Universitat Oberta de Catalunya

uoc.edu

Network-based Drug Repurposing for Multiple Myeloma using SAveRUNNER in R

Sergio Carracedo Huroz
Masters in Bioinformatics and Biostatistics
Area: Molecular Basis for Drug Development

Supervisor: Marta Enciso, PhD
Subject responsible: David Merino Arranz

January 2022


Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada 3.0 España de Creative Commons

FICHA DEL TRABAJO FINAL

Título del trabajo:	Network-based drug repurposing for mulitple myeloma
Nombre del autor:	Sergio Carracedo Huroz
Nombre del consultor/a:	Marta Enciso Carrasco
Nombre del PRA:	David Merino Arranz
Fecha de entrega (mm/aaaa):	01/2022
Titulación:	Máster en Bioinformática y Bioestadística
Área del Trabajo Final:	Bases Moleculares para el Desarrollo de Fármacos
Idioma del trabajo:	Inglés
Palabras clave	Drug repurposing, multiple myeloma, network-based, R, SAveRUNNER

Resumen del Trabajo (máximo 250 palabras): Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.

ANTECEDENTES: El mieloma múltiple (MM) es a día de hoy un cáncer letal, por lo que urge encontrar nuevos fármacos que permitan tratar esta enfermedad de manera más efectiva. Varios tipos de análisis *in silico* permiten encontrar fármacos actualmente disponibles para reposicionarlos en enfermedades diferentes de aquéllas para las que originalmente fueron diseñados, siendo el análisis computacional basado en redes uno de los más comunes.

METODOLOGÍA: Se ha hecho un estudio de reposicionamiento para MM mediante la implementación en R del algoritmo SAveRUNNER, el cual realiza un análisis basado en redes para generar listas de fármacos candidatos a reposicionamiento. Entre dichos candidatos, únicamente aquéllos validados por la herramienta ´Query´ de CMap a partir de los genes diferencialmente expresados en muestras de MM se consideraron como los más prometedores.

RESULTADOS: Se obtuvo una lista final de 22 candidatos a reposicionamiento para MM pertenecientes a diferentes categorías, muchos de los cuales se habían usado previamente con otros tipos de cánceres. Finalmente, se presentan análisis de acoplamiento molecular de los candidatos ponatinib y axitinib con la proteína KIT, sobreexpresada en MM según este estudio, con el fin de comparar sus afinidades y valorar cual sería preferible para una posible línea de tratamiento de MM.

CONCLUSIÓN: En este estudio se muestra la exactitud de SAveRUNNER para generar fármacos para tratar MM al sugerir candidatos que actualmente ya se usan para tratar esta enfermedad. Además, SAveRUNNER sugiere nuevos candidatos a reposicionamiento que podrían mejorar el actual mal prognosis del MM.

Abstract (in English, 250 words or less):

BACKGROUND: Multiple myeloma (MM) remains a lethal blood malignancy, so new drugs are necessary in order to treat this cancer more effectively. Different types of *in silico* analyses make it possible to repurpose currently available drugs to diseases other than those they were originally designed for, with network-based analyses being a commonly chosen approach.

METHODOLOGY: In this work, a drug repurposing study for MM was carried out by implementing in R the recently published algorithm SAveRUNNER, which performs network-based analyses to generate lists of potentially repurposable candidates for diseases of interest. Among the candidates to repurpose to MM suggested by SAveRUNNER, only those validated by differential gene expression analyses in MM samples followed by CMap queries were considered as most promising.

RESULTS: A final list of 22 drugs for MM repositioning belonging to different categories, such as enzyme inhibitors or steroids, was obtained, with many of them being already used to treat other types of cancers. Finally, molecular docking analyses of the potentially repurposable candidates ponatinib or axitinib with the KIT protein, overexpressed in MM according to this study, are presented to compare affinities of a protein for drugs of the same type in order to assess which would be preferable if included in a potential line of MM treatment.

CONCLUSION: This study shows the accuracy of SAveRUNNER by suggesting drugs currently used to treat MM, and suggests new candidates for repositioning that may improve MM's current poor prognosis.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1. Context and rationale for this thesis

Cancer causes one out of six deaths worldwide, and is thus the second leading cause of death globally[1]. As earlier mentioned, MM remains as one of the incurable types of cancer, with current treatments being able to modestly extend patients' lives[2]. Therefore, finding new therapeutics capable of improving existing approaches to treat this disease is imperative.

On the other hand, and unlike the resource and time consuming *de novo* drug discovery process, drug repurposing via computational analyses stand as a promising alternative to find new therapeutics with a considerably smaller budget and time frame.

Thus, considering the urge to find new therapeutics for MM and the potential of *in silico* drug repurposing to suggest new candidates to treat diseases in a relatively short time, carrying out a study that implements a computational algorithm to find approved drugs with novel indications for MM seems well justified. Although scarce, a couple of studies on this direction have actually already been performed in the past, but rather than using an algorithm to find candidates to repurpose, the approach in both instances found differentially expressed genes (DEGs) in MM to upload them to CMap in order to find compounds to repurpose to this disease[3,4]. In this thesis, however, a somewhat different methodology is used to find repurposable compounds for MM by implementing the novel algorithm SAveRUNNER, leaving the differential gene expression/CMap analyses as a validation/filtering step for the resulting list of drug candidates instead, as recently done for other diseases[5]. It is worth noting that SAveRUNNER seems to have better performance than previous top algorithms[5,6], which might raise the possibility of finding new drugs that can help improve the outcome of current MM treatments.

1.2. Objectives

1.2.1. Main Objectives

The main goal of thesis is finding drug candidates with reposition potential for MM that can be added to the currently available set of drugs used to treat this disease by implementing the network-based computational algorithm SAveRUNNER.

1.2.2. Specific Objectives

In order to achieve the main objective, the following specific objectives were accomplished:

1. Executing the following SAveRUNNER implementations in R:

- **Implementation 1**: Generation of a drug-disease subnetwork for MM by using this disease together with all blood related malignancies available in the

Phenopedia[7] database as input in the configuration file of SAveRUNNER (config.R file).

- **Implementation 2:** Generation of a drug-disease subnetwork for MM by using this disease together with diseases related to MM symptoms[8] available in the Phenopedia database as input in the config.R file.
- **Implementation 3:** Generation of a drug-disease subnetwork for a disease seemingly unrelated to MM (Obsessive Compulsive Disorder, OCD) as negative control for the two previous implementations by using OCD together with a set of diseases also unrelated to MM as input in the config.R file.

  Only compounds obtained simultaneously in both implementation 1 and implementation 2 will be selected for further validation and therefore with possibility to be considered as drug candidates with repurposing potential.

2. Full RNAseq DEA of the dataset GSE175384 containing samples from MM and healthy subjects by following the pipeline provided at https://github.com/ASPteaching/Omics_data_analysis-Case_study_2-RNA-seq , which was provided to students of the Master's degree in Bioinformatics and Biostatistics from the University Oberta of Catalonia and University of Barcelona during the first part of the virtual class 'Omics data analyses'. Among other results, this analysis yielded a list of annotated DEGs for MM subjects included in this dataset.
3. Full microarrray DEA of the dataset GSE47552 containing samples from MM and healthy subjects by following the pipeline provided at https://github.com/ASPteaching/Omics_Data_Analysis-Case_Study_1-Microarrays , which was provided to students of the Master's degree in Bioinformatics and Biostatistics from the University Oberta of Catalonia and University of Barcelona during the second part of the virtual class 'Omics data analyses'. Among other results, this analysis yielded a list of annotated DEGs for MM subjects included in this dataset.
4. Independent Queries of DEGs found in DEAs of GSE175384 and GSE47552 datasets by using the CMap Query tool (www.clue.io/query) in order to find compounds that counteract regulation of DEGs for these datasets, and would thereby have potential to treat patients included in these studies.
5. Generation of a final list of validated drugs with repurposable potential for MM by selecting compounds generated by SAveRUNNER for this disease that:

   a) Are also part of at least one of the DEA/CMap analyses performed in this study,

   b) At the time of writing this thesis, have not been found as part of any study related to MM in the literature.

6. An example of molecular docking analysis by using the KIT kinase domain protein (target) corresponding to a DEG found in GSE175384, and the validated repurposable candidates ponatinib and axitinib (ligands) with similar binding

sites to compare their affinities to their common target, thereby assessing which one would be preferable.

1.3. Planning with tasks, milestones, and calendar

This study was carried out during the first semester of the academic year 2021/22, with start date on the 15th of September, and going through different evaluation tests (pruebas de evaluación contínua (PECs)), until its public defense during the second/third week of January 2022. Each PEC was considered a milestone, and consisted of different tasks to be completed in order to achieve the project objectives successfully. The dates for the completion of the different tasks in each of the PECs/milestones can be found in the following chronogram (Figure 1):

Outline n...	Name	Begin date	End date	Duration
1	PEC0: MASTER'S FINAL PROJECT PROPOSAL	15/09/2021	22/09/2021	8
2	PEC1: WORK PLAN	23/09/2021	04/10/2021	12
2.1	Bibliography search on Multiple Mieloma	23/09/2021	25/09/2021	3
2.2	Bibliography search on drug repurposing	26/09/2021	28/09/2021	3
2.3	Familiarization with SAveRUNNER	29/09/2021	30/09/2021	2
2.4	Ellaboration of the PEC1 report	01/10/2021	04/10/2021	4
3	PEC2: WORK DEVELOPMENT - PHASE 1	05/10/2021	08/11/2021	35
3.1	Implementations of SAveRUNNER	05/10/2021	06/11/2021	33
3.1.1	Network with MM and blood related diseases	05/10/2021	28/10/2021	24
3.1.2	Network with MM and symptoms related diseases	29/10/2021	02/11/2021	5
3.1.3	Network without MM and with random diseases	03/11/2021	06/11/2021	4
3.2	Ellaboration of the PEC2 report	07/11/2021	08/11/2021	2
4	PEC3: WORK DEVELOPMENT - PHASE 2	09/11/2021	06/12/2021	28
4.1	Differential expression analyses (DEAs)	09/11/2021	22/11/2021	14
4.1.1	DEA for RNAseq GSE175384 dataset	09/11/2021	15/11/2021	7
4.1.1.1	Full gene expression analysis	09/11/2021	15/11/2021	7
4.1.1.2	List with DEGs	15/11/2021	15/11/2021	1
4.1.2	DEA for Microarray GSE47552 dataset	16/11/2021	22/11/2021	7
4.1.2.1	Full gene expression analysis	16/11/2021	22/11/2021	7
4.1.2.2	List with DEGs	22/11/2021	22/11/2021	1
4.2	CMap drug signatures	23/11/2021	26/11/2021	4
4.2.1	CMap drug signature for DEGs in GSE175384	23/11/2021	24/11/2021	2
4.2.2	CMap drug signature for DEGs in GSE47552	25/11/2021	26/11/2021	2
4.3	Lists of validated drug candidates	27/11/2021	28/11/2021	2
4.3.1	Drug candidates fully validated	27/11/2021	27/11/2021	1
4.3.2	Drug candidates partially validated	28/11/2021	28/11/2021	1
4.4	Examples of molecular docking to compare drug-target affinities	29/11/2021	02/12/2021	4
4.5	Ellaboration of the PEC3 report	03/12/2021	06/12/2021	4
5	PEC4: FINAL REPORT	07/12/2021	24/12/2021	18
6	PEC5a: PRESENTATION ELLABORATION	27/12/2021	03/01/2022	8
7	PEC5b: THESIS DEFENCE	13/01/2022	21/01/2022	9

Figure 1: Gantt chronogram for PECs (milestones) and corresponding tasks.
Made with GanttProject[9].

In the corresponding gantt chart (Figure 2), the different milestones with their respective tasks achieved during the development of this project are represented with different colors for better identification .

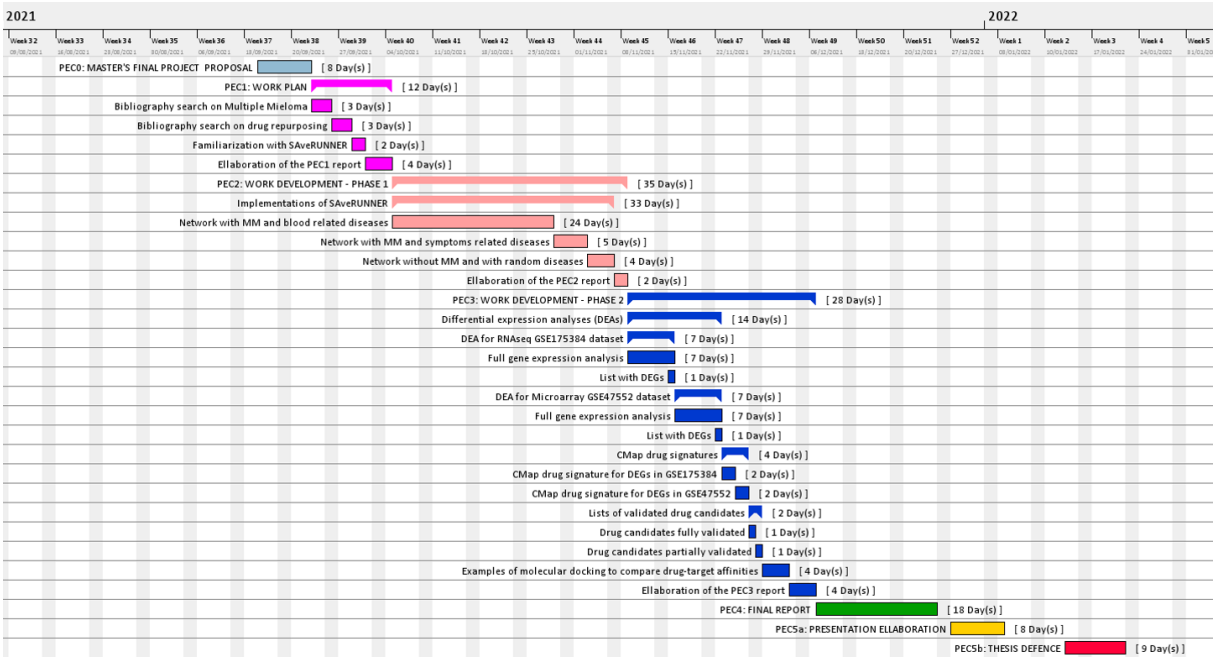


Figure 2: Gantt chart for PECs (milestones) and corresponding tasks.
Made with GanttProject[9].

1.4. Brief summary of obtained products

The following products have resulted upon completion of this master's final project:

- A **thesis (this document)**, which mainly includes introductions to MM and drug repurposing, followed methodology, as well as results and discussion related to the objectives mentioned in section 2.

- **Supporting information files** uploaded in github at https://github.com/appropiate/TFM_UoC, which consist of compound and DEG tables, DEA reports, as well as docking analyses that have all been generated as part of the results, but are too large to be completely included in the main text of this thesis.

- **Thesis presentation and defence (PEC5a and PEC5b)**: A brief introduction to the research topic of the project followed by a summary of the results has been ellaborated with LibreOffice Impress or similar tool to present and defend the thesis.

1.5. Brief description of other sections

Upon describing the rationale for this thesis, its objectives with their corresponding timelines, and a summary of the obtained products, the remaining sections of this work are briefly explained next:

- **Background** on MM and the SAveRUNNER algorithm to help the reader gain basic knowledge on concepts regarding the main topics of this thesis, thus facilitating its understanding.
- **Materials and methods**, where a description of the hardware, software, data, and methodology used in this work to produce the presented results is presented.
- **Results** showing drug candidates for reposition to MM generated by SAveRUNNER, validation of these using DEA/CMap analyses, selection of compounds that were not previously found associated to MM treatment in the literature, and a molecular docking example of ligands (ponatinib and axitinib) with a common binding site in a protein differentially expressed in a MM dataset used in this study (KIT kinase domain).
- **Conclusions**, where the main findings, changes on initial plans due to unforeseen difficulties, as well as brief ideas of future projects following up on this study are presented.
- **Glossary** with the meaning of all the abbreviations used throughout this thesis.
- **References,** consisting of a list of all the scientific articles, official websites, and repositories that have been referred to on the different sections of this work.
- **Supplementary information** listing the different files and directories with reports, tables and other files produced throughout this work that were too large to be completely included in this thesis, reason for which they have been made available via GitHub at https://github.com/appropiate/TFM_UoC .

2. BACKGROUND

This thesis is focused on implementing the network-based algorithm SAveRUNNER to find repurposable drug candidates for Multiple Myeloma (MM). Therefore, a general background for both MM and drug repurposing is presented in this section.

2.1. Background on Multiple Myeloma: A brief description

Although important advances have been made in recent years, MM is still an incurable and deadly type of blood cancer that encompasses 10% of all hematological malignancies[10] and 0.9% of all cancer diagnoses[11]. It is therefore paramount to keep elucidating new aspects and molecular mechanisms that may help create better and more specific therapeutical interventions in the future.

2.1.1. Diagnosis

Up to 74 years old, the cumulative risk of suffering MM worldwide is approximately 0.21 (approximately 1 in 500)[12], and the median age at diagnosis is 66-70 years[13]. In order to diagnose MM, the patient must meet certain criteria (Table 1). First, the population of clonal bone marrow plasma cells must be greater than 10% or, alternatively, a bony or extramedullar plasmacytoma must exist. The second criteria to be met is that the patient has either a) Calcium elevation, Renal dysfunction, Anemia and Bone disease (altogether known as 'CRAB' criteria), b) high percentage of clonal bone marrow cells, c) overabundance of involved serum-free light chains (secreted by myeloma cells), or d) lesions detected by magnetic resonance of at least 5 mm[14].

Multiple Myeloma	Both criteria must be met: • Clonal bone marrow plasma cells ≥10% or biopsy-proven bony or extramedullary plasmacytoma • Any one or more of the following myeloma defining events: – Evidence of end organ damage that can be attributed to the underlying plasma cell proliferative disorder, specifically: ♦ Hypercalcemia: serum calcium >0·25 mmol/L (>1 mg/dL) higher than the upper limit of normal or >2·75 mmol/L (>11 mg/dL) ♦ Renal insufficiency: creatinine clearance <40 mL per minute or serum creatinine >177 μmol/L (>2 mg/dL) ♦ Anemia: hemoglobin value of >2 g/dL below the lower limit of normal, or a hemoglobin value <10 g/dL ♦ Bone lesions: one or more osteolytic lesions on skeletal radiography, computed tomography (CT), or positron emission tomography-CT (PET-CT) – Clonal bone marrow plasma cell percentage ≥60% – Involved: uninvolved serum free light chain (FLC) ratio ≥100 (involved free light chain level must be ≥100 mg/L) – >1 focal lesions on magnetic resonance imaging (MRI) studies (at least 5mm in size)

Table 1. Diagnosis criteria for MM.
Adapted from[14].

2.1.2. Clinical manifestations

Symptoms experienced by MM patients include constipation, leg weakness/numbness, fatigue, reduced appetite, weight loss, or bone pain among others[8], with the latter being related to myeloma cells promoting the release of different factors, such as RANKL, TNF-a, IL-6, and VEGF, which all promote the activity of the osteoclast and its precursors, as well as the release of factors, such as DKK1, SFRP3, HGF, TGF-Beta, Sclerotin, or Activin A, that inhibit the activity of osteoblasts and precursors. Osteoclast activation and osteoblast inhibition cause disruption of the balance between bone formation versus bone resorption, and leads to osteolytic lesions[15] that MM patients often experience (Figure 3).

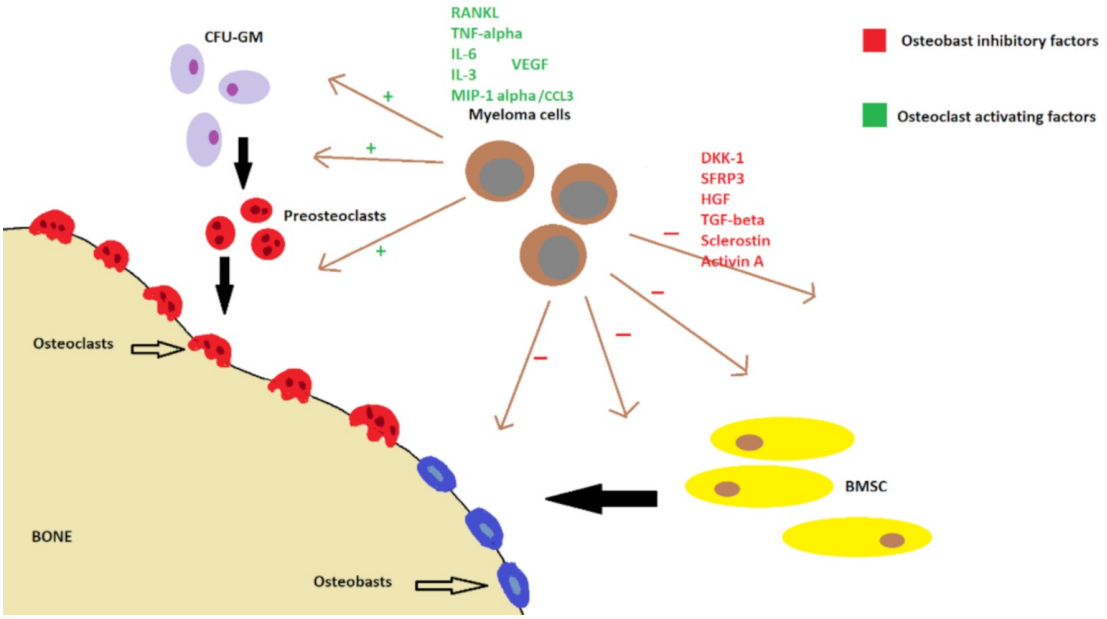


Figure 3: Bone activity influenced by plasma cells.
Osteoclast activity is enhanced whereas osteoblast activity is inhibited by abnormal plasma cells during MM[15].

2.1.3. MM classification, staging and stratification

MM is originated in the bone marrow and affects a type of white blood cell known as plasma cell, which normally makes antibodies to fight infections. An antibody is a molecule made of two equal heavy chain proteins (α, δ, ε, γ, and μ) and two light chain proteins (κ and λ), which all belong to the immunoglobulin (Ig) superfamily[16] (Figure 4).

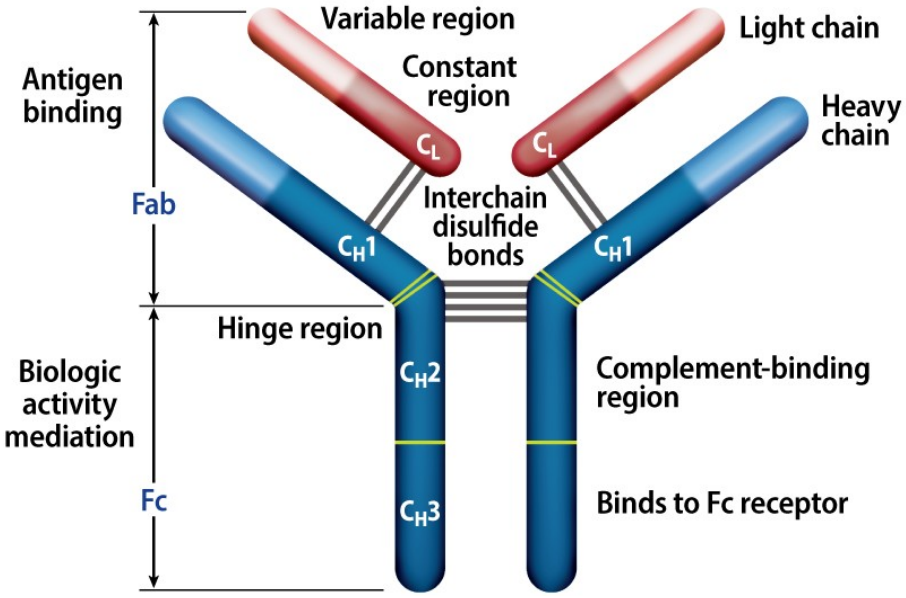


Figure 4: Antibody structure.
The structure of the antibody consists of two heavy and two light chains, each of which contains different regions with different functions[17].

Different combinations of heavy chains and light chains will result in different types of antibodies, with each plasma cell producing a specific type of antibody only. Therefore, having different plasma cells equips a healthy organism with a repertoire of different antibodies and enables it to fight infections. However, during MM, a plasma cell becomes dysregulated, proliferates abnormally, and over time outcrowds the rest of plasma cells, which makes the patient more susceptible to infections because abnormal plasma cells in a myeloma originally come from the same plasma cell clone and therefore make the same antibody or the same free light chain (FLC), which becomes predominant[18]. The secreted antibody or FLC type can in turn be used as a biological marker to help classify the type of MM[17], although non secretory myeloma also exist, and are harder to diagnose[19].

In order to classify MM attending to its stage, the International Staging System (ISS)[20] is commonly used as staging criteria, which divides MM cases in three categories attending to levels of albumin, β_2-microglobulin, and lactate dehydrogenase (LDH) or cytogenetics analyses (Table 2). However, other systems with different criteria, such as, myeloma cell mass[21], or serum free light chain ratio (sFLCR)[22] also have prognostic value.

Stage I
All of the following:
Serum albumin $\geq$ 3.5 gm/dL
Serum β-2-microglobulin $<$ 3.5 mg/L
No high-risk cytogenetics
Normal serum LDH
Stage II
Not fitting into stages I or III
Stage III
Both of the following:
Serum β-2-microglobulin $>$ 5.5 mg/L
High-risk cytogenetics [*t*(4;14), *t*(14;16), or del(17p)] or elevated serum LDH

Table 2. *Staging criteria for Multiple Myeloma*
Adapted from[20].

Different genetic alterations found in MM, such as deletions and translocations, are strongly associated to the aggressiveness of MM. Therefore, the prognosis for a patient will significantly depend on the type(s) of cytogenetic abnormality present in the plasma cell clone that is responsible for a given myeloma (Table 3).

Standard risk
Trisomies
t(11;14)
t(6;14)
Intermediate risk
t(4;14)
Gain(1q)
High risk
t(14;16)
t(14;20)
Del(17p)
TP53 mutation
High-risk GEP signature
R-ISS stage III
High plasma cell S-phase

Table 3. mSMART risk stratification. Adapted from[20].

In addition, MM can also be classified as newly diagnosed (NDMM) or relapsed/refractory MM (RRMM), with the latter referring to reappearance of MM signs and symptoms after a period of partial remission[23].

2.1.4. Current treatments

Different types of drugs are commonly included as part of MM treatment (Figure 5), including:

- **Proteosome inhibitors** (PIs), such as bortezomib (usually named in treatments as Velcade®) or cafilzomib (kyprolis®), which basically work by inhibiting degradation of proteins that need to be eliminated so that myeloma cells can thrive and keep proliferating[24,25].
- Derivatives of thalidomide, lenolinamide[26,27] (Revlimid®) and pomalydomide[28], which are known as **immunomodulatory drugs** (IMdDs). They seem to help fighting myeloma cells differently, with the mechanism of action of the latter yet to be elucidated[28].
- **Alkylating agents**, such as cyclophosphamide, whose cytotoxic activity involves DNA and RNA cross-linking and inhibition of protein synthesis[29].
- **Monoclonal antibodies** (immunotherapy) can also be part of a treatment. For example, Daratumumab[30] is used as part of second line treatment and targets

CD38, which is overexpressed by at least a subset of MM cells, thus causing them to go into apoptosis[31].

- **Glucocorticoids**, such as dexamethasone or prednisone, are usually also included in the combination of drugs to treat MM as anti-inflammatory, although dexamethasone is also used due to its cytotoxicity on myeloma cells[32,33].

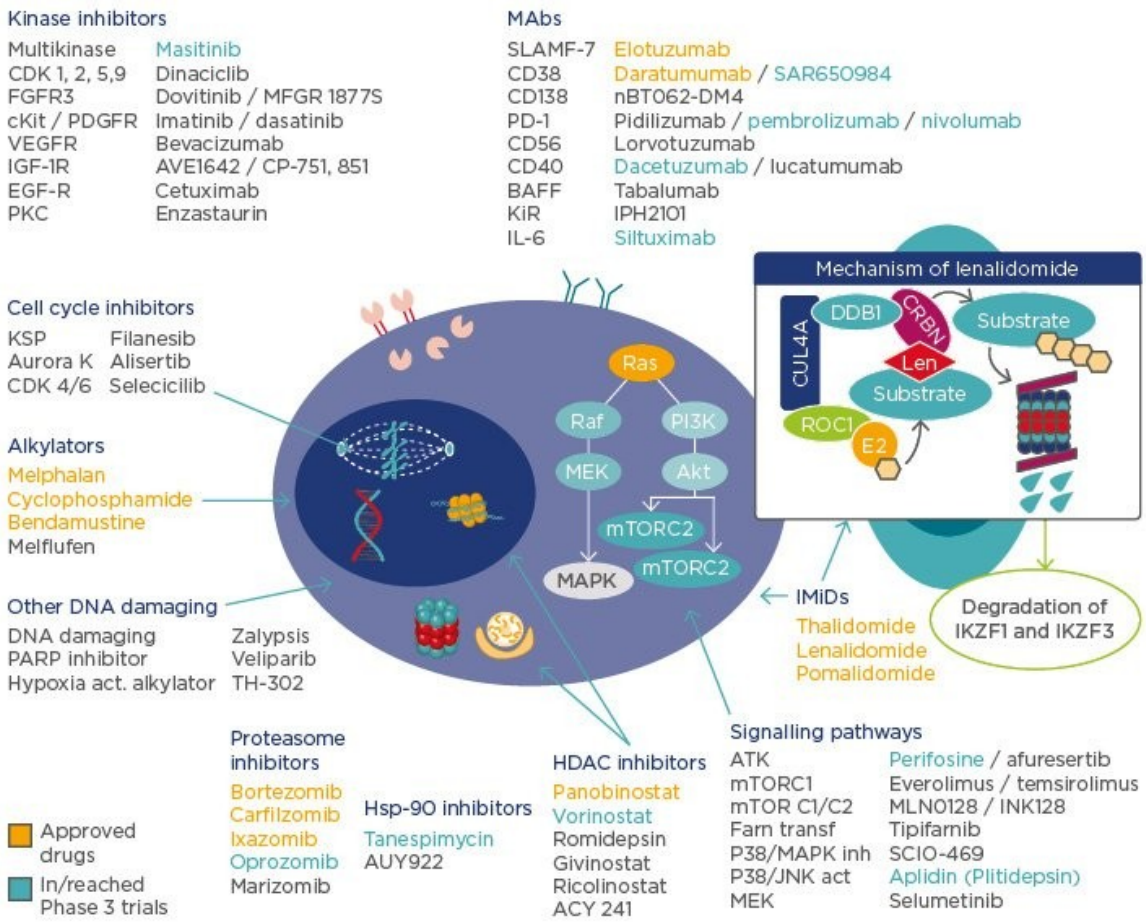


Figure 5: Compounds overview to treat MM.

IMiDs: immunomodulatory drugs; MAbs: monocolonal antibodies; PARP: poly A ribose polymerases; HDAC: histone deacetylase; Hsp-90: heat shock protein 90; IL: interleukin; FGFR3: fibroblast growth factor receptor 3; PDGFR: platelet-derived growth factor receptor; VEGFR: vascular endothelial growth factor receptor; IGF: insulin-like growth factor; EGF: epidermal growth factor; PD-1: programmed cell death protein 1; BAFF: B cell activating factor; KSP: kinesin spindle protein; MAPK: mitogen-activated protein kinase; MTORC: mammalian target of rapamycin complex. Adapted from [34].

Although the landscape to treat NDMM or RRMM is evolving relatively often, and once the possibility of a clinical trial is discarded, patients would normally undergo a bone marrow transplant as a first line of treatment, provided they are eligible (Figure 6). This procedure consists of destroying the patient's bone marrow cells with high-dose chemotherapy to eliminate myeloma and bone marrow cells[35]. The next step would consist of restoring bone marrow stem cells with cells from the patient that were collected before chemotherapy (autologus transplant[36]), or from a healthy and compatible donor (allogeneic transplant[37]). After considering this step, a specific

combination of the above mentioned drugs would usually follow depending on the case to maximize the efficacy of the treatment (Figure 6).

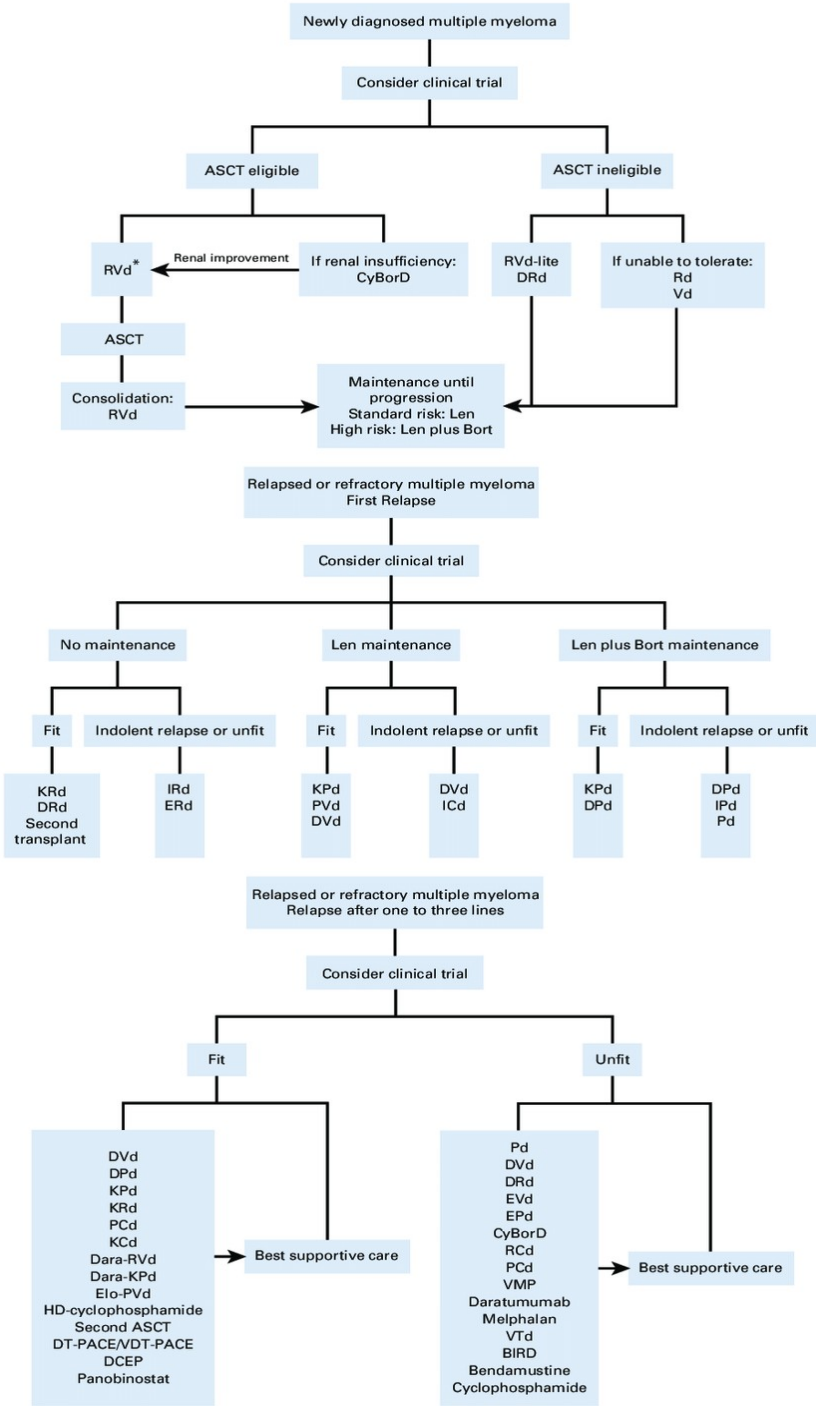


Figure 6: Treatment algorithm for NDMM and RRMM.

(*) KRd under investigation; four drug combinations under investigation (eg, Dara-IRd). ASCT, autologous stem-cell transplantation; BIRD, clarithromycin-lenalidomide-dexamethasone; Bort, bortezomib; CyBorD, cyclophophosphamide-bortezomib-dexamethasone; Dara-IRd, daratumumab-ixazomib-lenalidomide-dexamethasone; Dara-KPd, daratumumab-carfilzomib-pomalidomide-dexamethasone; Dara-RVd, daratumumab-lenalidomide-bortezomib-dexamethasone; DCEP, dexamethasone-cyclophosphamide-etoposide-cisplatin; DPd, daratumumab-pomalidomide-dexamethasone; DRd, daratumumab-lenalidomide-dexamethasone; DT-PACE, dexamethasone-thalidomide-cisplatin-doxorubicin-cyclophosphamide-etoposide; DVd, daratumumab-bortezomib-

dexamethasone; Elo-PVd, elotuzumab-pomalidomide-bortezomib-dexamethasone; EPd, elotuzumab-pomalidomide-dexamethasone; ERd, elotuzumab-lenalidomide-dexamethasone; EVd, elotuzumab-bortezomib-dexamethasone; HD-cyclophosphamide, high-dose cyclophosphamide; ICd, ixazomib-cyclophosphamide-dexamethasone; IPd, ixazomib-pomalidomide-dexamethasone; IRd, ixazomib-lenalidomide-dexamethasone; KCd, carfilzomib-cyclophosphamide-dexamethasone; KPd, carfilzomib-pomalidomide-dexamethasone; KRd, carfilzomib-lenalidomide-dexamethasone; Len, lenalidomide; PCd, pomalidomide-cyclophosphamide-dexamethasone; Pd, pomalidomide-dexamethasone; PVd, pomalidomide-bortezomib-dexamethasone; RCd, lenalidomide-cyclophosphamide-dexamethasone; Rd, lenalidomide-dexamethasone; RVd, lenalidomide-bortezomib-dexamethasone; Vd, bortezomib-dexamethasone; VDT-PACE, bortezomib-dexamethasone-thalidomide-cisplatin-doxorubicin-cyclophosphamide-etoposide; VMP, bortezomib-melphalan-prednisone; VTd, bortezomib-thalidomide-dexamethasone[35].

Although the current plethora of treatments has greatly contributed to increasing survival time for MM patients, finding new therapeutics that can further improve the prognosis of MM is necessary since this disease still remains lethal. One reason that could contribute to explain such lethality even under treatment is the fact that myeloma cells can become resistant to therapeutics due to genetic/epigenetic alterations, abnormal drug transport/metabolism, or dysregulation of apoptosis among other mechanisms[38]. These changes as myeloma cells proliferate lead to heterogeneous sub populations (hence the term multiple myeloma), some of which can be unaffected by the patient's treatment and become the more predominant sub-type leading to RRMM[39,40].

2.1.5. Prognosis

The median age at death for MM is 75 years old[41], but the survival rates vary significantly worldwide due to age, staging, ethnicity, lifestyle, and disparities in access to health care for different countries[2]. In general, the outcome of MM is better the earlier, (regarding age[42] (Figure 7) and stage) it is detected. In any case, the great increase in survival time achieved in the last decades[42] (Figure 7) sheds hope on finding a definite cure for this disease in a not so distant future.

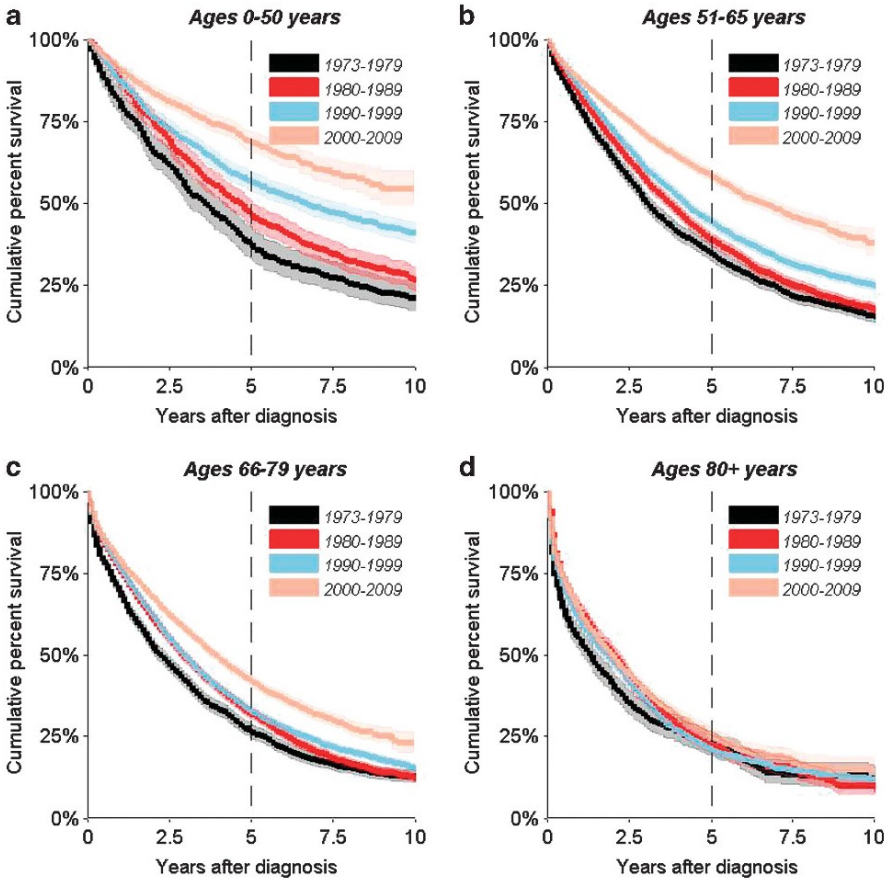


Figure 7: Evolution of MM prognosis.
Cumulative percent survival in the last decades by age[42].

2.2. Background on drug repurposing: A brief description

Drug repurposing consists of finding clinically available drugs with potential to treat diseases other than those they were initially approved for. The beginnings of drug repurposing consisted of using drugs on a new disease based on indirect empirical evidence, such as off target effects, or hypotheses about their potential to treat a different illness. However, rather than repurposing a drug at at time, recent advances in computational analyses now allow the processing of '*big data*' in a relatively short time, which can be used to generate lists of drug-target-interactions and thereby find therapeutics that could be novelly applied to a disease. Thus, *in silico* drug repurposing has contributed to a new avenue on the discovery of therapeutics that consists of shifting from the 'one drug -> one target -> one disease' approach to a 'several drugs -> several targets -> several disesases' paradigm[43,44]. This allows finding novel drug candidates without having to go through the '*de novo*' drug discovery process, thereby saving a considerable amount of time, human and financial resources, which makes it a desirable approach for both patients and pharmaceutical companies[43]. Depending on the data resources used in the analysis[45], *in silico* drug repurposing methods can be grouped into different categories, with docking-based, machine learning-based, and network-based methods being considered as major groups.

2.2.1. Docking-based methods

Molecular docking is a type of computational analysis that makes simulations and predictions about the best energetically and geometrically binding conformation between two or more molecules. Multiple analyses using different candidates thus give the possibility to assign a score to each docking and thereby obtain a final rank indicating the most suitable drug for a given molecule according to docking-related criteria[46]. Among other applications (Figure 8)[47], docking-based methods can be used to perform drug repurposing, where docking scores will suggest the most appropriate drug to be used for a protein of known 3D structure and that is suspected to be key in the development of the disease of interest. In this thesis, docking analyses will only be performed as a representative example of two potentially repurposable candidates to MM (ponatinib and axitinib) binding to a common target (KIT's kinase domain), and which one could be more suitable based on their respective affinity scores.

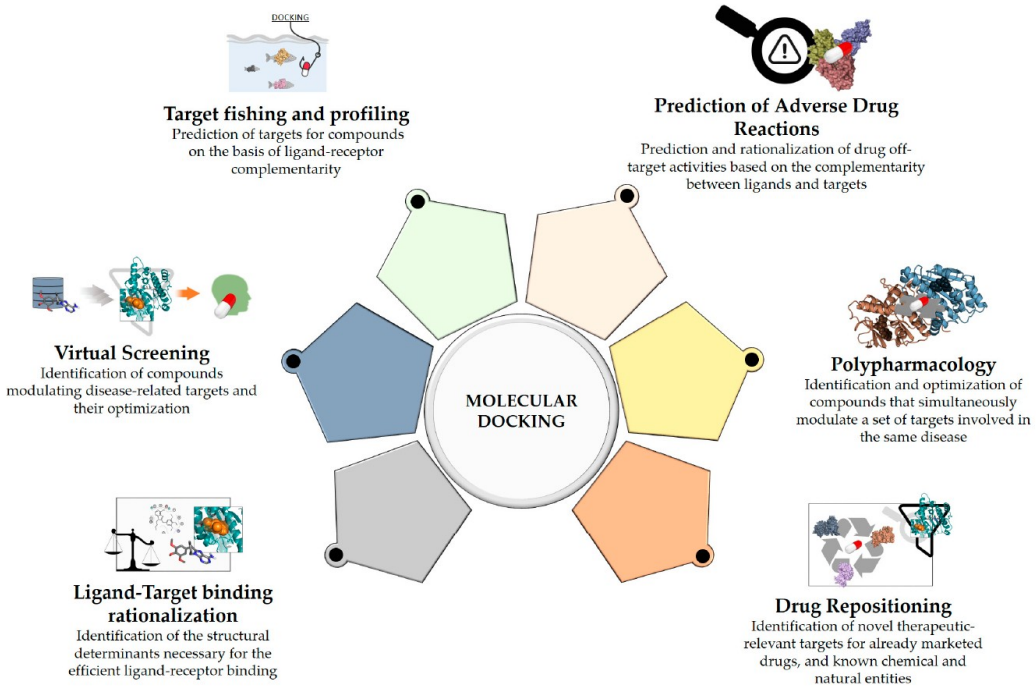


Figure 8: Applications of molecular docking include drug repositioning. Adapted from[47].

2.2.2. Machine learning-based methods

Machine learning (ML) consists of using computational algorithms and statistical models to find patterns in data that will allow making inferences and predictions. When it comes to drug repurposing, different types of data can be used to build ML models based on algorithms such as network propagation, matrix factorization, or deep learning, which have been used with relative success[45,48]. Suggested candidates are, as in other drug repurposing methods, further validated via *in vitro / in vivo* experimentation before using them on clinical trials[49] (Figure 9). ML models will generally be classified as supervised or unsupervised, with the former being characterized by using labeled data[45]

(i.e. with known values for the variable of interest) to build models, unlike the unsupervised model approach[50].

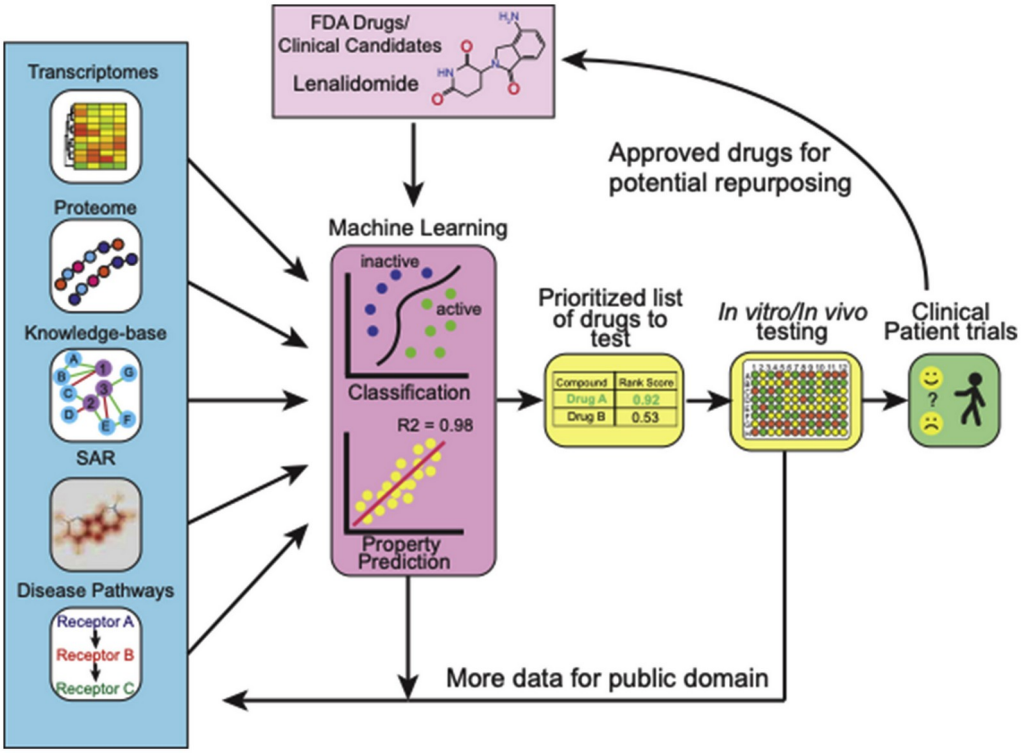


Figure 9: Overview of ML approaches for drug repurposing.
SAR: structure activity relationship[49].

2.2.3. Network-based methods

Different but yet complementary types of biological data (such as protein, metabolite, drug, or disease data) related to a given organism can be integrated into a network to model their interactions and gain information on their interconnection strength. Thus, molecular interaction networks, such as gene-protein interaction networks, metabolic networks, or protein-protein interaction networks, have been used for this purpose, and have lead to different types of network-based drug repurposing methods[51,52]. A recently published network-based algorithm known as SAveRUNNER (Searching off-lAbel dRUg aNd NEtwoRk), integrates drug-target interactions and disease-gene associations in the human interactome (the cellular network of all known physical molecular interactions) to generate a drug-disease network that suggests new repurposable drug candidates for diseases of interest. Based on premises of network medicine[51,53] that

1. Diseases are not usually caused by a single gene mutation, but rather the deregulation of a network of genes interconnected to each other.

2. The human interactome can be interpreted as a map, with diseases being local perturbations of it, and where genes associated with the a given disease therefore tend to aggregate nearby in the network, forming, “disease modules”.

3. Specific drugs can perturb common proteins and also act as local perturbations in the human interactome.

SAveRUNNER calculates proximity and similarity scores between disease and drug modules by using a similarity measure that gives priority to associations between drugs and diseases located nearby within the human interactome[5,6]. The material and methods section of this thesis contains a more thorough description on the different computational steps implemented by SAveRUNNER that lead to the generation of repurposable drug candidates (also summarized in Figure 10).

3. MATERIALS AND METHODS

3.1. Hardware

The computer (laptop) used to write this thesis and perform the necessary computations has the following specifications:

- Processor: Intel® Core™ i7-7700
- RAM memory: 32 GB DDR4
- Hard drive: SSD 512 GB

3.2. Software

The following tools will be used to generate the master's thesis:

- Windows 10 Home as operative system. (https://www.microsoft.com/en-gb/software-download/windows10)
- LibreOffice as text editor (https://www.libreoffice.org/discover/libreoffice/).
- Ganttproject to make the chronogram and ganttchart shown in this thesis. (https://www.ganttproject.biz/).
- Zotero for reference management (https://www.zotero.org/).
- R as programming language (https://www.r-project.org/).
- Rstudio as integrated development environment (IDE, https://www.rstudio.com/).
- SAveRUNNER for computational analyses of drug repurposing in R (https://github.com/giuliafiscon/SAveRUNNER.git).
- Different packages in R for microarray and RNAseq DEAs as well as generation of lists and Venn diagrams (provided as supplementary information in the corresponding reports).
- Connectivity Map (CMap) Query tool (https://clue.io/) to validate drug candidates generated by SAveRUNNER.

- BIOVIA Discovery Studio Visualizer (BDSV, https://discover.3ds.com/discovery-studio-visualizer-download) for graphical representation of protein-ligand interactions.
- Open Babel (http://openbabel.org/wiki/Main_Page) to convert files from .sdf format to .pdb format.
- AutoDock Vina and AutoDock Tools (https://autodock.scripps.edu/) for docking analyses.

3.3. Network-based drug repurposing algorithm: SAveRUNNER

3.3.1. Brief description

As described in the original articles[5,6], the SAveRUNNER algorithm hypothesizes that the drug-associated targets (drug module) and the disease-associated genes (disease module) should be nearby in the human interactome for a drug to be effective against a given disease. Briefly, it uses lists of drug-targets and disease-associated genes to create a network-based similarity measure in order to make predictions about drug-disease associations by performing the following steps:

1. Network proximity (*p*) computation by implementing the formula:

$$p(T,S) = \frac{1}{||T||}\sum_{t\in T}^{p} \min_{s\in S} d(t,s) \qquad \text{(Equation 3.1)}$$

where *p* represents the average shortest path length between drug targets *t* in the drug module *T* and the closest disease genes *s* in the disease module *S*[51].

2. The network proximity values are then z-score normalized, considering as proximal (significant) in this study only drug-disease associations with normalized z-score ≤ 1.65, which are used for further computations.

3. Translation of the z-score normalized network proximity measure into the network similarity measure within the range [0-1]:

$$similarity = \frac{max(p) - p}{max(p)} \qquad \text{(Equation 3.2)}$$

The greater the similarity measure the closer a drug and a disease module will be within the human interactome since the network proximity between them (*p*) will be smaller.

4. Cluster detection, using an algorithm based on greedy optimization of network modularity54 that groups drugs and diseases upon their similarity. To evaluate the quality of the clusters, SAveRUNNER computes a quality cluster (QC) score:

$$QC = \frac{W_{in}}{W_{in} + W_{out} + P} \qquad \text{(Equation 3.3)}$$

with W_{in} denoting the total weight of edges within a cluster, W_{out} denoting total weight of edges connecting different clusters, and P being the fraction of nodes within each cluster, which penalizes too large and not well define clusters.

5. Network similarity adjustment, by increasing similarity values for drug-disease associations by a factor proportional to QC:

$$similarity = (1 + QC) \cdot similarity \qquad \text{(Equation 3.4)}$$

Associations that fall within the same cluster will thus have a greater increase of their similarity values, highlighting the suitability of repurposing the corresponding drug for the respective disease.

6. Network similarity normalization by applying the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-c(x-d)}} \qquad \text{(Equation 3.5)}$$

where *x* represents adjusted similarity (Equation 3.4), *d* the sigmoid midpoint (max(x)/2), and *c* is the sigmoid steepness (set as 10).

Once completed the above steps (summarized in Figure 10, and more comprehensively described in the user guide at https://github.com/sportingCode/SAveRUNNER), SAveRUNNER generates a .txt file that contains a list of drug-disease associations as a weighted bipartite network, where nodes will correspond to either a disease or a drug. There will be an edge/link in the network between each disease and drug with a z-score proximity ≤ 1.65 ($p \leq 0.05$), with the corresponding normalized similarity value representing the weight of their interaction. SAveRUNNER will also generate additional files and folders, such as disease specific subnetwork, if specified in the config.R file. (see the section 3.3.3. on implementation below). This subnetwork also contains a .txt file consisting of a list with the associated drugs to a disease of interest, in the case of this study being MM. Given that the scope of this thesis is focused on finding drug candidates to be repurposed for MM, only the .txt file containing the list of drug candidates created for the MM subnetwork will be used to select drugs with statistically significant ($p \leq 0.05$) drug-MM association.

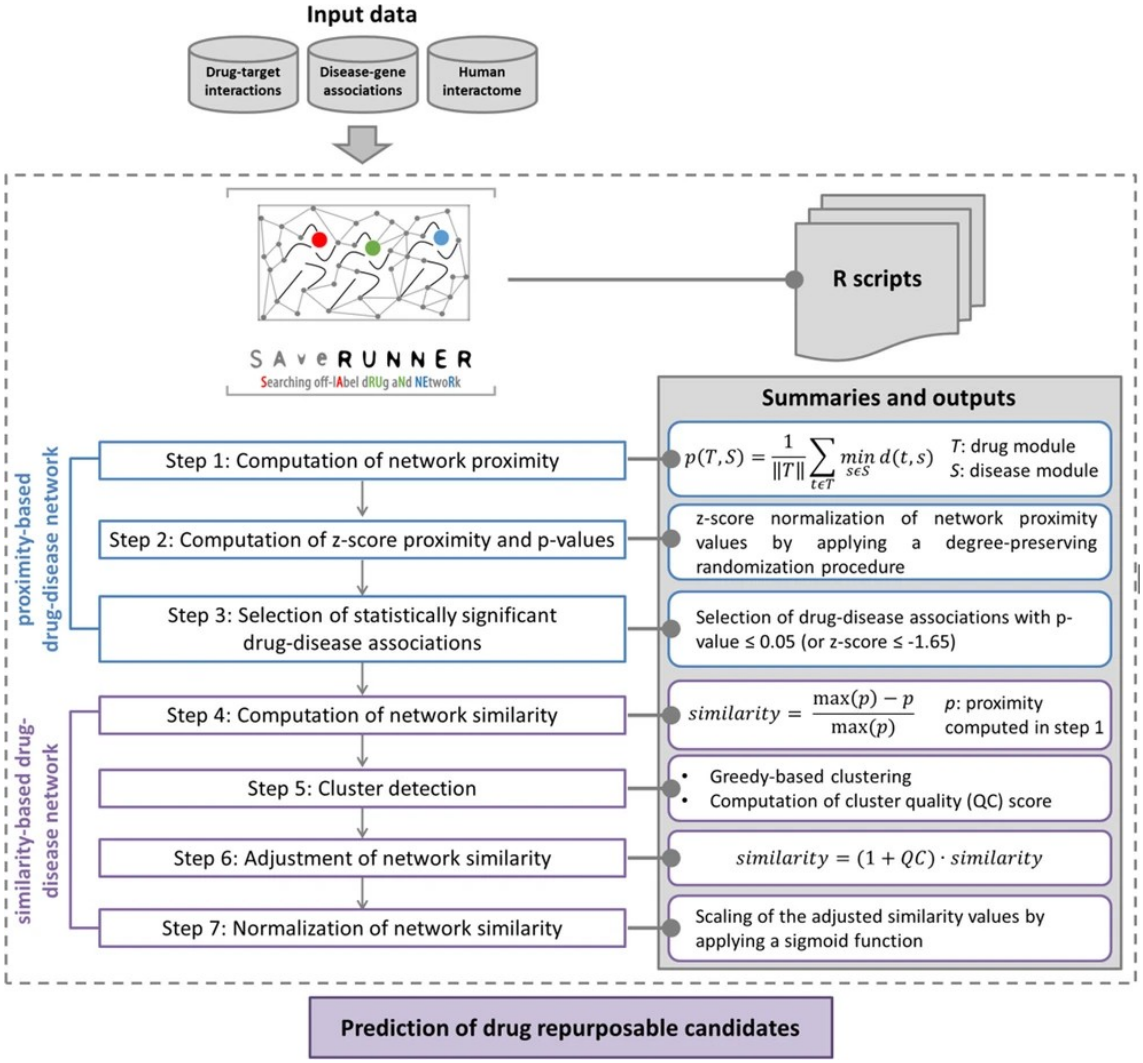


Figure 10: SAveRUNNER algorithm to generate lists of drug-disease associations. Adapted from[6].

3.3.2. Databases used by SAveRUNNER

All the databases used for the analyses performed in this thesis are included as part of the folder with the SAveRUNNER algorithm coded in R, and available as a .zip file in the GitHub repository specified in section 3.2. Briefly, SAveRUNNER comes with the following three databases

1. The disease-gene network was obtained from the Phenopedia database[7] which, as for 27-04-2020, contains gene associations for 3255 diseases. This database is currently part of the HuGe Navigator[55], and provides data about genes being linked to a given disease or phenotype

2. The drug-target database released from DrugBank[56] on 22-04-2020 consists of 13,563 compounds, of which 2627 are approved small molecule drugs, 1373 are approved biologics, 131 are nutraceuticals, and more than 6370 are experimental drugs. The Uniprot IDs for targets provided by DrugBank were mapped to Entrez gene IDs

with the BioMart – Ensembl tool (https://www.ensembl.org/). Drugs of interest without targets in DrugBank were integrated with drug-target interactions from the Therapeutic Target Database.

3. The human protein–protein interactome provided with SAveRUNNER consists of 217,160 protein–protein interactions connecting 15,970 unique proteins, and was obtained from Cheng and coauthors[51]. They assembled their own human protein–protein interactome and 15 commonly used databases containing several types of experimental evidence, such as binary PPIs from 3D protein structures; literature-curated PPIs identified by affinity purification followed by mass spectrometry; literature-derived PPIs from low-throughput experiments from BioGRID[57], HPRD[58], MINT, IntAct[59], or InnateDB[60]; signaling networks from literature-derived low-throughput experiments; and kinase-substrate interactions from literature-derived low-throughput and high-throughput experiments.

3.3.3. Implementation of SAveRUNNER

The user guide for SAveRUNNER, which can be downloaded at https://github.com/sportingCode/SAveRUNNER , explains in detail the steps to follow to implement the algorithm. Briefly, provided the working directory has been specified in the main.R file, and since SAveRUNNER includes by default the necessary drug-target, disease-gene, and human interactome lists as input files to perform the analysis, it is the file config.R where the user needs to configure the different parameters, such as disesases of interest to build the drug-disease network, p-value threshold, type to interaction to consider (proximity or similarity), whether or not adjust similarity, or the subnetwork with drug-disease entries for one of the specified diseases (MM in this case) among others (Figure 11). The subnetwork for MM will be created for this thesis, from where the best drug candidates for MM according to adjusted similarity values will be selected.

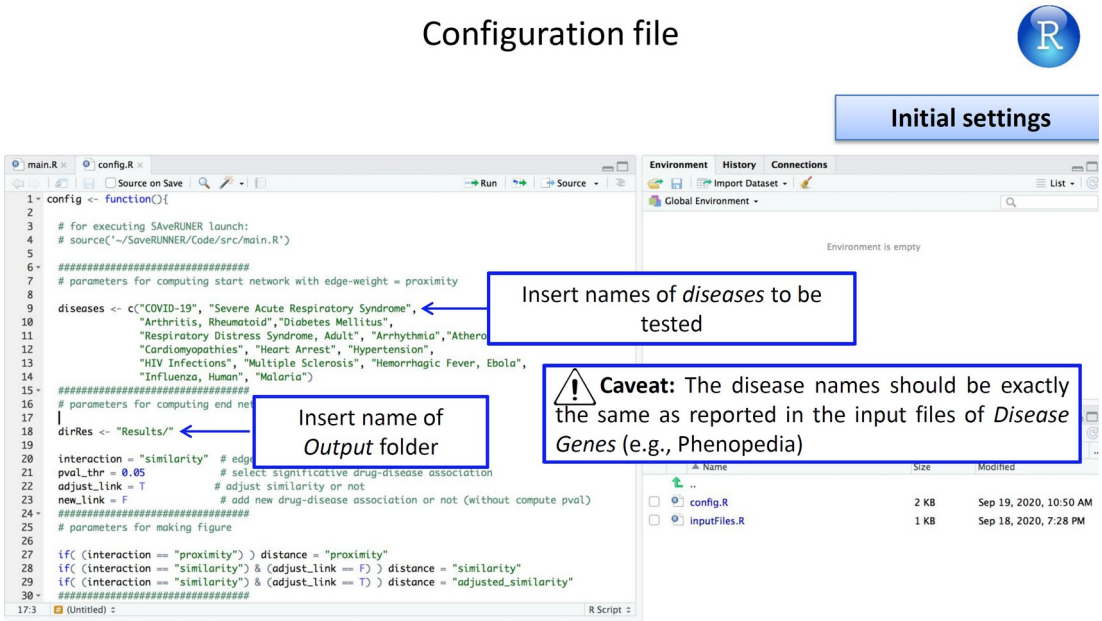


Figure 11: User guide for SAveRUNNER.
Representative slide explaining how to configure some parameters in the config.R file[61].

Once all the parameters are configured, SAveRUNNER can be launched by executing the file main.R.

As mentioned in the user guide, the creation of the main drug-disease network and creation of the corresponding file by SAveRUNNER is computationally intensive, so the diseases included for the different planned implementations may vary depending on their duration in order to fit the deadline of delivery for the corresponding PEC.

3.4. Differential expression analyses (DEAs) in MM samples

A dataset with microarray data as well as another one with RNAseq data from the Gene Expression Omnibus (GEO) will be used to perform DEA in R. The pipelines to be used in the respective analyses were provided to students of this master's degree in the class 'Omics data analyses', and are available at https://github.com/ASPteaching/Omics_Data_Analysis-Case_Study_1-Microarrays for microarray data, and https://github.com/ASPteaching/Omics_data_analysis-Case_study_2-RNA-seq for RNAseq data. Whereas lists with top DEGs in each dataset will be presented on this thesis, the corresponding full lists with DEGs as well as the respective DEAs reports will be available as supplementary information in github. At https://github.com/appropiate/TFM_UoC.

3.5. CMap query of DEGs in MM samples

DEGs for each of the MM datasets will be uploaded to the Query tool of CMap (https://clue.io/), which will yield lists containing different types of information for all the available compounds on their database (Touchstone). For this study, a relevant piece of information is how similarly/dissimilarly these compounds regulate the uploaded DEGs, which is shown through a parameter known as normalized connectivity score (ncs). Only compounds with negative ncs will be considered as they will tend to regulate the uploaded DEGs in the opposite manner as they were presented in the Query tool, i.e. they will theoretically counteract the regulation of at least some of those DEGs and thus treat MM. As indicated by the CMap Query tool tutorial (connectopedia)[62], to increase reliability of selected compounds to counteract uploaded DEGs, such compounds will need to have a negative ncs as well as values beyond a certain threshold for other parameters, such as significant adjusted p-value (fold discovery rate or fdr) or signature signal strength among others. In this work, the following parameters and threshold values will be used for the selection of compounds generated by the CMap Query tool:

- **ncs < 0** : Regulation of DEGs in opposite manner as uploaded in CMap Query tool.
- **pert_type = trt_cp**: Filtering of CMap list to show only experiments corresponding to cells treated with compounds.
- **fdr_q_nlog10 > 1**: Adjusted p-value < 0.1

- **cc_q75 > 0.2**: Replicate Correlation Coefficient, the higher the value, the more consistent the response induced by the compound.
- **ss_ngene > 200**: Signature Strength, representing the number of landmark genes with absolute z-score ≥ 2.

Threshold values for cc_q75 and ss_ngene are set so that only compounds with a strong and reproducible transcriptional activity, i.e. a high transcriptional activity score (TAS)[63], are selected (Figure 12).

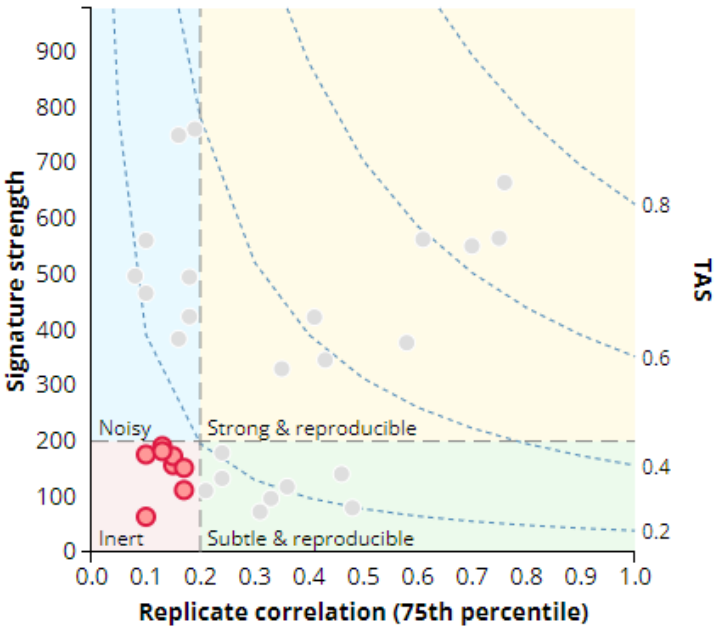


Figure 12: Transcription activity score (TAS).
TAS for penicilin and tetracycline in different cell lines (red dots) is low since these compounds are antibacterial drugs. Adapted from CMap Website[63].

Only compounds that meet the above criteria will be selected in order to validate drug candidates generated by SAveRUNNER.

3.6. Molecular Docking analyses

The 3D structure for the KIT kinase domain in a complex with ponatinib and ID 4u0i was downloaded in .pdb format from the protein data bank (PDB, https://www.rcsb.org/), whereas the compounds ponatinib and axitinib were downloaded in 3D-SDF format from drugbank (https://go.drugbank.com/structures/search/small_molecule_drugs/structure), and converted to .pdb format using Open Babel[64]. Protein and ligand preparation was performed with AutoDock Tools[65] or BDSV, 3D structures and interactions were visualized with BDSV, and docking analyses were performed using AutoDock Vina (see section 4.4.2 for a more detailed explanation and see the followed workflow).

4. RESULTS

4.1. Generation of drug candidates with SAveRUNNER

As earlier mentioned, in order to obtain drugs to be repurposed to MM using SAveRUNNER, MM together with several other diseases need to be selected from the Phenopedia database and included in the config.R file of SAveRUNNER to generate drug candidates with potential for repositioning on MM (Figure 11). To do so, and as mentioned in section 1.2.2, two different criteria have been followed in this thesis, which yielded two different lists of drug candidates with potential to be repurposed for MM. Only compounds included in both lists will be considered for further validation. A third implementation was also used as a negative control for the reliability of this method to generate specific candidates for to be repurposed to MM, as presented below.

4.1.1. Blood lineage diseases-based network

Since MM is a type of cancer affecting a cell type from the blood line (plasma cell), as a first criteria diseases related to blood cells were extracted in R from the Disease_gene database by Phenopedia (Figure 13), and included as references in SAveRUNNER's configuration in order to generate potential repurposable candidates.

```
>   #Filter diseases to use in the analysis:
>   disease_genes <- read.csv('C:/Users/Owner/Google Drive/Master Bioinformatica UoC/TFM/SAveRUNNER-main/code/input_files/Phenopedia.txt', header = T, sep = '\t')
>   disease_genes <- as.data.frame(unique(disease_genes$disease))
>   disease_genes <- disease_genes[grep('Lymphoma|kemia|yelo',disease_genes[,1]),] # Keywords to obtain desired diseases
> disease_genes
 [1] "Burkitt Lymphoma"                              "Leukemia"
 [3] "Leukemia, B-Cell"                              "Leukemia, B-Cell, Acute"
 [5] "Leukemia, B-Cell, Chronic"                     "Leukemia, Biphenotypic, Acute"
 [7] "Leukemia, Erythroblastic, Acute"               "Leukemia, Experimental"
 [9] "Leukemia, Hairy Cell"                          "Leukemia, Large Granular Lymphocytic"
[11] "Leukemia, Lymphocytic"                         "Leukemia, Lymphocytic, Acute"
[13] "Leukemia, Lymphocytic, Acute, L1"              "Leukemia, Lymphocytic, Acute, L2"
[15] "Leukemia, Lymphocytic, Chronic"                "Leukemia, Mast-Cell"
[17] "Leukemia, Megakaryocytic, Acute"               "Leukemia, Monocytic, Acute"
[19] "Leukemia, Myelocytic, Acute"                   "Leukemia, Myeloid"
[21] "Leukemia, Myeloid, Chronic"                    "Leukemia, Myeloid, Chronic-Phase"
[23] "Leukemia, Myelomonocytic, Acute"               "Leukemia, Myelomonocytic, Chronic"
[25] "Leukemia, Myelomonocytic, Juvenile"            "Leukemia, Neutrophilic, Chronic"
[27] "Leukemia, Nonlymphocytic, Acute"               "Leukemia, Plasma Cell"
[29] "Leukemia, Pre-B-Cell"                          "Leukemia, Prolymphocytic"
[31] "Leukemia, Promyelocytic, Acute"                "Leukemia, Radiation-Induced"
[33] "Leukemia, T-Cell, Acute"                       "Lymphoma"
[35] "Lymphoma, AIDS-Related"                        "Lymphoma, B-Cell"
[37] "Lymphoma, Extranodal NK-T-Cell"                "Lymphoma, Follicular"
[39] "Lymphoma, Large-Cell"                          "Lymphoma, Large-Cell, Diffuse"
[41] "Lymphoma, Large-Cell, Immunoblastic"           "Lymphoma, Large-Cell, Ki-1"
[43] "Lymphoma, Lymphoblastic"                       "Lymphoma, Mantle-Cell"
[45] "Lymphoma, Mucosa-Associated Lymphoid Tissue"   "Lymphoma, Non-Hodgkin"
[47] "Lymphoma, T-Cell"                              "Lymphoma, T-Cell, Cutaneous"
[49] "Lymphoma, T-Cell, Peripheral"                  "Lymphomatoid Granulomatosis"
[51] "Meningomyelocele"                              "Multiple Myeloma"
[53] "Myelodysplastic Syndromes"                     "Myelofibrosis"
[55] "Myeloproliferative Disorders"                  "Precursor T-Cell Lymphoblastic Leukemia-Lymphoma"
[57] "Preleukemia"                                   "Primary Myelofibrosis"
[59] "Pyelonephritis"
>
```

Figure 13: Blood related diseases to generate the first drug-disease network in SAveRUNNER. Diseases were obtained from the Phenopedia database by using the keyword 'Lymphoma' and the partial keywords 'kemia' and 'yelo' to include all possible leukemias and myeloma/myeloid related diseases, respectively.

The implementation of SAveRUNNER with these diseases as reference yielded a list of 360 candidates, among which most of the typical drugs used for MM treatment, such as lenalinomide[66], dexamethasone[67] or bortezomid[67], were included (Figure 14, Table 4, and Supplementary information 1)

```
> # SAveRUNNER compounds for MM using blood related diseases:
> mm_subnetwork <- read.csv('C:/Users/Owner/Google Drive/Master Bioinformatica UoC/TFM/SAveRUNNER-main/code/Results2/subnetwork/MultipleMyeloma/txtFile/Drug_Disease_network.txt', header = T, sep = '\t')
>
> # Order alphabetically
> mm_subnetwork <- mm_subnetwork[order(mm_subnetwork$adjusted_similarity,decreasing = T),]
>
> # First candidates in the list:
> head(mm_subnetwork)
            disease            drug proximity        pval similarity adjusted_similarity
256 Multiple Myeloma     cantharidin         0 0.003356958          1            0.960023
257 Multiple Myeloma       denosumab         0 0.003497314          1            0.960023
258 Multiple Myeloma     ibandronate         0 0.002037600          1            0.960023
259 Multiple Myeloma       lepirudin         0 0.014299424          1            0.960023
260 Multiple Myeloma pamidronic acid         0 0.001863521          1            0.960023
261 Multiple Myeloma      proflavine         0 0.001317948          1            0.960023
>
> # Number of drug candidates:
> length(mm_subnetwork$drug)
[1] 360
>
>
> all (c('bortezomib', 'cyclophosphamide','carfilzomib','dexamethasone','lenalidomide') %in% mm_subnetwork$drug)
[1] TRUE
>
```

Figure 14: Extract of drug candidates generated by the first implementation of SAveRUNNER. 360 compounds, such as bortezomib or lenalidomide, were generated by SAveRUNNER when using blood -related diseases to generate the disease-drug network. The full list of candidates can be found as supplementary information 1.

Disease	Drug	Proximity	P val	Similarity	Adj. similarity
Multiple Myeloma	cantharidin	0	0.0033570	1	0.9600230
Multiple Myeloma	denosumab	0	0.0034973	1	0.9600230
Multiple Myeloma	ibandronate	0	0.0020376	1	0.9600230
Multiple Myeloma	lepirudin	0	0.0142994	1	0.9600230
Multiple Myeloma	pamidronic acid	0	0.0018635	1	0.9600230
Multiple Myeloma	proflavine	0	0.0013179	1	0.9600230
Multiple Myeloma	risedronic acid	0	0.0022835	1	0.9600230
Multiple Myeloma	romosozumab	0	0.0000004	1	0.9600230
Multiple Myeloma	ximelagatran	0	0.0066476	1	0.9600230
Multiple Myeloma	abiraterone	0	0.0026319	1	0.8884007
Multiple Myeloma	adalimumab	0	0.0174099	1	0.8884007
Multiple Myeloma	alclofenac	0	0.0034104	1	0.8884007
Multiple Myeloma	anakinra	0	0.0086742	1	0.8884007
Multiple Myeloma	argatroban	0	0.0072982	1	0.8884007
Multiple Myeloma	axicabtagene ciloleucel	0	0.0033690	1	0.8884007
Multiple Myeloma	belimumab	0	0.0000861	1	0.8884007
Multiple Myeloma	bivalirudin	0	0.0043944	1	0.8884007
Multiple Myeloma	brentuximab vedotin	0	0.0000861	1	0.8884007
Multiple Myeloma	brolucizumab	0	0.0035184	1	0.8884007
Multiple Myeloma	calcifediol	0	0.0007785	1	0.8884007

Table 4: Top 20 drugs for MM generated by SAveRUNNER when blood related diseases were used to generate the drug-disease network.
Candidates ordered in terms of adjusted similarity. The full list of candidates can be found as supplementary information 1.

4.1.2. MM symptoms-based network

As a second analysis, diseases related to symptoms often experienced during MM were selected for a new implementation of SAveRUNNER (Figure 15).

```
config <- function(){

  # for executing SAveRUNER launch:
  # source("~/SAveRUNNER/code/src/main.R")

  ##################################
  # parameters for computing start network with edge-weight = proximity


    diseases <- c(
                  "Multiple Myeloma",
                  "Osteoporosis" ,
                  "Leukopenia",
                  "Anemia",
                  "Kidney Failure",
                  "Constipation",
                  "Dehydration",
                  "Hypercalcemia"
                  )
  ##################################
```

Figure 15: Diseases related to MM symptoms used for the second implementation of SAveRUNNER.
Diseases were included in the config.R file to generate the corresponding disease-drug network.

In this case, 354 drug candidates were generated, with many of them being also present in the previous list (Figure 16, Table 5, and supplementary information 2).

```
> # SAveRUNNER compounds for MM using MM-related symptoms:
> mm_subnetwork2 <- read.csv('C:/Users/Owner/Google Drive/Master Bioinformatica UoC/TFM/Supplementary inform
ation/SAveRUNNER/ResultsSymptoms/subnetwork/MultipleMyeloma/txtFile/Drug_Disease_network.txt', header = T, s
ep = '\t')
>
> # Order on adjusted similarity:
> mm_subnetwork2 <- mm_subnetwork2[order(mm_subnetwork2$adjusted_similarity,decreasing = T),]
>
>
> # First candidates in the list:
> head(mm_subnetwork2)
          disease                  drug proximity         pval similarity adjusted_similarity
4  Multiple Myeloma           abiraterone         0 4.065967e-03          1           0.9987696
11 Multiple Myeloma            alclofenac         0 6.402749e-03          1           0.9987696
20 Multiple Myeloma              anakinra         0 1.217169e-02          1           0.9987696
26 Multiple Myeloma             argatroban         0 9.366160e-03          1           0.9987696
31 Multiple Myeloma axicabtagene ciloleucel         0 8.154217e-03          1           0.9987696
38 Multiple Myeloma              belimumab         0 8.221664e-07          1           0.9987696
>
> # Number of drug candidates:
> length(mm_subnetwork2$drug)
[1] 354
>
> all(c('bortezomib','lenalidomide') %in% mm_subnetwork2$drug)
[1] TRUE
>
```

Figure 16: Extract of drug candidates generated by the second implementation of SAveRUNNER.
354 drug candidates to be repurposed in MM, such as bortezomib or lenalidomide, were generated by SAveRUNNER when using MM-related symptoms to generate the disease-drug network. The full list of candidates can be found as supplementary information 2.

Disease	Drug	Proximity	P val	Similarity	Adj. similarity
Multiple Myeloma	abiraterone	0	0.0040660	1	0.9987696
Multiple Myeloma	alclofenac	0	0.0064027	1	0.9987696
Multiple Myeloma	anakinra	0	0.0121717	1	0.9987696
Multiple Myeloma	argatroban	0	0.0093662	1	0.9987696
Multiple Myeloma	axicabtagene ciloleucel	0	0.0081542	1	0.9987696
Multiple Myeloma	belimumab	0	0.0000008	1	0.9987696
Multiple Myeloma	bivalirudin	0	0.0205429	1	0.9987696
Multiple Myeloma	brentuximab vedotin	0	0.0013691	1	0.9987696
Multiple Myeloma	brolucizumab	0	0.0034973	1	0.9987696
Multiple Myeloma	capecitabine	0	0.0012342	1	0.9987696
Multiple Myeloma	carbocisteine	0	0.0044159	1	0.9987696
Multiple Myeloma	cefazolin	0	0.0000000	1	0.9987696
Multiple Myeloma	cenegermin	0	0.0000974	1	0.9987696
Multiple Myeloma	cyclophosphamide	0	0.0035843	1	0.9987696
Multiple Myeloma	dabigatran etexilate	0	0.0077595	1	0.9987696
Multiple Myeloma	denosumab	0	0.0024347	1	0.9987696
Multiple Myeloma	econazole	0	0.0100851	1	0.9987696
Multiple Myeloma	emapalumab	0	0.0006167	1	0.9987696
Multiple Myeloma	encorafenib	0	0.0000481	1	0.9987696
Multiple Myeloma	enoxacin	0	0.0035254	1	0.9987696

Table 5: Top 20 drugs for MM generated by SAveRUNNER when diseases related to MM symptoms were used to generate the drug-disease network.
Candidates are ordered by adjusted similarity. The full list of candidates can be found as supplementary information 2.

4.1.3. Network without MM as negative control

A new implementation of SAveRUNNER as a negative control was executed with the aim of showing that the generation of drugs by SAveRUNNER seems disease specific. For this purpose, diseases with a similar number of genes associated as MM and that, at the same time, have as few genes in common with MM as possible were included in the config.R file of SAveRUNNER to run the implementation that generates the main drug-disease network. These diseases, ordered from less to more common genes with MM, were COVID-19 with 2 genes in common, obsessive compulsive disorder OCD with 12, panic disorder with 15, attention deficit disorder with hyperactivity with 23, ataxia with 26, and psychotic disorders with 34 genes in common with MM (Figure 17). Due to having the lowest number of genes in common with MM, COVID-19 was first selected to generate the corresponding specific drug-disease subnetwork, yielding only around 100 candidates vs approximately 360 generated for each of the implementations for MM (data not shown). That is why a drug-disease subnetwork for OCD was next generated (Figure 17), which yielded 736 candidates (Figure 18 and 19), of which 45 (approximately 6%) were also generated for the two specific subnetworks for MM with the corresponding SAveRUNNER implementations (Figure 19). Therefore, the subnetwork generated for OCD using MM unrelated diseases contains mostly diseases that were not part of the subnetworks generated for MM, which demonstrates that SAveRUNNER mainly generates disease-specific candidates with potential for reposition.

```
  ################################
  # parameters for computing start network with edge-weight = proximity

diseases <- c("COVID-19",
              "Obsessive-Compulsive Disorder",
              "Panic Disorder",
              "Attention Decifit Disorder with Hyperactivity",
              "Ataxia",
              "Psychotic Disorders"
              )
  ################################
  # parameters for computing end network

  dirRes <- "ResultsControl3/"


  interaction = "similarity"  # edge-weight = similarity or proximity
  pval_thr = 0.05             # select significative drug-disease association
  adjust_link = T            # adjust similarity or not
  new_link = F                # add new drug-disease association or not (without compute pval)
  ################################
  # parameters for making figure

  if( (interaction == "proximity") ) distance = "proximity"
  if( (interaction == "similarity") & (adjust_link == F) ) distance = "similarity"
  if( (interaction == "similarity") & (adjust_link == T) ) distance = "adjusted_similarity"
  ################################
  # parameters for computing subnetwork

  # sel_drug = "tocilizumab"
  # sel_disease = "Severe Acute Respiratory Syndrome"
  sel_drug = NULL
  sel_disease = "Obsessive-Compulsive Disorder"
  ################################
```

Figure 17: Negative control for SAveRUNNER implementation.
Diseases unrelated to MM were selected to generate a disease-drug network as well as drug-disease subnetwork with candidates to be repurposed for OCD.

```
> ### SAveRUNNER compounds for Obsessive-Compulsive Disorder and unrelated diseases to MM:
> mm_subnetwork7 <- read.csv('C:/Users/Owner/Google Drive/Master Bioinformatica UoC/TFM/SAveRUNNER-m
ain/code/ResultsControl3/subnetwork/Obsessive-CompulsiveDisorder/txtFile/Drug_Disease_network.txt',h
eader = T, sep = '\t')
>
> # Order on adjusted similarity:
> mm_subnetwork7 <- mm_subnetwork7[order(mm_subnetwork7$adjusted_similarity,decreasing = T),]
>
>
> # First candidates in the list:
> head(mm_subnetwork7)
                        disease        drug proximity         pval similarity adjusted_similarity
1  Obsessive-Compulsive Disorder    abarelix         0 5.233279e-08          1           0.9992862
3  Obsessive-Compulsive Disorder  acebutolol         0 8.118189e-04          1           0.9992862
9  Obsessive-Compulsive Disorder   adenosine         0 4.043644e-14          1           0.9992862
14 Obsessive-Compulsive Disorder almotriptan         0 1.487887e-07          1           0.9992862
15 Obsessive-Compulsive Disorder   alosetron         0 3.436059e-06          1           0.9992862
19 Obsessive-Compulsive Disorder   alvimopan         0 5.571210e-07          1           0.9992862
>
> # Number of drug candidates:
> length(mm_subnetwork7$drug)
[1] 736
>
```

Figure 18: Implementation of SAveRUNNER as negative control.
736 drug candidates for OCD obtained when using MM-unrelated diseases to generate the disease-drug subnetwork. The full list of candidates can be found as supplementary information 3.

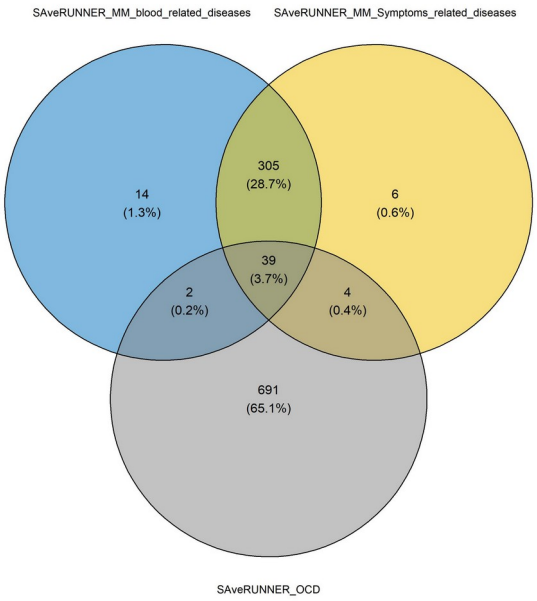


Figure 19: Venn diagram for compounds generated by SAveRUNNER implementations.
The first two analyses (supplementary information 1 and supplementary information 2, respectively) involved diseases somewhat related to MM and MM symptoms, and yielded similar lists of drug candidates, unlike the negative control implementation, where MM was not included in the configuration file to generate both the general drug-disease network and the subnetwork for OCD.

Thus, 344 compounds have been commonly generated by the first two implementations of SAveRUNNER, and will be further validated/filtered out to obtain a final list of candidates by using DEA/CMap analyses, outlined in the next section.

4.2. Validation of candidates via DEA/CMap analyses

The workflow followed in this study to validate candidates for MM generated by SAveRUNNER is described in the next subsections.

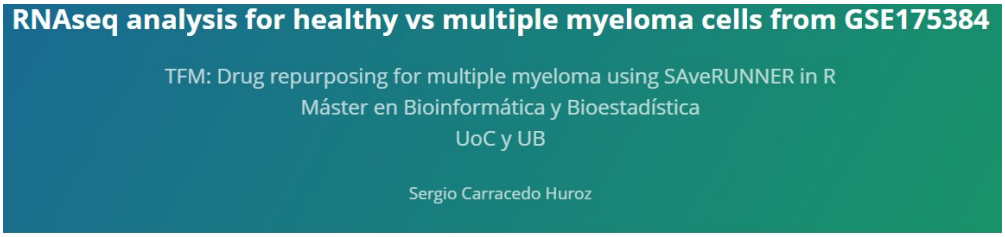
4.2.1. Differential expression analysis (DEA)

This step consists of:

1. **Collection of data** on MM vs healthy plasma cells from the gene expression omnibus (GEO) repository. RNAseq data under the accession number GSE175384 (read counts table for 41 healthy adults and 32 MM patients[68]) and mRNA microarray data under GSE47552 (5 Normal plasma cell samples and 41 clonal plasma cell samples[69]) were collected for this work.

2. **Implementation of DEA in R** to find DEGs for the GSE175384 RNAseq dataset (Figure 20) and GSE47552 microarray dataset (Figures 21), which will represent the gene signature or transcription profile for the respective datasets. For an adjusted p-value < 0.05 and log2 fold change >1, 4424 DEGs were found for GSE175384 (Table 6 and supplementary information 4 and 5) whereas 768 DEGs were obtained for

GSE47552 (Table 7 and supplementary information 6 and 7). The difference in the number of DEGs between RNAseq and microarray analyses is somewhat expected given the superiority of RNAseq technology to detect DEGs[70].



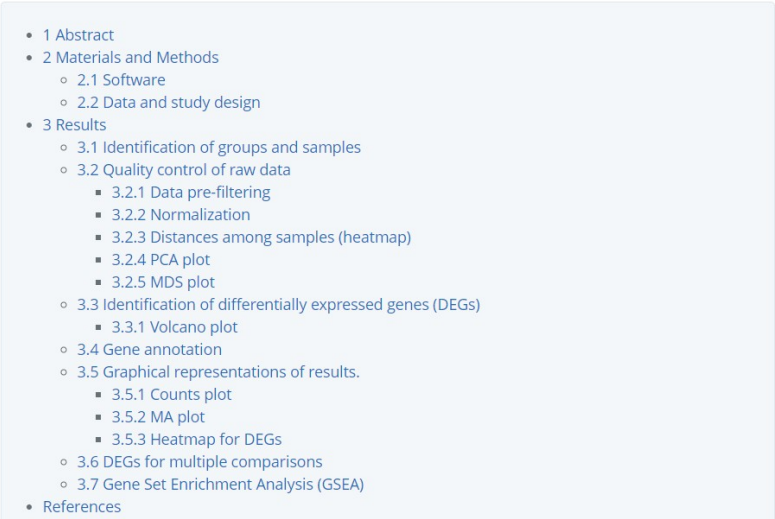


Figure 20: DEA for the RNAseq dataset GSE175384.
The full report and corresponding R code is available as suplementary information 4.

ENTREZID	SYMBOL	GENENAME	log2FoldChange	padj
28825	IGLV1-40	immunoglobulin lambda variable 1-40	-7.443295	0
28820	IGLV1-51	immunoglobulin lambda variable 1-51	-6.686168	0
28784	IGLV4-69	immunoglobulin lambda variable 4-69	-6.900870	0
28774	IGLV8-61	immunoglobulin lambda variable 8-61	-5.619471	0
81873	ARPC5L	actin related protein 2/3 complex subunit 5 like	1.929150	0
28775	IGLV7-46	immunoglobulin lambda variable 7-46 (gene/pseudogene)	-6.556131	0
28822	IGLV1-47	immunoglobulin lambda variable 1-47	-5.670269	0
28772	IGLV10-54	immunoglobulin lambda variable 10-54	-7.080012	0
28797	IGLV3-19	immunoglobulin lambda variable 3-19	-5.696330	0
28776	IGLV7-43	immunoglobulin lambda variable 7-43	-6.108259	0
5970	RELA	RELA proto-oncogene, NF-kB subunit	1.767951	0
53942	CNTN5	contactin 5	6.712467	0
28434	IGHV3-33	immunoglobulin heavy variable 3-33	-4.921755	0
3507	IGHM	immunoglobulin heavy constant mu	-6.412635	0
7866	IFRD2	interferon related developmental regulator 2	1.287167	0
939	CD27	CD27 molecule	-3.126071	0
28448	IGHV3-15	immunoglobulin heavy variable 3-15	-5.517794	0
28461	IGHV1-69	immunoglobulin heavy variable 1-69	-5.564210	0
219285	SAMD9L	sterile alpha motif domain containing 9 like	-3.907078	0
51028	VPS36	vacuolar protein sorting 36 homolog	-1.855504	0

Table 6: The 20 most significant DEGs in MM samples for the GSE175384 dataset.
The complete DEA and a full list of DEGs in MM can be found as supplementary information 4 and 5, respectively.

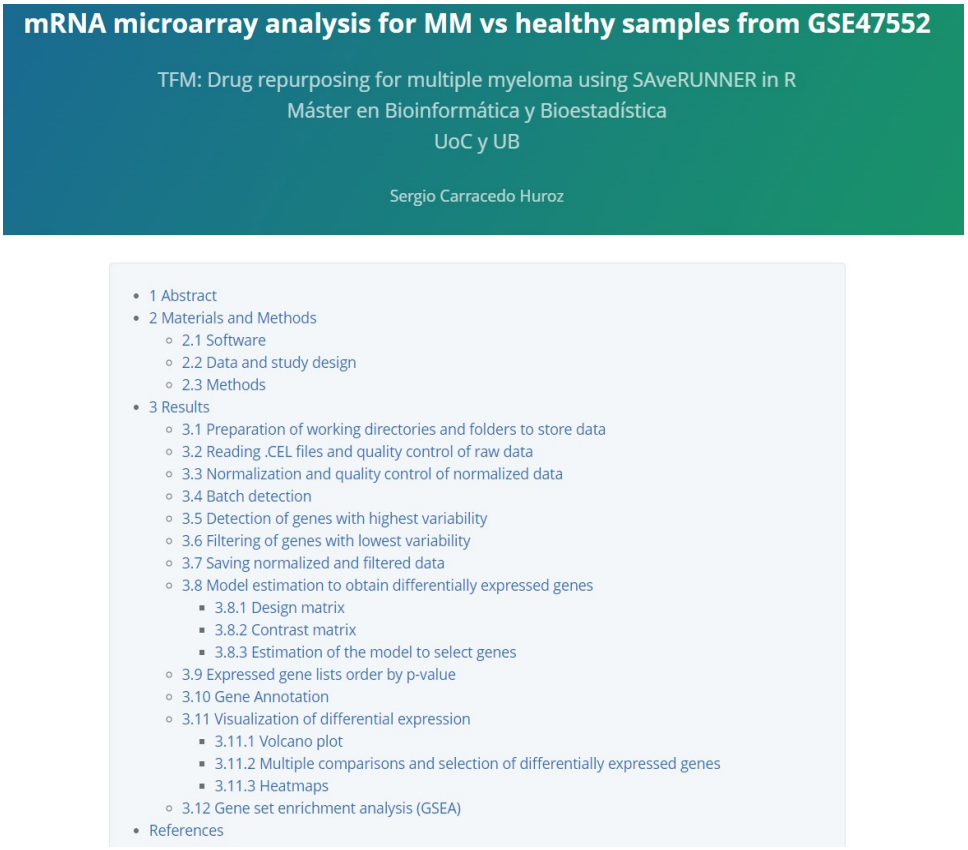


Figure 21: DEA for the mRNA microarray dataset GSE47552.
The full report and corresponding R code is available as suplementary information 6.

ENTREZID	SYMBOL	GENENAME	logFC	adj.P.Val
28904	IGKV1D-8	immunoglobulin kappa variable 1D-8	-3.968053	0e+00
55252	ASXL2	ASXL transcriptional regulator 2	-2.413374	0e+00
644714	LIMD1-AS1	LIMD1 antisense RNA 1	-2.777854	0e+00
91768	CABLES1	Cdk5 and Abl enzyme substrate 1	-1.991807	0e+00
11276	SYNRG	synergin gamma	-2.079914	0e+00
3738	KCNA3	potassium voltage-gated channel subfamily A member 3	-2.572281	2e-07
158747	MOSPD2	motile sperm domain containing 2	-2.420246	2e-07
167838	TXLNB	taxilin beta	-2.427357	2e-07
23595	ORC3	origin recognition complex subunit 3	-2.066035	2e-07
28923	IGKV2-24	immunoglobulin kappa variable 2-24	-5.035069	2e-07
9321	TRIP11	thyroid hormone receptor interactor 11	-2.504162	2e-07
344787	ZNF860	zinc finger protein 860	-2.080445	2e-07
84272	YIPF4	Yip1 domain family member 4	-1.569583	3e-07
55763	EXOC1	exocyst complex component 1	-1.716421	4e-07
56242	ZNF253	zinc finger protein 253	-1.743651	4e-07
6218	RPS17	ribosomal protein S17	1.336183	4e-07
3660	IRF2	interferon regulatory factor 2	-1.826822	6e-07
2820	GPD2	glycerol-3-phosphate dehydrogenase 2	-1.998799	7e-07
9847	C2CD5	C2 calcium dependent domain containing 5	-1.632152	8e-07
11196	SEC23IP	SEC23 interacting protein	-1.843214	9e-07

Table 7: The 20 most significant DEGs in MM samples for the GSE47552 dataset.
The complete DEA of the corresponding microarray data, and a full list of DEGs can be found as supplementary information 6 and 7, respectively.

4.2.2. Query of DEGs in Connectivity Map (CMap)

Drug candidates to be repurposed on a new disease should be able to counteract at least some of the disease's dysregulated genes (which would correspond here to MM DEGs) and pathways in order to treat such disease effectively. In this regard, one way to find the strongest drug candidates generated by SAveRUNNER is to select those that *oppositely* regulate the highest number of differentially DEGs in MM. For example, if only genes A, B and C were *upregulated* in MM, then a good candidate to be repurposed to MM would be expected to *downregulate* at least one of those genes. The best candidates would thus be those counteracting more MM DEGs. This is an *in silico* procedure used in previous studies to validate drugs that have been obtained via computational analyses[5]. In this study, the compounds in the lists generated by SAveRUNNER which meet this validation criteria in at least one of the datasets containing MM samples from the GEO repository (GSE175384 and GSE47552) will be considered as the strongest candidates for drug repurposing in MM.

4.2.2.1. Use of the Query tool from CMap

In order to implement the above described criteria to validate drug candidates, DEGs obtained for a given dataset of healthy vs MM samples (together considered as gene signature or transcription profile) would be uploaded to the query tool offered by CMap (https://clue.io/query). Once the set of genes of interest have been uploaded and the query has been executed, a table containing several parameters for compounds that belong to the CMap database and regulate different genes and pathways is generated[71]. One of these parameters is known as connectivity score (CS), which reflects how similarly/dissimilarly a compound regulates the set of genes the user has entered as input (DEGs for each MM dataset in this case). The more positive the CS is for a given compound, the more similarly the compound regulates the DEGs the user has provided, i.e the compound will cause a similar transcription profile upon treatment and will therefore tend to upregulate the majority of the genes that are upregulated in the uploaded DEG set, and downregulate the majority of genes that are dowregulated on the uploaded DEG set. In contrast, compounds with a negative CS would cause an opposite transcription profile upon treatment, i.e. they will tend to downregulate the majority of the genes that are upregulated in the uploaded DEG set and upregulate the majority of genes that are downregulated in the uploaded DEG set. Therefore, according to the goal of this thesis, the compounds of interests will be those with negative CS since they will tend to counteract the uploaded DEGs (dysregulated genes) in MM, i.e. they will have most potential to treat MM. A normalized version of CS (ncs) together with other parameters is actually a better criteria to use since it accounts for cell type and other conditions in which transcription profiles caused by these compounds were obtained by CMap[72]. Thus, for this thesis, the query tool in the CMap website was used two times by uploading DEGs for either GSE175384 (RNAseq) or GSE47552 (microarray) in order to find drugs with gene signatures that are most dissimilar to these sets of DEGs (Figure 22), i.e. we would be selecting compounds based on the criteria earlier specified in to come up with a set of drug candidates that would counteract the transcription profile for

each of the MM used in this thesis (Tables 8 and 9 and supplementary information 8 and 9).

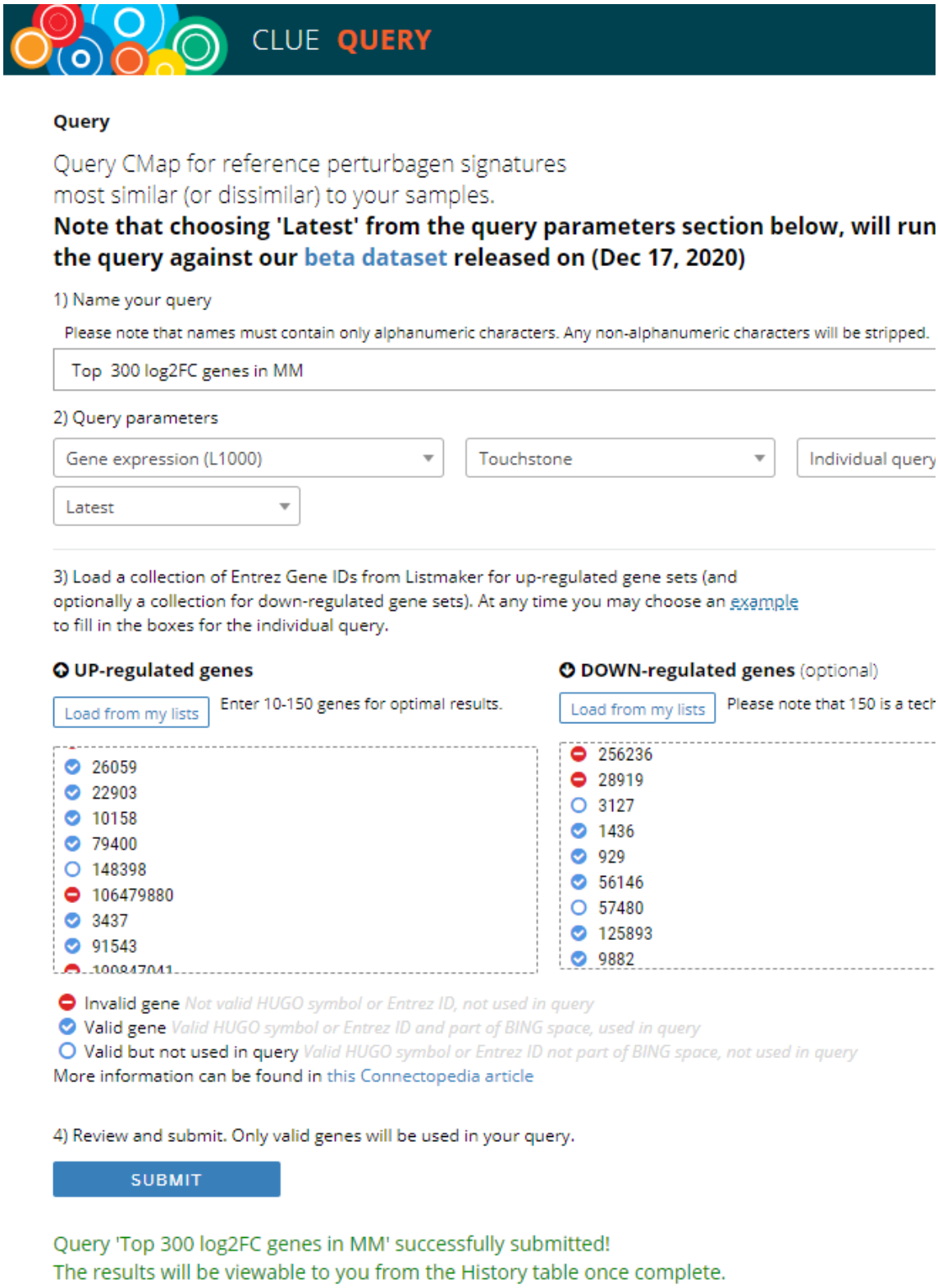


Figure 22: CMap Query tool.
Top upregulated and downregulated genes for each DEA (GSE175384 and GSES47552) were introduced in the GUI of the CMap query tool to generate the corresponding lists of compounds with their corresponding drug signatures and connectivity scores (supplementary information 8 and 9, respectively).

pert_iname	pert_type	ss_ngene	cc_q75	fdr_q_nlog10	norm_cs
BRD-K02581333	trt_cp	316	0.2300	15.6536	-1.5901
BRD-K79797751	trt_cp	237	0.3600	15.6536	-1.5751
HG-6-64-01	trt_cp	219	0.4500	15.6536	-1.5629
BRD-K20986251	trt_cp	209	0.3600	15.6536	-1.5044
AS-605240	trt_cp	278	0.5114	15.6536	-1.4929
BRD-A19037878	trt_cp	210	0.4900	15.6536	-1.4820
BRD-U43867373	trt_cp	280	0.3469	15.6536	-1.4769
BRD-K04465546	trt_cp	334	0.4700	15.6536	-1.4724
BRD-K58547240	trt_cp	210	0.2500	15.6536	-1.4722
MW-STK33-2A	trt_cp	209	0.4000	15.6536	-1.4710
triamcinolone	trt_cp	221	0.3400	15.6536	-1.4681
ML-179	trt_cp	309	0.3700	15.6536	-1.4548
BRD-K48225424	trt_cp	233	0.2200	15.6536	-1.4436
erlotinib	trt_cp	251	0.4200	15.6536	-1.4376
sparfloxacin	trt_cp	310	0.2576	15.6536	-1.4373
BRD-K64447917	trt_cp	252	0.2300	15.6536	-1.4276
GSK-269962	trt_cp	263	0.4400	15.6536	-1.4215
dexamethasone-acetate	trt_cp	264	0.5700	15.6536	-1.4107
diltiazem	trt_cp	203	0.3000	15.3525	-1.4059
BRD-K75999307	trt_cp	286	0.3600	15.3525	-1.4025

Table 8: Top 20 compounds with most dissimilar normalized connectivity scores (norm_cs) for DEGs in GSE175384.
The full list is available as supplementary information 8.

pert_iname	pert_type	ss_ngene	cc_q75	fdr_q_nlog10	norm_cs
alvocidib	trt_cp	396	0.6500	15.6536	-1.8626
BRD-K67778494	trt_cp	212	0.2278	15.6536	-1.8146
BRD-K20718732	trt_cp	314	0.2500	15.6536	-1.7691
dactinomycin	trt_cp	664	0.5500	15.6536	-1.7634
BRD-K31108633	trt_cp	278	0.2700	15.6536	-1.7463
AZD-5438	trt_cp	249	0.3049	15.6536	-1.7370
WH-4023	trt_cp	268	0.4300	15.6536	-1.7203
radicicol	trt_cp	304	0.5300	15.6536	-1.7182
BRD-K64062072	trt_cp	246	0.3200	15.6536	-1.7115
GS-9973	trt_cp	248	0.5300	15.6536	-1.7086
fludarabine	trt_cp	519	0.6000	15.6536	-1.6926
BRD-K96072942	trt_cp	415	0.3500	15.6536	-1.6826
rebastinib	trt_cp	210	0.4741	15.6536	-1.6754
BRD-K16826857	trt_cp	598	0.5900	15.6536	-1.6591
BRD-K34974324	trt_cp	300	0.3200	15.6536	-1.6460
ibrutinib	trt_cp	313	0.5175	15.6536	-1.6412
BRD-K27161987	trt_cp	322	0.3100	15.6536	-1.6367
BRD-K17722419	trt_cp	676	0.7800	15.6536	-1.6354
dicyclohexylurea	trt_cp	357	0.4200	15.6536	-1.6338
dantron	trt_cp	232	0.2400	15.6536	-1.6284

Table 9: The 20 compounds with most dissimilar normalized connectivity scores (norm_cs) for DEGs in GSE47552.
The full list is available as supplementary information 9.

4.2.3. Intersection of compounds lists found via CMap and SAveRUNNER

Drugs present in all the lists earlier presented, i.e. compounds generated by both SAveRUNNER implementations for MM and further validated by both DEAs/CMap analyses, would be considered as the strongest candidates for drug repurposing to MM in this study. Seven drugs met this criteria (Figure 23 and Table 10), of which only the antibiotic sparfloxacin could be considered as candidate for drug repurposing in MM because the rest of these validated drugs have already been used in at least one published study aiming at treating MM (Supplementary information 10).

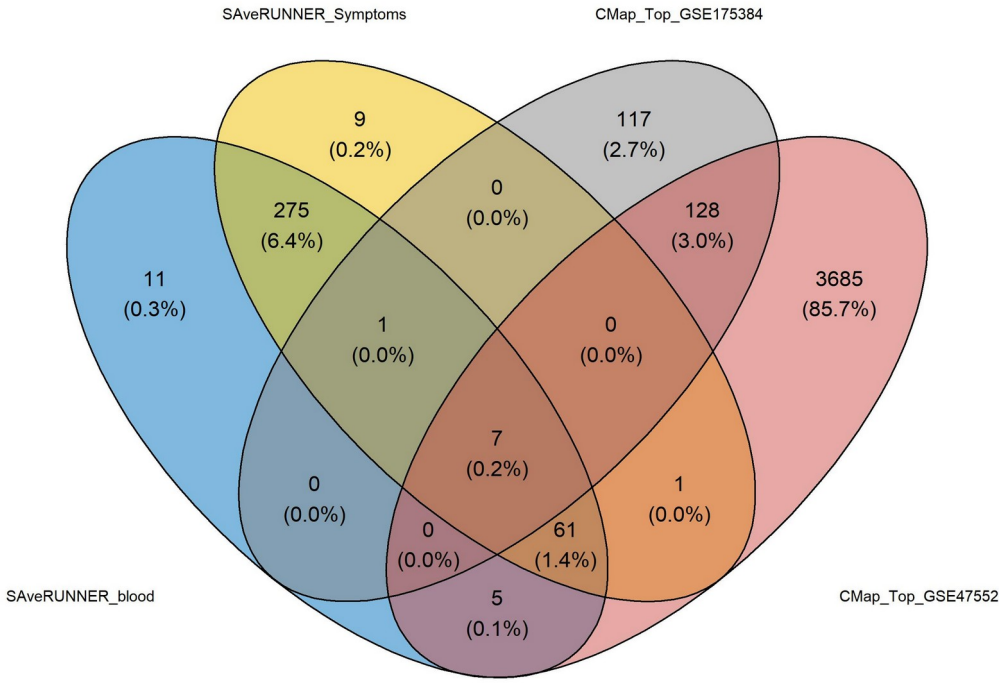


Figure 23: Venn diagram for compounds obtained with SAveRUNNER and CMap. DEGs for GSE175384 and GSE47552 datasets were used in the CMap Query tool to obtain drug signatures with ncs values.

Disease	Drug	Adj. similarity_A	Adj. similarity_B	mean_adj_similarity
Multiple Myeloma	mitoxantrone	0.8884007	0.9987696	0.9435852
Multiple Myeloma	sparfloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	lenalidomide	0.8284254	0.9969081	0.9126668
Multiple Myeloma	erlotinib	0.8884007	0.8728060	0.8806034
Multiple Myeloma	etoposide	0.7454534	0.9922521	0.8688527
Multiple Myeloma	sunitinib	0.6398028	0.5127472	0.5762750
Multiple Myeloma	dexamethasone	0.6164503	0.4815084	0.5489794

Table 10: Totally validated candidates generated by SAveRUNNER. Drugs resulting from the intersection of lists generated by the two SAveRUNNER implementations and the two DEG/CMap analyses using the GSE175384 and GSE47552 datasets.

In order to increase the number of candidates generated by SAveRUNNER that could be considered for repurposing in MM, a validation by one DEA/CMap analysis, i.e. partial validation, was also performed. In this case, 62 compounds were validated (Figure 23, Table 11, and supplementary information 11). As in the case of totally

validated drugs, partially validated candidates included several compounds used in MM treatments, such as polimalidomide[67], daunorubicin[73], or ifosfamide[74] (Table 11 and supplementary information 11).

Disease	Drug	Adj. similarity_A	Adj. similarity_B	mean_adj_similarity
Multiple Myeloma	argatroban	0.8884007	0.9987696	0.9435852
Multiple Myeloma	cyclophosphamide	0.8884007	0.9987696	0.9435852
Multiple Myeloma	etoricoxib	0.8884007	0.9987696	0.9435852
Multiple Myeloma	floxuridine	0.8884007	0.9987696	0.9435852
Multiple Myeloma	idarubicin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	ifosfamide	0.8884007	0.9987696	0.9435852
Multiple Myeloma	levofloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	moxifloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	norfloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	rilpivirine	0.8884007	0.9987696	0.9435852
Multiple Myeloma	trifluridine	0.8884007	0.9987696	0.9435852
Multiple Myeloma	trovafloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	valrubicin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	bortezomib	0.8877775	0.9922521	0.9400148
Multiple Myeloma	mesalazine	0.8180334	0.9964737	0.9072536
Multiple Myeloma	flutamide	0.8034253	0.9957985	0.8996119
Multiple Myeloma	cefdinir	0.8884007	0.8728060	0.8806034
Multiple Myeloma	oxiconazole	0.8884007	0.8728060	0.8806034
Multiple Myeloma	pomalidomide	0.8884007	0.8728060	0.8806034
Multiple Myeloma	rifampicin	0.8884007	0.8728060	0.8806034

Table 11: Top twenty candidates validated by either GSE175384 or GSE47552 datasets. The full list can be found as supplementary information 11.

However, in this case 21 of those 62 compounds were not found to be used in studies related to MM at the time this study was carried out, and could thus be considered as repurposable candidates for MM (Table 12 and supplementary information 11)

Disease	Drug	Adj. similarity_A	Adj. similarity_B	mean_adj_similarity
Multiple Myeloma	etoricoxib	0.8884007	0.9987696	0.9435852
Multiple Myeloma	floxuridine	0.8884007	0.9987696	0.9435852
Multiple Myeloma	levofloxacin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	rilpivirine	0.8884007	0.9987696	0.9435852
Multiple Myeloma	valrubicin	0.8884007	0.9987696	0.9435852
Multiple Myeloma	flutamide	0.8034253	0.9957985	0.8996119
Multiple Myeloma	cefdinir	0.8884007	0.8728060	0.8806034
Multiple Myeloma	oxiconazole	0.8884007	0.8728060	0.8806034
Multiple Myeloma	balsalazide	0.7454534	0.9922521	0.8688527
Multiple Myeloma	parecoxib	0.7454534	0.9922521	0.8688527
Multiple Myeloma	tenoxicam	0.7454534	0.9922521	0.8688527
Multiple Myeloma	doxylamine	0.7226858	0.9528403	0.8377631
Multiple Myeloma	ketoprofen	0.6772528	0.9857551	0.8315039
Multiple Myeloma	carbidopa	0.5186185	0.9528403	0.7357294
Multiple Myeloma	triclosan	0.8636750	0.6049181	0.7342966
Multiple Myeloma	meclizine	0.7454534	0.6628453	0.7041493
Multiple Myeloma	oxaprozin	0.7454534	0.6628453	0.7041493
Multiple Myeloma	axitinib	0.6772528	0.5644724	0.6208626
Multiple Myeloma	mifepristone	0.6398028	0.5127472	0.5762750
Multiple Myeloma	ponatinib	0.5844759	0.4401220	0.5122990
Multiple Myeloma	pregnenolone	0.5186185	0.3603143	0.4394664

Table 12: Novelly repurposable candidates validated by either the GSE175384 or GSE47552 datasets. Table ordered according to their mean adjusted similarity.

4.3. Repurposable drugs for MM generated by SAveRUNNER

Twenty two candidates generated by SAveRUNNER have thus met the criteria of total or partial validation approach, and are therefore considered as drugs with potential to be novelly repurposed to MM (Table 13).

Drug	Target(s)
axitinib	KDR|FLT1|FLT4|CYP2C19|CYP3A5|KIT|PDGFRB|CSF1|PLK4
balsalazide	PTGS1|PTGS2|ALOX5|PPARG
carbidopa	DDC
cefdinir	MPO
doxylamine	HRH1|CHRM1
etoricoxib	PTGS2
floxuridine	TYMS
flutamide	-
ketoprofen	PTGS2|PTGS1|CXCR1|SLC5A8
levofloxacin	TOP2A
meclizine	NR1I3|HRH1
mifepristone	PGR|NR3C1|AR|CYP2B6|CYP2C8|CYP3A5|CYP3A7|NR1I2
oxaprozin	PTGS1|PTGS2
oxiconazole	-
parecoxib	PTGS2|LTF
ponatinib	ABL1|BCR|FLT3|RET|FGFR1|FGFR2|FGFR3|FGFR4|KIT|TEK|CYP2C8|CYP3A5|FGF2|FLT1|KDR|LCK|LYN|PDGFRA|SRC
pregnenolone	PGR|CYP17A1|SULT2B1
rilpivirine	CYP2C19|CYP3A5|NR1I2|SCN10A
sparfloxacin	TOP2A
tenoxicam	PTGS2|PTGS1
triclosan	DNMT1
valrubicin	TOP2A

Table 13: Repurposable candidates validated by GSE175384 and/or GSE47552 datasets and their annotated targets in CMap.

The above compounds can be grouped attending to their function:

- **Non-steroid anti-inflammatory drugs (NSAIDs)**, which are generally used to treat pain or fever among other inflammatory events[75], but have also been shown to help in several types of cancer[76]. However, in the case of MM, they should be handled with caution due to potential complications related to kidney failure[77]. According to SAveRUNNER implementations and subsequent validated approach followed in this thesis, balsalazide, ketoprofen, oxaprozin and tenoxicam are the NSAID candidates that may be repurposed for MM patients to help treat this disease (Figure 24).

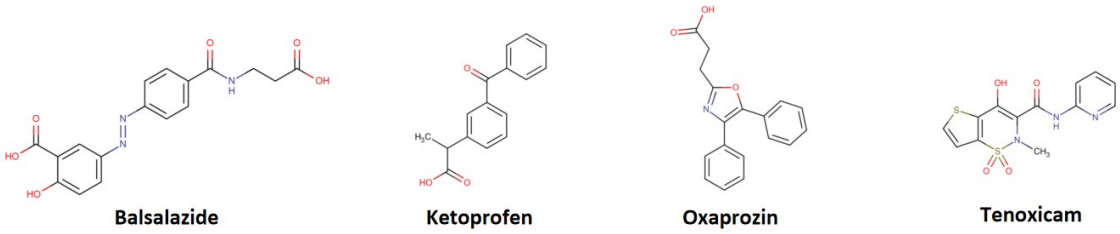


Figure 24: Molecular structures of the indicated NSAIDs.
Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- **Antibiotics**, which are widely used to selectively treat bacterial infections given their capacity to target bacterial specific processes, such as the building of their cell wall, provided that the bacteria has not become resistant to the antibiotic[79]. In addition to this role, some antibiotics have previously been used to treat

cancer, hence the possibility that those obtained in this thesis might be eligible for MM treatment. However, the potential adverse effects of antibiotics, such as changes in the intestinal microbiota, needs to be taken into account when considering their use in cancer patients[80]. According to the performed bioinformatic analyses in this thesis, the antibiotics sparfloxacin, cefdinir, levofloxacin, and triclosan (Figure 25) are suggested to have potential to be repurposed to MM.

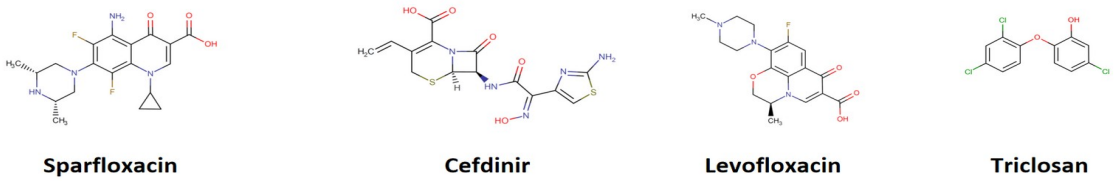


Figure 25: Molecular structures of the indicated antibiotics.
Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- **Antifungal agents**, which are compounds that specifically target pathogens of fungal origin, and are divided into different classes according to their molecular structure and targets[81]. In the context of cancer, Itraconazole, which belongs to the azole antifungal subgroup, and has recently been repurposed to treat this disease[82]. Thus, another azole compound, oxiconazole, which has come up in this study together with triclosan as a potential agent, might be relevant if repurposed to MM (Figure 26).

Figure 26: Molecular structures of the indicated antifungals.
Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- **Antihistamines,** often used to alleviate symptoms, have in some cases also been repurposed for cancer treatments given their capacity to revert multidrug resistance[83]. Therefore, the antihistamines doxylamine and meclizine (Figure 27) obtained upon implementation of SAveRUNNER might be candidates to consider when testing for drugs novelly applied to MM treatment. In the case of meclizine, it actually has been already used against several types of cancer[84,85].

Figure 27: Molecular structures of the indicated antihistamines. Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- **Steroids**, are compounds that include both laboratory-synthesized (corticosteroids) and naturally produced hormones with anti-inflammatory properties and modulatory effects on the immune system[86], which makes them useful on a variety of diseases, such as multiple sclerosis or autoimmune diseases[87]. They can also be used in cancer treatment, with prednisone[88] and dexamethasone[33] actually being currently used for MM therapy. In this drug category, SAveRUNNER analyses implemented in this study suggests the steroids mifepristone and pregnenolone as potential candidates to be repurposed to MM (Figure 28), which would expand the applications these compounds currently have in cancer[89,90].

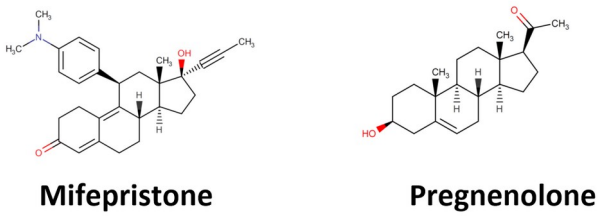


Figure 28: Molecular structures of the indicated steroids. Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- **Enzyme inhibitors**, which usually block the biological reactions their named after, can also be used for cancer treatment provided they target enzymes involved in cancer key events, such as cell cycle dysregulation[91] or apoptosis inhibition[92] among other events. In this study, the decarboxylase inhibitor carbidopa, the selective COX-2 inhibitors etoricoxib and parecoxib, and the tyrosine kinase inhibitors axitinib and ponatinib (Figure 29) have emerged as potential candidates to be repurposed for MM treatment. Thus, these compounds may have capacity to treat other cancers than those they have been already used against[93–97].

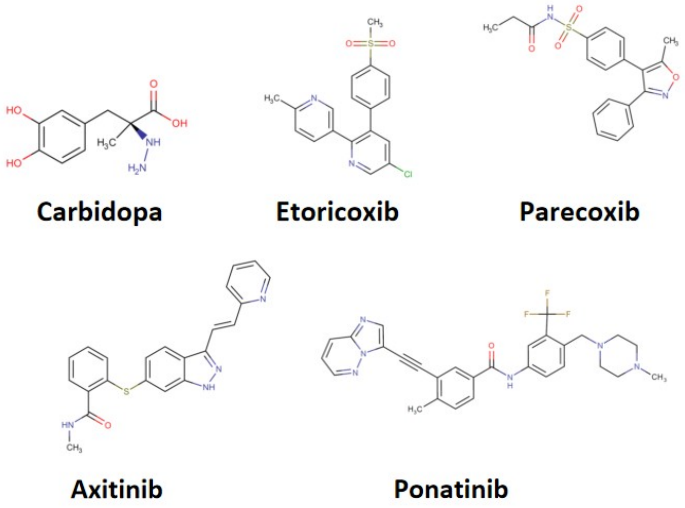


Figure 29: Molecular structures of the indicated enzyme inhibitors . Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

- Lastly, **Floxuridine, rilpivirine, valrubicin, and flutamide**, not belonging to any of the above mentioned categories, have also been obtained in this study as repurposable candidates to MM (Figure 30). Floxuridine inhibits cell division, thus being useful to target cells that divide rapidly, such as cancer cells[98]. This may make floxuridine potentially beneficial in the treatment of MM. Rilpivirine is a non-nucleoside reverse transcriptase inhibitor typically used to treat patients infected with the HIV-1 virus that has also shown toxic effects on pancreatic cancer cells[99]. Thus, MM cells may also benefit from this drug, and experimental validation might be worth a try. Valrubicin has been used for bladder carcinoma, and is one of the topoisomerase inhibitors (TIs) used in cancer treatment[100]. Since TOP2A is upregulated in a subset of MM patients[101], TIs appear useful when treating this MM subgroup. However, the DEAs for the patient samples included in the datasets analyzed in this thesis show downregulation TOP2A (supplementary information 5 and 7), which may make valrubicin less appropriate for these patients. Flutamide is a drug that prevent testosterone to bind the target cell receptor and has thus been used in the treatment of advanced prostate cancer[102]. Since MM patients seem to have low levels of testosterone, it seems unclear how flutamide could be beneficial in this scenario.

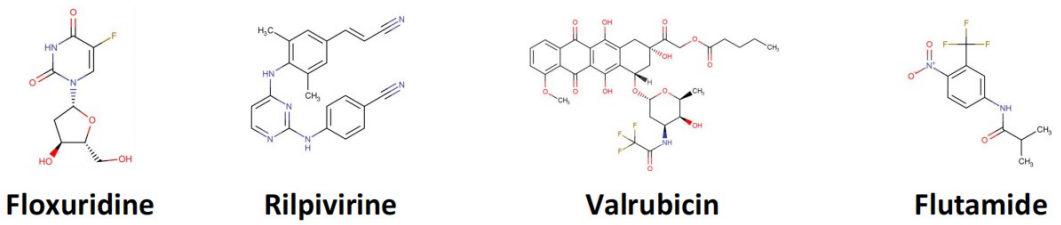


Figure 30: Molecular structures of the indicated compounds. Compounds generated by both SAveRUNNER and validated via DEA/CMap analyses. Structures were obtained from DrugBank[78].

4.4. Molecular affinity example: docking of ponatinib or axitinib with KIT

4.4.1. Selection criteria for drugs and target

As discussed earlier, when entering DEGs obtained via DEAs (Tables 6 and 7, and supplementary information 5 and 7, respectively) as input in the query tool of CMap, the corresponding .gct files were generated (supplementary information 8 and 9, respectively). These files contain different parameters, such as ncs scores, for which threshold values have been used as filters in order to select compounds that potentially counteract the transcription profiles (DEGs) of the MM datasets used for DEAs in this thesis. This means that such compounds would have the potential to treat MM. In addition to these parameters (specified in section 3.5), these .gct files also contain annotated genes for each compound, i.e genes that can be regulated by each compound according to what is published on the literature, and independently of the query performed using the DEGs introduced by the user (Table 13 and supplementary information 8 and 9). These annotated genes for each compound present in the final list of candidates to repurpose to MM has been used as a filter of all the DEGs found via DEAs in order to select a gene (target) to perform molecular docking as a representative example of which drugs could have more affinity for this given target. The selected gene was KIT, since it is involved in cell survival/proliferation[103], promotes cancer in its mutated forms[104], it appeared significantly upregulated in the GSE175384 dataset (supplementary information 5), is expressed in plasma cells in at least a subset of MM patients[105], and, according to the CMap analyses performed in this thesis as well as to previous studies[106,107], is a target for the anticancer drugs axitinib and ponatinib, both present in the final table of repurposable candidates (Table 13). Axitinib has been used to inhibit tumor growth in renal carcinoma[108] or breast cancer[109], whereas ponatinib is better known as a drug against leukemia[110]. Although both drugs have already been shown to interact with the KIT's kinase domain via docking studies[106,107], a question about which one could have more affinity and might therefore be used on a lower dose could be clinically relevant in MM. This could be addressed by comparing the corresponding docking analyses, which is shown in the next subsection as a representative example of an approach to follow when selecting a candidate among available drugs with common targets.

4.4.2. Molecular Docking analyses: Workflow and results

The crystal structure of the KIT kinase domain as a complex with ponatinib was downloaded in .cif format from the protein data bank (PDB, ID 4u0i), and further processed in BIOVIA Discovery Studio Visualizer software (BDSVS, https://discover.3ds.com/discovery-studio-visualizer-download) by removing ponatinib and water molecules to be left with only the KIT kinase domain. In addition, BDSVS detected three active sites for KIT in the .cif file, with Active Site 1 containing residues that belong to the ATP, selectivity, and DFG pockets (Figure 31), which are directly involved in KIT-ponatinib[107] and KIT-axitinib[106] interactions to explain the KIT kinase domain inhibition. However, no information on ponatinib or axitinib interaction with the

other active sites was found in the literature. For this reason, the docking analyses were only performed for active site 1, whose center coordinates were assigned based on a sphere generated by BDSVS that spanned this site (Figure 32).

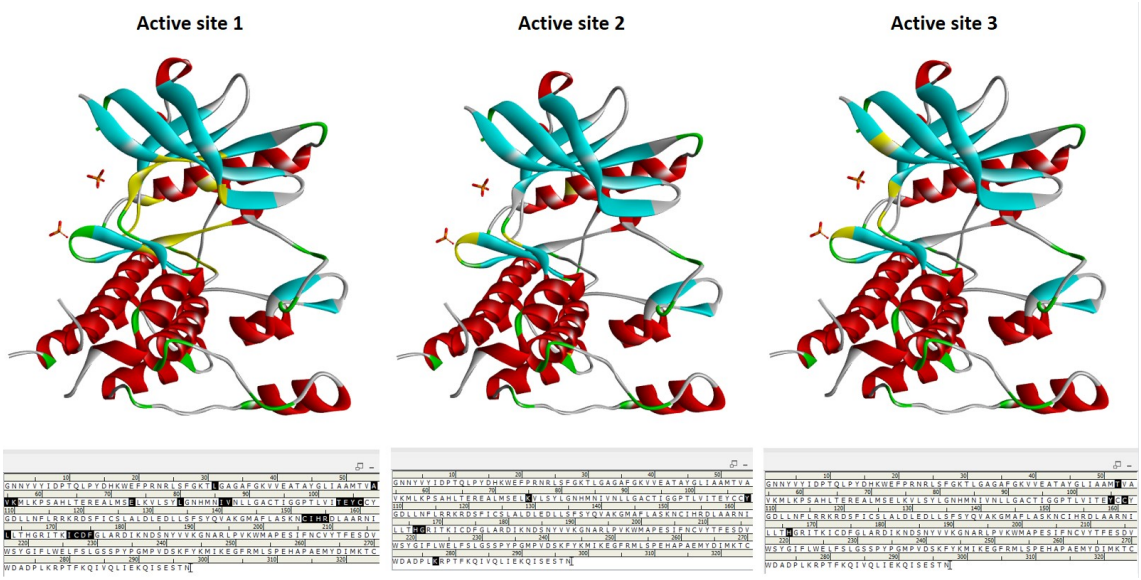


Figure 31: Active sites for KIT's kinase domain detected by BDSVS.
Active sites are highlighted in yellow (3D structure) and black (amino acid sequence). Only active site 1 appears to interact with ponatinib and axitinib in the current literature.

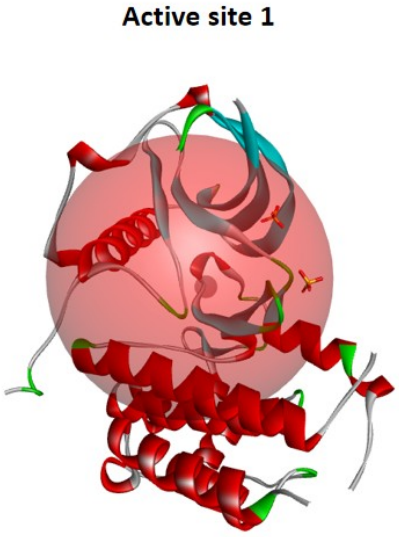


Figure 32: Active site 1 coordinates determined by BDSV.
The sphere radius and center determine coordinates for the active site 1 in human KIT kinase domain, which were used as reference for molecular docking analyses in ATS.

The center coordinates for the sphere together with 2 times the respective radius used to set the size of each dimension of the grid-box used as reference in the configuration

files (Figure 33) needed by Autodock Vina and AutoDock Tools Software (ATS)[65] to perform the docking analyses of KIT with either ponatinib or axitinib (Figure 34).

```
size_x = 36
size_y = 36
size_z = 36

center_x = 34.6539
center_y = 10.825
center_z = 45.7105



energy_range = 4
exhaustiveness= 8
```

Figure 33: Config.txt file used by Autodock Vina for docking analyses.
Each file with the respective coordinates and grid box size can be found as supplementary information 12.

Prior to running docking analyses, preparation of KIT, poxatinib and axitinib as well as subsequent creation of the corresponding .pdbqt files was carried out. KIT was prepared mostly in ATS (unless specified otherwise), and consisted of:

- Removing water molecules and ponatinib including in the initial .cis file and saving output as .pdb file for further use in Autodock Vina and ATS (done in BDSVS).
- Adding polar hydrogens.
- Adding Kolmann charges and computing Gaisteger charges.
- Set atoms in format AD4.
- Save molecule in .pdbqt format.

Regarding preparation of the ligands, ponatinib and axitinib structures were downloaded from Drugbank[78] in 3D-SDF format, transformed into .pdb format with Open Babel[64], and loaded onto ATS as ligands, where:

- Gasteiger charges were added.
- Non-polar hydrogens were merged.
- Rotatable bonds were detected.
- Number of torsional degrees of freedom in the ligand (TORSDOF) were detected.
- Root was automatically detected and selected to set the torsion tree.

- Molecules were saved in .pdbqt format.

Docking analyses were all performed from the Windows command line using vina.exe and the corresponding parameters, which yielded different poses with their respective affinity scores for each docking analysis (Figure 34 and supplementary information 12).

```
C:\Docking>vina.exe --config config.txt --log log.txt --out output.txt --receptor kit.pdbqt --ligand ponatinib.pdbqt
#################################################################
# If you used AutoDock Vina in your work, please cite:          #
#                                                               #
# O. Trott, A. J. Olson,                                        #
# AutoDock Vina: improving the speed and accuracy of docking    #
# with a new scoring function, efficient optimization and       #
# multithreading, Journal of Computational Chemistry 31 (2010)  #
# 455-461                                                       #
#                                                               #
# DOI 10.1002/jcc.21334                                         #
#                                                               #
# Please see http://vina.scripps.edu for more information.      #
#################################################################

WARNING: The search space volume > 27000 Angstrom^3 (See FAQ)
Detected 8 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: -559531552
Performing search ...
0%   10   20   30   40   50   60   70   80   90   100%
|----|----|----|----|----|----|----|----|----|----|
**************************************************
done.
Refining results ... done.

mode |   affinity | dist from best mode
     | (kcal/mol) | rmsd l.b.| rmsd u.b.
-----+------------+----------+----------
   1        -11.9      0.000      0.000
   2        -11.4      1.882      2.862
   3        -11.3      1.655      2.516
   4        -11.1      2.079      3.006
   5        -11.1      1.980      2.742
   6        -10.8      2.378      3.464
   7        -10.5      1.814      2.663
   8        -10.1     11.736     17.358
   9        -10.1     11.574     14.237
Writing output ... done.
```

Figure 34: Execution of AutoDock Vina in Windows Command line.
Representative example of molecular docking analysis yielding affinity scores for different poses between active site 1 in KIT and ponatinib. Docking analyses for both ponatinib and axitinib with active site 1 in KIT are available as supplementary information 12.

The output file for each docking analysis contained affinity scores for different ligand poses (Figure 34), and was further splitted into .pdqt files, each corresponding to a pose, by using vina_split.exe (Figure 35) in order to be able to use the best ponatinib and axitinib poses individually with KIT's active site 1 for docking visual representation (Figure 36 and Figure 37, respectively).

```
C:\Docking>vina_split.exe --input output.txt
Prefix for ligands will be output.txt_ligand_
Prefix for flexible side chains will be output.txt_flex_
```

Figure 35: Creation of pdbqt.files for individual poses in each docking analysis.
All .pdbqt files for individual poses can be found as supplementary information 12.

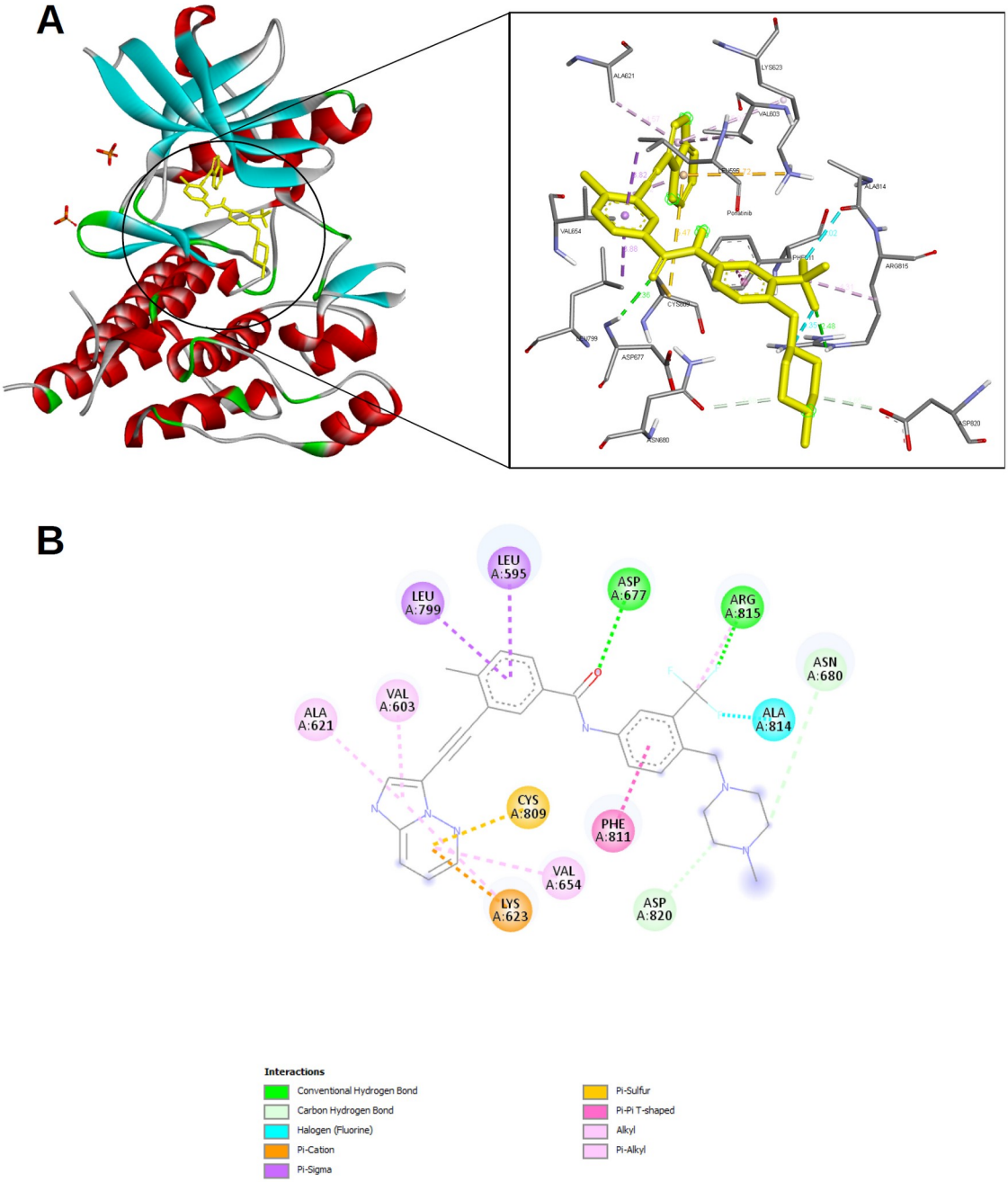


Figure 36: Molecular docking between KIT and ponatinib.
A,3D representation for the best pose (lowest affinity score) between residues in active site 1 of KIT kinase domain and ponatinib (yellow). B, 2D representation of A. Images generated with BDSVS.

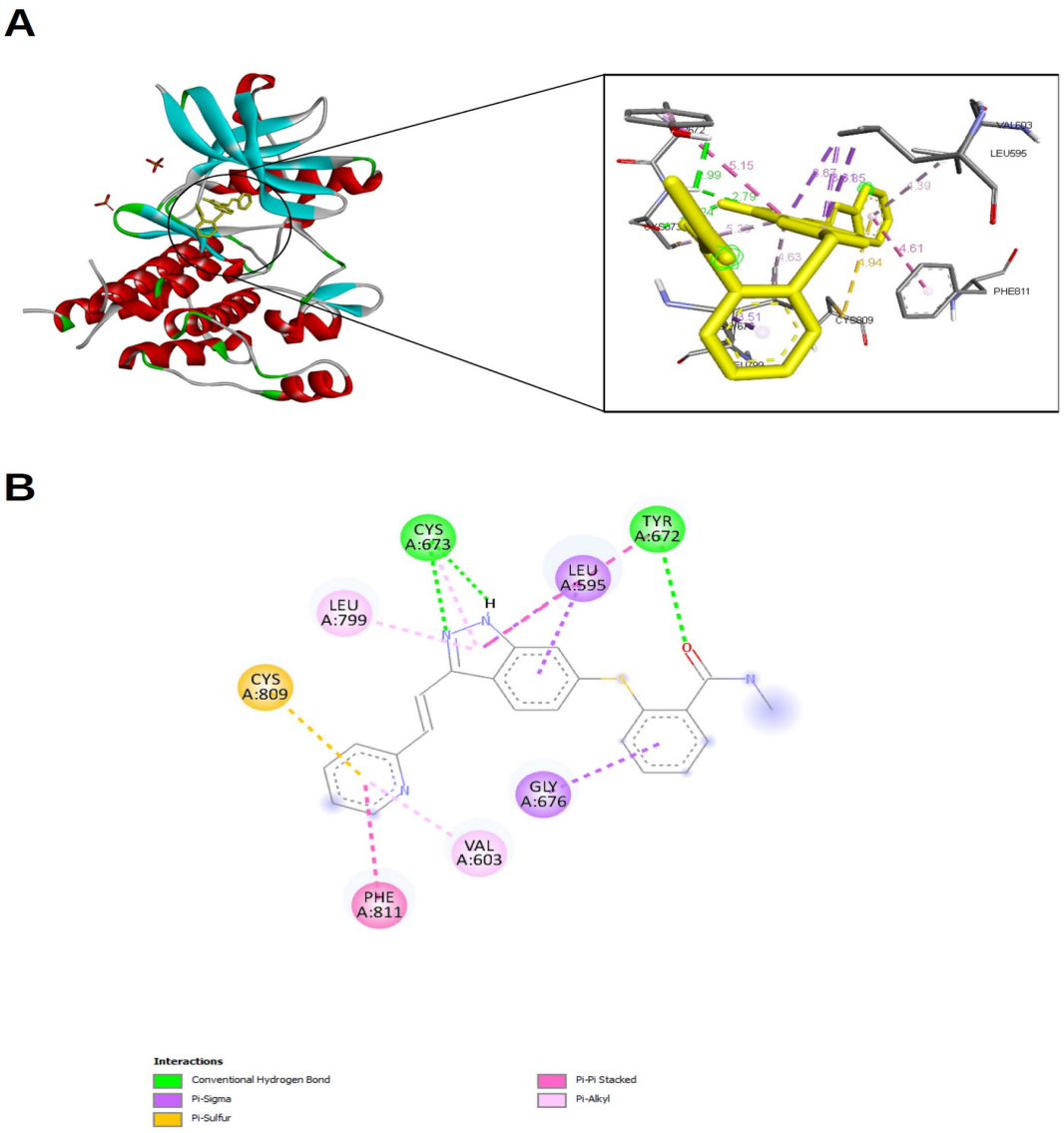


Figure 37: Molecular docking between KIT and axitinib.
A,3D representation for the best pose (lowest affinity score) between residues in active site 1 of KIT kinase domain and axitinib (yellow). B, 2D representation of A. Images generated with BDSVS.

The best (lowest) affinity score corresponded to the Active Site 1-ponatinib interaction (Figure 34 and 36, Table 14, and supplementary information 12). Therefore, due to its greater affinity, ponatinib might be preferable over axitinib to inhibit KIT's kinase domain according to this analysis. In this case, since KIT appeared differentially expressed in the GSE175384 dataset used in this study (supplementary information 5), at least some MM patients might find more beneficial to undergo a treatment that includes ponatinib rather than axitinib for KIT inhibition, i.e. they may have less side effects with ponatinib than with axitinib since a lower dose would be needed. It is however worth noting that, even though the active site 1 contains the residues involved in both KIT-ponatinib and KIT-axitinib interactions available in the literature[106,107], the interactions generated by BDSVS (Figure 36 and Figure 37) upon running docking analysis with ATS (Figure 34 and supplementary information 12) do not exactly match what is shown in the corresponding publications[106,107]. Thus, further optimization of the

obtained structures, a more detailed docking analysis as well as experimental validation (all out of the scope of this work) are possibly necessary in order to address this question with more certainty.

	Active Site 1
Ponatinib	-11.9
Axitinib	-9.4

Table 14: Protein-ligand Affinity scores (Kcal/mol)
Values for the best pose of each docking analysis between KIT's active site in the kinase domain and axitinib or ponatinib. Scores for all poses in each docking analysis are available as supplementary information 12.

5. CONCLUSIONS

New drugs to treat MM are necessary given its currently poor prognosis. *In silico* drug repurposing stands as a valuable tool since it helps with the reposition of drugs that are currently being used for other diseases in a relatively fast and inexpensive manner. In this regard, this thesis presents some candidates that might be repurposed to MM, thereby helping improve the, as for today, lethal final outcome of this malignancy.

5.1. Goals achieved

In this study, the initial aim of finding candidates to repurpose to MM has been accomplished: A final list of twenty two candidates belonging to different drug classes are suggested to have the potential to be repurposed to MM. This has been achieved by:

1. Implementing SAveRUNNER, a network-based algorithm to that performs *in silico* analyses in R that yields candidates with repurposing potential for a given disease[5,6]. These candidates appear to be specific for MM (and perhaps also for MM related diseases) when compared to drugs obtained through an independent SAveRUNNER implementation that served as negative control (Figure 19).

2. Validating/filtering the compounds generated by SAveRUNNER via DEA/CMap analyses in order to select those that counteract the dysregulation of DEGs in MM.

In addition, molecular docking analyses between two compounds of the final list, ponatinib and axitinib, and the KIT's kinase domain was performed as an example of candidates with a common target that could be important in MM. Indeed, KIT was significantly upregulated in the GSE175384 dataset (supplementary information 5), favors cancer in its mutated forms[104], was shown to be expressed in plasma cells in some MM patients[105], and is a target for the anticancer drugs axitinib and ponatinib according to previous studies[106,107].

5.2. Planning and methodology

Although the objectives for each milestone (PEC) have been accomplished, the work plan had to be somewhat modified along the way. A step that has taken longer than expected was the implementation of the SAveRUNNER algorithm in R, which took several weeks of computational processing when nearly 60 diseases (blood related diseases) were used to generate the final disease-drug network and MM-drug subnetwork. Therefore, far fewer diseases were used for the second implementation of SAveRUNNER (using diseases related to MM symptoms) in order to meet the initially planned timeline. Interestingly, the implementation with this second implementation yielded a similar number of drug candidates. Therefore, future implementations for other diseases could probably generate relatively good results when few diseases are used together with the malignancy of interest in the corresponding config.R file. It is also worth noting that alternative databases can be used with SAveRUNNER provided they are placed in the corresponding folder ('input files' folder) and in the required format. This would allow to change outdated databases or using others that might be more comprehensive in the future.

DEA/CMap analyses, however, could be performed faster than anticipated, which allowed to perform molecular docking analyses that were considered as an extra task in case there was extra time.

5.3. Future perspectives

The final list of repurposable candidates suggested by this study can serve as the basis for future projects aiming at experimentally and/or clinically validating their suitability for reposition to MM. For example, a logical next step could consist of treating a panel of MM cell lines[111] with some of these candidates to assess their ability to inhibit cell growth and/or induce apoptosis. At this stage, combinations of different candidates, perhaps also including currently available drugs for MM treatment, could also be tested to try to find potent cocktails that may help fight this plasma cell malignancy. As a next step, successful candidates/cocktails could be used in an *in vivo* scenario, such as an adequate MM mouse model[112], to verify their potentially beneficial effects, or in a small clinical trial since all the candidates are already approved for use in patients with diseases other than MM, once potential side effects have been taken into consideration. Furthermore, if the repurposable drugs suggested by SAveRUNNER in this study were indeed proven to experimentally/clinically help treat MM, then perhaps lines of research similar to those above mentioned could also be done for COVID-19 and Amyotrophic Lateral Sclerosis, since SAveRUNNER has also suggested some candidates for these diseases[5,113].

As a different approach, SAveRUNNER could also be modified in different ways in order to obtain new candidates given that SAveRUNNER's code is readily available online[61]. For example, as earlier mentioned, different drug-target or disease-genes databases could be used provided they are specified in the corresponding directory

(SAveRUNNER-main\code\input_files), or the proximity mathematical formula to calculate proximity scores could be changed, which could modify the final list of compounds to be considered for repositioning generated by SAveRUNNER.

6. GLOSSARY

Akt	Protein kinase B
ASCT	Autologous Stem Cell Transplant
ATS	Autodock Tools Software
BAFF	B cell Activation Factor
BDSVS	BIOVIA Discovery Studio Visualizer software
BiRd	Biaxin-Revlimid-dexamethasone
BMSC	Bone Marrow Stem Cells
Bort	Bortezomib
CCL3	Chemokine (C-C motif) Ligand 3
CD38	Cluster of Differentiation 38
CDK	Cyklin D Kinase
CFU-GM	Colony Forming Unit-Granulocyte-Macrophage
cKit	Tyrosine-Protein Kinase Kit
CRBN	Cereblon
CS	Connectivity score
CUL4A	Cullin 4 A
BDSVS	BIOVIA Discovery Studio Visualizer software
CyBorD	Cyclophophosphamide-Bortezomid-dexamethasone
Dara-IRd	Daratumumab-ixazomib-Revlimid-dexamethasone
Dara-KPd	Daratumumab-Kyprolis-Pomalidomide-dexamethasone
Dara-RVd	Daratumumab-Revlimid-Velcade-dexamethasone
DCEP	Dexamethasone-Cyclophosphamide-Etoposide-Platinum
DDB1	Damage Specific DNA Binding Protein 1
DKK1	Dickkopf WNT Signaling Pathway Inhibitor 1
DNA	Deoxyribonucleic Acid
DPd	Daratumumab-Pomalidomide-dexamethasone
DRd	Daratumumab-Revlimid-dexamethasone
DTI	Drug-Target Interaction
DT-PACE	Dexamethasone-thalidomide-Platinum-Adriamycin -Cyclophosphamide-Etoposide
BDSVS	BIOVIA Discovery Studio Visualizer software
DVd	Daratumumab-Velcade-dexamethasone
EGF-R	Epidermal Growth Factor Receptor
Elo-PVd	Elotuzumab-Pomalidomide-Velcade-dexamethasone
EPd	Elotuzumab-Pomalidomide-dexamethasone
ERd	Elotuzumab-Revlimid-dexamethasone
EVd	Elotuzumab-Velcade-dexamethasone

Fab	Fragment antigen-binding
Fc	Fragment, crystallizable
FGFR3	Fibroblast Growth Factor Receptor 3
FLC	Free Light Chain
GEO	Gene Expression Omnibus
GUI	Graphic user interface
HD-cyclophosphamide	High-dose cyclophosphamide
HGF	Hepatocyte Growth Factor
Hsp	Heat shock protein
ICd	Ixazomib-Cyclophosphamide-dexamethasone
Ig	Immunoglobulin
IGF-1R	Insulin Growth Factor 1 Receptor
IKZF1/3	IKaros Zinc Finger
IL-6	Interleukin-6
IMdDs	Immunomodulatory drugs
IPd	Ixazomib-Pomalidomide-dexamethasone
IRd	Ixazomib-Revlimid-dexamethasone
ISS	International Staging System
JNK	c-Jun N-terminal Kinase
KCd	Kyprolis-cyclophosphamide-dexamethasone
KIR	Killer cell Immunoglobulin-like Receptor
KPd	Kyprolis-Pomalidomide-dexamethasone
KRd	Kyprolis-Revlimid-dexamethasone
KSP	Kinesin Spindle Protein
LDH	Lactate Dehydrogenase
Len	Lenalidomide
MAPK	Mitogen-Activated Protein Kinase
MEK	Mitogen-activated protein Kinase kinase
MM	Multiple Myeloma
MPI-1 alpha	MacroPhage Inflammatory protein 1-alpha
mTORC	Mammalian Target of Rapamycin Complex
ncs	Normalized connectivity score
NDMM	Newly Diagnosed Multiple Myeloma
PARP	Poly ADP-Ribose Polymerase
PCd	Pomalidomide-Cyclophosphamide-dexamethasone
Pd	Pomalidomide-dexamethasone
PD-1	Programmed cell Death protein 1
PDGFR3	Platelet-Derived Growth Factor Receptor 3
PEC	Prueba de evaluación contínua
PI	Proteasome Inhibitor
PDB	Protein Data Bank
PKC	Protein Kinase C
PVd	Pomalidomide-Velcade-dexamethasone
Raf	Rapidly accelerated fibrosarcoma

RANKL	Receptor Activator of Nuclear factor kappa beta (NFkB ligand)
RCd	Revlimid-Cyclophosphamide-dexamethasone
Rd	Revlimid-dexamethasone
RNA	Rybonucleic Acid
ROC1	Rotamase CYP 1
RRMM	Relapsed/refractory multiple myeloma
RVd	Revlimid-Velcade-dexamethasone
SAveRUNNER	Searching off-lAbel dRUg aNd NEtwoRk
sFLCR	Serum free light chain ratio
SFRP3	Secreted frizzled-related protein 3
SLAMF-7	Signaling Lymphocytic Activation Molecule Family 7
TAS	Transcription activity score
TGF-Beta	Transforming growth factor-Beta
TI	Topoisomerase inhibitor
TNF-a	Tumor necrosis factor-alpha
Vd	Velcade-dexamethasone
VDT-PACE	Velcade-Dexamethasone-Thalidomide-Platinum-Adriamycin-Cyclophosphamide-Etoposide
VEGF	Vascular Endothelial Growth Factor
VEGFR	Vascular Endothelial Growth Factor Receptor
VMP	Velcade-Melphalan-Prednisone
VTd	Velcade-Thalidomide-dexamethasone

7. REFERENCES

1. World Health Organization (WHO), 2021. All Rights Reserved. Cancer worldwide. Accessed in July 2021. https://www.who.int/health-topics/cancer#tab=tab_1.
2. Ludwig, H., Novis Durie, S., Meckl, A., Hinke, A. & Durie, B. Multiple Myeloma Incidence and Mortality Around the Globe; Interrelations Between Health Access and Quality, Economic Resources, and Patient Empowerment. *The Oncologist* **25**, e1406–e1413 (2020).
3. Yan, H. *et al.* Identification of key candidate genes and pathways in multiple myeloma by integrated bioinformatics analysis. *J. Cell. Physiol.* **234**, 23785–23797 (2019).

4. Gysi, D. M. *et al.* Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc. Natl. Acad. Sci.* **118**, (2021).
5. Fiscon, G., Conte, F., Farina, L. & Paci, P. SAveRUNNER: A network-based algorithm for drug repurposing and its application to COVID-19. *PLOS Comput. Biol.* **17**, e1008686 (2021).
6. Fiscon, G. & Paci, P. SAveRUNNER: an R-based tool for drug repurposing. *BMC Bioinformatics* **22**, 150 (2021).
7. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/26/1/145/182533.
8. Albagoush, S. A. & Azevedo, A. M. Multiple Myeloma. in *StatPearls* (StatPearls Publishing, 2021).
9. s.r.o, B. S. GanttProject: free project management tool for Windows, macOS and Linux. *GanttProject* https://www.ganttproject.biz.
10. Rajkumar, S. V. Multiple Myeloma: 2016 update on Diagnosis, Risk-stratification and Management. *Am. J. Hematol.* **91**, 719–734 (2016).
11. Medical Sciences | Free Full-Text | Epidemiology, Staging, and Management of Multiple Myeloma. https://www.mdpi.com/2076-3271/9/1/3.
12. 35-Multiple-myeloma-fact-sheet. Accessed 17/07/2021. https://gco.iarc.fr/today/data/factsheets/cancers/35-Multiple-myeloma-fact-sheet.pdf.
13. Kazandjian, D. Multiple myeloma epidemiology and survival, a unique malignancy. *Semin. Oncol.* **43**, 676–681 (2016).
14. Rajkumar, S. V. Updated Diagnostic Criteria and Staging System for Multiple Myeloma. *Am. Soc. Clin. Oncol. Educ. Book* e418–e423 (2016) doi:10.1200/EDBK_159009.
15. Mukkamalla, S. K. R. & Malipeddi, D. Myeloma Bone Disease: A Comprehensive Review. *Int. J. Mol. Sci.* **22**, 6208 (2021).

16. Chiu, M. L., Goulet, D. R., Teplyakov, A. & Gilliland, G. L. Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies* **8**, 55 (2019).
17. Types of Multiple Myeloma | International Myeloma Foundation. Accessed 21/7/2021. https://www.myeloma.org/types-of-myeloma.
18. What Is Multiple Myeloma? Accessed in July, 2021. https://www.cancer.org/cancer/multiple-myeloma/about/what-is-multiple-myeloma.html (2021).
19. Corso, A. & Mangiacavalli, S. Non-Secretory Myeloma: Ready for a new Definition? *Mediterr. J. Hematol. Infect. Dis.* **9**, e2017053 (2017).
20. Berbari, H. E. & Kumar, S. K. Initial Therapeutic Approaches to Patients with Multiple Myeloma. *Adv. Ther.* **38**, 3694–3711 (2021).
21. A Comparison of Different Staging Systems for Multiple Myeloma: Can the MRI Pattern Play a Prognostic Role? : American Journal of Roentgenology : Vol. 209, No. 1 (AJR). https://www.ajronline.org/doi/10.2214/AJR.16.17219.
22. Prognostic Value of Serum Free Light Chains Measurements in Multiple Myeloma Patients. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0166841.
23. Yang, Y., Li, Y., Gu, H., Dong, M. & Cai, Z. Emerging agents and regimens for multiple myeloma. *J. Hematol. Oncol.J Hematol Oncol* **13**, 150 (2020).
24. Ito, S. Proteasome Inhibitors for the Treatment of Multiple Myeloma. *Cancers* **12**, 265 (2020).
25. Fogli, S., Galimberti, S., Gori, V., Del Re, M. & Danesi, R. Pharmacology differences among proteasome inhibitors: Implications for their use in clinical practice. *Pharmacol. Res.* **167**, 105537 (2021).
26. Breitkreutz, I. *et al.* Lenalidomide inhibits osteoclastogenesis, survival factors and bone-remodeling markers in multiple myeloma. *Leukemia* **22**, 1925–1932 (2008).
27. Fink, E. C. & Ebert, B. L. The novel mechanism of lenalidomide activity. *Blood* **126**, 2366–2369 (2015).

28. Lentzsch, S. *et al.* S-3-Amino-phthalimido-glutarimide Inhibits Angiogenesis and Growth of B-Cell Neoplasias in Mice. *Cancer Res.* **62**, 2300–2305 (2002).
29. Cyclophosphamide_monograph_1June2013_formatted.pdf. http://www.bccancer.bc.ca/drug-database-site/Drug%20Index/Cyclophosphamide_monograph_1June2013_formatted.pdf.
30. Research, C. for D. E. and. Daratumumab (DARZALEX). *FDA* (2019).
31. Roccatello, D. *et al.* CD38 and Anti-CD38 Monoclonal Antibodies in AL Amyloidosis: Targeting Plasma Cells and beyond. *Int. J. Mol. Sci.* **21**, 4129 (2020).
32. Dexamethasone-induced apoptotic mechanisms in myeloma cells investigated by analysis of mutant glucocorticoid receptors | Blood | American Society of Hematology. https://ashpublications.org/blood/article/112/4/1338/25237/Dexamethasone-induced-apoptotic-mechanisms-in.
33. Kervoëlen, C. *et al.* Dexamethasone-induced cell death is restricted to specific molecular subgroups of multiple myeloma. *Oncotarget* **6**, 26922–26934 (2015).
34. Navigating the Changing Multiple Myeloma Treatment Landscape. *European Medical Journal* https://www.emjreviews.com/hematology/symposium/navigating-the-changing-multiple-myeloma-treatment-landscape/ (2016).
35. Current Treatment Strategies for Multiple Myeloma | JCO Oncology Practice. https://ascopubs.org/doi/10.1200/JOP.19.00244?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub++0pubmed&.
36. Autologous Transplantation and Maintenance Therapy in Multiple Myeloma | NEJM. https://www.nejm.org/doi/full/10.1056/nejmoa1402888.
37. Luoma, S. *et al.* Long-term outcome after allogeneic stem cell transplantation in multiple myeloma. *Ann. Hematol.* **100**, 1553–1567 (2021).
38. Pinto, V. *et al.* Multiple Myeloma: Available Therapies and Causes of Drug Resistance. *Cancers* **12**, 407 (2020).

39. Lohr, J. G. *et al.* Widespread Genetic Heterogeneity in Multiple Myeloma: Implications for Targeted Therapy. *Cancer Cell* **25**, 91–101 (2014).
40. Bolli, N. *et al.* Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat. Commun.* **5**, 2997 (2014).
41. Multiple myeloma epidemiology and demographics - wikidoc. Accessed 29.7.2021. https://www.wikidoc.org/index.php/Multiple_myeloma_epidemiology_and_demographics.
42. Improved long-term survival in multiple myeloma up to the age of 80 years | Leukemia. https://www.nature.com/articles/leu201423.
43. Drug repurposing: progress, challenges and recommendations | Nature Reviews Drug Discovery. https://www.nature.com/articles/nrd.2018.168.
44. Sadeghi, S. S. & Keyvanpour, M. R. An Analytical Review of Computational Drug Repurposing. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**, 472–488 (2021).
45. Artificial intelligence, machine learning, and drug repurposing in cancer. https://www.tandfonline.com/doi/full/10.1080/17460441.2021.1883585.
46. Morris, G. M. & Lim-Wilby, M. Molecular docking. *Methods Mol. Biol. Clifton NJ* **443**, 365–382 (2008).
47. Pinzi, L. & Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *Int. J. Mol. Sci.* **20**, 4331 (2019).
48. Zhao, K. & So, H.-C. Using Drug Expression Profiles and Machine Learning Approach for Drug Repurposing. in *Computational Methods for Drug Repurposing* (ed. Vanhaelen, Q.) 219–237 (Springer, 2019). doi:10.1007/978-1-4939-8955-3_13.
49. Recent advances in drug repurposing using machine learning - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1367593121000818.

50. Hameed, P. N., Verspoor, K., Kusljic, S. & Halgamuge, S. A two-tiered unsupervised clustering approach for drug repositioning through heterogeneous data integration. *BMC Bioinformatics* **19**, 129 (2018).
51. Cheng, F. *et al.* Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **9**, 2691 (2018).
52. Lotfi Shahreza, M., Ghadiri, N., Mousavi, S. R., Varshosaz, J. & Green, J. R. A review of network-based approaches to drug repositioning. *Brief. Bioinform.* **19**, 878–892 (2018).
53. Molecular networks in Network Medicine: Development and applications - Silverman - 2020 - WIREs Systems Biology and Medicine - Wiley Online Library. https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wsbm.1489.
54. Clauset, A., Newman, M. E. J. & Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **70**, 066111 (2004).
55. Yu, W., Gwinn, M., Clyne, M., Yesupriya, A. & Khoury, M. J. A navigator for human genome epidemiology. *Nat. Genet.* **40**, 124–125 (2008).
56. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
57. Chatr-aryamontri, A. *et al.* The BioGRID interaction database: 2015 update. *Nucleic Acids Res.* **43**, D470–D478 (2015).
58. Peri, S. *et al.* Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Res.* **32**, D497–D501 (2004).
59. Orchard, S. *et al.* The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–D363 (2014).
60. Breuer, K. *et al.* InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res.* **41**, D1228–D1233 (2013).
61. sportingCode. *sportingCode/SAveRUNNER Accessed 13.10.2021 https://github.com/sportingCode/SAveRUNNER*. (2021).

62. CONNECTOPEDIA [clue.io]. Accessed on 29/9/2021. https://clue.io/connectopedia/.
63. CONNECTOPEDIA [clue.io]. Accessed on 01.10.2021. https://clue.io/connectopedia/signature_quality_metrics.
64. O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminformatics* **3**, 33 (2011).
65. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041641/.
66. Alonso, R. *et al.* Prolonged lenalidomide maintenance therapy improves the depth of response in multiple myeloma. *Blood Adv.* **4**, 2163–2171 (2020).
67. Dimopoulos, M. *et al.* Pomalidomide, bortezomib, and dexamethasone for multiple myeloma previously treated with lenalidomide (OPTIMISMM): outcomes by prior treatment at first relapse. *Leukemia* **35**, 1722–1731 (2021).
68. GSE175384: GEO Accession viewer. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE175384.
69. GSE47552: GEO Accession viewer. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE47552.
70. Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K. & Liu, X. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLOS ONE* **9**, e78644 (2014).
71. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease - PubMed. https://pubmed.ncbi.nlm.nih.gov/17008526/.
72. CONNECTOPEDIA [clue.io] Accessed 12/12/2021. https://clue.io/connectopedia/cmap_algorithms.
73. Liposomal daunorubicin (DaunoXome) in multiple myeloma: a mo... : Anti-Cancer Drugs.

https://journals.lww.com/anti-cancerdrugs/Abstract/2003/11000/Liposomal_daunorubicin__DaunoXome__in_multiple.4.aspx.

74. Ishii, H. Combination chemotherapy for multiple myeloma with melphalan, ifosfamide, prednisolone, nitrosourea and vincristine. *Acta Med. Okayama* **42**, 175–182 (1988).

75. Ghlichloo, I. & Gerriets, V. Nonsteroidal Anti-inflammatory Drugs (NSAIDs). in *StatPearls* (StatPearls Publishing, 2021).

76. Rayburn, E. R., Ezell, S. J. & Zhang, R. Anti-Inflammatory Agents for Cancer Therapy. *Mol. Cell. Pharmacol.* **1**, 29–43 (2009).

77. Cullis, J. Haematology: Multiple Myeloma. *Clin. Med.* **19**, 188 (2019).

78. Chemical Structure Search | DrugBank Online. Accessed 30.10.2021. https://go.drugbank.com/structures/search/small_molecule_drugs/structure.

79. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *J. Anaesthesiol. Clin. Pharmacol.* **33**, 300–305 (2017).

80. Gao, Y. *et al.* Antibiotics for cancer treatment: A double-edged sword. *J. Cancer* **11**, 5135–5149 (2020).

81. Dixon, D. M. & Walsh, T. J. Antifungal Agents. in *Medical Microbiology* (ed. Baron, S.) (University of Texas Medical Branch at Galveston, 1996).

82. Pounds, R., Leonard, S., Dawson, C. & Kehoe, S. Repurposing itraconazole for the treatment of cancer. *Oncol. Lett.* **14**, 2587–2597 (2017).

83. Ellegaard, A.-M. *et al.* Repurposing Cationic Amphiphilic Antihistamines for Cancer Treatment. *EBioMedicine* **9**, 130–139 (2016).

84. Armaghany, T. *A Window of Opportunity Trial With Meclizine in Hepatocellular Carcinoma*. https://clinicaltrials.gov/ct2/show/NCT03253289 (2021).

85. Lin, J.-C. *et al.* Induction of apoptosis and cell-cycle arrest in human colon cancer cells by meclizine. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* **45**, 935–944 (2007).

86. Ericson-Neilsen, W. & Kaye, A. D. Steroids: Pharmacology, Complications, and Practice Delivery Issues. *Ochsner J.* **14**, 203–207 (2014).

87. Steroid Treatment - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/steroid-treatment.

88. Berenson, J. R. *et al.* Maintenance therapy with alternate-day prednisone improves survival in multiple myeloma patients. *Blood* **99**, 3163–3168 (2002).

89. Goyeneche, A. A., Carón, R. W. & Telleria, C. M. Mifepristone Inhibits Ovarian Cancer Cancer Cell Growth In Vitro and In Vivo. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **13**, 3370–3379 (2007).

90. XIAO, X. *et al.* Pregnenolone, a cholesterol metabolite, induces glioma cell apoptosis via activating extrinsic and intrinsic apoptotic pathways. *Oncol. Lett.* **8**, 645–650 (2014).

91. Otto, T. & Sicinski, P. Cell cycle proteins as promising targets in cancer therapy. *Nat. Rev. Cancer* **17**, 93–115 (2017).

92. Hunter, A. M., LaCasse, E. C. & Korneluk, R. G. The inhibitors of apoptosis (IAPs) as cancer targets. *Apoptosis Int. J. Program. Cell Death* **12**, 1543–1568 (2007).

93. Chen, Z. *et al.* Carbidopa suppresses prostate cancer via aryl hydrocarbon receptor-mediated ubiquitination and degradation of androgen receptor. *Oncogenesis* **9**, 1–13 (2020).

94. Saini, M. K., Sharma, P., Kaur, J. & Sanyal, S. N. The cyclooxygenase-2 inhibitor etoricoxib is a potent chemopreventive agent of colon carcinogenesis in the rat model. *J. Environ. Pathol. Toxicol. Oncol. Off. Organ Int. Soc. Environ. Toxicol. Cancer* **28**, 39–46 (2009).

95. Pan, C. *et al.* Down Regulation of the Expression of ELMO3 by COX2 Inhibitor Suppresses Tumor Growth and Metastasis in Non-Small-Cell Lung Cancer. *Front. Oncol.* **9**, 363 (2019).

96. Tsironis, G. *et al.* Axitinib as a third or further line of treatment in renal cancer: a single institution experience. *BMC Urol.* **20**, 60 (2020).

97. Tan, F. H., Putoczki, T. L., Stylli, S. S. & Luwor, R. B. Ponatinib: a novel multi-tyrosine kinase inhibitor against human malignancies. *OncoTargets Ther.* **12**, 635–645 (2019).

98. Hohn, D. C. *et al.* A randomized trial of continuous intravenous versus hepatic intraarterial floxuridine in patients with colorectal cancer metastatic to the liver: the Northern California Oncology Group trial. *J. Clin. Oncol.* **7**, 1646–1654 (1989).

99. Hecht, M. *et al.* Efavirenz Has the Highest Anti-Proliferative Effect of Non-Nucleoside Reverse Transcriptase Inhibitors against Pancreatic Cancer Cells. *PLOS ONE* **10**, e0130277 (2015).

100. Hevener, KirkE., Verstak, T. A., Lutat, K. E., Riggsbee, D. L. & Mooney, J. W. Recent developments in topoisomerase-targeted cancer chemotherapy. *Acta Pharm. Sin. B* **8**, 844–861 (2018).

101. Reale, A. *et al.* TOP2A expression predicts responsiveness to carfilzomib in myeloma and informs novel combinatorial strategies for enhanced proteasome inhibitor cell killing. *Leuk. Lymphoma* **62**, 337–347 (2021).

102. Goldspiel, B. R. & Kohler, D. R. Flutamide: an antiandrogen for advanced prostate cancer. *DICP Ann. Pharmacother.* **24**, 616–623 (1990).

103. Regan, J. L. *et al.* c-Kit is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. *Oncogene* **31**, 869–883 (2012).

104. Stankov, K., Popovic, S. & Mikov, M. C-KIT signaling in cancer treatment. *Curr. Pharm. Des.* **20**, 2849–2880 (2014).

105. Montero, J. C., López-Pérez, R., San Miguel, J. F. & Pandiella, A. Expression of c-Kit isoforms in multiple myeloma: differences in signaling and drug sensitivity. *Haematologica* **93**, 851–859 (2008).

106. Liu, F. *et al.* Axitinib overcomes multiple imatinib resistant cKIT mutations including the gatekeeper mutation T670I in gastrointestinal stromal tumors. *Ther. Adv. Med. Oncol.* **11**, 1758835919849757 (2019).
107. Garner, A. P. *et al.* Ponatinib Inhibits Polyclonal Drug-Resistant KIT Oncoproteins and Shows Therapeutic Potential in Heavily Pretreated Gastrointestinal Stromal Tumor (GIST) Patients. *Clin. Cancer Res.* **20**, 5745–5755 (2014).
108. Rini, B. I. *et al.* Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **380**, 1116–1127 (2019).
109. Ma, Y. *et al.* Antitumor effect of axitinib combined with dopamine and PK-PD modeling in the treatment of human breast cancer xenograft. *Acta Pharmacol. Sin.* **40**, 243–256 (2019).
110. Shamroe, C. L. & Comeau, J. M. Ponatinib: A new tyrosine kinase inhibitor for the treatment of chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia. *Ann. Pharmacother.* **47**, 1540–1546 (2013).
111. Sarin, V. *et al.* Evaluating the efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic correlation analysis. *Leukemia* **34**, 2754–2765 (2020).
112. Rossi, M. *et al.* Mouse models of multiple myeloma: technologic platforms and perspectives. *Oncotarget* **9**, 20119–20133 (2018).
113. Fiscon, G., Conte, F., Amadio, S., Volonté, C. & Paci, P. Drug Repurposing: A Network-based Approach to Amyotrophic Lateral Sclerosis. *Neurother. J. Am. Soc. Exp. Neurother.* **18**, 1678–1691 (2021).

SUPPLEMENTARY INFORMATION

All the supplementary information cited throughout this thesis is available in the Github repository created for this TFM at https://github.com/appropiate/TFM_UoC, and consist of:

- **Supplementary information 1**
  Full list of drug candidates for repurposing in MM generated by SAveRUNNER when using a network based on blood related diseases:
  Drug_Disease_network_blood.txt

- **Supplementary information 2**
  Full list of drug candidates for repurposing in MM generated by SAveRUNNER when using a network based on MM symptom related diseases:
  Drug_Disease_network_symptoms.txt

- **Supplementary information 3**
  Full list of drug candidates for repurposing in ocular albinism generated by SAveRUNNER when using a network based on random diseases (negative control for previous MM based networks generated by SAveRUNNER) :
  Drug_Disease_network_control.txt

- **Supplementary information 4**
  Full report and pipeline with code to do GEA for the GSE175384 dataset:
  Carracedo_Huroz_Sergio_DEGs_in_MM_GSE175384.html

- **Supplementary information 5**
  Full list of DEGs with log2FC > 1 and padj < 0.05 for the GSE175384 dataset:
  Annotated_DEGs_Healthy_vs_MM_Log2FC_1_Padj0.05_GSE175384.csv

- **Supplementary information 6**
  Full report and pipeline with code to do GEA for the GSE47552 dataset:
  Carracedo_Huroz_Sergio_DEGs_MM_GSE47552.html

- **Supplementary information 7**
  Full list of DEGs with log2FC > 1 and padj < 0.05 for the GSE47552 dataset:
  Annotated_DEGs_Healthy_vs_MM_Log2FC_1_Padj0.05_GSE47552.csv

- **Supplementary information 8**
  Full list of compounds in CMap database for GSE175384 gene signature:
  query_result_CMap_DEGs_GSE175384.gct

- **Supplementary information 9**
  Full list of compounds in CMap database for GSE47552 gene signature:
  query_result_Cmap_DEGs_GSE47552.gct

- **Supplementary information 10**
  Full list of compounds totally validated by DEAs/CMap analyses:
  totally_validated_drug_candidates.csv

- **Supplementary information 11**
  Full list of compounds partially validated by DEAs/CMap analyses:
  Partially_validated_drug_candidates_new2.csv

- **Supplementary information 12**
  Molecular docking analyses of the active site 1 of the KIT's kinase domain as protein with the candidates for reposition to MM ponatinib and axitinib as ligands.