of each resistant group from GSE635 before and after the normalization step.

As can be seen in Figure 5, RMA normalization was performed successfully across all of the four resistant subtypes. For each of the expression sets, the RMA

reduced background noise, variation and any batch effects that may have occurred from the different datasets for each of the samples, bringing their expression values to a cohesive set in each resistant group. Although not pictured here, the same process was done for the sensitive samples. In the following steps of DEG analysis, the normalized values were used for all samples.

3.1.2. Determination of the DEGs in GSE635

Once the normalized gene expression levels were obtained, LIMMA was used for the DEG analysis step (Section 2.2.2). To determine the DEGs, the list of upregulated and downregulated probes of resistant groups were compared to the sensitive groups. Those that were upregulated in both the resistant and sensitive groups of a single group (such as asparaginase resistant against asparaginase sensitive) were eliminated from the list of DEGs. Similarly, the results of the Pred resistant group were compared to the prednisolone sensitive patient data of GSE19143 for further validation. As such, only those that were uniquely upregulated for the resistant groups were selected as upregulated DEGs, and a similar process was done with downregulated DEGs. Using the "EnhancedVolcano" package, volcano plots were generated for all four resistant groups to visualize the DEG variation of resistant against healthy groups (Figure 6).

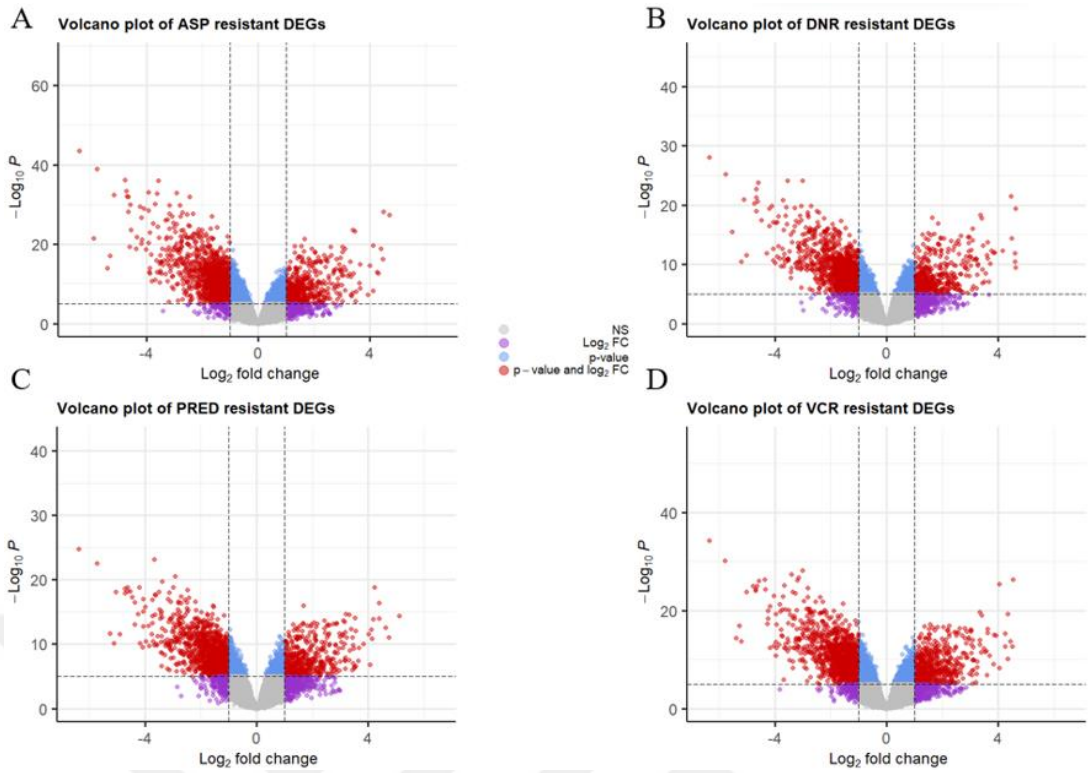

Figure 6. Volcano plots of Asp resistant (A) Pred resistant (B) Dnr resistant (C) and Vcr resistant groups. Blue indicates p-value < 0.00001, purple indicates Log2FC < -1 and > 1, and red indicates both. Gray depicts insignificant genes.

According to Figure 6, each of the resistant groups possessed a higher degree of downregulated genes, and overall they all had a high number of DEGs that satisfied the p-value of less than 0.00001, selected for the purpose of better visualization. Additionally, heatmaps of each resistant group were generated using Heatmapper (www.heatmapper.ca) (Babicki et al., 2016) to illustrate 100 upregulated and 20 downregulated DEGs compared to their sensitive counterparts (Figure 7).

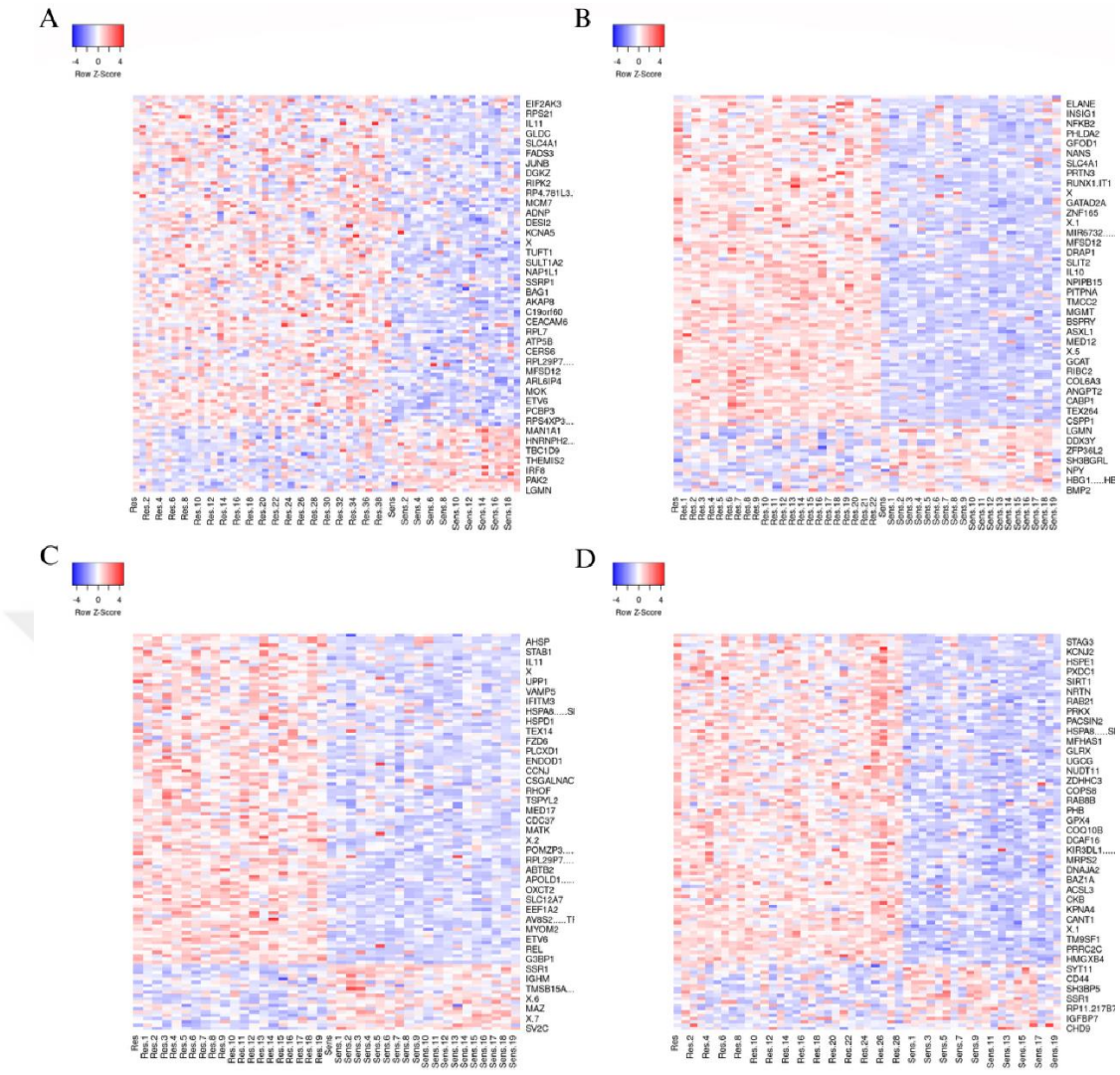

Figure 7. Heatmaps of asparaginase (A), daunorubicin (B), prednisolone (C), and vincristine (D) resistant subtypes. Red depicts upregulated genes with blue depicting downregulated genes, whereas white indicates no significant changes in expression.

Separately, 660 upregulated probes were detected in the asparaginase resistant, 565 in the daunorubicin resistant, 613 in prednisolone resistant and 589 in vincristine resistant groups. A venn diagram was drawn to visualize the commonality found in the upregulated probes of each resistant group using the "VennDiagram" package (Chen and Boutros, 2011). Although sixteen probes were determined to be upregulated in all of the four resistant groups, after annotation some of these were discovered to have no gene IDs or were mapped to the same genes, and removing these left twelve genes in common (Figure 8).

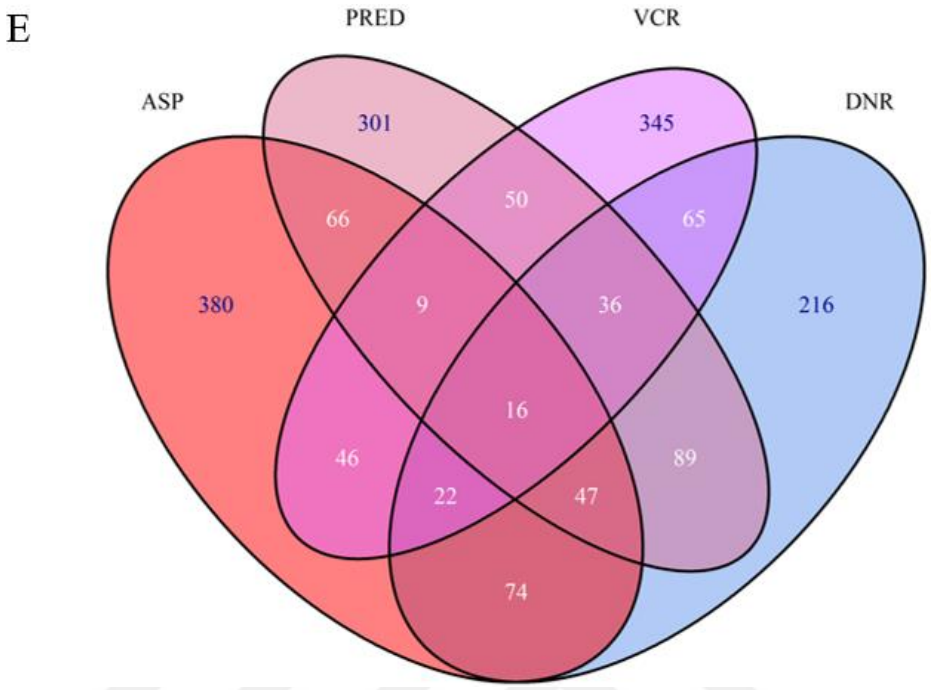

Figure 8. Venn diagram of the number of probes found to be upregulated in each resistant subtype, separately and in common.

The full list of common genes was obtained after using ShinyGo for annotation. This list can be seen in Table 2, along with each of the genes' probe and ENSEMBL IDs.

Table 2. Twelve genes found in common in all four resistant groups, with probe and ENSEMBL IDs.

Probe ID	Gene Name	ENSEMBL ID
215295_AT	DTNB	ENSG00000138101
207450_S_AT	POU6F2	ENSG00000106536
205425_AT	HIP1	ENSG00000127946
203559_S_AT	AOC1	ENSG00000002726

Table 2 (Continued). Twelve genes found in common in all four resistant groups, with probe and ENSEMBL IDs.

204529_S_AT	TOX	ENSG00000198846
213858_AT	ZNF250	ENSG00000196150
221891_X_AT	HSPA8	ENSG00000109971
215384_S_AT	MAP1A	ENSG00000166963
214635_AT	CLDN9	ENSG00000213937
219985_AT	HS3ST3A1	ENSG00000153976
210037_S_AT	NOS2	ENSG00000007171
210960_AT	ADRA1D	ENSG00000171873

3.1.3. Gene Ontology and Pathway Analysis

Also referred to as gene enrichment analysis, this process is mainly performed to ascertain whether a set of genes identified via genomic analysis had genes from distinct pathways or functional categories using statistical methods (Ge et al., 2020). As mentioned in Section 2.2.3, GO and pathway analysis was performed using the ShinyGo V0.77 webtool. The list of probes were put into ShinyGo and a list of annotated upregulated DEGs were obtained.

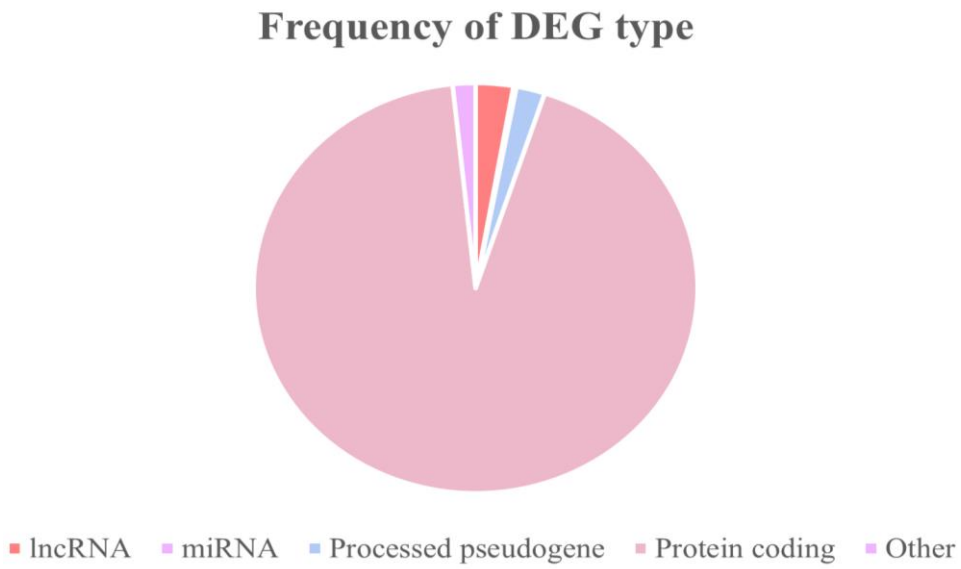

Figure 9. Pie chart of the types of upregulated DEGs, as obtained from the ShinyGO annotation.

As can be seen in Figure 9, ShinyGO mapped the upregulated probes to several types of genes. A total of 1338 probes were mapped by ShinyGO. Among these, there were 38 lncRNAs, 3 miRNAs, 29 processed pseudogenes and 1294 protein coding genes. The other section contained 23 several types of genes, including transcribed processed pseudogenes and unprocessed pseudogenes.

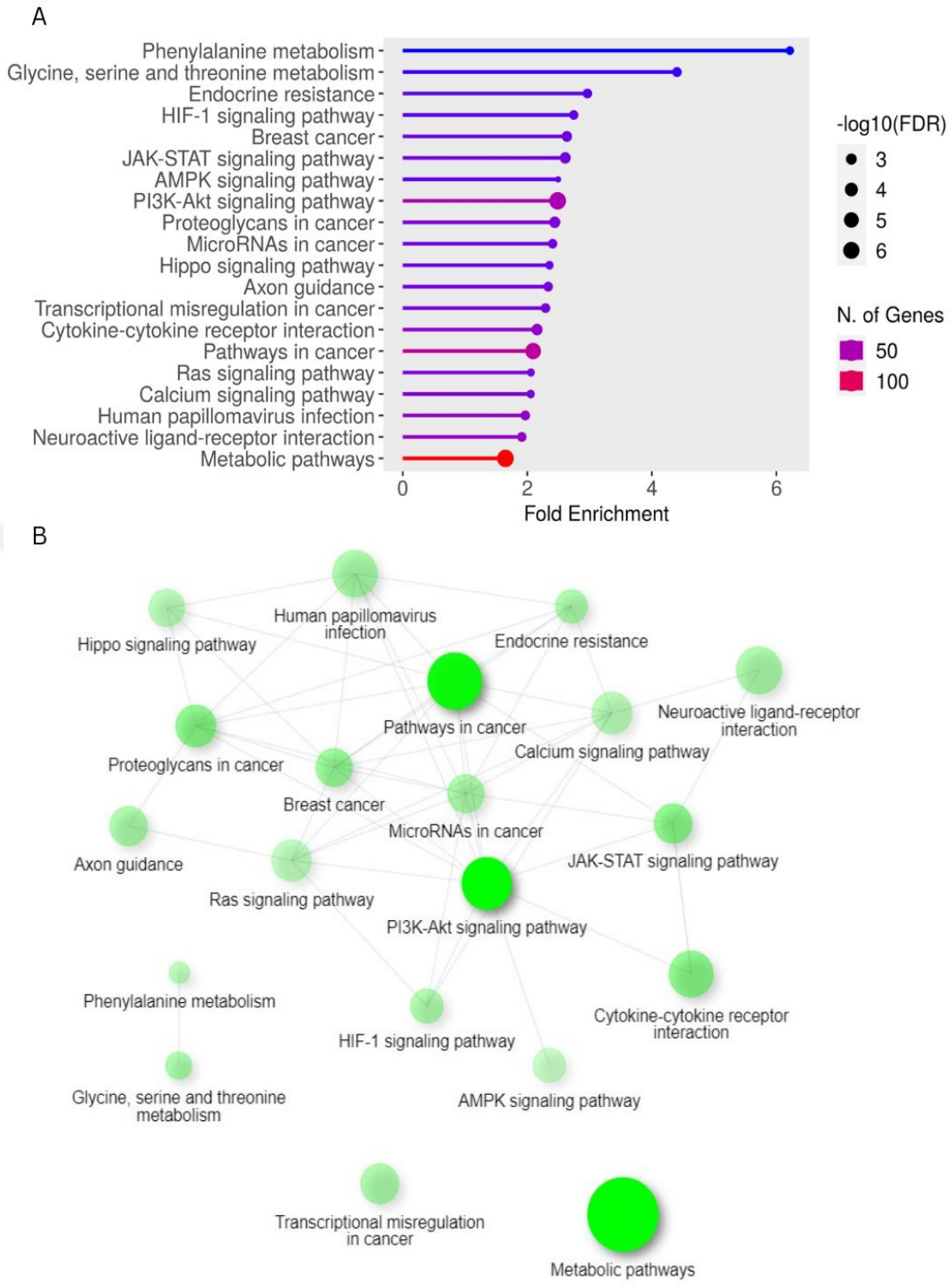

Figure 10. Results of Kegg Pathway Enrichment analysis. The lollipop chart of KEGG pathway enrichment with an FDR-corrected p-value threshold of 0.05, with colors illustrating the number of genes, and size -log10(FDR) (A). Network diagram of Kegg enrichment with FDR-corrected p-value and edge cutoff thresholds at 0.05 and 0.1, respectively (B). Each node is a pathway, significantly enriched genes are shown in

more intense colors and node size changes depending on the number of genes found in the pathway. Two nodes are linked if they share over 20% genes.

First pathway that was checked through ShinyGO was the KEGG pathways, which illustrated that the upregulated DEGs showed the most significant enrichment in the PI3K-Akt and pathways in cancer, which according to the KEGG database, includes oncogenic pathways such as WNT, VEGF, cAMP, mTOR, TGF-Beta and several more. Additionally, microRNAs in cancer, the JAK-STAT pathway, Hippo and Ras pathways were among the enriched pathways (Figure 10). In addition to having high fold enrichments, the PI3K-Akt pathway and pathways in cancer also contained a higher number of genes, as can be inferred from both Figure 10A and 10B. The genes in each of the enriched oncogenic pathways can be seen in Table 3.

Table 3. Genes that belong to the oncogenic pathways obtained from KEGG enrichment, alongside their FDR-corrected p-values and fold enrichment levels.

Pathway	FDR	Fold Enrichment	Genes
JAK-STAT	8.4E-04	2.6	CREBBP IL20RA PIAS2 IL11 IL23A GHR IL4 STAM2 LEPR CCND2 PDGFRA CSH1 IL10 STAM AOX1 IFNA16 IL6R LEP IFNW1 SOCS3 EPOR PDGFA CSF2RA GH1
AMPK	7.3E-03	2.5	PPP2R3A SIRT1 RAB2A CPT1A CAMKK2 LEPR STK11 FOXO3 ADRA1A PFKFB2 PCK1 PRKAA1 PPP2R1B LEP RPS6KB2 ADIPOQ PPP2R2A

Table 3. (Continued) Genes that belong to the oncogenic pathways obtained from KEGG enrichment, alongside their FDR-corrected p-values and fold enrichment levels.

PIK3-Akt	4.2E-07	2.5	BRCA1 LPAR2 PKN2 FGF22 PPP2R3A GNB1 ANGPT2 FLT1 CDC37 PIK3CG ITGB8 FGF6 VEGFA GHR IL4 FGF1 FN1 STK11 FOXO3 CCND2 TEK PCK1 GNG13 PRKAA1 PDGFRA MDM2 CSH1 PPP2R1B THBS1 COL2A1 ERBB2 ITGA9 CREB5 IFNA16 NTRK2 ITGB1 IL6R COL6A3 COL1A2 YWHAZ PTK2 LAMB2 RPS6KB2 EFNA5 CSF1 EPOR MAGI2 PDGFA PPP2R2A GH1
HIPPO	4.5E-03	2.4	DVL2 WWTR1 CTNNA2 FGF1 WNT5A CCND2 TGFB3 PARD6B DLG4 APC PPP2R1B CSNK1D SMAD4 FZD7 FZD5 YWHAZ FZD6 SOX2 WNT7B BMPR2 PPP2R2A
Pathways in Cancer	2.6E-06	2.1	DVL2 CREBBP NOS2 DAPK2 LPAR2 CTNNA2 FGF22 NOTCH3 RARB NFKB2 GNB1 GNAS ESR1 ABL1 CRKL CBL IL23A FGF6 VEGFA MSH3 IL4 FGF1 KNG1 WNT5A HES1 FN1 NFE2L2 CCND2 TGFB3 PAX8 GNG13 PDGFRA APC HEY2 MDM2 EDNRB ADCY3 ESR2 PML SMAD4 ERBB2 IFNA16 NOTCH1 ITGB1 EDNRA FZD7 IL6R ADCY9 FZD5 FZD6 PTK2 CXCL8 PTGER4 LAMB2 RPS6KB2 CALML5 JAG2 PTCH1 EPOR WNT7B DAPK1 PDGFA CSF2RA
Ras	5.3E-03	2.1	RASGRF1 FGF22 GNB1 ANGPT2 ABL1 PAK5 FLT1 GAB1 FGF6 VEGFA FGF1 PLA2G4A TEK PLA2G5 GNG13 AFDN PDGFRA PAK6 NTRK2 SHC1 PLA2G1B CALML5 EFNA5 CSF1 PLA2G6 NF1 PDGFA

Next, the upregulated DEGs were analyzed for GO pathways. All three were rich in number of genes, although had lower fold enrichments compared to the KEGG pathways. For BP (Figure 11A), CC (Figure 11B) and MF (Figure 11C), the pathways

that showed the most enrichment with statistically significant p-values were anatomical structure morphogenesis, cell adhesion, and cell junction, respectively.

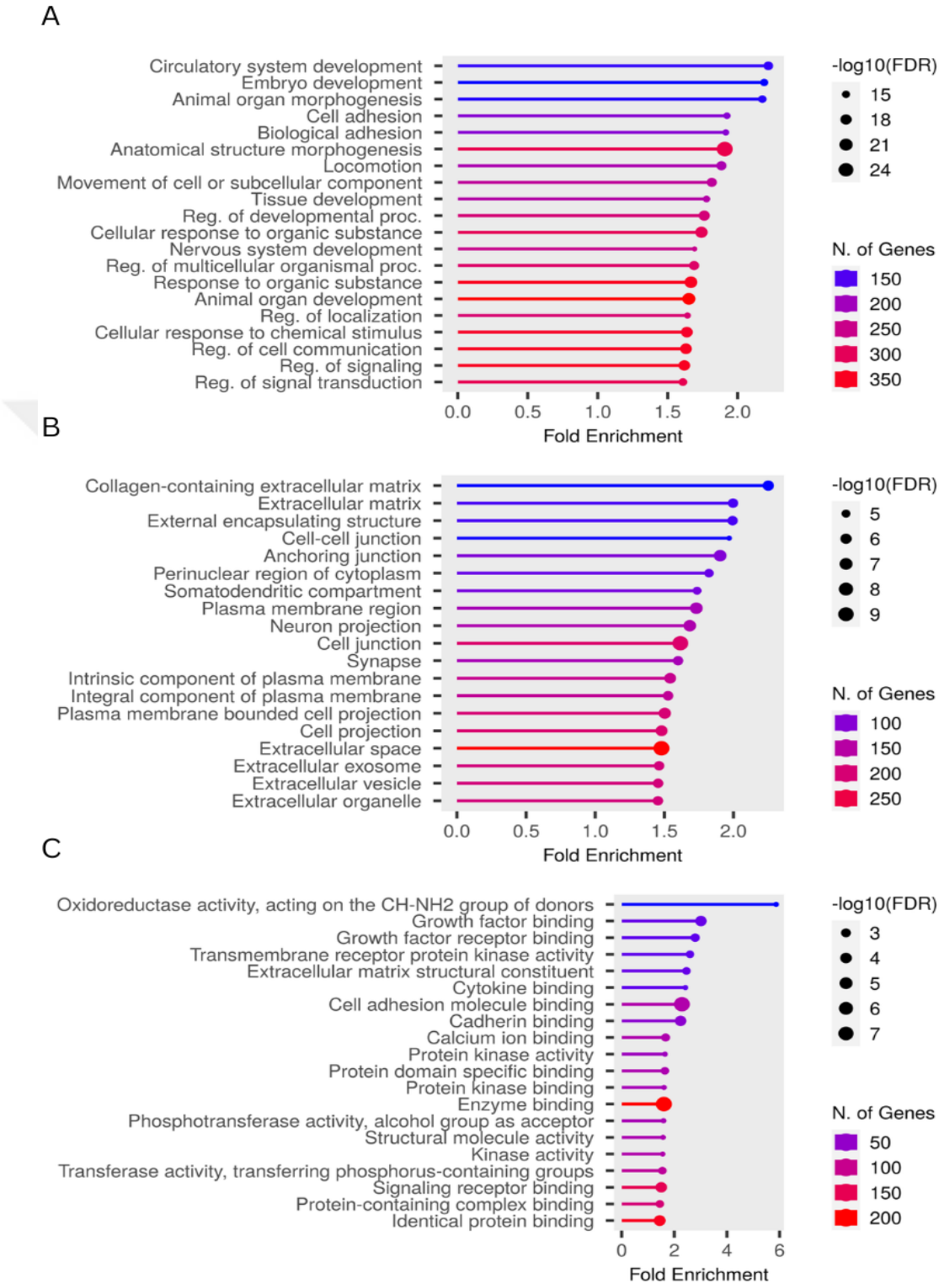

Figure 11. The lollipop charts of GO enrichment of BP (A) CC (B) MF (C) pathways, with an FDR-corrected p-value threshold of 0.05, colors illustrating the number of genes, and size -log10(FDR).

3.1.4. Protein-Protein Interaction Networks

Explained in further detail in Section 2.2.4, the annotated list of upregulated protein coding genes, obtained from ShinyGo was first put into String V12 (Szklarczyk et al., 2019), a total of 1294 genes, and the network file obtained from the website was uploaded to CytoScape (Shannon et al., 2003) to visualize the PPI network and determine hub genes using CytoHubba (Chin et al., 2014). The entire PPI network had 1207 nodes (proteins) and 7778 edges (interactions).

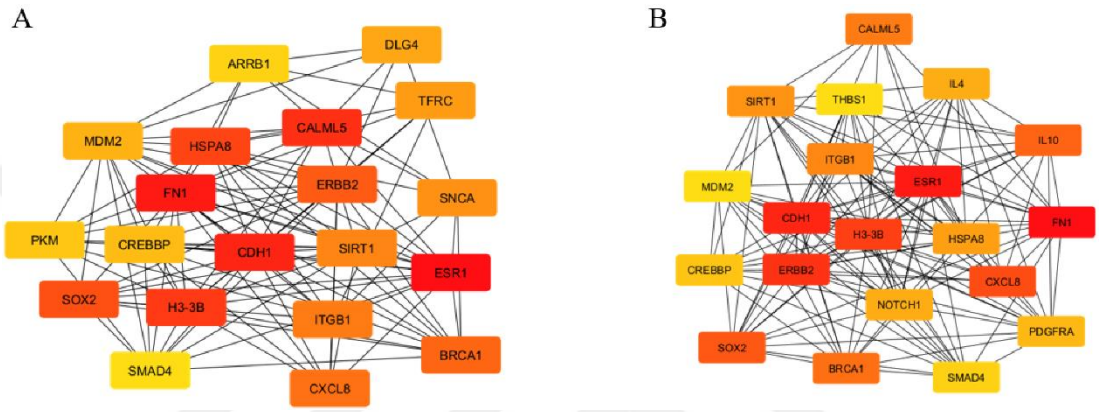

Figure 12. Two networks of twenty hub genes calculated through betweenness (A) and degree (B) ranking methods of Cytohubba.

A plug-in of Cytoscape, Cytohubba can be used to determine the hub genes of a biological network based on eleven node ranking methods, including degree and betweenness, as well as several others (Chin et al., 2014). In this thesis, two sets of hub genes were generated depending on betweenness (Figure 12A) and degree (Figure 12B). Each set contained 20 genes, with 15 found in common in both sets, resulting in a total of 25 hub genes identified. In both of the sets, CDH1, FN1, H3-3B and ESR1 had consistently intense color, related to their high scores in both calculations.

3.2. In Silico Screening

3.2.1. Validation of the Crystal Structures

As described in Section 2.3.2, Procheck and Prosa were used to validate the crystal structures of CLDN9 and HS3ST3A1 that were preprocessed using UCSF Chimera (Pettersen et al., 2004). Using PROCHECK (Laskowski et al., 1996), Ramachandran plots of the 3D structures were generated, and ProsA (Wiederstein and Sippl, 2007) was utilized to calculate the z-scores of both structures, which is processed by comparing the target model with publicly available similar PDB models (Wiederstein and Sippl, 2007).

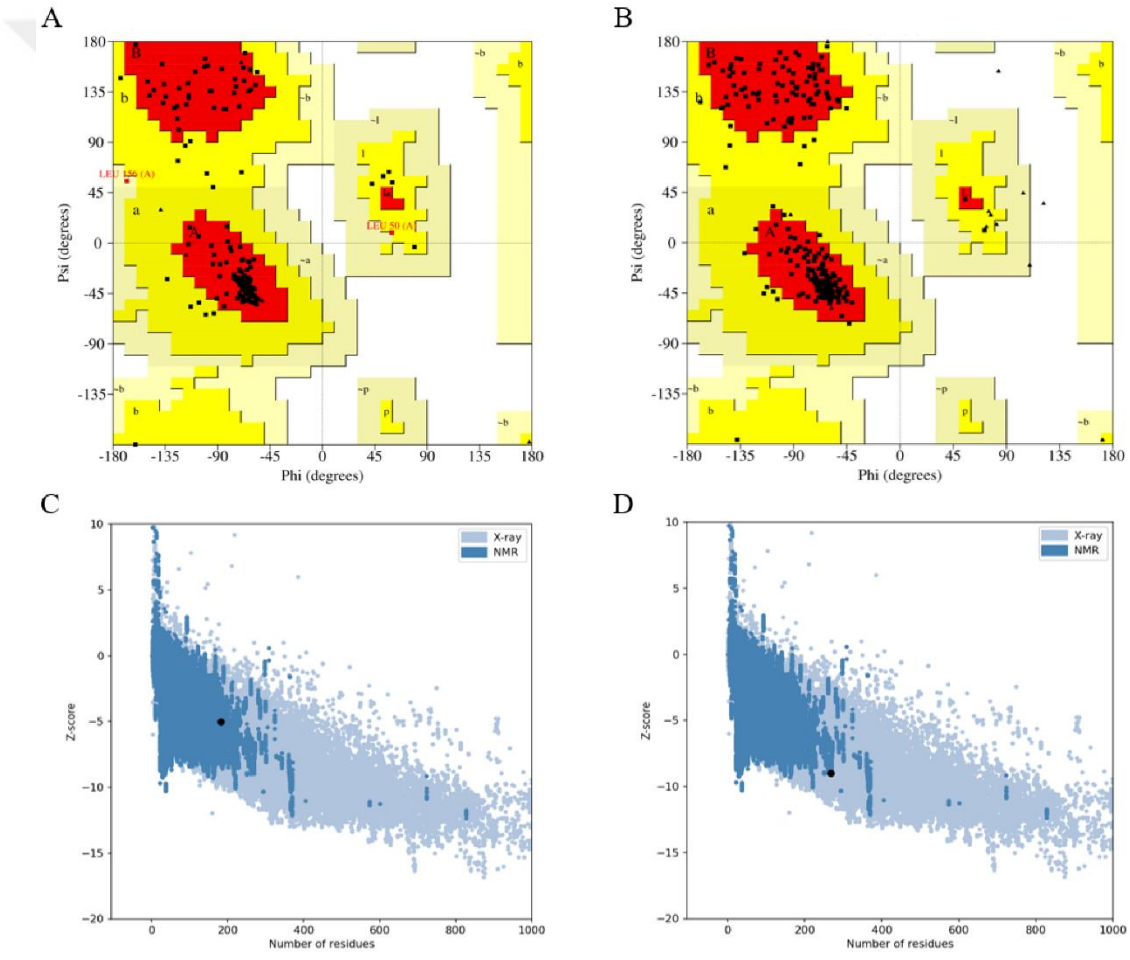

Figure 13. Validation of the preprocessed crystal structure of CLDN9 and HS3ST3A1. Ramachandran validation of CLDN9 (A) via PROCHECK and its z-score(B) via PROSA along with the Ramachandran validation (C) and z-score of HS3ST3A1 (D).

First discovered in the 1960s, the Ramachandran plot is one of the roots of macromolecular structural analysis and crystallography (Hollingsworth and Karplus, 2010). The plot works by calculating the psi and phi angles of aminoacid residues and especially in the last thirty years, it's been one of the go-to methods of validation for 3D protein structures generated through crystallography, NMR spectroscopy, and recently popular computational modeling (Carugo and Djinovic-Carugo, 2013). The ProSA web server is another popular tool for the validation of crystal structures. The z-scores of a given structure are calculated based on comparison with crystal structures of other similar native proteins (Wiederstein and Sippl, 2007). In this thesis, for CLDN9 86. 9% and 11.9% of residues were found in the most favored regions and additionally allowed regions, respectively, with 1.2% of the residues in the generously allowed regions (Figure 13A). The crystal structure also had a z-score of -5.03 (Figure 13B). Whereas HS3ST3A1 had 91.3% of residues in the most favored regions and the remaining 8.7% in the additional allowed regions (Figure 13C), the structure also had a z-score of -8.99 (Figure 13D). For both of the crystal structures, the z-score calculated by ProSA are within the range of scores found for similarly sized native proteins.

3.2.2. Structure based ligand screening and molecular docking

Structure based drug repurposing is based on developing affinity estimations of docking simulations of a 3D structure and its potential ligands (Vyas et al., 2008). As can be read in further detail in Section 2.3.3, PyRx, and its docking program AutoDock Vina, was the preferred program to perform this step in this thesis. The proteins CLDN9 and HS3ST3A1 were obtained from the results seen in Section 3.1.2, both among the proteins found in common in all four resistant groups.

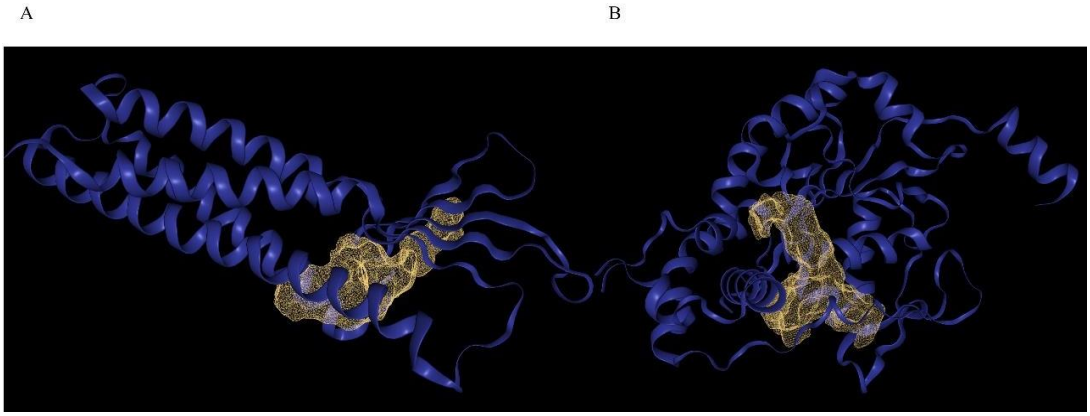

Figure 14. Druggable binding pockets determined by DogSiteScorer (Volkamer et al., 2012). CLDN9 (A) and HS3ST3A1 (B).

After preprocessing, the 3D structures were uploaded to DogSiteScorer to determine binding pockets (Volkamer et al., 2012). Figure 14A shows the selected binding pocket for CLDN9, with a drug score of 0.81, found in its extracellular domain. Figure 14B shows the selected binding pocket for HS3ST3A1, with a drug score of 0.81. Once the binding pockets were identified, both proteins were docked against a drug library of 3556 molecules obtained from ZINC15 (Figure 15), and the results of this process were analyzed based on the binding affinity and RMSD of the binding poses of each drug and protein.

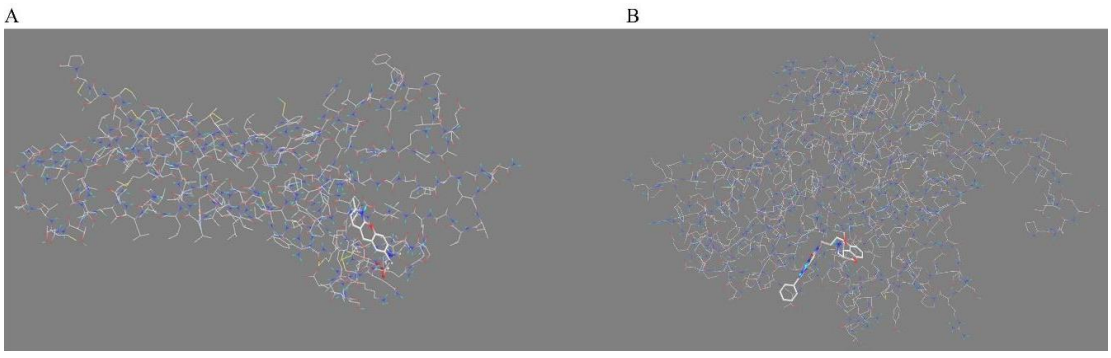

Figure 15. Docking in PyRx. Docking simulation of CLDN9 (A) and HS3ST3A1 (B) in AutoDock Vina.

Previous research has proved that negative binding affinities are associated with better interactions between proteins and ligands (Dallakyan and Olson, 2015). As such, with the results of both CLDN9 and HS3ST3A1 taken into consideration, the binding thresholds were selected as -8.0 kcal/mol and -9.0 kcal/mol, respectively, to limit the number of molecules for the following MD simulations as setting the binding threshold of HS3ST3A1 resulted in over three hundred extra molecules. Additionally, the molecules with RMSD values of 0 that also satisfied the binding thresholds of both proteins in this step were selected for the ADMET analysis. The list of small molecules with the best twenty binding affinities can be seen below in Table 4.

Table 4. Top 20 list of small molecules obtained by docking and their binding energies, ranked by the binding energy of CLDN9.

Zinc ID	**Compound name**	**Binding energy (kcal/mol)**	
		CLDN9	**HS3ST3A1**
ZINC3978005	Dihydroergotamine	-9.6	-9.2
ZINC18710085	Chir-265	-9.6	-9.1
ZINC43208634	Omipalisib	-9.3	-10.5
ZINC1612996	Irinotecan	-9.3	-9.2
ZINC11679756	**Eltrombopag**	**-9.1**	**-9.0**
ZINC59676426	Implitapide	-9.0	-9.4
ZINC1844627	**Talniflumate**	**-8.7**	**-10.2**
ZINC13831810	Mizolastine	-8.6	-10.2

Table 4. (Continued) Top 20 list of small molecules obtained by docking and their binding energies, ranked by the binding energy of CLDN9.

ZINC95493347	Efonidipine	-8.6	-9.6
ZINC537805	Diabeta	-8.6	-9.5
ZINC3775812	Pagoclone	-8.5	-9.9
ZINC1542146	Pranlukast	-8.5	-9.5
ZINC12503187	Conivaptan	-8.5	-9.1
ZINC100016063	Tivantinib	-8.4	-11.1
ZINC1494900	Enzastaurin	-8.4	-10.3
ZINC1482077	Gliquidone	-8.4	-9.1
ZINC896717	Accolate	-8.3	-10.2
ZINC52509366	Zelboraf	-8.3	-9.7
ZINC68267814	Rimegepant	-8.2	-10.3
ZINC19360739	**Flunarizine**	**-8.2**	**-10.2**

3.2.3. ADMET analysis

In this step, as explained in detail in Section 2.3.4, the selected molecules were analyzed to further investigate their efficacy, drug-likeness and safety. SwissADME and AdmetSAR, two popular online analysis tools, were used for this analysis. SwissADME was used first, and the selective parameters were chosen as

bioavailability ≥ 0.55 (Martin, 2005), Ghose, Lipinski and Muegge filters, which are related to pharmacokinetic principles a molecule needs to have to be counted as a potential drug candidate, such as rotatable bonds, molecular weight and lipophilicity (Benet et al., 2016).

Once the list of twelve molecules with acceptable SwissADME scores were obtained, they were put into AdmetSar to analyze for toxicity and drug efficacy. In this step the selective parameters were oral toxicity scores < 2.5 (Zhu et al., 2009), human intestinal absorption and human oral bioavailability, important factors for the efficacy of oral drugs. This analysis resulted in four molecules with suitable AdmetSAR scores, and three were selected: Flunarizine, Talniflumate and Eltrombopag. The list of ADMET scores of the twelve molecules can be seen in Table 5.

Table 5. ADMET properties of candidate molecules obtained from both SwissADME and AdmetSAR, with abbreviations Bio. (Bioavailability), H. Oral Bio. (Human oral bioavailability), H. Int. Abs. (Human intestinal absorption) and A. Oral Tox. (Acute oral toxicity).

Zinc ID	**Compound Name**	**Lipinski**	**Muegge**	**Ghose**	**Bio.**	**H. Oral Bio.**	**H. Int. Abs.**	**A. Oral Tox.**
ZINC537877	Ketanserin	0	0	0	55	+	-	2.41
ZINC596951	Sr-2640	0	1	0	85	+	+	2.51
ZINC775812	Pagoclone	0	0	0	55	+	-	1.82
ZINC1481805	Balaglitazone	0	0	0	55	+	+	2.32
ZINC537928	Loperamide	1	1	1	55	+	-	3.22
ZINC11681534	Nebivolol	0	0	0	55	+	-	3.73

Table 5. (Continued) ADMET properties of candidate molecules obtained from both SwissADME and AdmetSAR, with abbreviations Bio. (Bioavailability), H. Oral Bio. (Human oral bioavailability), H. Int. Abs. (Human intestinal absorption) and A. Oral Tox. (Acute oral toxicity).

ZINC13831810	Mizolastine	0	0	0	55	+	+	2.59
ZINC19360739	**Flunarizine**	**1**	**1**	**0**	**55**	**+**	**+**	**2.05**
ZINC538194	Pirenperone	0	0	0	55	+	+	2.82
ZINC1844627	**Talniflumate**	**0**	**1**	**1**	**55**	**+**	**+**	**2.12**
ZINC11679756	**Eltrombopag**	**0**	**0**	**1**	**56**	**+**	**+**	**2.15**
ZINC100016063	Tivantinib	0	0	0	55	+	+	3.33

3.3. Molecular Dynamics

In broad terms, MD simulations are computational techniques that are used to analyze the dynamic behaviors of atoms and molecules. As such, MD simulations were performed to elucidate the stability, flexibility and bond strength of the three small molecules, Flunarizine, Eltrombopag and Talniflumate, obtained in the previous sections, against the protein HS3ST3A1 using GROMACS 2021.2. As was explained in further detail in Section 2.4, during the simulations the temperature was 300 K and pressure was 1 bar throughout all 100 ns, and the trajectory data was captured every 10 ps.

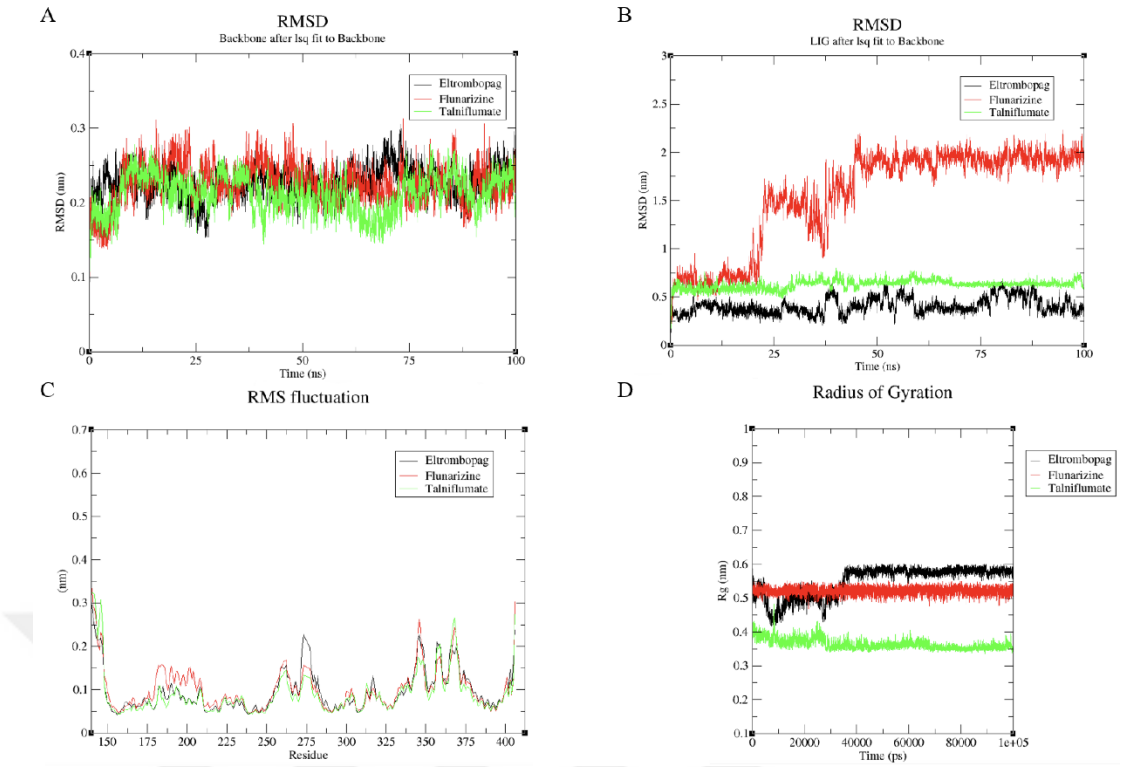

Figure 16. Results of thorough MD analysis of HS3ST3A1 and its potential inhibitors. RMSD values according to the backbone of HS3ST3A1 upon ligand binding against simulation time (A). RMSD values of ligands after binding to protein against simulation time (B). RMSF values upon ligand binding against residues (C). Radius of Gyration of potential inhibitors after binding (D).

RMSD, RSMF, Rg values and H-bond formation were all analyzed as a result of this simulation (Figure 16). The RMSD value denotes the net of the movements each atom makes at each time interval (Taghizadeh et al., 2022) and a lower RMSD is linked to better stability of the complex (Esmaili and Shahlaei, 2015). In light of this, the RMSD of the HS3ST3A1 backbone atoms upon three separate ligand bindings, and the RMSD of the ligands themselves were inspected to analyze conformational changes. The backbone RMSD upon binding is similar in all three drug molecules (Figure 16A). RMSD of Eltrombopag and Talniflumate are in similar ranges, although Eltrombopag has more fluctuations, while Flunarizine dramatically increases between 25 and 50 ns before reaching a stabler pattern (Figure 16B). The RMSF values of the C-alpha atoms of the protein were calculated to gain insight into its flexibility throughout the simulation (Saxena et al., 2009). Upon the binding of all three

molecules, the RMSF was in close range and showed peaks in the same residues (Figure 16C). The Rg values are calculated to gain an understanding of the compactness and stability of the protein complex (Sneha and George Priya Doss, 2016). Although molecules lead to Rg values in similar ranges, Talniflumate exhibited the smallest Rg, with it and Flunarizine mostly stable, whereas Eltrombopag showed fluctuation until 40 ns, before steadying (Figure 16D).

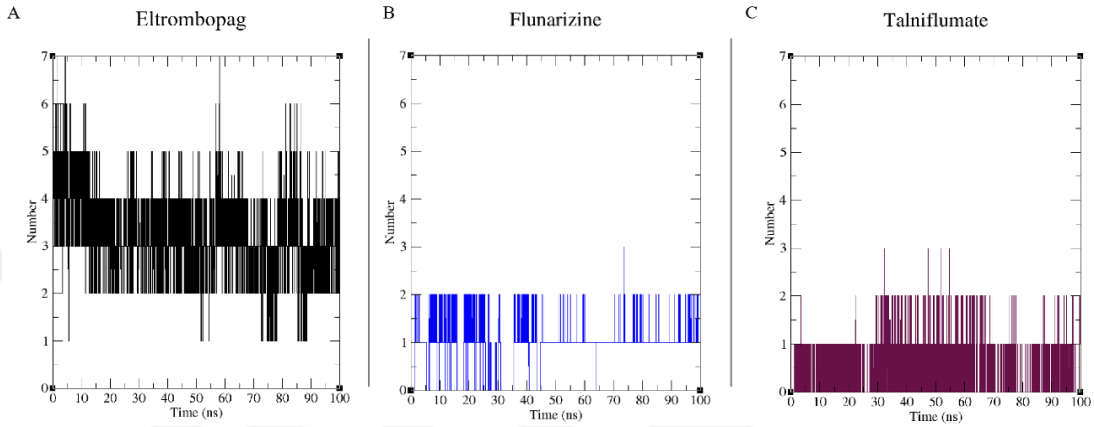

Figure 17. Number of hydrogen bonds that form between HS3ST3A1 and the three ligands: Eltrombopag (A) Flunarizine (B) Talniflumate (C).

Furthermore, investigations into H-bonding of each molecule with HS3ST3A1 was performed to gain insight into the stability of the complex (Sneha and George Priya Doss, 2016). Eltrombopag had a substantially higher number of H-Bonds, with a mean 3.254 (Figure 17A), whereas Flunarizine (17B) and Talniflumate (17C) had 0.620 and 0.415, respectively.

3.4. Cytotoxicity Assays

3.4.1. Cytotoxicity Analysis on SUP-B15 Cell Lines

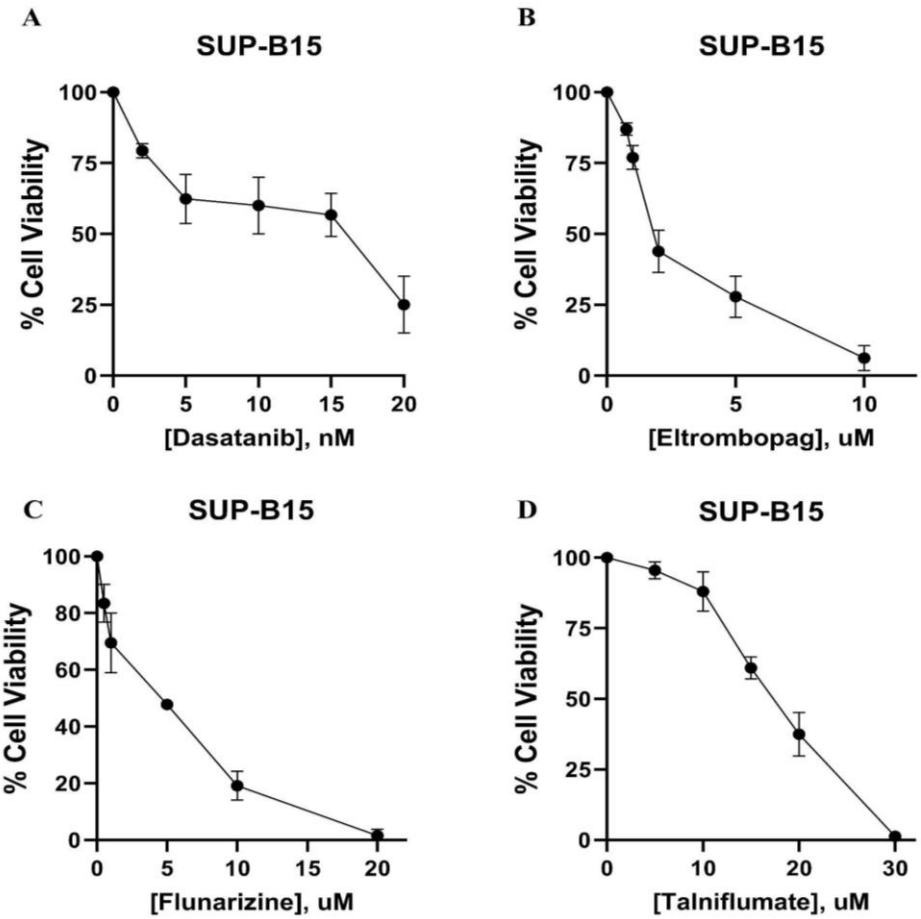

Figure 18. Results of cytotoxicity analysis of SUP-B15 after administering (A) Datasanib, (B) Eltrombopag, (C) Flunarizine and (D) Talniflumate.

The three drugs that were identified from the previous molecular docking and dynamics analyses were tested on Ph+ ALL cells, SUP-B15 to determine their cytotoxic activities. To this end, each drug was administered to the cell line separately and incubated for 72 h, then analyzed using MTT assays. For additional comparison, Dasatinib was used as a control drug as it is a known targeted therapy agent used for Ph+ ALL (Figure 18A). The results demonstrated that, when compared to the other two candidates, Eltrombopag (Figure 18B) was more effective, and its IC50 dose was

determined as 1.78 uM, whereas Talniflumate had an IC50 of 15.36 uM (Figure 18C) and Flunarizine of 4.5 uM (Figure 18D).

3.4.2. Cytotoxicity Analysis on Jurkat Cell Lines

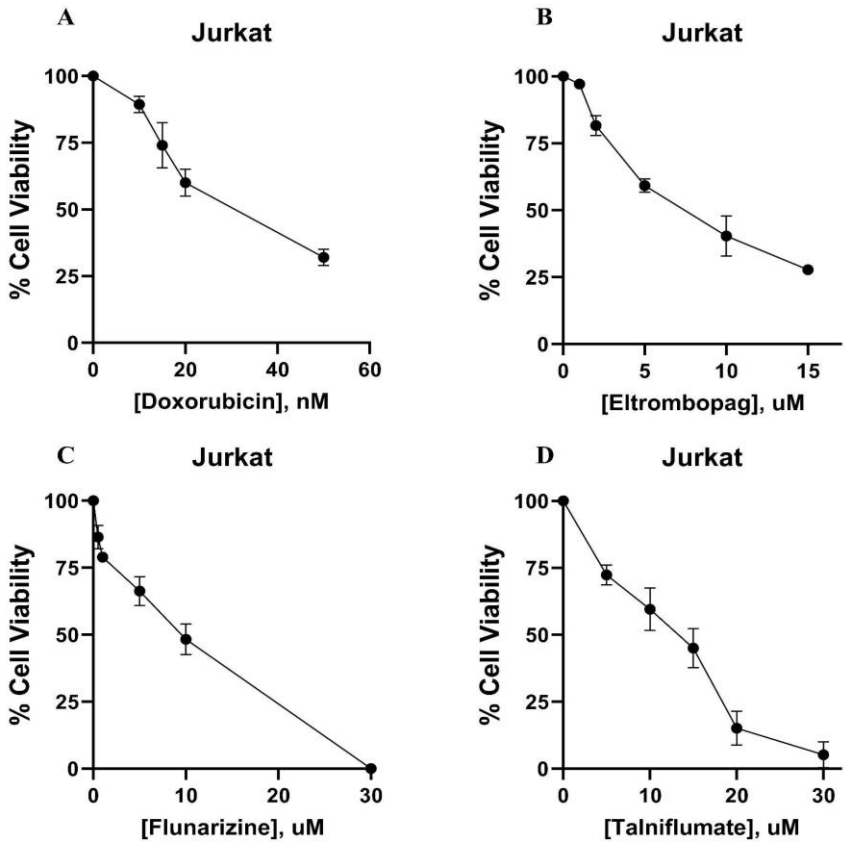

Figure 19. Results of cytotoxicity analysis of Jurkat cells after administering (A) Doxorubicin, (B) Eltrombopag, (C) Flunarizine and (D) Talniflumate.

The drugs were further tested on T-ALL cell line Jurkat using to test their cytotoxic effects on Ph- cells. Each drug was administered separately and incubated for 48h. The analysis was done using MTT assays. Doxorubicin was used as a control, as it is a known chemotherapy drug used in ALL treatment (Figure 19A). Eltrombopag was again more effective compared to the other two drugs, with an IC50 of 7.5 uM (Figure 19B), whereas Flunarizine had IC50 of 8.65 uM (Figure 19C) and Talniflumate of 16.6 uM (Figure 19D).

3.5. Antiproliferative effects on HUVEC Cell lines

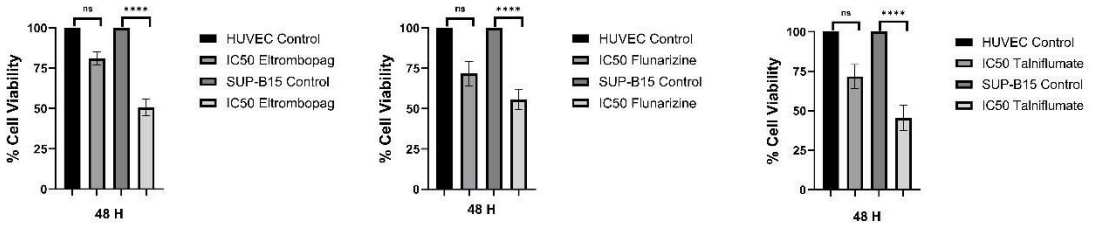

Figure 20. Results of Trypan Blue analysis on HUVEC cells of (A) Eltrombopag and (B) Talniflumate compared to SUP-B15.

Lastly, all of the candidate drugs were administered to the healthy cell line HUVEC, to determine cytotoxic activity on healthy cells. The cells were administered the drugs separately; they were given 1.78 uM Eltrombopag (Figure 20A), 4.5 uM Flunarizine (Figure 20B) and 16.36 uM Talniflumate (Figure 20C). The results showed that all of the drugs, showed a significantly lesser degree of anti-proliferative effect on healthy cells compared to SUP-B15 cells with Eltrombopag once again shining through as a comparatively better candidate.

CHAPTER 4: DISCUSSION

Although the survival rates of childhood ALL in high-income countries is around 90%, the remaining patients are deemed incurable due to being unresponsive to treatment, and 20% of those that do respond end up relapsing, typically two years after treatment (Jędraszek et al., 2022). As targeted therapy is not yet viable for 90% of the pediatric ALL population (Inaba and Pui, 2021) and both immunotherapy and targeted therapy are too high-priced to be feasibly used for all patients (Sleire et al., 2017), overcoming drug resistance towards first-line chemotherapy drugs is still a must. This study focuses on four essential drugs used in chemotherapy treatment of childhood ALL, and aims to identify novel genes that are related to the development of resistance against these drugs by determining possible inhibitors through in silico methods and drug repurposing. To this end, drug repurposing can be used as a critical tool, a much more cost-effective and time-saving method compared to the classical drug discovery process (Jourdan et al., 2020). As such, we obtained gene expression data of Asp, Pred, Dnr and Vcr resistant patients from the GEO dataset GSE635, and upon thorough DEG and pathway analysis, identified 1294 upregulated protein-coding DEGs, twenty five hub genes, and twelve DEGs found in common across the four resistant subgroups. The pathway analysis revealed that, among KEGG pathways, the DEGs showed significant enrichment in the PI3K-Akt pathway, which previous studies showed to be related to drug resistance in various cancers (Rascio et al., 2021). Other oncogenic pathways such as Jak-Stat, Hippo and RAS, the latter found in relation to relapse in B-ALL (Irving et al., 2014), were among those enriched. Additional GO analysis also pointed to anatomical structure morphogenesis, cell adhesion, and cell junction as the most significantly enriched pathways of BP, CC, MF, respectively.

A protooncogene that blocks TP53, MDM2 was found among the hub genes as a result of PPI network analysis. Recent studies suggest that it's involved in drug resistance in several types of cancer (Hou et al., 2019), including multiple myeloma and CLL (Faruq et al., 2022). CREBBP has already been associated with relapse in ALL (Inaba and Mullighan, 2020). IL-10 and IL-4 are both cytokines, a part of the blood marrow microenvironment which plays a role in differentiation, proliferation and drug resistance (Jones et al., 2016). In both sets of hub genes that were determined

based on the betweenness and degree parameters, fifteen were found in common. Out of these fifteen, CDH1, FN1, H3-3B and ESR1 were found consistently in high rank. CDH1, cadherin-1 or E-cadherin, is a member of the cadherin family, which are transmembrane glycoproteins that work in calcium dependent adhesion (Shenoy, 2019). Recent studies correlated CDH1 overexpression with bladder cancer, and it was suggested as an oncogene in breast cancer (Fan et al., 2022; Xie et al., 2022), although no study of CDH1 has been done for ALL. FN1, or fibronectin 1, is associated with proliferation and migration, as well as both healthy and malignant extracellular matrix changes through integrins (Cai et al., 2018). A previous study on cerebrospinal fluid analysis of ALL patients has pointed to FN1 as a possible prognostic marker (Guo et al., 2019), though no further research has been done as of yet. HSPA8 was found both among the hub genes and common genes. A member of the heat shock protein 70 family, increased HSPA8 was found to be related with adverse outcomes in AML (Li and Ge, 2021), and past studies point to its relation to increased chemoresistance and proliferation of cancer cells in CML (José-Enériz et al., 2008; Liu et al., 2021).

CLDN9 was also found among the common genes, and it is a member of the claudin family of proteins, with specific expression patterns in different cells. A transmembrane protein found in tight junctions, CLDN9 itself is a poorly researched protein, and its expression in normal tissue is much lower while it is overexpressed in various cancers (Liu et al., 2019; Sharma et al., 2016), with a recent study indicating it as a biomarker of endometrial cancer (Endo et al., 2022). Other members of its family have been found upregulated in various cancers and related to metastasis and invasion, with some evidence pointing to their association with drug resistance (Kwon, 2013; Singh et al., 2010). Another one of the common genes, HS3ST3A1 is a member of the heparan sulfate modifying enzyme family, and it is also a protein that has little research. Again, studies on its family show that its enzyme activity may be associated with elevated drug resistance and proliferation of cancer cells both in vivo and in vitro in MM (Baert et al., 2023). Due to both genes having the potential to be compelling, novel targets of chemoresistance in ALL, they were selected as the drug repositioning targets of this thesis.

Drug repurposing is an area that has been gaining more popularity over the years, especially with the outbreak of SARS-CoV-2 and the urgent treatment needs

that arose from it, and several drugs were suggested as treatment (Chakraborty et al., 2022). Since the estimated number of cancer patients increases every year (Soerjomataram and Bray, 2021), the need for effective cancer treatment rises accordingly. Although the drug discovery process is lengthy and high-cost, in recent years the number of approved drugs has been getting lower, with the approval rate of cancer drugs to phase I trials going down to 5% (Sleire et al., 2017). One of the biggest drug repurposing successes is lenalidomide, which was derived from a sedative that was taken out of the market due to its side effects in pregnant women, and was later found successful in MM treatment (Pushpakom et al., 2019). As such, we performed *in silico* screening of a small compound library containing drugs that are both FDA-approved and in trials in order to determine potential inhibitors for CLDN9 and HS3ST3A1. To this end, we used PyRx and its docking tool AutoDock Vina, and identified a number of molecules which were then tested for their ADMET properties. Consequently, three molecules passed these analyses: Eltrombopag, Flunarazine and Talniflumate. These drugs then went through further analysis using MD simulations, as it can be used as a superior binding validation tool.

The initial MD analysis was performed against HS3ST3A1, as it had better binding affinities with the drug molecules. According to the MD results, RMSD backbone values are similar across the three complexes (Figure 16A Section 3.3), pointing to the comparable stability of the protein upon ligand binding in all three cases. In comparison, Figure 16B, Section 3.3 shows that the RMSD ligand value of Flunarizine has a sharp increase between 25 to 50 ns, before it stabilizes, whereas the other two ligands, Eltrombopag and Talniflumate, were similar, with lower RMSD values. Considering that lower RMSD values confer better stability of the complex, (Taghizadeh et al., 2022), Flunarizine's RMSD of 2 nM signifies much less stability of the molecule when compared to Eltrombopag and Talniflumate. Additionally, the complex was analyzed based on its RMSF values upon three separate bindings. Increased RMSF values are related to higher flexibility, while low RMSF suggests a more rigid movement of the residues during the simulation. Similarly to Figure 16A, RMSF results seen in Figure 16C Section 3.3, indicate that all three molecules have similar fluctuations and have peaks in generally the same residues to approximately the same degree, such as in residues 250 to 270 or 340 to 375. The increased flexibility in these regions might be due to the fact that these are residues of HS3ST3A1 that are

known to form B-strands, which are reported to be more flexible compared to other secondary structures (Kopeć et al., 2019). Rg analysis of the three complexes illustrated that while all three molecules exhibited Rg values within close range of each other, Eltrombopag and Flunarizine showed more similarity, although the former displayed more fluctuations (Figure 16D, Section 3.3). Increased Rg indicates that the molecules of the outermost areas are further away from the center of mass, implying a structure that is more loosely folded, thereby suggesting reduced compactness (Taghizadeh et al., 2022). As such, it can be said that the protein complex doesn't exhibit differing structural changes upon ligand binding across three different molecules. Moreover, hydrogen bonding between each molecule and HS3ST3A1 was investigated to determine the strength and stability of binding, as hydrogen bond formations have been associated with better binding affinities and drug efficacy (Sneha and George Priya Doss, 2016). However, although Figure 17 Section 3.3 shows that Eltrombopag had a considerably higher number of hydrogen bonds with an average of 3.254 compared to Flunarizine and Talniflumate with averages 0.620 and 0.415, respectively, its binding affinity with HS3ST3A1 is comparatively less than the other two at -9.0 kcal/mol to the -10.2 kcal/mol of the other two molecules. As such, it must be considered that Flunarazine and Talniflumate may have increased hydrophobic and elevated Van der Waals interactions that might have led to their elevated affinities. In consideration with all the MD results, while all molecules exhibit similar results, Eltrombopag might be more a promising inhibitor of HS3ST3A1 due to having a high number of hydrogen bonds and the lowest RMSD ligand values.

In light of the MD results, further analysis into the three drugs was made to investigate their use in literature. Eltrombopag is a thrombopoietin receptor agonist, and works through the activation of the JAK-STAT and MAPK pathways to promote platelet differentiation (Bussel and Pinheiro, 2011). Used for chronic idiopathic thrombocytopenic purpura, it's in being clinically investigated for thrombocytopenia related to chemotherapy and chronic liver disease, and hepatitis C treatment, Eltrompobag was recently investigated for its effects on AML and T-cell leukemia, resulting in reduced proliferation (Erickson-Miller et al., 2010). An anti-inflammatory drug, Talniflumate binds to GCNT3, a mucin type transferase, to inhibit the production of mucin and its glycosylation (Walker et al., 2006). Although it was originally intended to be used against cystic fibrosis, asthma and chronic pulmonary issues,

recent studies suggest its possible use in Parkinson's and pancreatic cancer (Liu et al., 2023; Rao et al., 2016). One study in particular shines a light on Talniflumate's effect on mucin-specific O-Glycosylation, which was found to be associated with increased proliferation, and demonstrates that the use of Talniflumate reduces the immune escape of pancreatic cancer cells (Agostini et al., 2023). Additionally, GCNT3 was suggested to lower survival and increased chemoresistance in ovarian and colon cancers (Fernández et al., 2018). The last potential inhibitor, Flunarizine, is used against migraines, vertigo and occasionally in epilepsy treatment regimens. It is a selective Ca+ entry blocker, and works through calmodulin binding and blocks histamine and dopamine D2 (Chen et al., 2021). Past studies on its potential in drug repurposing have shown that it shows promise in lymphoma and MM cells (Schmeel et al., 2015), as well as T-cell leukemia and glioblastoma cells (Chen et al., 2021; Conrad et al., 2010). Interestingly, Flunarizine was found to degrade NRAS through the autophagy mechanism in basal like breast cancer (Zheng et al., 2018). NRAS is related to a number of oncogenic pathways such as PI3K-Akt, MAPK and RAF-MEK-ERK, and RAS mutations were found to be related to increased relapse in pediatric ALL (Qian et al., 2022). In summary, all three molecules show potential against various cancers in in vitro studies, although none of them have been studied in B-cell leukemia, nor were they analyzed for their effectiveness against ALL chemotherapy resistance.

To this end, the three candidates were tested on three cell lines; SUP-B15, Jurkat and HUVEC. The analysis showed that Eltrombopag was significantly reduced in cytotoxic activity in Jurkat cells, with its IC50 increasing from 1.78 uM (Figure 18B) to 7.45 uM (Figure 19B). Similarly, Flunarizine lost effectives as well, its IC50 increased from 4.5 uM in SUP-B15 (Figure 18C) to 8.65 uM in Jurkat cells (Figure 19C). Talniflumate did not show a significant difference in IC50 values, and its 15.36 uM IC50 in SUP-B15 (Figure 18D) cells was increased to 16.57 uM in Jurkat Figure 19D). Considering that SUP-B15 is a Ph+ ALL cell line with potentially increased chemoresistance (Wieduwilt, 2022), it is possible that the three candidate inhibitors, especially Eltrombopag and Flunarizine, may be able to sensitize chemoresistant cells to cytotoxic drugs considering that they work more efficiently in SUP-B15 cells. Although further tests are necessary for confirmation of this hypothesis.

Trypan Assay was employed to determine the cytotoxic effects of each drug in healthy cells, and HUVEC cells were administered the IC50 doses determined in SUP-B15 cells. Each of the drugs, Eltrombopag (Figure 20A), Flunarizine (Figure 20B) and Talniflumate (Figure 20C) showed significantly lowered cytotoxic effects in HUVEC cells when compared to SUP-B15, displaying their potential as inhibitors. In line with the bioinformatics results, Eltrombopag's cytotoxic effects shine through among the three candidates in both ALL cell lines, but specifically in SUP-B15, as Eltrombopag has the Eltrombopag has the lowest IC50 doses in both cell lines. Additionally, it has the least antiproliferative effect in healthy HUVEC cells, making it an even more attractive inhibitor candidate.

CHAPTER 5: CONCLUSION

The aim of this thesis was to analyze asparaginase, vincristine, prednisolone and daunorubicin resistant ALL data for DEGs related specifically to resistance, in order to determine genes that were common in all four types, and then, using the common genes, offer possible inhibitors that could be used to overcome resistance regardless of the drug type. To this end, resistant and sensitive ALL data was acquired from GEO, and DEG analysis was performed using LIMMA, comparing both resistant and sensitive data with healthy expression data. The resulting sensitive and resistant DEGs were compared to one another, and those that were found in common were eliminated as disease but not resistance related. At the end of DEG analysis, 1294 DEGs were found in total, and twelve of these were common in the four resistant subtypes. The majority of DEGs were protein coding, with 38 lncRNAs and 3 miRNAs. Subsequent KEGG pathway analysis revealed that pathways such as PIK-Akt, oncogenic pathways such as WNT, mTOR, cAMP, as well as JAK-STAT, and RAS were enriched. GO analysis showed high enrichment in pathways related to cell adhesion and morphology. PPI network analysis revealed 25 hub genes, including CREBBP, FN1 and HSPA8, along with IL4 and IL10. Among the common genes, CLDN9 and HS3ST3A1 were selected for in silico screening due to their novelty in cancer research and recent studies associating their gene families with resistance in various cancers. Their crystal structures were obtained from PDB and preprocessed with USCF Chimera, and they were docked against a small molecule library of 3556 acquired from ZINC15. The molecules with the lowest binding affinities in both analyses were selected and further investigated based on their ADMET properties, specifically absorption and toxicity. The three molecules that satisfied these constraints, Eltrombopag, Talniflumate and Flunarizine, were investigated with MD simulations for further validation. According to MD, all three molecules exhibited similar results, with Eltrombopag showing comparatively better potential as an inhibitor. Additional cytotoxicity analysis on ALL cell lines SUP-B15 and Jurkat, as well as healthy cell line HUVEC demonstrate again that among all three drugs, Eltrombopag has the potential to be most effective in ALL treatment with its low IC50 values, especially in SUP-B15.

While further in vitro and in vivo analyses of the effectiveness of these drugs and the transcription and protein expression levels of the genes are necessary to corroborate the findings, this study reveals novel genes associated with chemoresistance in ALL and points to three potential inhibitors in order to overcome the current relapse and chemoresistance rates.

REFERENCES

Aberuyi, N., Rahgozar, S., Ghodousi, E. S. and Ghaedi, K. (2019) *Drug Resistance Biomarkers and Their Clinical Applications in Childhood Acute Lymphoblastic Leukemia*. Frontiers in Oncology, Vol. 9, pp. 1496.

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. and Lindahl, E. (2015) *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers*. SoftwareX, Vol. 1-2, pp. 19–25.

Agostini, A., Guerriero, I., Piro, G., Quero, G., Roberto, L., Esposito, A., Caggiano, A., Priori, L., Scaglione, G., De Sanctis, F., Sistigu, A., Musella, M., Larghi, A., Rizzatti, G., Lucchetti, D., Alfieri, S., Sgambato, A., Bria, E., Bizzozero, L., … Carbone, C. (2023) *Talniflumate abrogates mucin immune suppressive barrier improving efficacy of gemcitabine and nab-paclitaxel treatment in pancreatic cancer*. Journal of Translational Medicine, Vol. 21(1), pp. 843.

Agu, P. C., Afiukwa, C. A., Orji, O. U., Ezeh, E. M., Ofoke, I. H., Ogbu, C. O., Ugwuja, E. I. and Aja, P. M. (2023) *Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management.* Scientific Reports, Vol. 13(1), pp. 1–18.

Aifantis, I., Raetz, E. and Buonamici, S. (2008) *Molecular pathogenesis of T-cell leukaemia and lymphoma.* Nature Reviews. Immunology, Vol. 8(5), pp. 380–390.

Allen, W. J., Balius, T. E., Mukherjee, S., Brozell, S. R., Moustakas, D. T., Lang, P. T., Case, D. A., Kuntz, I. D. and Rizzo, R. C. (2015) *DOCK 6: Impact of new features and current docking performance*. Journal of Computational Chemistry, Vol. 36(15), pp. 1132–1156.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M. and Sherlock, G. (2000) *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nature Genetics, Vol. 25(1), pp. 25–29.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A. and Wishart, D. S. (2016) *Heatmapper: web-enabled heat mapping for all.* Nucleic Acids Research, Vol. 44(W1), pp. W147–W153.

Bacci, M. R., Santos, J. A. B., Zing, N. C. P. and Barros, D. M. (2013) *Acute lymphocitic leukaemia and AIDS.* BMJ Case Reports, Vol. 2013.

Baert, L., Manfroi, B., Quintero, M., Chavarria, O., Barbon, P. V., Clement, E., Zeller, A., Van Kuppevelt, T., Sturm, N., Moreaux, J., Tveita, A., Bogen, B., McKee, T. and Huard, B. (2023) *3-O sulfation of syndecan-1 mediated by the sulfotransferase HS3ST3a1 enhances myeloma aggressiveness.* Matrix Biology: Journal of the International Society for Matrix Biology, Vol. 120, pp. 60–75.

Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngau, W.-C., Ledoux, P., Rudnev, D., Lash, A. E., Fujibuchi, W. and Edgar, R. (2005) *NCBI GEO: mining millions of expression profiles—database and tools.* Nucleic Acids Research, Vol. 33(suppl_1), pp. D562–D566.

Beer, D. G., Kardia, S. L. R., Huang, C.-C., Giordano, T. J., Levin, A. M., Misek, D. E., Lin, L., Chen, G., Gharib, T. G., Thomas, D. G., Lizyness, M. L., Kuick, R., Hayasaka, S., Taylor, J. M. G., Iannettoni, M. D., Orringer, M. B. and Hanash, S. (2002) *Gene-expression profiles predict survival of patients with lung adenocarcinoma.* Nature Medicine, Vol. 8(8), pp. 816–824.

Belver, L. and Ferrando, A. (2016) *The genetics and mechanisms of T cell acute lymphoblastic leukaemia.* Nature Reviews. Cancer, Vol. 16(8), pp. 494–507.

Benet, L. Z., Hosey, C. M., Ursu, O. and Oprea, T. I. (2016) *BDDCS, the Rule of 5 and drugability*. Advanced Drug Delivery Reviews, Vol. 101, pp. 89–98.

Benjamini, Y. and Hochberg, Y. (1995) *Controlling the false discovery rate: A practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society, Vol. 57(1), pp. 289–300.

Berendsen, H. J. C., van der Spoel, D. and van Drunen, R. (1995) *GROMACS: A message-passing parallel molecular dynamics implementation*. Computer Physics Communications, Vol. 91(1), pp. 43–56.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) *The Protein Data Bank*. Nucleic Acids Research, Vol. 28(1), pp. 235–242.

Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B. and Lindahl, E. (2010) *Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites and Water Models*. Journal of Chemical Theory and Computation, Vol. 6(2), pp. 459–466.

Blighe, K. (2018) *EnhancedVolcano*. Bioconductor.

Bolstad, B. M., Irizarry, R. A., Astrand, M. and Speed, T. P. (2003) *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.* Bioinformatics , Vol. 19(2), pp. 185–193.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983) *CHARMM: A program for macromolecular energy, minimization and dynamics calculations. Journal of Computational Chemistry*, Vol. 4(2), pp.187–217.

Bullinger, L., Döhner, K., Bair, E., Fröhling, S., Schlenk, R. F., Tibshirani, R., Döhner, H. and Pollack, J. R. (2004) *Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia*. The New England Journal of Medicine, Vol. 350(16), pp. 1605–1616.

Bussel, J. B. and Pinheiro, M. P. (2011) *Eltrombopag.* Cancer Treatment and Research, Vol. 157, pp. 289–303.

Caiazzo, M., Dell'Anno, M. T., Dvoretskova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T. D., Menegon, A., Roncaglia, P., Colciago, G., Russo, G., Carninci, P., Pezzoli, G., Gainetdinov, R. R., Gustincich, S., Dityatev, A. and Broccoli, V. (2011) *Direct generation of functional dopaminergic neurons from mouse and human fibroblasts*. Nature, Vol. 476(7359), pp. 224–227.

Cai, X., Liu, C., Zhang, T.-N., Zhu, Y.-W., Dong, X. and Xue, P. (2018) *Down-regulation of FN1 inhibits colorectal carcinogenesis by suppressing proliferation, migration and invasion*. Journal of Cellular Biochemistry, Vol. 119(6), pp. 4717–4728.

Carroll, W. L. and Bhatla, T. (2016) *Chapter 18 - Acute Lymphoblastic Leukemia*. In P. Lanzkowsky, J. M. Lipton and J. D. Fish (Eds.), *Lanzkowsky's Manual of Pediatric Hematology and Oncology (Sixth Edition)*, pp. 367–389. Academic Press.

Carugo, O. and Djinovic-Carugo, K. (2013) *Half a century of Ramachandran plots.* Acta Crystallographica. Section D, Biological Crystallography, Vol. 69(Pt 8), pp. 1333–1341.

Carvalho, B. S. and Irizarry, R. A. (2010) *A framework for oligonucleotide microarray preprocessing.* Bioinformatics, Vol. 26(19), pp. 2363–2367.

Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr, Onufriev, A., Simmerling, C., Wang, B. and Woods, R. J. (2005) *The Amber biomolecular simulation programs.* Journal of Computational Chemistry, Vol. 26(16), pp. 1668–1688.

Casella, G. (1992) *Illustrating empirical Bayes methods*. Chemometrics and Intelligent Laboratory Systems, Vol. 16(2), pp. 107–125.

Chakraborty, R., Bhattacharje, G., Baral, J., Manna, B., Mullick, J., Mathapati, B. S., Abraham, P., J, M., Hasija, Y., Ghosh, A. and Das, A. K. (2022) *In-silico screening and in-vitro assay show the antiviral effect of Indomethacin against SARS-CoV-2.* Computers in Biology and Medicine, Vol. 147, 105788.

Chautard, E., Fatoux-Ardore, M., Ballut, L., Thierry-Mieg, N. and Ricard-Blum, S. (2011*) MatrixDB, the extracellular matrix interaction database*. Nucleic Acids Research, Vol. 39(Database issue), pp. D235–D240.

Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., Salzberg, A. C. and Huang, E. S. (2007) *Structure-based maximal affinity model predicts small-molecule druggability*. Nature Biotechnology, Vol. 25(1), pp. 71–75.

Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., Lee, P. W. and Tang, Y. (2012) *admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties.* Journal of Chemical Information and Modeling, Vol. 52(11), pp. 3099–3105.

Chen, H. and Boutros, P. C. (2011) *VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R*. BMC Bioinformatics, Vol. 12, pp. 35.

Chen, L., Zhang, B., Wu, Z., Liu, G., Li, W. and Tang, Y. (2023) *In Silico discovery of aptamers with an enhanced library design strategy.* Computational and Structural Biotechnology Journal, Vol. 21, pp. 1005–1013.

Chen, S.-H., Chao, C.-N., Chen, S.-Y., Lin, H.-P., Huang, H.-Y. and Fang, C.-Y. (2021) *Flunarizine, a drug approved for treating migraine and vertigo, exhibits cytotoxicity in GBM cells*. European Journal of Pharmacology, Vol. 892, pp. 173756.

Chen, W.-J., Tang, R.-X., He, R.-Q., Li, D.-Y., Liang, L., Zeng, J.-H., Hu, X.-H., Ma, J., Li, S.-K. and Chen, G. (2017) *Clinical roles of the aberrantly expressed lncRNAs in lung squamous cell carcinoma: a study based on RNA-sequencing and microarray data mining.* Oncotarget, Vol. 8(37), pp. 61282–61304.

Chiaretti, S., Zini, G. and Bassan, R. (2014) *Diagnosis and subclassification of acute lymphoblastic leukemia.* Mediterranean Journal of Hematology and Infectious Diseases, Vol. 6(1), pp. e2014073.

Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T. and Lin, C.-Y. (2014) *cytoHubba: identifying hub objects and sub-networks from complex interactome*. BMC Systems Biology, Vol. 8 Suppl 4(Suppl 4), pp. S11.

Clappier, E., Cuccuini, W., Cayuela, J. M., Vecchione, D., Baruchel, A., Dombret, H., Sigaux, F. and Soulier, J. (2006) *Cyclin D2 dysregulation by chromosomal translocations to TCR loci in T-cell acute lymphoblastic leukemias*. Leukemia, Vol. 20(1), pp. 82–86.

Clough, E. and Barrett, T. (2016) *The Gene Expression Omnibus Database*. Methods in Molecular Biology, Vol. 1418, pp. 93–110.

Cobaleda, C. and Sánchez-García, I. (2009) *B-cell acute lymphoblastic leukaemia: towards understanding its cellular origin. BioEssays: News and Reviews in Molecular*, Cellular and Developmental Biology, Vol. 31(6), pp. 600–609.

Conrad, D. M., Furlong, S. J., Doucette, C. D., West, K. A. and Hoskin, D. W. (2010) *The Ca(2+) channel blocker flunarizine induces caspase-10-dependent apoptosis in Jurkat T-leukemia cells.* Apoptosis: An International Journal on Programmed Cell Death, Vol. 15(5), pp. 597–607.

Conter, V., Bartram, C. R., Valsecchi, M. G., Schrauder, A., Panzer-Grümayer, R., Möricke, A., Aricò, M., Zimmermann, M., Mann, G., De Rossi, G., Stanulla, M., Locatelli, F., Basso, G., Niggli, F., Barisone, E., Henze, G., Ludwig, W.-D., Haas, O. A., Cazzaniga, G., … Schrappe, M. (2010) *Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study*. Blood, Vol. 115(16), pp. 3206–3214.

Cortes, J., Thomas, D., Koller, C., Giles, F., Estey, E., Faderl, S., Garcia-Manero, G., McConkey, D., Ruiz, S. L., Guerciolini, R., Wright, J. and Kantarjian, H. (2004) *Phase I study of bortezomib in refractory or relapsed acute leukemias.* Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, Vol. 10(10), pp. 3371–3376.

Curtis, R. K., Oresic, M. and Vidal-Puig, A. (2005) *Pathways to the analysis of microarray data.* Trends in Biotechnology, Vol. 23(8), pp. 429–435.

Daina, A., Michielin, O. and Zoete, V. (2017) *SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules*. Scientific Reports, Vol. 7, pp. 42717.

Dallakyan, S. and Olson, A. J. (2015) *Small-molecule library screening by docking with PyRx*. Methods in Molecular Biology, Vol. 1263, pp. 243–250.

Dawe, G. B., Musgaard, M., Aurousseau, M. R. P., Nayeem, N., Green, T., Biggin, P. C. and Bowie, D. (2016) *Distinct structural pathways coordinate the activation of AMPA receptor-auxiliary subunit complexes*. Neuron, Vol. 89(6), pp. 1264–1276.

Duffield, A. S., Mullighan, C. G. and Borowitz, M. J. (2023) *International Consensus Classification of acute lymphoblastic leukemia/lymphoma*. Virchows Archiv: An International Journal of Pathology, Vol. 482(1), pp. 11–26.

Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., Ji, X.-L. and Liu, S.-Q. (2016) *Insights into Protein-Ligand Interactions: Mechanisms, Models and Methods.* International Journal of Molecular Sciences, Vol. 17(2).

Edgar, R., Domrachev, M. and Lash, A. E. (2002) *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository*. Nucleic Acids Research, Vol. 30(1), pp. 207–210.

Egler, R. A., Ahuja, S. P. and Matloub, Y. (2016*) L-asparaginase in the treatment of patients with acute lymphoblastic leukemia*. Journal of Pharmacology & Pharmacotherapeutics, Vol. 7(2), pp. 62–71.

Endo, Y., Sugimoto, K., Kobayashi, M., Kobayashi, Y., Kojima, M., Furukawa, S., Soeda, S., Watanabe, T., Higashi, A. Y., Higashi, T., Hashimoto, Y., Fujimori, K. and Chiba, H. (2022) *Claudin-9 is a novel prognostic biomarker for endometrial cancer.* International Journal of Oncology, Vol. 61(5).

Erickson-Miller, C. L., Kirchner, J., Aivado, M., May, R., Payne, P. and Chadderton, A. (2010) *Reduced proliferation of non-megakaryocytic acute myelogenous leukemia and other leukemia and lymphoma cell lines in response to eltrombopag*. Leukemia Research, Vol. 34(9), pp. 1224–1231.

Esmaili, E. and Shahlaei, M. (2015) *Analysis of the flexibility and stability of the structure of magainin in a bilayer and in aqueous and nonaqueous solutions using molecular dynamics simulations*. Journal of Molecular Modeling, Vol. 21(4), pp. 73.

Fadrná, E., Hladecková, K. and Koca, J. (2005) *Long-range electrostatic interactions in molecular dynamics: an endothelin-1 case study*. Journal of Biomolecular Structure & Dynamics, Vol. 23(2), pp. 151–162.

Fan, T., Xue, L., Dong, B., He, H., Zhang, W., Hao, L., Ma, W., Zang, G., Han, C. and Dong, Y. (2022) *CDH1 overexpression predicts bladder cancer from early stage and inversely correlates with immune infiltration.* BMC Urology, Vol. 22(1), pp. 156.

Farber, C. R. and Mesner, L. D. (2016) *Chapter 3 - A Systems-Level Understanding of Cardiovascular Disease through Networks.* In A. Rodriguez-Oquendo (Ed.), *Translational Cardiometabolic Genomic Medicine*, pp. 59–81. Academic Press.

Faruq, O., Zhao, D., Shrestha, M., Vecchione, A., Zacksenhaus, E. and Chang, H. (2022) *Targeting an MDM2/MYC Axis to Overcome Drug Resistance in Multiple Myeloma. Cancers*, Vol. 14(6).

Fernández, L. P., Sánchez-Martínez, R., Vargas, T., Herranz, J., Martín-Hernández, R., Mendiola, M., Hardisson, D., Reglero, G., Feliu, J., Redondo, A. and Ramírez de Molina, A. (2018) *The role of glycosyltransferase enzyme GCNT3 in colon and ovarian cancer prognosis and chemoresistance*. Scientific Reports, Vol. 8(1), pp. 8485.

First MIC Cooperative Study Group. (1986) *Morphologic, immunologic and cytogenetic (MIC) working classification of acute lymphoblastic leukemias*: Report of the Workshop held in Leuven, Belgium, April 22–23, 1985. Cancer Genetics and Cytogenetics, Vol. 23(3), pp. 189–197.

Garland, W., Benezra, R. and Chaudhary, J. (2013) *Chapter Fifteen - Targeting Protein–Protein Interactions to Treat Cancer—Recent Progress and Future Directions.* In M. C. Desai (Ed.), *Annual Reports in Medicinal Chemistry*, Vol. 48, pp. 227–245. Academic Press.

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B. and Overington, J. P. (2012) *ChEMBL: a large-scale bioactivity database for drug discovery*. Nucleic Acids Research, Vol. 40(Database issue), pp. D1100–D1107.

Gautier, L., Cope, L., Bolstad, B. M. and Irizarry, R. A. (2004) *affy--analysis of Affymetrix GeneChip data at the probe level*. Bioinformatics, Vol. 20(3), pp. 307–315.

Ge, S. X., Jung, D. and Yao, R. (2020) *ShinyGO: a graphical gene-set enrichment tool for animals and plants.* Bioinformatics, Vol. 36(8), pp. 2628–2629.

Gewirtz, D. A. (1999*) A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin*. Biochemical Pharmacology, Vol. 57(7), pp. 727–741.

Ghose, A. K., Viswanadhan, V. N. and Wendoloski, J. J. (1999) *A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug*

discovery. 1. A qualitative and quantitative characterization of known drug databases. Journal of Combinatorial Chemistry, Vol. 1(1), pp. 55–68.

Gidding, C. E., Kellie, S. J., Kamps, W. A. and de Graaf, S. S. (1999) *Vincristine revisited*. Critical Reviews in Oncology/hematology, Vol. 29(3), pp. 267–287.

Goll, J., Rajagopala, S. V., Shiau, S. C., Wu, H., Lamb, B. T. and Uetz, P. (2008) *MPIDB: the microbial protein interaction database*. Bioinformatics, Vol. 24(15), pp. 1743–1744.

González, M. A. (2011) *Force fields and molecular dynamics simulations*. École Thématique de La Société Française de La Neutronique, Vol. 12, pp. 169–200.

Grossfield, A. and Zuckerman, D. M. (2009) *Quantifying uncertainty and sampling quality in biomolecular simulations*. Annual Reports in Computational Chemistry, Vol. 5, pp. 23–48.

Gungormez, C., Gumushan Aktas, H., Dilsiz, N. and Borazan, E. (2019) *Novel miRNAs as potential biomarkers in stage II colon cancer: microarray analysis*. Molecular Biology Reports, Vol. 46(4), pp. 4175–4183.

Guo, L., Ren, H., Zeng, H., Gong, Y. and Ma, X. (2019) *Proteomic analysis of cerebrospinal fluid in pediatric acute lymphoblastic leukemia patients: a pilot study*. OncoTargets and Therapy, Vol. 12, pp. 3859–3868.

Gupta, M. K., Gouda, G., Sabarinathan, S., Donde, R., Dash, G. K., Vadde, R. and Behera, L. (2021) *Gene Ontology and Pathway Enrichment Analysis*. In M. K. Gupta and L. Behera (Eds.), *Bioinformatics in Rice Research: Theories and Techniques*, pp. 257–279. Springer Singapore.

Gutierrez, A., Jr, Tschumper, R. C., Wu, X., Shanafelt, T. D., Eckel-Passow, J., Huddleston, P. M., 3rd, Slager, S. L., Kay, N. E. and Jelinek, D. F. (2010) *LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis*. Blood, Vol. 116(16), pp. 2975–2983.

Haas, O. A. and Borkhardt, A. (2022*) Hyperdiploidy: the longest known, most prevalent and most enigmatic form of acute lymphoblastic leukemia in children.* Leukemia, Vol. 36(12), pp. 2769–2783.

Hansson, T., Oostenbrink, C. and van Gunsteren, W. (2002) *Molecular dynamics simulations*. Current Opinion in Structural Biology, Vol. 12(2), pp. 190–196.

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R. and Friesner, R. A. (2016) *OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins*. Journal of Chemical Theory and Computation, Vol. 12(1), pp. 281–296.

Harrison, C. J. and Johansson, B. (2015) *Acute lymphoblastic leukemia*. In Cancer Cytogenetics, pp. 198–251. John Wiley & Sons, Ltd.

Henegar, C., Cancello, R., Rome, S., Vidal, H., Clément, K. and Zucker, J.-D. (2006) *Clustering biological annotations and gene expression data to identify putatively co-regulated biological processes*. Journal of Bioinformatics and Computational Biology, Vol. 4(4), pp. 833–852.

Holleman, A., Cheok, M. H., den Boer, M. L., Yang, W., Veerman, A. J. P., Kazemier, K. M., Pei, D., Cheng, C., Pui, C.-H., Relling, M. V., Janka-Schaub, G. E., Pieters, R. and Evans, W. E. (2004) *Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment*. The New England Journal of Medicine, Vol. 351(6), pp. 533–542.

Hollingsworth, S. A. and Dror, R. O. (2018) *Molecular Dynamics Simulation for All*. Neuron, Vol. 99(6), pp. 1129–1143.

Hollingsworth, S. A. and Karplus, P. A. (2010) *A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins*. Biomolecular Concepts, Vol. 1(3-4), pp. 271–283.

Hospital, A, Goñi, J. R., Orozco, M. and Gelpi, J. L. (2015) *Molecular dynamics simulations: advances and applications.* Advances and Applications in Bioinformatics and Chemistry: AABC, Vol. 8, pp. 37–47.

Hou, H., Sun, D. and Zhang, X. (2019) *The role of MDM2 amplification and overexpression in therapeutic resistance of malignant tumors*. Cancer Cell International, Vol. 19, pp. 216.

Housman, G., Byler, S., Heerboth, S., Lapinska, K., Longacre, M., Snyder, N. and Sarkar, S. (2014*) Drug resistance in cancer: an overview*. Cancers, Vol. 6(3), pp. 1769–1792.

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C. and Lempicki, R. A. (2007) *DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists.* Nucleic Acids Research, Vol. 35(Web Server issue), pp. W169–W175.

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmüller, H. and MacKerell, A. D., Jr. (2017) *CHARMM36m: an improved force field for folded and intrinsically disordered proteins*. Nature Methods, Vol. 14(1), pp. 71–73.

Huang, W., Nussinov, R. and Zhang, J. (2017) *Computational Tools for Allosteric Drug Discovery: Site Identification and Focus Library Design*. Methods in Molecular Biology, Vol. 1529, pp. 439–446.

Iacobucci, I. and Mullighan, C. G. (2017) *Genetic Basis of Acute Lymphoblastic Leukemia*. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, Vol. 35(9), pp. 975–983.

Inaba, H. and Mullighan, C. G. (2020) *Pediatric acute lymphoblastic leukemia*. Haematologica, Vol. 105(11), pp. 2524–2539.

Inaba, H. and Pui, C.-H. (2010) *Glucocorticoid use in acute lymphoblastic leukaemia*. The Lancet Oncology, Vol. 11(11), pp. 1096–1106.

Inaba, H. and Pui, C.-H. (2021) *Advances in the Diagnosis and Treatment of Pediatric Acute Lymphoblastic Leukemia*. Journal of Clinical Medicine Research, Vol. 10(9).

Irving, J., Matheson, E., Minto, L., Blair, H., Case, M., Halsey, C., Swidenbank, I., Ponthan, F., Kirschner-Schwabe, R., Groeneveld-Krentz, S., Hof, J., Allan, J., Harrison, C., Vormoor, J., von Stackelberg, A. and Eckert, C. (2014) *Ras pathway*

mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. Blood, Vol. 124(23), pp. 3420–3430.

Jain, N., Lamb, A. V., O'Brien, S., Ravandi, F., Konopleva, M., Jabbour, E., Zuo, Z., Jorgensen, J., Lin, P., Pierce, S., Thomas, D., Rytting, M., Borthakur, G., Kadia, T., Cortes, J., Kantarjian, H. M. and Khoury, J. (2016) *Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype*. Blood, Vol. 127(15), pp. 1863–1869.

Jędraszek, K., Malczewska, M., Parysek-Wójcik, K. and Lejman, M. (2022) *Resistance Mechanisms in Pediatric B-Cell Acute Lymphoblastic Leukemia*. International Journal of Molecular Sciences, Vol. 23(6).

Jones, V. S., Huang, R.-Y., Chen, L.-P., Chen, Z.-S., Fu, L. and Huang, R.-P. (2016) *Cytokines in cancer drug resistance: Cues to new therapeutic strategies*. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, Vol. 1865(2), pp. 255–265.

José-Enériz, E. S., Román-Gómez, J., Cordeu, L., Ballestar, E., Gárate, L. andreu, E. J., Isidro, I., Guruceaga, E., Jiménez-Velasco, A., Heiniger, A., Torres, A., Calasanz, M. J., Esteller, M., Gutiérrez, N. C., Rubio, A., Pérez-Roger, I., Agirre, X. and Prósper, F. (2008) *BCR-ABL1-induced expression of HSPA8 promotes cell survival in chronic myeloid leukaemia.* British Journal of Haematology, Vol. 142(4), pp. 571–582.

Jourdan, J.-P., Bureau, R., Rochais, C. and Dallemagne, P. (2020) *Drug repositioning: a brief overview*. The Journal of Pharmacy and Pharmacology, Vol. 72(9), pp. 1145–1151.

Kakaje, A., Alhalabi, M. M., Ghareeb, A., Karam, B., Mansour, B., Zahra, B. and Hamdan, O. (2020) *Rates and trends of childhood acute lymphoblastic leukaemia: an epidemiology study*. Scientific Reports, Vol. 10(1), pp. 6756.

Kamijo, T., Weber, J. D., Zambetti, G., Zindy, F., Roussel, M. F. and Sherr, C. J. (1998) *Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 95(14), pp. 8292–8297.

Karplus, M. and Petsko, G. A. (1990) *Molecular dynamics simulations in biology.* Nature, Vol. 347(6294), pp. 631–639.

Kar, S. and Leszczynski, J. (2020) *Open access in silico tools to predict the ADMET profiling of drug candidates.* Expert Opinion on Drug Discovery, Vol. 15(12), pp. 1473–1487.

Kawedia, J. D. and Rytting, M. E. (2014) *Asparaginase in acute lymphoblastic leukemia*. Clinical Lymphoma, Myeloma & Leukemia, Vol. 14 Suppl, pp. S14–S17.

Klaus, B. and Reisenauer, S. (2016) *An end to end workflow for differential gene expression using Affymetrix microarrays*. F1000Research, Vol. 5, pp. 1384.

Kopeć, K., Pędziwiatr, M., Gront, D., Sztatelman, O., Sławski, J., Łazicka, M., Worch, R., Zawada, K., Makarova, K., Nyk, M. and Grzyb, J. (2019) *Comparison of α-Helix and β-Sheet Structure Adaptation to a Quantum Dot Geometry: Toward the Identification of an Optimal Motif for a Protein Nanoparticle Cover.* ACS Omega, Vol. 4(8), pp. 13086–13099.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W. and Ma'ayan, A. (2016) *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update*. Nucleic Acids Research, Vol. 44(W1), pp. W90–W97.

Kwon, M. J. (2013) *Emerging roles of claudins in human cancer*. International Journal of Molecular Sciences, Vol. 14(9), pp. 18148–18180.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR*. Journal of Biomolecular NMR, Vol. 8(4), pp. 477–486.

Lejman, M., Chałupnik, A., Chilimoniuk, Z. and Dobosz, M. (2022) *Genetic Biomarkers and Their Clinical Implications in B-Cell Acute Lymphoblastic Leukemia in Children*. International Journal of Molecular Sciences, Vol. 23(5).

Liang, J., Edelsbrunner, H. and Woodward, C. (1998) *Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design*. Protein Science: A Publication of the Protein Society, Vol. 7(9), pp. 1884–1897.

Li, J. and Ge, Z. (2021) *High HSPA8 expression predicts adverse outcomes of acute myeloid leukemia*. BMC Cancer, Vol. 21(1), pp. 475.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. and Shaw, D. E. (2010) *Improved side-chain torsion potentials for the Amber ff99SB protein force field*. Proteins, Vol. 78(8), pp. 1950–1958.

Lin, M., Ye, M., Zhou, J., Wang, Z. P. and Zhu, X. (2019) *Recent Advances on the Molecular Mechanism of Cervical Carcinogenesis Based on Systems Biology Technologies.* Computational and Structural Biotechnology Journal, Vol. 17, pp. 241–250.

Lin, X., Li, X. and Lin, X. (2020) *A Review on Applications of Computational Methods in Drug Screening and Design.* Molecules, Vol. 25(6).

Lipinski, C. A., Lombardo, F., Dominy, B. W. and Feeney, P. J. (2001*) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. Advanced Drug Delivery Reviews, Vol. 46(1-3), pp. 3–26.

Li, Q. and Shah, S. (2017) *Structure-Based Virtual Screening*. Methods in Molecular Biology, Vol. 1558, pp. 111–124.

Liu, H., Wang, M., Liang, N. and Guan, L. (2019) *Claudin-9 enhances the metastatic potential of hepatocytes via Tyk2/Stat3 signaling*. The Turkish Journal of Gastroenterology, Vol. 30(8), pp. 722–731.

Liu, Y., Liu, T., Zhou, Y., Li, W., Wang, M., Song, N., Zhang, W., Jiang, J., Yuan, S., Ding, J., Hu, G. and Lu, M. (2023) *Impeding the combination of astrocytic ASCT2 and NLRP3 by talniflumate alleviates neuroinflammation in experimental models of Parkinson's disease*. Acta Pharmaceutica Sinica. B, Vol. 13(2), pp. 662–677.

Liu, Z., Zheng, W., Liu, Y., Zhou, B., Zhang, Y. and Wang, F. (2021) *Targeting HSPA8 inhibits proliferation via downregulating BCR-ABL and enhances chemosensitivity in imatinib-resistant chronic myeloid leukemia cells*. Experimental Cell Research, Vol. 405(2), pp. 112708.

Loghavi, S., Kutok, J. L. and Jorgensen, J. L. (2015) *B-acute lymphoblastic leukemia/lymphoblastic lymphoma.* American Journal of Clinical Pathology, Vol. 144(3), pp. 393–410.

Lu, S., Li, S. and Zhang, J. (2014) *Harnessing allostery: a novel approach to drug discovery*. Medicinal Research Reviews, Vol. 34(6), pp. 1242–1285.

Malard, F. and Mohty, M. (2020) *Acute lymphoblastic leukaemia*. The Lancet, Vol. 395(10230), pp. 1146–1162.

Malczewska, M., Kośmider, K., Bednarz, K., Ostapińska, K., Lejman, M. and Zawitkowska, J. (2022) *Recent Advances in Treatment Options for Childhood Acute Lymphoblastic Leukemia*. Cancers, Vol. 14(8).

Malouf, C. and Ottersbach, K. (2018) *Molecular processes involved in B cell acute lymphoblastic leukaemia.* Cellular and Molecular Life Sciences: CMLS, Vol. 75(3), pp. 417–446.

Manglik, A., Lin, H., Aryal, D. K., McCorvy, J. D., Dengler, D., Corder, G., Levit, A., Kling, R. C., Bernat, V., Hübner, H., Huang, X.-P., Sassano, M. F., Giguère, P. M., Löber, S., Duan, D., Scherrer, G., Kobilka, B. K., Gmeiner, P., Roth, B. L. and Shoichet, B. K. (2016) *Structure-based discovery of opioid analgesics with reduced side effects*. Nature, Vol. 537(7619), pp. 185–190.

Martin, Y. C. (2005) *A bioavailability score*. Journal of Medicinal Chemistry, Vol. 48(9), pp. 3164–3170.

Masuda, M. and Yamada, T. (2015) *Signaling pathway profiling by reverse-phase protein array for personalized cancer medicine*. Biochimica et Biophysica Acta, Vol. 1854(6), pp. 651–657.

McCammon, J. A., Gelin, B. R. and Karplus, M. (1977) *Dynamics of folded proteins*. Nature, Vol. 267(5612), pp. 585–590.

McCorvy, J. D., Butler, K. V., Kelly, B., Rechsteiner, K., Karpiak, J., Betz, R. M., Kormos, B. L., Shoichet, B. K., Dror, R. O., Jin, J. and Roth, B. L. (2018) *Structure-inspired design of β-arrestin-biased ligands for aminergic GPCRs*. Nature Chemical Biology, Vol. 14(2), pp. 126–134.

Messinger, Y. H., Gaynon, P. S., Sposto, R., van der Giessen, J., Eckroth, E., Malvar, J., Bostrom, B. C. and Therapeutic Advances in Childhood Leukemia & Lymphoma (TACL) Consortium. (2012) *Bortezomib with chemotherapy is highly active in advanced B-precursor acute lymphoblastic leukemia: Therapeutic Advances in Childhood Leukemia & Lymphoma (TACL) Study.* Blood, Vol. 120(2), pp. 285–290.

Mohs, R. C. and Greig, N. H. (2017) *Drug discovery and development: Role of basic biological research.* Alzheimer's & Dementia: The Journal of the Alzheimer's Association, Vol. 3(4), pp. 651–657.

Moon, A. F., Edavettal, S. C., Krahn, J. M., Munoz, E. M., Negishi, M., Linhardt, R. J., Liu, J. and Pedersen, L. C. (2004) *Structural analysis of the sulfotransferase (3-o-sulfotransferase isoform 3) involved in the biosynthesis of an entry receptor for herpes simplex virus 1*. The Journal of Biological Chemistry, Vol. 279(43), pp. 45185–45193.

Moore, A. and Pinkerton, R. (2009) *Vincristine: Can its therapeutic index be enhanced?* Pediatric Blood & Cancer, Vol. 53(7), pp. 1180–1187.

Moorman, A. V., Harrison, C. J., Buck, G. A. N., Richards, S. M., Secker-Walker, L. M., Martineau, M., Vance, G. H., Cherry, A. M., Higgins, R. R., Fielding, A. K., Foroni, L., Paietta, E., Tallman, M. S., Litzow, M. R., Wiernik, P. H., Rowe, J. M., Goldstone, A. H., Dewald, G. W. and Adult Leukaemia Working Party, Medical Research Council/National Cancer Research Institute. (2007) *Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial*. Blood, Vol. 109(8), pp. 3189–3197.

Morita, K., Jain, N., Kantarjian, H., Takahashi, K., Fang, H., Konopleva, M., El Hussein, S., Wang, F., Short, N. J., Maiti, A., Sasaki, K., Garcia-Manero, G., Konoplev, S., Ravandi, F., Khoury, J. D. and Jabbour, E. (2021) *Outcome of T-cell acute lymphoblastic leukemia/lymphoma: Focus on near-ETP phenotype and differential impact of nelarabine.* American Journal of Hematology, Vol. 96(5), pp. 589–598.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) *AutoDock4 and AutoDockTools4: Automated docking with*

selective receptor flexibility. Journal of Computational Chemistry, Vol. 30(16), pp. 2785–2791.

Morshed, A. K. M. H., Al Azad, S., Mia, M. A. R., Uddin, M. F., Ema, T. I., Yeasin, R. B., Srishti, S. A., Sarker, P., Aurthi, R. Y., Jamil, F., Samia, N. S. N., Biswas, P., Sharmeen, I. A., Ahmed, R., Siddiquy, M. and Nurunnahar. (2023) *Oncoinformatic screening of the gene clusters involved in the HER2-positive breast cancer formation along with the in silico pharmacodynamic profiling of selective long-chain omega-3 fatty acids as the metastatic antagonists.* Molecular Diversity, Vol. 27(6), pp. 2651–2672.

Muegge, I., Heald, S. L. and Brittelli, D. (2001) *Simple selection criteria for drug-like chemical matter.* Journal of Medicinal Chemistry, Vol. 44(12), pp. 1841–1846.

Müller, C. E., Schiedel, A. C. and Baqi, Y. (2012) *Allosteric modulators of rhodopsin-like G protein-coupled receptors: opportunities in drug development.* Pharmacology & Therapeutics, Vol. 135(3), pp. 292–315.

Nisar, M., Paracha, R. Z., Arshad, I., Adil, S., Zeb, S., Hanif, R., Rafiq, M. and Hussain, Z. (2021) *Integrated Analysis of Microarray and RNA-Seq Data for the Identification of Hub Genes and Networks Involved in the Pancreatic Cancer.* Frontiers in Genetics, Vol. 12, pp. 663787.

Nussinov, R. and Tsai, C.-J. (2014) *Unraveling structural mechanisms of allosteric drug action.* Trends in Pharmacological Sciences, Vol. 35(5), pp. 256–264.

Onciu, M. (2009) *Acute lymphoblastic leukemia*. Hematology/oncology Clinics of North America, Vol. 23(4), pp. 655–674.

Opo, F. A. D. M., Rahman, M. M., Ahammad, F., Ahmed, I., Bhuiyan, M. A. and Asiri, A. M. (2021) *Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein.* Scientific Reports, Vol. 11(1), pp. 4049.

Pagadala, N. S., Syed, K. and Tuszynski, J. (2017*) Software for molecular docking: a review.* Biophysical Reviews, Vol. 9(2), pp. 91–102.

Pejovic, T. and Schwartz, P. E. (2002) *Leukemias*. Clinical Obstetrics and Gynecology, Vol. 45(3), pp. 866–878.

Petit, C., Gouel, F., Dubus, I., Heuclin, C., Roget, K. and Vannier, J. P. (2016) *Hypoxia promotes chemoresistance in acute lymphoblastic leukemia cell lines by modulating death signaling pathways*. BMC Cancer, Vol. 16(1), pp. 746.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) *UCSF Chimera--a visualization system for exploratory research and analysis.* Journal of Computational Chemistry, Vol. 25(13), pp. 1605–1612.

Plimpton, S. (1995) *Fast Parallel Algorithms for Short-Range Molecular Dynamics*. Journal of Computational Physics, Vol. 117(1), pp. 1–19.

Price, D. J. and Brooks, C. L., 3rd. (2004) *A modified TIP3P water potential for simulation with Ewald summation.* The Journal of Chemical Physics, Vol. 121(20), pp. 10096–10103.

Pufall, M. A. (2015) *Glucocorticoids and Cancer.* Advances in Experimental Medicine and Biology, Vol. 872, pp. 315–333.

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D. and Pirmohamed, M. (2019) *Drug repurposing: progress, challenges and recommendations*. Nature Reviews. Drug Discovery, Vol. 18(1), pp. 41–58.

Qian, J., Li, Z., Pei, K., Li, Z., Li, C., Yan, M., Qian, M., Song, Y., Zhang, H. and He, Y. (2022) *Effects of NRAS Mutations on Leukemogenesis and Targeting of Children With Acute Lymphoblastic Leukemia*. Frontiers in Cell and Developmental Biology, Vol. 10, pp. 712484.

Raetz, E. A. and Teachey, D. T. (2016) *T-cell acute lymphoblastic leukemia*. Hematolog, American Society of Hematology Education Program, 2016(1), pp. 580–588.

Rao, C. V., Janakiram, N. B., Madka, V., Kumar, G., Scott, E. J., Pathuri, G., Bryant, T., Kutche, H., Zhang, Y., Biddick, L., Gali, H., Zhao, Y. D., Lightfoot, S. and

Mohammed, A. (2016) *Small-Molecule Inhibition of GCNT3 Disrupts Mucin Biosynthesis and Malignant Cellular Behaviors in Pancreatic Cancer.* Cancer Research, Vol. 76(7), pp.1965–1974.

Rao, V. S., Srinivas, K., Sujini, G. N. and Kumar, G. N. S. (2014) *Protein-protein interaction detection: methods and analysis.* International Journal of Proteomics, Vol. 2014, pp.147648.

Rascio, F., Spadaccino, F., Rocchetti, M. T., Castellano, G., Stallone, G., Netti, G. S. and Ranieri, E. (2021) *The Pathogenic Role of PI3K/AKT Pathway in Cancer Onset and Drug Resistance: An Updated Review*. Cancers, Vol. 13(16).

Remke, M., Pfister, S., Kox, C., Toedt, G., Becker, N., Benner, A., Werft, W., Breit, S., Liu, S., Engel, F., Wittmann, A., Zimmermann, M., Stanulla, M., Schrappe, M., Ludwig, W.-D., Bartram, C. R., Radlwimmer, B., Muckenthaler, M. U., Lichter, P. and Kulozik, A. E. (2009) *High-resolution genomic profiling of childhood T-ALL reveals frequent copy-number alterations affecting the TGF-beta and PI3K-AKT pathways and deletions at 6q15-16.1 as a genomic marker for unfavorable early treatment response*. Blood, Vol. 114(5), pp. 1053–1062.

Ren, Y., Roy, S., Ding, Y., Iqbal, J. and Broome, J. D. (2004) *Methylation of the asparagine synthetase promoter in human leukemic cell lines is associated with a specific methyl binding protein*. Oncogene, Vol. 23(22), pp. 3953–3961.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015) *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Research, Vol. 43(7), pp. e47.

Roberts, K. G. and Mullighan, C. G. (2020) *The Biology of B-Progenitor Acute Lymphoblastic Leukemia.* Cold Spring Harbor Perspectives in Medicine, Vol. 10(7).

Rognan, D. (2017) *The impact of in silico screening in the discovery of novel and safer drug candidates*. Pharmacology & Therapeutics, Vol. 175, pp. 47–66.

Samosir, S. M., Utamayasa, I. K. A. andarsini, M. R., Rahman, M. A., Ontoseno, T., Hidayat, T., Ugrasena, I. D. G., Larasati, M. C. S. and Cahyadi, A. (2021) *Risk Factors of Daunorubicine Induced Early Cardiotoxicity in Childhood Acute Lymphoblastic*

Leukemia: A Retrospective Study. Asian Pacific Journal of Cancer Prevention: APJCP, Vol. 22(5), pp. 1407–1412.

Saxena, A., Wong, D., Diraviyam, K. and Sept, D. (2009) *Chapter 12 - The Basic Concepts of Molecular Modeling*. In M. L. Johnson and L. Brand (Eds.), *Methods in Enzymology*, Vol. 467, pp. 307–334. Academic Press.

Schmeel, L. C., Schmeel, F. C., Kim, Y., Blaum-Feder, S., Endo, T. and Schmidt-Wolf, I. G. H. (2015) *Flunarizine exhibits in vitro efficacy against lymphoma and multiple myeloma cells*. Anticancer Research, Vol. 35(3), pp. 1369–1376.

Schrappe, M., Bleckmann, K., Zimmermann, M., Biondi, A., Möricke, A., Locatelli, F., Cario, G., Rizzari, C., Attarbaschi, A., Valsecchi, M. G., Bartram, C. R., Barisone, E., Niggli, F., Niemeyer, C., Testi, A. M., Mann, G., Ziino, O., Schäfer, B., Panzer-Grümayer, R., … Conter, V. (2018) *Reduced-Intensity Delayed Intensification in Standard-Risk Pediatric Acute Lymphoblastic Leukemia Defined by Undetectable Minimal Residual Disease: Results of an International Randomized Trial (AIEOP-BFM ALL 2000).* Journal of Clinical Oncology, Vol. 36(3), pp. 244–253.

Shaker, B., Ahmad, S., Lee, J., Jung, C. and Na, D. (2021) *In silico methods and tools for drug discovery.* Computers in Biology and Medicine, Vol. 137, pp. 104851.

Shandilya, M., Sharma, S., Prasad Das, P. and Charak, S. (2021) *Molecular-level understanding of the anticancer action mechanism of anthracyclines*. In *Advances in Precision Medicine Oncology*. IntechOpen.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) *Cytoscape: a software environment for integrated models of biomolecular interaction networks*. Genome Research, Vol. 13(11), pp. 2498–2504.

Sharma, R. K., Chheda, Z. S., Das Purkayastha, B. P., Gomez-Gutierrez, J. G., Jala, V. R. and Haribabu, B. (2016) *A spontaneous metastasis model reveals the significance of claudin-9 overexpression in lung cancer metastasis*. Clinical & Experimental Metastasis, Vol. 33(3), pp. 263–275.

Shenoy, S. (2019) CDH1 (E-Cadherin) *Mutation and Gastric Cancer: Genetics, Molecular Mechanisms and Guidelines for Management*. Cancer Management and Research, Vol. 11, pp. 10477–10486.

Siegel, R. L., Miller, K. D., Wagle, N. S. and Jemal, A. (2023) *Cancer statistics, 2023*. CA: A Cancer Journal for Clinicians, Vol. 73(1), pp. 17–48.

Singh, A. B., Sharma, A. and Dhawan, P. (2010) *Claudin family of proteins and cancer: an overview*. Journal of Oncology, Vol. 2010, pp. 541957.

Slayton, W. B., Schultz, K. R., Silverman, L. B. and Hunger, S. P. (2020) *How we approach Philadelphia chromosome-positive acute lymphoblastic leukemia in children and young adults.* Pediatric Blood & Cancer, Vol. 67(10), pp. e28543.

Sleire, L., Førde, H. E., Netland, I. A., Leiss, L., Skeie, B. S. and Enger, P. Ø. (2017) *Drug repurposing in cancer*. Pharmacological Research, Vol. 124, pp. 74–91.

Slonim, D. K. and Yanai, I. (2009) *Getting started in gene expression microarray analysis*. PLoS Computational Biology, Vol. 5(10), pp. e1000543.

Smyth, G. K. (2005) *limma: Linear Models for Microarray Data*. In R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry and S. Dudoit (Eds.), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, pp. 397–420. Springer New York.

Sneha, P. and George Priya Doss, C. (2016) *Chapter Seven - Molecular Dynamics: New Frontier in Personalized Medicine*. In R. Donev (Ed.), *Advances in Protein Chemistry and Structural Biology*, Vol. 102, pp. 181–224. Academic Press.

Soerjomataram, I. and Bray, F. (2021) *Planning for tomorrow: global cancer incidence and the role of prevention 2020–2070*. Nature Reviews. Clinical Oncology, Vol. 18(10), pp. 663–672.

Sotillo, E., Barrett, D. M., Black, K. L., Bagashev, A., Oldridge, D., Wu, G., Sussman, R., Lanauze, C., Ruella, M., Gazzara, M. R., Martinez, N. M., Harrington, C. T., Chung, E. Y., Perazzelli, J., Hofmann, T. J., Maude, S. L., Raman, P., Barrera, A., Gill, S., … Thomas-Tikhonenko, A. (2015) *Convergence of Acquired Mutations and*

Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy. Cancer Discovery, Vol. 5(12), pp. 1282–1295.

Stam, R. W., Den Boer, M. L., Schneider, P., de Boer, J., Hagelstein, J., Valsecchi, M. G., de Lorenzo, P., Sallan, S. E., Brady, H. J. M., Armstrong, S. A. and Pieters, R. (2010) *Association of high-level MCL-1 expression with in vitro and in vivo prednisone resistance in MLL-rearranged infant acute lymphoblastic leukemia*. Blood, Vol. 115(5), pp. 1018–1025.

Stank, A., Kokh, D. B., Fuller, J. C. and Wade, R. C. (2016) *Protein Binding Pocket Dynamics*. Accounts of Chemical Research, Vol. 49(5), pp. 809–815.

Sterling, T. and Irwin, J. J. (2015) *ZINC 15--Ligand Discovery for Everyone*. Journal of Chemical Information and Modeling, Vol. 55(11), pp. 2324–2337.

Sun, D., Gao, W., Hu, H. and Zhou, S. (2022) *Why 90% of clinical drug development fails and how to improve it?* Acta Pharmaceutica Sinica. B, Vol. 12(7), pp. 3049–3062.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021) *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer Journal for Clinicians, Vol. 71(3), pp. 209–249.

Szczepański, T., van der Velden, V. H. J. and van Dongen, J. J. M. (2003) *Classification systems for acute and chronic leukaemias.* Best Practice & Research Clinical Haematology, Vol. 16(4), pp. 561–582.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J. and Mering, C. von. (2019) *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets*. Nucleic Acids Research, Vol. 47(D1), pp. D607–D613.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J. and von Mering, C. (2023) *The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest.* Nucleic Acids Research, Vol. 51(D1), pp. D638–D646.

Taghizadeh, M. S., Niazi, A., Moghadam, A. and Afsharifar, A. (2022) *Experimental, molecular docking and molecular dynamic studies of natural products targeting overexpressed receptors in breast cancer*. PloS One, Vol. 17(5), pp. e0267961.

Tao, Z., Shi, A., Li, R., Wang, Y., Wang, X. and Zhao, J. (2017) *Microarray bioinformatics in cancer- a review.* Official Journal of the Balkan Union of Oncology, Vol. 22(4), pp. 838–843.

Tasian, S. K., Loh, M. L. and Hunger, S. P. (2017) *Philadelphia chromosome–like acute lymphoblastic leukemia*. Blood, Vol. 130(19), pp. 2064–2072.

Teachey, D. T., Devidas, M., Wood, B. L., Chen, Z., Hayashi, R. J., Annett, R. D., Asselin, B. L., August, K. J., Cho, S. Y., Dunsmore, K. P., Fisher, B. T., Freedman, J. L., Galardy, P. J., Harker-Murray, P., Hermiston, M. L., Horton, T. M., Jaju, A. I., Lam, A., Messinger, Y. H., … Raetz, E. A. (2020) *Cranial Radiation Can be Eliminated in Most Children with T-Cell Acute Lymphoblastic Leukemia (T-ALL) and Bortezomib Potentially Improves Survival in Children with T-Cell Lymphoblastic Lymphoma (T-LL): Results of Children's Oncology Group (COG) Trial AALL1231*. Blood, Vol. 136(Supplement 1), pp. 11–12.

Terwilliger, T. and Abdul-Hay, M. (2017) *Acute lymphoblastic leukemia: a comprehensive review and 2017 update.* Blood Cancer Journal, Vol. 7(6), pp. e577.

Trott, O. and Olson, A. J. (2010) *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading*. Journal of Computational Chemistry, Vol. 31(2), pp. 455–461.

Vadillo, E., Dorantes-Acosta, E., Pelayo, R. and Schnoor, M. (2018) *T cell acute lymphoblastic leukemia (T-ALL): New insights into the cellular origins and infiltration mechanisms common and unique among hematologic malignancies.* Blood Reviews, Vol. 32(1), pp. 36–51.

Valk, P. J. M., Verhaak, R. G. W., Beijen, M. A., Erpelinck, C. A. J., Barjesteh van Waalwijk van Doorn-Khosrovani, S., Boer, J. M., Beverloo, H. B., Moorhouse, M. J., van der Spek, P. J., Löwenberg, B. and Delwel, R. (2004) *Prognostically useful gene-expression profiles in acute myeloid leukemia*. The New England Journal of Medicine, Vol. 350(16), pp. 1617–1628.

Van Vlierberghe, P., Ambesi-Impiombato, A., De Keersmaecker, K., Hadler, M., Paietta, E., Tallman, M. S., Rowe, J. M., Forne, C., Rue, M. and Ferrando, A. A. (2013) *Prognostic relevance of integrated genetic profiling in adult T-cell acute lymphoblastic leukemia*. Blood, Vol. 122(1), pp. 74–82.

Van Vlierberghe, P. and Ferrando, A. (2012) *The molecular basis of T cell acute lymphoblastic leukemia.* The Journal of Clinical Investigation, Vol. 122(10), pp. 3398–3406.

Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., Harris, N. L., Le Beau, M. M., Hellström-Lindberg, E., Tefferi, A. and Bloomfield, C. D. (2009) *The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes*. Blood, Vol. 114(5), pp. 937–951.

Vecchio, A. J. and Stroud, R. M. (2019) *Claudin-9 structures reveal mechanism for toxin-induced gut barrier breakdown.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 116(36), pp. 17817–17824.

Volkamer, A., Kuhn, D., Rippmann, F. and Rarey, M. (2012) *DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment.* Bioinformatics, Vol. 28(15), pp. 2074–2075.

Vyas, V., Jain, A., Jain, A. and Gupta, A. (2008) *Virtual Screening: A Fast Tool for Drug Design.* Scientia Pharmaceutica, Vol. 76(3), pp. 333–360.

Walker, N. M., Simpson, J. E., Levitt, R. C., Boyle, K. T. and Clarke, L. L. (2006) *Talniflumate increases survival in a cystic fibrosis mouse model of distal intestinal obstructive syndrome*. The Journal of Pharmacology and Experimental Therapeutics, Vol. 317(1), pp. 275–283.

Wang, K., Li, M. and Bucan, M. (2007) *Pathway-based approaches for analysis of genomewide association studies.* American Journal of Human Genetics, Vol. 81(6), pp. 1278–1283.

Wang, Y., Zeng, H.-M. and Zhang, L.-P. (2018) *ETV6/RUNX1-positive childhood acute lymphoblastic leukemia in China: excellent prognosis with improved BFM protocol.* Italian Journal of Pediatrics, Vol. 44(1), pp. 94.

Wiederstein, M. and Sippl, M. J. (2007) *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Research, Vol. 35(Web Server issue), pp. W407–W410.

Wieduwilt, M. J. (2022) *Ph+ ALL in 2022: is there an optimal approach*? Hematology American Society of Hematology Education Program, Vol. 2022(1), pp. 206–212.

Willemze, R., Hillen, H., Hartgrink-Groeneveld, C. A. and Haanen, C. (1975) *Treatment of acute lymphoblastic leukemia in adolescents and adults: A retrospective study of 41 patients (1970-1973)* Blood, Vol. 46(6), pp. 823–834.

Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B. and Hassanali, M. (2008*) DrugBank: a knowledgebase for drugs, drug actions and drug targets.* Nucleic Acids Research, Vol. 36(Database issue), pp. D901–D906.

Wolff, A., Bayerlová, M., Gaedcke, J., Kube, D. and Beißbarth, T. (2018) *A comparative study of RNA-Seq and microarray data analysis on the two examples of rectal-cancer patients and Burkitt Lymphoma cells*. PloS One, Vol. 13(5), pp. e0197162.

Worton, K. S., Kerbel, R. S. and Andrulis, I. L. (1991) *Hypomethylation and reactivation of the asparagine synthetase gene induced by L-asparaginase and ethyl methanesulfonate*. Cancer Research, Vol. 51(3), pp. 985–989.

Wu, P., Castner, D. G. and Grainger, D. W. (2008) *Diagnostic devices as biomaterials: a review of nucleic acid and protein microarray surface performance issues.* Journal of Biomaterials Science. Polymer Edition, Vol. 19(6), pp. 725–753.

Xie, D., Chen, Y., Wan, X., Li, J., Pei, Q., Luo, Y., Liu, J. and Ye, T. (2022) *The Potential Role of CDH1 as an Oncogene Combined With Related miRNAs and Their Diagnostic Value in Breast Cancer*. Frontiers in Endocrinology, Vol. 13, pp. 916469.

Yang, H., Lou, C., Sun, L., Li, J., Cai, Y., Wang, Z., Li, W., Liu, G. and Tang, Y. (2019) *admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties*. Bioinformatics, Vol. 35(6), pp. 1067–1069.

Yoda, S., Lin, J. J., Lawrence, M. S., Burke, B. J., Friboulet, L., Langenbucher, A., Dardaei, L., Prutisto-Chang, K., Dagogo-Jack, I., Timofeevski, S., Hubbeling, H.,

Gainor, J. F., Ferris, L. A., Riley, A. K., Kattermann, K. E., Timonina, D., Heist, R. S., Iafrate, A. J., Benes, C. H., … Shaw, A. T. (2018) *Sequential ALK Inhibitors Can Select for Lorlatinib-Resistant Compound ALK Mutations in ALK-Positive Lung Cancer*. Cancer Discovery, Vol. 8(6), pp. 714–729.

Zhang, J., Adrián, F. J., Jahnke, W., Cowan-Jacob, S. W., Li, A. G., Iacob, R. E., Sim, T., Powers, J., Dierks, C., Sun, F., Guo, G.-R., Ding, Q., Okram, B., Choi, Y., Wojciechowski, A., Deng, X., Liu, G., Fendrich, G., Strauss, A., … Gray, N. S. (2010) *Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors*. Nature, Vol. 463(7280), pp. 501–506.

Zhang, L., Fu, L., Zhang, S., Zhang, J., Zhao, Y., Zheng, Y., He, G., Yang, S., Ouyang, L. and Liu, B. (2017) *Discovery of a small molecule targeting ULK1-modulated cell death of triple negative breast cancer in vitro and in vivo*. Chemical Science, Vol. 8(4), pp. 2687–2701.

Zhang, Y., Szustakowski, J. and Schinke, M. (2009) *Bioinformatics Analysis of Microarray Data*. In K. DiPetrillo (Ed.*), Cardiovascular Genomics: Methods and Protocols*, pp. 259–284. Humana Press.

Zheng, Z.-Y., Li, J., Li, F., Zhu, Y., Cui, K., Wong, S. T., Chang, E. C. and Liao, Y.-H. (2018) *Induction of N-Ras degradation by flunarizine-mediated autophagy*. Scientific Reports, Vol. 8(1), pp. 16932.

Zhu, H., Martin, T. M., Ye, L., Sedykh, A., Young, D. M. and Tropsha, A. (2009) *Quantitative structure-activity relationship modeling of rat acute toxicity by oral exposure*. Chemical Research in Toxicology, Vol. 22(12), pp. 1913–1921.

Zoete, V., Cuendet, M. A., Grosdidier, A. and Michielin, O. (2011) *SwissParam: a fast force field generation tool for small organic molecules*. Journal of Computational Chemistry, Vol. 32(11), pp. 2359–2368.