REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

NOVEL METHODS FOR DRIVER GENE AND ANTI-CANCER PHARMACOTHERAPY PRIORITIZATION IN PERSONALIZED ONCOLOGY

EGE ÜLGEN, M.D.

DOCTORATE THESIS

DEPARTMENT OF BIOSTATISTICS AND MEDICAL INFORMATICS

SUPERVISOR
Prof. Osman Uğur Sezerman

ISTANBUL – 2022

REPUBLIC OF TURKEY
ACIBADEM MEHMET ALİ AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

NOVEL METHODS FOR DRIVER GENE AND ANTI-CANCER PHARMACOTHERAPY PRIORITIZATION IN PERSONALIZED ONCOLOGY

EGE ÜLGEN, M.D.

DOCTORATE THESIS

DEPARTMENT OF BIOSTATISTICS AND MEDICAL INFORMATICS

SUPERVISOR
Prof. Osman Uğur Sezerman

ISTANBUL – 2022

Department:	Department of Biostatistics and Medical Informatics
Program:	Biostatistics and Bioinformatics
Thesis Title:	Novel Methods for Driver Gene and Anti-Cancer Pharmacotherapy Prioritization in Personalized Oncology
Student's Name and Surname:	Ege Ülgen
Date of Defense:	11/04/2022

This is to certify that I have examined this copy of Ph.D. thesis. I have found that he prepared after fulfilling the specified requirements in the associated legislations before the final examining committee, whose signatures are below.

Jury Member (Head of the Defense)	Title, Name Surname Instituition	Signature
Jury Member (Thesis Supervisor)	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature
Jury Member	Title, Name Surname Instituition	Signature

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

Date: 11/04/2022

Ege Ülgen

PREFACE AND ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor Prof. Osman Uğur Sezerman, Ph.D., who has been the best supervisor and mentor I could ask for with immense patience, motivation, enthusiasm, and excellent knowledge and guidance. He has been nothing but supportive and constructive and has helped move my career always forward since my first day.

For their invaluable evaluation, I would like to thank the jury members and substitutive jury members for my dissertation, Prof. Koray Özduman, MD, Assoc. Prof. Burcu Güngör, Ph.D., Prof. Uğur Özbek, MD, Asst. Prof. A. Ercüment Çiçek, Ph.D., Assoc. Prof. Emel Timuçin, Ph.D., Assoc. Prof. Can Alkan, Ph.D.

Since my first years in medical school, I have had the privilege to work alongside Prof. Koray Özduman, MD. I'm incapable of finding the words to express my gratitude for his mentorship. Throughout my academic journey, he has been the best role model, the wisest scientific mentor, in addition to being one of the kindest people I know.

I would like to express my deepest gratitude for my colleague, my friend Ozan Özışık, Ph.D. Without our extensive brainstorming sessions and his endless patience in responding to my inquiries, it would have been near impossible for me to finish this dissertation.

For her support during the preparation of this thesis and for answering my countless questions, I would like to thank Assoc. Prof. Emel Timuçin.

I would like to thank the mischievous duo Deniz Ece Kaya and Nogayhan Seymen for all the fun as well as the insightful discussions we shared. This journey would have been unbearable without you.

I feel grateful to have worked with many Sezermanlab friends and alumni and lucky to call them friends. For supporting me on every turn of this journey, I'd like to

thank Sıla Gerlevik, Umut Gerlevik, Berk Gürdamar, Tuğçe Bozkurt, Begüm Özemek Güner, Rüçhan Ekren, Pınar Altıner, Ceren Saygı, Hüseyin Okan Soykam, Aslı Yenenler Kutlu, Zehra Hazal Sezer, Tayyip Karaman, Narod Kebabçı, Eray Şahin.

For their never-ending, unconditional support, I would like to thank my parents, Ayla and Berkant. I would not be able to achieve any of my accomplishments without your boundless efforts.

To Kıvanç, we might not remember how we exactly met, but the memories we collected along the way will remain. Thank you for always being there and putting up with all my anxiety.

This dissertation was supported by the TUBITAK 2244 project 118C039.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

α	alpha
AUC	Area Under the Curve
AUROC	Area Under the Receiver Operating Characteristic curve
CaDrA	Candidate Driver Analysis
cBioPortal	cBio Cancer Genomics Portal
CCLE	Cancer Cell Line Encyclopedia
CGC	Cancer Gene Census
CGP	Cancer Genome Project
CIViC	Clinical Interpretation of Variants in Cancer
COSMIC	Catalogue of Somatic Mutations in Cancer
CTD	Comparative Toxicogenomics Database
DEG	Differentially Expressed Gene
DoCM	Database of Curated Mutations
GDSC	Genomics of Drug Sensitivity in Cancer
iCAGES	integrated CAncer GEnome Score
ICGC	International Cancer Genome Consortium
indel	Insertion/Deletion
IntOGen	The Integrative OncoGenomics platform
KEGG	Kyoto Encyclopedia of Genes and Genomes
λ	Lambda
LOTUS	Learning Oncogenes and TUmor Suppressors
MCR	Minimal Common Region
MEMo	Mutual Exclusivity Modules in Cancer
MTL	Multi-Task Learning
MuSiC	Mutational Significance in Cancer
MutSigCV	Mutation Significance Covariates
NGS	Next-Generation Sequencing
OncoVar	Oncogenic driver Variants
PANACEA	PersonAlized Network-based Anti-Cancer therapy EvaluAtion
Phenolyzer	Phenotype Based Gene Analyzer
PIN	Protein Interaction Network
PRODIGY	Personalized Ranking of DrIver Genes analYsis
ReDO_DB	Repurposing Drugs in Oncology Database
ROC	Receiver Operating Characteristic
RWR	Random Walk with Restart
SCNA	Somatic Copy Number Alteration
SCS	Single-sample Controller Strategy
SVM	Support Vector Machine
TALC	Targeted Agents in Lung Cancer

TARGET	Tumor Alterations Relevant for GEnomics-driven Therapy
TCGA	The Cancer Genome Atlas
TEND	Trends in the Exploitation of Novel Drug targets
TieDie	Tied Diffusion through Interacting Events
WES	Whole-Exome Sequencing
WGS	Whole-Genome Sequencing

LIST OF FIGURES

LIST OF TABLES

ÖZET

Kişiselleştirilmiş Onkolojide Sürücü Gen ve Anti-Kanser Farmakoterapi Önceliklendirilmesi İçin Yeni Yöntemler

Kişiselleştirilmiş onkolojideki en önemli zorluklar arasında (i) bir tümör örneğindeki sürücü genlerin tespiti ve (ii) bu klinik açıdan alakalı sürücülerin rasyonel, etkin ve kişiselleştirilmiş tedavi önerilmesinde kullanılması sayılabilir. Bu tezde, bahsi geçen zorlukların üstesinden gelebilmek amacıyla yeni sürücü gen ve ilaç önceliklendirme metotları geliştirilmiştir. Sürücü gen önceliklendirme metodu, driveR, genomik veri ve biyolojik bilgi birikimi harmanlamakta ve her bir gene sürücülük olasılığı atamak amacıyla birçok görevli öğrenme sınıflandırma modeli kullanmaktadır. Model, hem tümör kohortlarında hem de bireysel tümör örneklerinde 21 farklı kanser türü için sürücü genleri tespit edebilmektedir. Bu yeni yönteminin, önceki sürücü önceliklendirme yöntemlerinden daha iyi performans gösterdiği gösterilmiştir. Bunu takiben, driveR kullanılarak elde edilmiş sürücü önceliklendirme sonuçlarıyla birlikte ilaç-gen etkileşimleriyle genişletilmiş bir etkileşim ağını kullanarak, PANACEA olarak isimlendirilen bireysel tümör örneklerinde kanser önleyici ilaçları önceliklendirmek için yeni ağ tabanlı yöntemler geliştirilmiştir. "Mesafeye dayalı" yöntem, sürücülük olasılıkları ve ağ üzerindeki ilaçlar ve genler arasındaki mesafeyi kullanarak ilaçları skorlamakta ve önceliklendirmektedir. "Sürücü yayılımı" yöntemi, aynı amaçla yeniden başlatma ile rastgele yürüyüş algoriması yoluyla sürücülük olasılıklarını yaymaktadır. Sürücülük olasıklarına ek olarak, diferansiyel olarak eksprese edilmiş gen (DEG) bilgisini yayan "sürücü ve DEG yayılımı" yöntemi, bir ilacın puanını, sürücü yayılma puanının ve DEG yayılma puanının çarpımı olarak belirlemektedir. Titiz bir değerlendirmeyle, bu yeni yöntemlerin onkoterapiyle ilgili ve potansiyel olarak etkin ilaçlara öncelik verdiği gösterilmiştir. Bu tezde önerilen yeni yöntemlerin kişiselleştirilmiş onkoloji araştırmalarını ilerletmesi umulmaktadır.

Anahtar Sözcükler

Neoplazmlar, Sürücü gen, Farmakoterapi, Önceliklendirme metodu, Genom bilimi

ABSTRACT

Novel Methods For Driver Gene And Anti-Cancer Pharmacotherapy Prioritization In Personalized Oncology

Two of the greatest challenges in personalized oncology are (i) identification of driver genes in a tumor sample and (ii) using these clinically relevant drivers to recommend rational, effective, and personalized treatments. In this dissertation, novel methods for driver gene and drug prioritization methods were developed to address these challenges. The driver gene prioritization method, driveR, incorporates genomics data and prior biological information and uses a multi-task learning classification model to assign driverness probabilities to each somatically altered gene. The model can detect driver genes for 21 cancer types in both tumor cohorts and individual tumor samples. The novel method driveR was shown to outperform previous driver prioritization methods. Successively, utilizing driver prioritization results from driveR and an interaction network extended by drug-gene interactions, novel network-based methods for prioritizing anti-cancer drugs for individual tumor samples, PANACEA, were developed. The "distance-based" method uses driverness probabilities and distances between drugs and altered genes in the network to score and prioritize the drugs. The "driver propagation" method propagates driverness probabilities via a random walk with restart framework for the same purpose. Additionally propagating differentially expressed gene (DEG) information, the "driver and DEG propagation" method determines the score of a drug as the product of its driver propagation score and its DEG propagation score. Through meticulous evaluation, it was demonstrated that these novel methods prioritize potentially effective drugs relevant to oncotherapy. The novel methods proposed in this thesis will hopefully advance personalized oncology research.

Keywords

Neoplasms, Driver Gene, Pharmacotherapy, Prioritization method, Genomics

1. INTRODUCTION and AIM

Cancer is a heterogeneous collection of diverse diseases characterized by numerous “Hallmarks of Cancer” acquired during their development (1-3). Cancer has a substantial genetic component. Somatic alterations acquired throughout an individual’s lifetime are the foundation of oncogenesis and cancer progression (4). Therefore, during the past decade, the advancement of next-generation sequencing technologies ushered a transformation in the oncology field by enabling the identification of crucial genetic alterations, genes, and oncogenic processes involved in various types of cancer (5-7). Significantly, with the aid of these technologies, individualized analysis of cancer genome sequences was made possible, and the novel field of personalized oncology was born with the intent of more precise, personalized diagnoses and tailored treatments (8-11).

While cancers contain large amounts of somatic alterations, most alterations are “passengers” developing merely due to genomic instability in the course of oncogenesis. In contrast, a minority of somatic alterations are “drivers” involved in oncogenesis and cancer progression (12-14). Identification of drive genes harboring such somatic alterations is crucial in the setting of personalized oncology because accurate detection of driver genes in a tumor sample will result in more precise diagnosis, additionally informing the clinician to potentially identify personalized therapeutic targets (15, 16). As discussed in the “Background” section, another substantial challenge in personalized oncology is the identification of suitable pharmacotherapy options based on somatic genomic events. The methods proposed in this thesis aim to address these challenges in personalized oncology.

The first aim of this thesis was to devise a method for personalized prioritization of cancer driver genes (Figure 1 top). For this purpose, a novel approach, driveR, enabling prioritization of altered genes in both cohorts of tumor samples and a single tumor sample, was developed (17). This method utilizes the genomic landscape and prior biological knowledge to prioritize driver genes via a multi-task learning model.

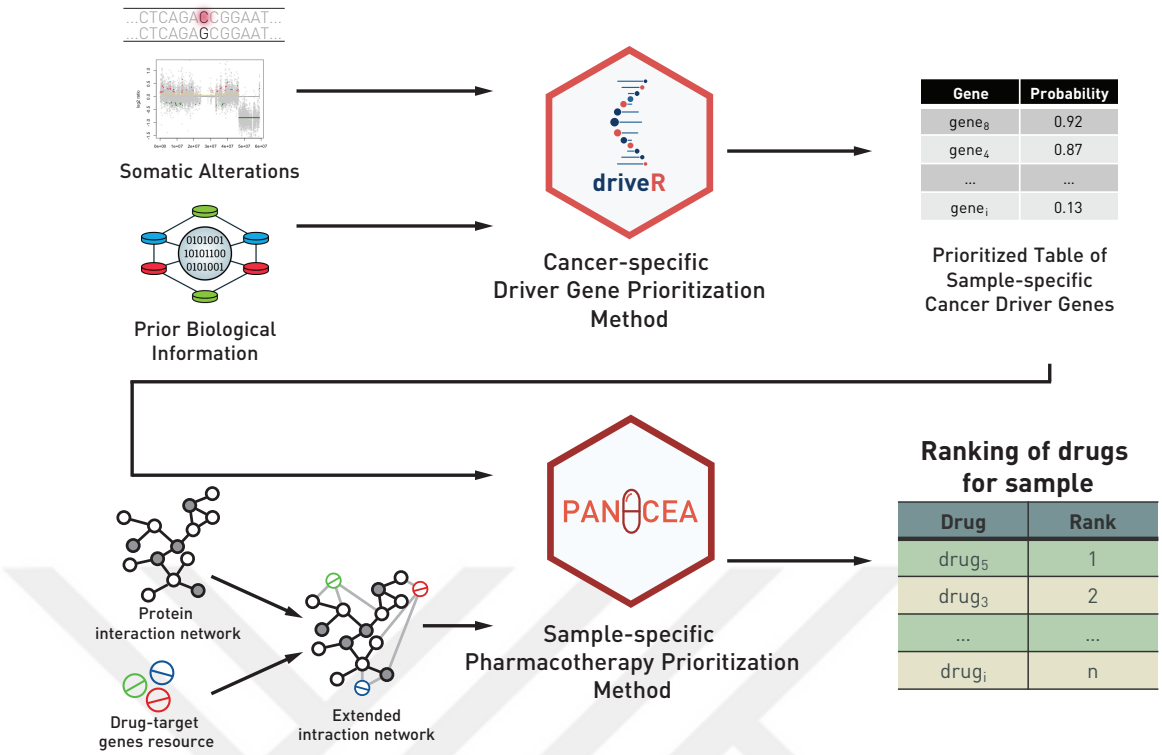


Figure 1: Overview of the personalized driver gene and anti-cancer pharmacotherapy prioritization methods proposed in this thesis

The second aim of this thesis was to devise network-oriented methods for personalized prioritization of anti-cancer pharmacotherapy options (Figure 1 bottom). For this purpose, various network-based methods, collectively named PANACEA, were developed. These methods enable the prioritization of drugs from a curated source, using the driver gene prioritization results and a protein interaction network extended with curated drug-gene interactions. The "distance-based" method scores each drug based on its distance to somatically altered genes and the genes' driverness probabilities. The "driver propagation" method scores drugs by propagating the driverness probabilities of altered genes. Lastly, the "driver and DEG propagation" method scores drugs by propagating the driverness probabilities of altered genes as well the $|\log_2\text{-fold-change}|$ values of differentially expressed genes in the sample.

2. BACKGROUND

2.1. Cancer

Cancer is a collection of many distinct diseases characterized by relatively unrestrained proliferation of cells that can invade beyond normal tissue boundaries and metastasize to distant organs. It causes about 1 in every six deaths and is the second-leading cause of death worldwide (18).

Numerous capabilities are acquired during the development of cancers. Hanahan and Weinberg suggested the following hallmarks of cancer to provide a framework for better understanding the complex biology of cancer: sustaining proliferative signaling, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, tumor-promoting inflammation, activating invasion & metastasis, inducing angiogenesis, genome instability & mutation, resisting cell death, and deregulating cellular energetics (1, 2). By 2022, Hanahan expanded these hallmarks to include the following emerging characteristics: unlocking phenotypic plasticity, nonmutual epigenetic reprogramming, senescent cells, and polymorphic microbiomes (3). Studying these molecular and cellular traits characterizing cancer allows a better understanding of the disease and the development of novel cancer treatments.

There is a broad consensus that cancer is, in essence, a genetic disease (4). Cancer develops due to the accumulation of changes in the DNA sequence of the cancer cells' genomes. In other words, somatic mutations acquired during an individual's lifetime are the basis of cancer formation and progression. These somatic mutations may include several different types of change, namely: nucleotide substitutions, small insertion or deletions (indels), rearrangements, and copy number alterations (amplifications or deletions). Additionally, the cancer genome may have acquired entirely new DNA sequences, for example, from viruses such as the human papillomavirus, Epstein Barr virus, hepatitis B virus, human T lymphotropic virus 1, and human herpesvirus 8 (19). Epigenetic alterations, altering chromatin structure and gene expression, also play essential roles in the development and progression of cancers (20, 21).

In the last decade, the progress of DNA sequencing technologies provided new insights into cancer and transformed the field of oncology (22). To name a few benefits: (i) tumor taxonomies are moving from a histologic- to a genetic-based level, (ii) these new technologies identified somatic genetic alterations that have been shown to be highly sensitive biomarkers for cancer detection and monitoring, (iii) some genetic alterations have been demonstrated to be legitimate targets for therapy, and (iv) these technologies allowed individualized analysis of cancer genomes, resulting in more precise diagnosis and tailored treatments.

2.2. Analysis of cancer genomes

The completion of the Human Genome Project in 2004, which provided the full sequence-based map of the normal human genome, ushered significant advancements in all fields of biomedical science (23). As cancer is essentially a genetic disease, oncology was one of the first to be affected by this pivotal advancement. Facilitated by the availability of the human genome sequence, several large-scale projects, including whole-genome or whole-exome sequencing of cancer samples, were undertaken. These large-scale projects necessitated the development of new massive parallel sequencing methods, known as Next-Generation Sequencing (NGS) technologies (24-26). These NGS approaches lowered the cost and time required for analysis and are more sensitive than Sanger sequencing, detecting even lower-frequency mutations in subclones. These developments generated an extraordinary throughput in cataloging somatic mutations in cancer.

The repertoire of oncogenic mutations is exceptionally heterogeneous (27). Therefore, it is difficult for independent initiatives to generate comprehensive catalogs of oncogenic mutations in a broad spectrum of distinct cancers. Accordingly, several different networks to coordinate cancer genome sequencing projects worldwide were established, including the Cancer Genome Project (CGP) of the Wellcome Trust Sanger Institute, the Cancer Genome Atlas (TCGA), and the International Cancer Genome Consortium (ICGC). These coordinated projects provided new insights into the catalog of oncogenic mutations oncogenic signaling pathways and unveiled

specific signatures of mutagenesis (27-31). Additionally, these cancer-genome studies contributed to defining clinically relevant subtypes to better estimate prognosis and inform treatment (32-36).

The first analysis of an individual cancer genome by Pleasance et al. identified the comprehensive catalog of somatic mutations in a malignant melanoma sample and a lymphoblastoid cell line from the same person (37). Following this milestone, after the NGS methods have become more advanced and more quickly and affordably accessible, leveraging knowledge obtained through large-scale projects, individualized analysis of cancer genomes has become routine in the clinical setting (38-40). Currently, various NGS approaches, occasionally a combination of different approaches, are being used in oncology (41, 42). The most widely utilized techniques are targeted sequencing, whole-exome sequencing (WES), and whole-genome sequencing (WGS), each with its advantages and limitations (11, 43, 44). Target sequencing panels are designed to investigate curated cancer-related alterations, provide excellent depth, and are well-suited for working with formalin-fixed paraffin-embedded tumor samples, but they are only limited to a few targeted alterations (45, 46). WES and WGS provide more comprehensive genomics data, identifying all potentially relevant alterations (somatic mutations, copy number alterations) in the coding region and the whole genome, respectively. This more comprehensive genomics data, in turn, allows for the assessment of additional genomics information, including the direct measurement of tumor mutational burden and the exploration of the signatures of mutational processes, and the evaluation of microsatellite instability (28, 47-51). WES is an especially appealing choice in the clinical setting because its cost is lower than WGS, its computational analysis and interpretation are less challenging, and most clinically relevant alterations lie within the exome.

The most challenging aspect of WES/WGS of cancer genomics analysis is the clinical interpretation of findings, akin to “finding a needle in a haystack” (52). Several approaches have been put forward to organize, prioritize, and report such findings (53-57). Most approaches utilize databases that collect knowledge about cancer genes and alterations and their relationships to oncogenesis and clinical utility and present

findings in tiers according to clinical relevance. These approaches generally follow specific standards and guidelines to interpret and report alterations (44, 58).

2.3. Cancer driver genes

Cancer genomes frequently contain considerable numbers of mutations. Most genes with somatic mutations are “passengers”, not implicated in cancer formation or progression but instead arising due to genomic instability. On the other hand, a small proportion is “drivers” involved in oncogenesis or cancer progression, driving cancer evolution by giving cells a selective advantage (12-14).

Identifying cancer driver genes is a pressing challenge for personalized oncology because accurate identification of personalized driver genes will result in a precise understanding of oncogenesis and aid the clinicians to potentially devise personalized therapeutic approaches (15, 16).

Several computational approaches have been developed to identify driver genes in cancer cohorts (59, 60). Some noteworthy approaches for these batch analyses include MuSiC, MutSigCV, MutPanning, MEMo, Hierarchical HotNet, TieDie, DriverNet, CaDrA, OncodriveFML, and LOTUS (61-70). MuSiC uses overall mutation rate, pathway mutation rate, and association with clinical features to detect drivers. Given background mutation processes, MutSigCV determines potential driver genes that were mutated more frequently than expected. In addition to such “traditional” features, MutPanning establishes distinctive sequence contexts around passenger mutations and uses deviation from these contexts to prioritize drivers. Utilizing a network-based method, MEMo aims to identify subnetworks of genes involved in the same pathway with mutual exclusivity. Another method that uses protein interaction networks (PINs) is Hierarchical HotNet, which identifies a hierarchy of driver subnetworks containing frequently mutated genes. TieDie incorporates interaction and transcriptomics data and applies heats diffusion to find overlapping subnetworks exhibiting high mutation rates and expression values. Considering driver genes’ effect on expression, DriverNet aims to identify driver

genes by assessing genes whose alterations are linked to genes with dysregulated expression levels in the PIN. Essentially a heuristic approach, CaDrA utilizes a stepwise search method to identify subsets of alterations that are maximally associated with an outcome (such as drug sensitivity, expression, pathway/gene set activity) and likely drivers. OncodriveFML determines drivers by analyzing the patterns of mutations across tumor samples to identify positive selection signals bias. LOTUS utilizes mutation frequency, functional impact, and pathway-based information and applies a single- and multi-task learning algorithm to identify drivers.

Of interest, there are various platforms containing curated drivers for multiple cancer cohorts. Liu et al. analyzed several cancer cohorts using some of the batch driver identification methods above and established a database named DriverDBv3 (71). OncoVar, an integrated database and analysis platform, also utilized some of the methods discussed above to identify driver mutations, genes, and pathways (72). Another example is the Integrative OncoGenomics (IntOGen) platform, which compiles alterations, genes, and pathways implicated in oncogenesis (73). The latest release (Feb 1, 2020) identified cancer drivers analyzing 221 tumor cohorts by combining seven different driver identification methods.

As reviewed above, there are many methods for driver identification for cohorts of tumor samples, i.e., batch analysis approaches. However, methods for driver identification for individual tumor samples, i.e., personalized analysis approaches, are still in their infancy. Such personalized approaches are crucial for several reasons: (i) some tumors do not harbor any alterations in any known driver genes, which may necessitate the identification of putative driver genes, (ii) because some tumors have alterations in several known driver genes, identification of the actual driver or drivers for the particular tumor, (iii) because the number of therapeutical agents that can be safely administered simultaneously to a patient is limited due potential toxicity and adverse events, driver genes, alterations of which can suggest response or resistance to such treatments, need to be prioritized (74, 75).

Below, some personalized analysis approaches for driver identification are briefly described. DawnRank prioritizes altered genes in a single sample via a PageRank algorithm (76). Utilizing a directed PIN, the single-sample controller strategy (SCS) determines a set of altered genes interacting with downstream differentially expressed genes (77). By incorporating genomics information and prior biological knowledge, iCAGES aims to identify drivers considering the coding, non-coding, and structural variants (78). Using the prize-collecting Steiner tree approach, PRODIGY investigates an individual tumor's mutation and expression profiles, using known pathways and PIN information to quantify the impacts of altered genes on every altered pathway (79).

2.4. Personalized Pharmacotherapy Prioritization in Cancer

Another one of the most significant challenges in personalized oncology is the identification of the most rational, effective, and tailored treatment option based on the genomic landscape of the tumor (80). However, computational approaches for personalized prioritization of anti-cancer pharmacotherapies are lacking.

In their pioneering 2015 study, Rubio-Perez et al. proposed an in-silico drug prescription strategy, revealing that while only a minority of cancers were treatable by strictly following clinical guidelines, many more could benefit from their proposed strategy (81). The in-silico drug prescription strategy involved finding drivers per tumor and evaluating their druggability options, considering direct targeting, indirect targeting, and gene therapies. This study was one of the first to demonstrate the potential benefit of utilizing genomics data for personalized anti-cancer pharmacotherapy identification.

Also discussed in the previous subsection, iCAGES is a statistical framework analyzing personal cancer genomics data to prioritize driver variants, driver genes, and potentially effective drugs (78). For drug prioritization, iCAGES calculates drug scores integrating (i) the gene scores estimated in the previous step of the framework, (ii) a relatedness score from the BioSystems database, and (iii) the marginal activity score of the drug from the PubChem database (82, 83).

Another personalized drug prioritization approach was proposed by PanDrugs, which utilizes (i) a gene score, evaluating the biological significance of the gene in cancer and its clinical importance, and (ii) a drug score estimating drug response and suitability (84).

Of note, interest in network-based methods for drug repositioning purposes in cancer has been increasing in the past years (85). Moreover, network-based methods have been successfully utilized to infer personalized drug target genes and prioritize candidate drugs against specific types of cancer (86, 87). Therefore, it can be proposed that by extending the use of potential drug targets beyond direct interactions, network-oriented approaches can be utilized for devising effective personalized drug prioritization methods.

3. MATERIALS and METHODS

The purpose of this dissertation was to develop novel methods for (i) personalized driver gene prioritization and (ii) sample-specific anti-cancer pharmacotherapy prioritization. To address these aims, firstly, a metapredictor model for assessing the impact of coding variants was created. Then, a model-based approach utilizing this metapredictor along with other genomics data and prior biological information was developed. Lastly, network-based methods were devised and evaluated using driver prioritization results and curated drug-gene interactions. The details of each method are described below.

3.1. Coding Variant Impact Metaprediction

A metapredictor model that utilizes impact scores from multiple tools was established to estimate the probability of damaging impact of coding variants. For this purpose, multiple machine learning models were trained and assessed. The final metapredictor model was selected as the model with the highest performance. It was

later used to generate one of the features of the multi-task learning classification model for driver gene prioritization.

As the dataset for training and testing the coding impact metapredictor models, the benchmarking dataset from a study by Martelotto et al. was obtained (88). ANNOVAR was then used to annotate these benchmarking variants with precalculated predictions from twelve impact predictors using dbNSFP v3.0 (89, 90). These 12 impact predictor tools were: SIFT, PolyPhen-2 (HumDiv scores), LRT, MutationTaster, Mutation Assessor, FATHMM, GERP++, PhyloP, CADD, VEST, SiPhy, and DANN (91-102). Variants with missing predictor scores or the "uncertain" label were excluded. The processed dataset contained 814 "non-neutral" and 135 "neutral" variants.

To assess each predictor's predictive strength and to detect any collinearity issue, pairwise Pearson correlations between all predictors' scores as well as the outcome variable were evaluated. Moreover, the distributions of all 12 predictors' scores by outcome (either "neutral" or "non-neutral") were assessed to detect any outliers or similar issues.

The training dataset was formed by randomly selecting 75% of both "neutral" and "non-neutral" variants from the dataset. The remaining 25% was used as the test dataset.

Finally, six classification models were built and evaluated: (i) logistic regression, (ii) naïve Bayes, (iii) support vector machine (SVM) with linear kernel, (iv) SVM with radial kernel, (v) random forest, and (vi) gradient boosting machine. Training each model involved using 10-fold 3-times-repeated cross-validation, maximizing the area under the receiver operating characteristic curve (AUROC) metric. Each model was evaluated using the test dataset.

3.2. Driver gene prioritization

Figure 2 displays the overview diagram of how the driver gene prioritization model was established and how its performance was assessed. Briefly, data were obtained for 21 The Cancer Genome Atlas (TCGA) datasets available on the International Cancer Genome Consortium (ICGC) data portal (103). This data was split into the training (75% of patients) and test (25% of patients) datasets. Additional test datasets were obtained from various resources. Using the training data, a multi-task learning (MTL) classification model was trained. The performance of this model was then compared using the test datasets. The performance of the novel approach was also compared with other widely utilized methods.

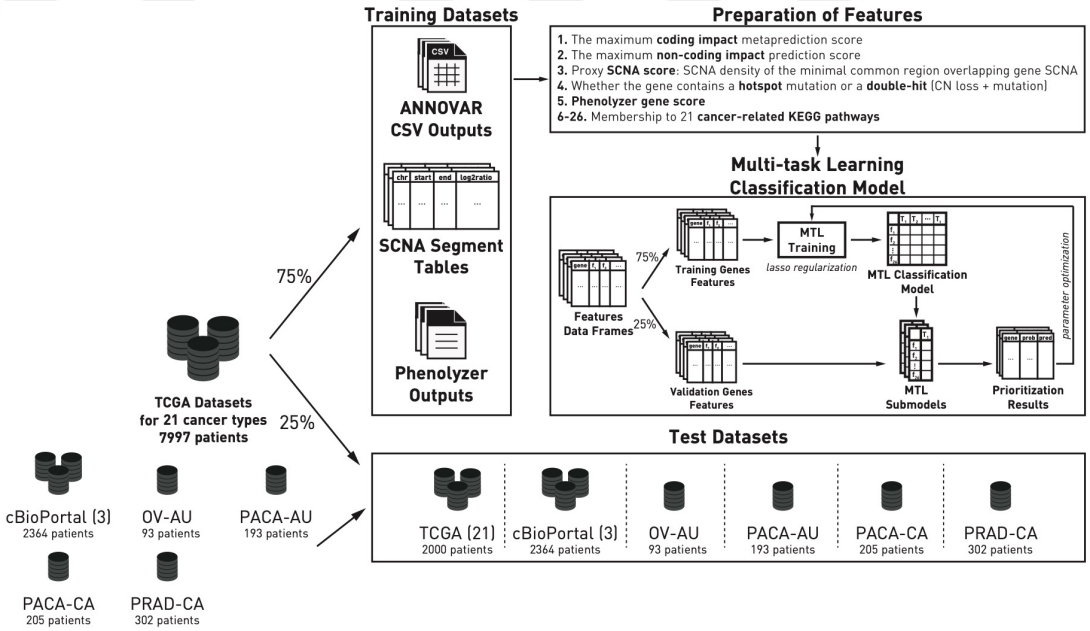


Figure 2. The overview of the approach for driver gene prioritization.

Detailed information on the datasets, model training, and performance evaluation are provided in the following subsections.

3.2.1. Training and testing data

Table 1 displays information on the datasets used for training and testing the MTL classification model.

Table 1. Training and testing datasets for the multi-task learning classification model.

Type	Dataset ID	Cancer Type	Source	Original Source
Test	BLCA-US	Bladder Urothelial Cancer	ICGC Data Portal Release 28	TCGA, US
Test	BRCA_METABRIC	Breast Cancer	cBioPortal - Datasets	METABRIC
Test	BRCA-US	Breast Cancer	ICGC Data Portal Release 28	TCGA, US
Test	CESC-US	Cervical Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	COAD_CPTAC	Colon Adenocarcinoma	cBioPortal - Datasets	CPTAC-2 Prospective
Test	COAD-US	Colon Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	GBM-US	Brain Glioblastoma Multiforme	ICGC Data Portal Release 28	TCGA, US
Test	HNSC-US	Head and Neck Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	KIRC-US	Kidney Renal Clear Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	KIRP-US	Kidney Renal Papillary Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	LAML-US	Acute Myeloid Leukemia	ICGC Data Portal Release 28	TCGA, US
Test	LGG-US	Brain Lower Grade Glioma	ICGC Data Portal Release 28	TCGA, US
Test	LIHC-US	Liver Hepatocellular carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	LUAD_ONCOSG	Lung Adenocarcinoma	cBioPortal - Datasets	OncoSG
Test	LUAD-US	Lung Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	LUSC-US	Lung Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	OV-AU	Ovarian Cancer	ICGC Data Portal Release 28	AU
Test	OV-US	Ovarian Serous Cystadenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	PAAD-US	Pancreatic Cancer	ICGC Data Portal Release 28	TCGA, US
Test	PACA-AU	Pancreatic Cancer	ICGC Data Portal Release 28	AU
Test	PACA-CA	Pancreatic Cancer	ICGC Data Portal Release 28	CA
Test	PRAD-CA	Prostate Adenocarcinoma	ICGC Data Portal Release 28	CA
Test	PRAD-US	Prostate Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	READ-US	Rectum Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	SKCM-US	Skin Cutaneous melanoma	ICGC Data Portal Release 28	TCGA, US
Test	STAD-US	Gastric Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Test	THCA-US	Head and Neck Thyroid Carcinoma	ICGC Data Portal Release 28	TCGA, US
Test	UCEC-US	Uterine Corpus Endometrial Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	BLCA-US	Bladder Urothelial Cancer	ICGC Data Portal Release 28	TCGA, US
Training	BRCA-US	Breast Cancer	ICGC Data Portal Release 28	TCGA, US
Training	CESC-US	Cervical Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	COAD-US	Colon Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	GBM-US	Brain Glioblastoma Multiforme	ICGC Data Portal Release 28	TCGA, US
Training	HNSC-US	Head and Neck Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	KIRC-US	Kidney Renal Clear Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	KIRP-US	Kidney Renal Papillary Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	LAML-US	Acute Myeloid Leukemia	ICGC Data Portal Release 28	TCGA, US

Training	LGG-US	Brain Lower Grade Glioma	ICGC Data Portal Release 28	TCGA, US
Training	LIHC-US	Liver Hepatocellular carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	LUAD-US	Lung Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	LUSC-US	Lung Squamous Cell Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	OV-US	Ovarian Serous Cystadenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	PAAD-US	Pancreatic Cancer	ICGC Data Portal Release 28	TCGA, US
Training	PRAD-US	Prostate Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	READ-US	Rectum Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	SKCM-US	Skin Cutaneous melanoma	ICGC Data Portal Release 28	TCGA, US
Training	STAD-US	Gastric Adenocarcinoma	ICGC Data Portal Release 28	TCGA, US
Training	THCA-US	Head and Neck Thyroid Carcinoma	ICGC Data Portal Release 28	TCGA, US
Training	UCEC-US	Uterine Corpus Endometrial Carcinoma	ICGC Data Portal Release 28	TCGA, US

As the first step, all genomics data for all 21 TCGA cohorts (of different cancer types) available on the ICGC data portal (release 28) was obtained. This data consisted of somatic mutation and somatic copy number alteration (SCNA) data. Next, genomics data for 75% of patients in each cohort were randomly selected as the training data, whereas data for the remaining 25% of patients in each cohort were designated as the test data. Further, as additional test data, genomics data for three different cohorts were obtained from the cBio Cancer Genomics Portal (cBioPortal) (104). Moreover, genomics data for four different cohorts were obtained from ICGC. The cBioPortal cohorts were BRCA-METABRIC (breast cancer), COAD-CPTAC (colorectal adenocarcinoma), and LUAD-ONCOSG (lung adenocarcinoma) (105-107). The additional ICGC cohorts were PACA-AU (pancreatic cancer), PACA-CA (pancreatic cancer), OV-AU (ovarian cancer), and PRAD-CA (prostate adenocarcinoma).

3.2.2. Features for training the driver gene classification model

The 723 experimentally validated driver genes curated by the Cancer Gene Census (CGC, v92) were defined as true driver genes (108). During training, these were labeled as “drivers”, and the remaining genes were labeled “non-drivers”. The driver gene classification model utilized twenty-six features to predict whether a gene with any somatic alteration is a driver or not. These features were generated using annotations obtained via ANNOVAR, SCNA data, and Phenotype Based Gene

Analyzer (Phenolyzer) gene scores (90, 109). In the following subsections, these features are described.

3.2.2.1. The maximum coding variant impact score

To assess the functional impact of coding variants, the coding variant impact metapredictor model trained and selected as described in section 3.1 was used. The maximum metaprediction score of all coding variants in a gene was used as one of the features of the driver gene classification model.

3.2.2.2. The maximum non-coding variant impact score

The Phred-scaled CADD scores were used to estimate the functional impact of non-coding variants (99). The maximum score of all non-coding variants in a gene was used as one of the features of the driver gene classification model.

3.2.2.3. Proxy SCNA score

To assess the impact of SCNA events in a gene, a proxy score was used as one of the features of the driver gene classification model.

The minimal region of SCNAs across several cancer samples under examination is called a Minimal Common Region (MCR) (110). To assign proxy scores to each gene with SCNA, pan-cancer MCR data was obtained from a study by Kim et al., in which SCNAs in a large number of cancer genomes were analyzed (111). The MCR data consisted of SCNA density (i.e., SCNA/Megabase). The proxy score was assigned as the SCNA density of the MCR region which a gene overlaps and has the same direction of SCNA event as the MCR region.

The gene-level SCNA events were established from the segment-level SCNA data: first, the overlap between an SCNA segment and each gene was assessed. If a segment had at least 25% overlap with a gene, the gene-level SCNA was determined.

Next, the $\log_2$ (tumor/normal depth) ratio for a gene with a called SCNA event was assigned as the maximum $|\log_2$ ratio$|$ value among all segments overlapping it. Gene-level SCNA events that are on sex chromosomes or with $|\log_2$ ratio$| < 0.25$ were excluded.

3.2.2.4. Hotspot or double-hit gene condition

Used as evidence of oncogene behavior, hotspot mutations were identified using Catalogue of Somatic Mutations in Cancer (COSMIC) v92 data (112). Hotspot mutations were defined as mutations with an occurrence greater than five in COSMIC.

Additionally, used as evidence of tumor suppressor gene behavior, a simultaneous non-synonymous somatic mutation and a homozygous somatic copy number loss was defined as a double-hit event. Homozygous somatic copy number loss was defined as an SCNA event with a $\log_2$ ratio < -1.

Whether or not a gene contained a hotspot or double-hit event was used as one of the features of the driver gene classification model.

3.2.2.5. Phenotype based gene analyzer score

All genes with somatic alterations were scored utilizing prior biological knowledge on the specific cancer type via Phenotype Based Gene Analyzer (Phenolyzer) (109). These cancer-type-specific scores were used as one of the features of the driver gene classification model.

3.2.2.6. Membership to cancer-related pathways

The final set of features of the driver gene classification model were memberships to a selected set of cancer-related Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (113). The cancer-associated KEGG pathways were chosen as 21 pathways related to “Pathways in cancer” (Table 2).

Table 2. Cancer-associated pathways related to KEGG hsa05200 - Pathways in cancer.

KEGG ID	Description
hsa03320	PPAR signaling pathway
hsa04010	MAPK signaling pathway
hsa04020	Calcium signaling pathway
hsa04024	cAMP signaling pathway
hsa04060	Cytokine-cytokine receptor interaction
hsa04066	HIF-1 signaling pathway
hsa04110	Cell cycle
hsa04115	p53 signaling pathway
hsa04150	mTOR signaling pathway
hsa04151	PI3K-Akt signaling pathway
hsa04210	Apoptosis
hsa04310	Wnt signaling pathway
hsa04330	Notch signaling pathway
hsa04340	Hedgehog signaling pathway
hsa04350	TGF-beta signaling pathway
hsa04370	VEGF signaling pathway
hsa04510	Focal adhesion
hsa04512	ECM-receptor interaction
hsa04520	Adherens junction
hsa04630	JAK-STAT signaling pathway
hsa04915	Estrogen signaling pathway

3.2.3. Training the MTL classification model

As mentioned above, an MTL model was built for driver gene classification. For this purpose, the R package RMTL was used (114). The MTL model was trained with a sparse structure (using lasso regularization) framework (115). The RMTL algorithm uses the framework provided in Equation 1:

$$\min_{W} \sum_{i=1}^{t} L(W_i|X_i, Y_i) + \lambda_1||W||_1 + \lambda_2||W||_F^2 \quad (1)$$

In this classification case, *L()* is the logistic loss function. There are *t* tasks (in this specific case, different cancer types). *W* is the coefficient matrix where W_i is the i^{th}

column of W, referring to the coefficient vector of task i. X is the set of predictor matrices of the t tasks, where X_i is the predictor matrix of task i. Y is the set of response vectors (in this specific case, "driver" vs. "non-driver"), where Y_i is the response vector of task i. $||W||_1$ indicates the L_1 norm of, and $||W||_F$ indicates the Frobenius norm of the coefficient matrix W. The penalty term λ_1 controls the effect of cross-task regularization and λ_2 is used to improve generalization performance by stabilizing the numerical results.

The initial training dataset described above was further divided into training genes and validation genes by randomly allocating 75% of somatically altered genes into training and the remaining 25% into validation datasets. The optimal λ_2 was determined to be 10^{-4} via the validation dataset performance of different options. The final classification model was built using the optimal λ_2 of 10^{-4} and the optimal λ_1 of 10^{-5} (determined via 10-fold cross-validation).

The final MTL driver gene classification model estimates the probability of being a driver gene per cancer type. Cancer-type-specific thresholds were determined as the probability value maximizing precision on the corresponding test dataset. A sample-specific threshold was also assessed for personalized analyses, defined as the 99th percentile of all predicted probabilities).

3.2.4. Assessment of performance

As described above, the MTL driver gene classification model was devised to analyze both tumor cohorts and individual tumor samples. Hence, the performance of the model was evaluated in both settings.

Each subtask of the MTL model corresponded to a different cancer type. The batch analysis performance was evaluated by obtaining the AUROC value per test dataset using the corresponding subtask. Further, the numbers of predicted driver genes per test dataset were determined using different thresholds of probabilities of being a driver gene. Moreover, the percentage of true drivers (i.e., genes curated by CGC) per cancer type was evaluated among all predicted driver genes (per different thresholds). Similarly, the percentage of actionable genes per cancer type was assessed using the tumor alterations relevant for the genomics-driven therapy (TARGET) database, containing a curated gene set of 135 clinically actionable genes (53).

The personalized analysis performance of the MTL classification model was evaluated by calculating the AUROC per test sample using the corresponding subtask. Once again, the numbers of predicted driver genes (per different thresholds) per test sample were determined, and the percentages of true drivers and actionable genes among the predicted drives were evaluated.

3.2.5. Comparison with other batch analysis approaches

By comparing AUROC values on the test datasets, the performance of the novel driver gene prioritization approach was compared with other batch analysis

approaches, namely, MutSigCV, DriverNet, OncodriveFML, and MutPanning (62, 63, 67, 69).

MutSigCV (version 1.3.5) and MutPanning (version 2.0) analyses were executed with the default settings on the GenePattern platform (116). DriverNet (version 1.28.0) analyses were conducted with the BioGRID Homo sapiens PIN (version 4.0.189) with the default settings (117, 118). OncodriveFML (version 2.3.0) analyses were executed using the default settings.

3.2.6. Comparison with other personalized analysis approaches

Additionally, the personalized analysis performance of the novel driver gene prioritization approach was compared with other personalized analysis approaches, namely, DawnRank and PRODIGY (76, 79). For this purpose, the AUROC values for the test samples were compared.

As both DawnRank and PRODIGY require expression data for normal tissue, for personalized analyses, only 16 datasets were used: BLCA-US (bladder cancer), BRCA-US (breast cancer), CESC-US (cervical squamous cell carcinoma), COAD-US (colon adenocarcinoma), HNSC-US (head and neck squamous cell carcinoma), KIRC-US (kidney renal clear cell carcinoma), KIRP-US (kidney renal papillary cell carcinoma), LIHC-US (liver hepatocellular carcinoma), LUAD-US (lung adenocarcinoma), LUSC-US (lung squamous cell carcinoma), PAAD-US (pancreatic cancer), PRAD-US (prostate adenocarcinoma), READ-US (rectum adenocarcinoma),

STAD-US (gastric adenocarcinoma), THCA-US (thyroid cancer), and UCEC-US (uterine corpus endometrial carcinoma).

PRODIGY (version 1.0) analyses were executed with the default settings using the STRING Homo sapiens PIN (version 11.0) and curated KEGG pathways (119). For comparison, PRODIGY's gene scores per pathway were aggregated into a single gene score. DawnRank (version 1.2) analyses were executed using the default settings with the BioGRID Homo sapiens PIN (version 4.0.189).

3.3. Sample-specific prioritization of anti-cancer pharmacotherapy

One of the aims of this thesis was to devise methods for personalized prioritization of anti-cancer pharmacotherapy options. For this purpose, two network-oriented approaches, described in detail below, were designed and evaluated. Each approach begins by assigning the "driverness" probabilities (the probability that the given gene is a driver gene, estimated via driveR v0.3.0) of all altered genes in the sample onto a PIN that is extended by curated drug-target interactions (Figure 3). The methods are collectively named PANACEA: PersonAlized Network-based Anti-Cancer therapy EvaluAtion.

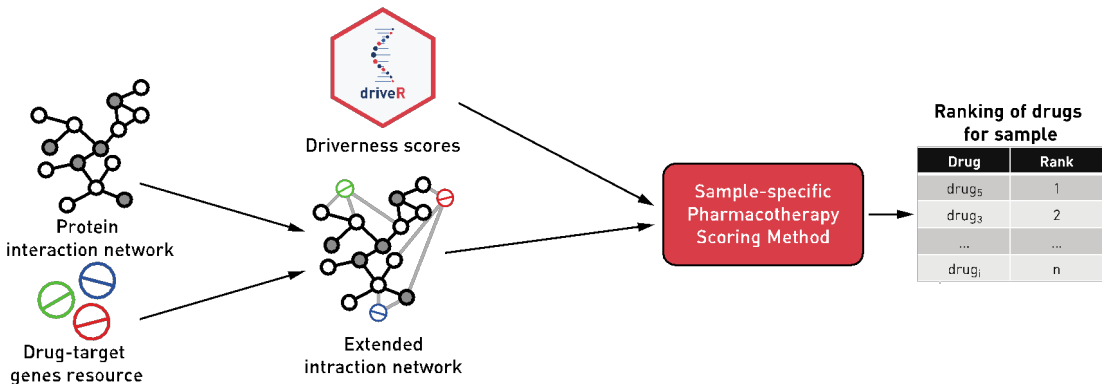


Figure 3. Overview of the sample-specific anti-cancer pharmacotherapy scoring approaches

The performances of PANACEA methods were assessed using cancer data for (i) The Cancer Genome Atlas (TCGA), (ii) Genomics of Drug Sensitivity in Cancer (GDSC), and (iii) two cancer cases with clinically demonstrated drug sensitivity (120-122). The evaluation schemes are described in detail below.

3.3.1. Protein Interaction Network and Extension by Drugs

All sample-specific anti-cancer pharmacotherapy scoring methods were performed using the STRING v11.5 PIN (119). Only interactions with a combined score of greater than 0.4 were kept for moderate confidence. All TCGA analyses were additionally performed and assessed using the BioGRID 4.4.205 PIN to illustrate that the methods are robust to the change of PIN (118).

The methods integrate drug-target gene information into analysis by extending the PIN with drugs, by adding drugs as nodes connected to their target genes. This will be referred to as the "extended PIN". In the Genomics of Drug Sensitivity in Cancer (GDSC) data analyses, the PIN was extended with the drugs available in the GDSC data. In the remaining analyses on TCGA data and the two cases with known drug response, the PIN was extended using drug-gene interactions from expert-curated sources on DGIdb (May 2021 release): Cancer Commons, Cancer Genome Interpreter, ChEMBL Interactions, Clinical Interpretation of Variants in Cancer (CIViC), ClearityFoundationBiomarkers, ClearityFoundationClinicalTrial, COSMIC, Database of Curated Mutations (DoCM), My Cancer Genome, MyCancerGenomeClinicalTrial, Targeted Agents in Lung Cancer (TALC), TdgClinicalTrial, Trends in the exploitation of novel drug targets (TEND) (112, 123-131). Drugs targeting the same target genes were merged before scoring in all analyses.

3.3.2. Sample-specific drug scoring methods

3.3.2.1. Distance-based method

For the distance-based method, firstly, the driverness probabilities of all altered genes from driveR are mapped onto the extended PIN. A score incorporating the driverness probability and distance between each drug and each altered gene is calculated, as formulated in Equation 2.

$$s(d,g) = \frac{1}{(dist(d,g)+1)^2} driver_g \quad (2)$$

where *d* is a drug, *g* is an altered gene, *dist(d,g)* is the distance of the shortest path between *d* and *g*, and $driver_g$ is the driveR probability for *g*. In determining shortest paths, drug nodes are never used as intermediate nodes.

The aggregated score for the drug *d* is then calculated using Equation 3.

$$score(d) = \frac{1}{|genes|} \sum_{g \in genes} s(d,g) \quad (3)$$

where *genes* is the set of all altered genes with a driverness probability greater than 0.05 in the sample.

3.3.2.2. Driver propagation

As the first step, the driverness probabilities of all altered genes are mapped onto the extended PIN. Using the driverness probabilities as prior knowledge (*Y*), a random walk with restart (RWR) framework propagates this information. Before use, the adjacency matrix (*W*) of the extended PIN is normalized using Laplacian normalization as $W' = D^{-1/2} W D^{-1/2}$, where *D* is the diagonal matrix such that $D_{ii} = \sum_j W_{ij}$. The iterative procedure described below is used to compute the propagation scores *F*. Starting with $F^{(0)}=Y$, *F* is updated at each iteration *t* as follows:

$$F^{(t)} = (1-\alpha) W'^{F^{(t-1)}} + \alpha Y \quad (4)$$

where α is the restart parameter. The iterative procedure is repeated until the number of iterations exceeds 1000 or until convergence when the following is met:

$$\left\| F^{(t)} - F^{(t-1)} \right\|_2 < 10^{-4} \quad (5)$$

The score of a drug is its final propagation score.

The procedure for selecting the restart parameter α is described below in the "Analysis of TCGA data" subsection.

3.3.2.3. Driver and differentially expressed gene propagation

For the driver and differentially expressed gene propagation method, initially, the driverness probabilities of all altered genes are mapped onto the extended PIN. Separately, differentially expressed genes (DEGs) in the tumor are identified via DESeq2, comparing the expression values in the tumor against normal samples(132). DEGs were defined as genes with $p < 0.05$ and $|\log_2\text{-fold-change}| > 2$. Next, using the RWR framework described above, driverness probabilities and the $|\log_2\text{-fold-change}|$ values are propagated. The score of a drug is the product of its final propagation scores.

The procedure for selecting the restart parameter α is described below in the "Analysis of TCGA data" subsection.

3.3.3. Analysis of TCGA data

3.3.3.1. Selection of cohorts for analyses

To determine the propagation methods' restart parameter α and evaluate the methods, dissimilar datasets from TCGA were selected. First, somatic mutation, somatic copy number alteration, and transcriptomics data for all cohorts available on the ICGC data portal were obtained (103). The driverness probabilities per sample for each cohort were estimated using driveR (v0.3.0), and the median driveR probability per gene for each cohort was determined. Lastly, Spearman's rank correlation

coefficients of shared genes between each pair of cohorts were calculated to cluster the cohorts (Figure 4). The minimum correlation was observed between LAML-US and THCA-US. For the “Driver propagation” method, LAML-US (Acute Myeloid Leukemia, n=135) was used for parameter selection, while THCA-US (Thyroid Cancer, n=486) was used to evaluate the methods' performances. The next minimum correlation was observed between PRAD-US and THCA-US. Because LAML-US did not have more than one normal sample with expression data, for the “Driver and DEG propagation” method, PRAD-US (Prostate Adenocarcinoma) was used for parameter selection, while THCA-US (Thyroid Cancer) was used to evaluate the methods' performances. For all analyses using the “Driver and DEG propagation” method, only samples with complete data: somatic mutation, SCNA, and transcriptomics data (with greater than one normal sample with expression data in the cohort).

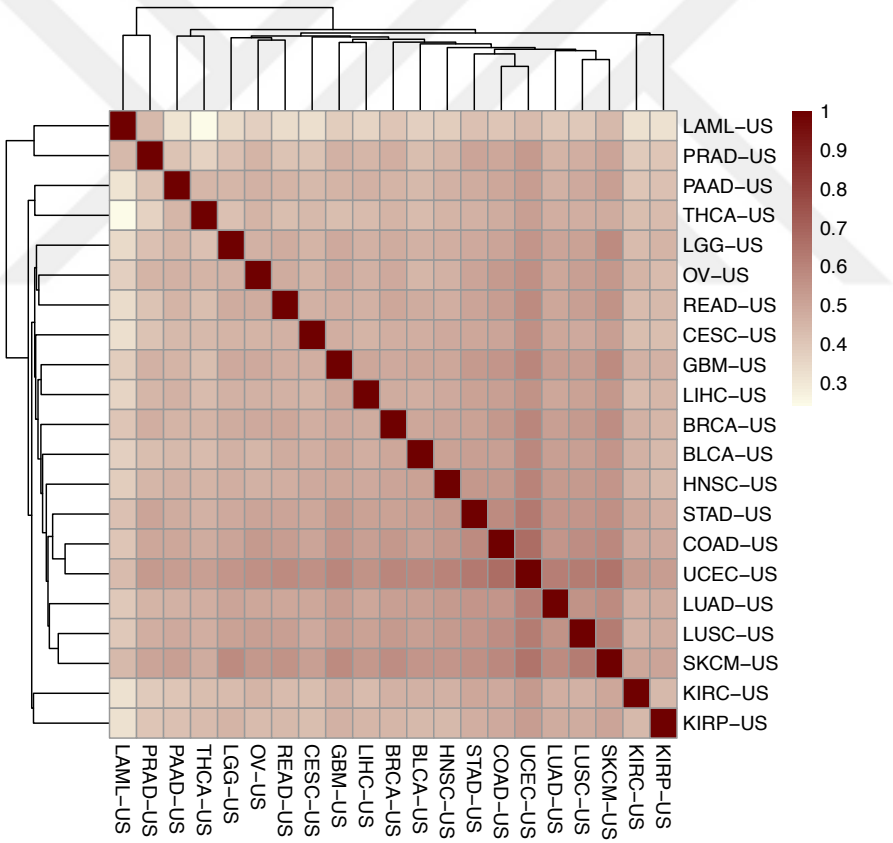


Figure 4. Clustered heatmap of the Spearman correlation coefficient between each pair of selected TCGA cohorts

3.3.3.2. Evaluation

For evaluating the relevancy of each method's ranking, each drug was classified into one of 10 tiers per the specific cancer type (Table 3). For tiers 1A/B and 2A/B, approved drugs and drugs in clinical trials were retrieved using the R package oncoPharmaDB (v0.6.6), which uses data from the Open Targets Platform (133). For tiers 3A/B and 4A/B, data for MeSH terms d002277 (carcinoma) and d009369 (neoplasm) were retrieved on Dec 16, 2021, from the Comparative Toxicogenomics Database (CTD) (134). For tier 5, drugs repurposed for cancer from the Repurposing Drugs in Oncology Database (ReDO_DB, 24 Nov 2021 update) were used (135). For all these sources, drug name matching with DGIdb was performed via the PubChem Identifier Exchange Service (https://pubchem.ncbi.nlm.nih.gov/idexchange/idexchange.cgi).

Table 3. Cancer-relatedness tiers for classifying drugs

Tier	Description
1A	approved for the cancer type under study
1B	approved for other cancer types
2A	in clinical trial(s) for the cancer type under study
2B	in clinical trial(s) for other cancer types
3A	Comparative Toxicogenomics Database (CTD) – curated for the cancer type under study
3B	CTD – curated for other cancer types
4A	CTD – inferred for the cancer type under study
4B	CTD – inferred for other cancer types
5	repurposed for cancer (in ReDO_DB)
6	No support

Per method, for each sample, the top 5, 10, and 25 drugs were identified, including drugs that have tying scores. After the tier of each drug was assigned, the proportions of drugs in each tier were calculated. Finally, these proportions were evaluated to estimate the support level of the drugs selected for each sample, comparing the cumulative proportion of Tiers 1-5 vs. Tier 6.

Moreover, the proportion of samples with any supported drugs directly targeting any driver genes in the sample was computed. This proportion was then compared with the proportion of supported drugs per each method.

3.3.3.3. Selection of the restart parameter

For the "Driver propagation" method, to determine the restart parameter α, the LAML-US cohort was used to rank drugs for each sample using α = 0.05, 0.1, 0.25, 0.5, 0.75, and 0.9. The same approach was performed for determining the restart parameter for the "Driver and DEG propagation" method. Next, the distributions of the supported drug proportions (Tiers 1-5) in the top 5, 10, and 25 were evaluated (Figure 5). The optimal α was determined to be 0.05 for both methods. This value of α was used for all analyses using the propagation methods.

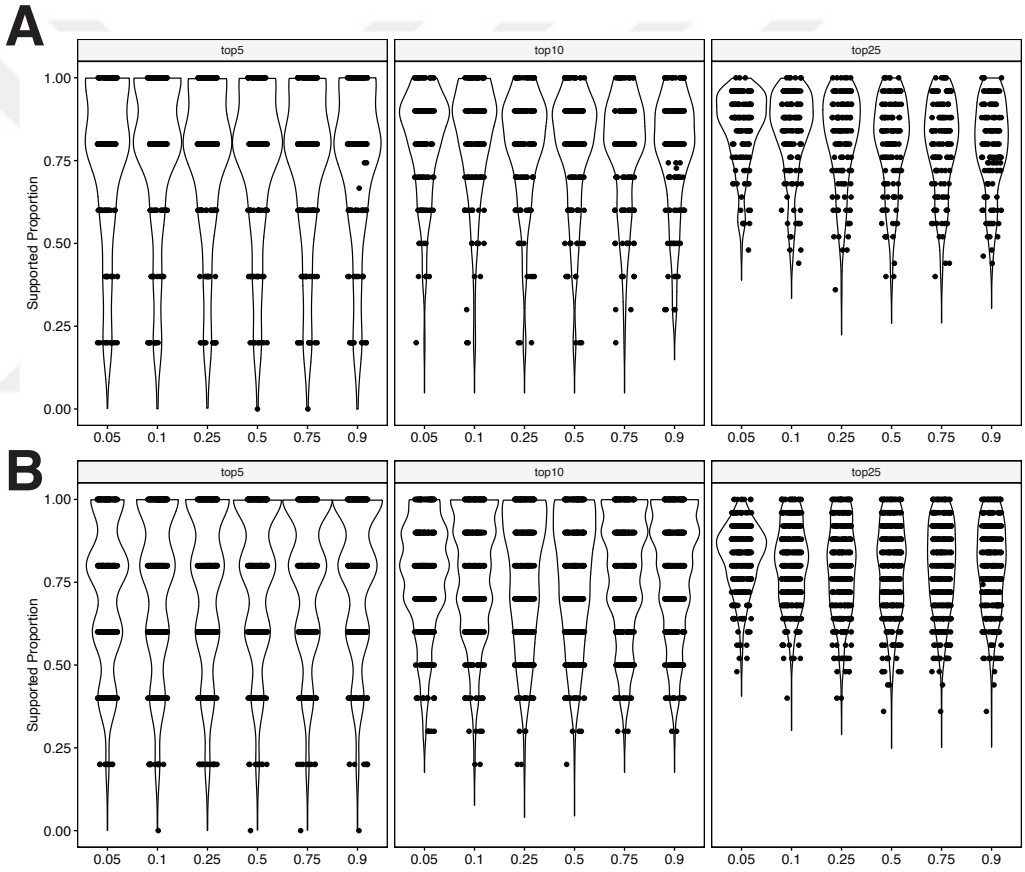


Figure 5. For the (A) Driver and (B) Driver and DEG propagation methods, violin plots displaying the proportion of selected drugs in tiers 1-5 for each choice of the restart parameter (α) per top 5, 10, and 25.

3.3.4. Analysis of GDSC data

3.3.4.1. Data

Dose-response data for Sanger GDSC1 and GDSC2 (Feb 2020) were obtained from the GDSC website (https://www.cancerrxgene.org/) (136, 137). From DepMap Public 21Q4 release, Cancer Cell Line Encyclopedia (CCLE) mutation, copy number alteration, expression, and sample information data were obtained (138). DGIdb was used as the drug-target genes resource. For GDSC drugs that were not in DGIdb, target genes were obtained from GDSC data. Only 515 cell lines of 18 cancer types suited for driveR analysis were included in the GDSC analyses. The "Driver and DEG propagation" method was performed on the 401 samples that had expression data available.

3.3.4.2. Evaluation

After each method was performed on the cell lines., the top 5, 10, and 25 drugs were identified. Per top n, the median and maximum Area Under the Curve (AUC) distributions for the drug-response curve per sample were evaluated.

3.3.5. Analyses of cases with clinically demonstrated drug response

For comparing the performance of the "Distance-based" and "Driver propagation" methods with existing methods, two cases with clinically demonstrated drug response were analyzed: (i) an Everolimus and Pazopanib sensitive bladder cancer case, and (ii) a lung adenocarcinoma case responsive to Sorafenib (121, 122).

Required data were obtained from the corresponding article for both cases. Using driveR results, the "Distance-based" and "Driver propagation" methods were performed to rank DGIdb drugs, and the ranks of the effective drugs were established. The "Driver and DEG propagation" method could not be performed on these cases because of missing expression data. Furthermore, drug prioritization results of iCAGES for both cases were acquired from the original article's supplementary material. Lastly, driveR results were saved as gene ranking files, and PanDrugs pan-

cancer and cancer-specific analyses were performed. The ranks of the effective drugs were compared between all novel and previous methods.

4. RESULTS

4.1. Coding Variant Impact Metaprediction

Prior to training the coding impact metapredictor models, pairwise Pearson correlations between the scores of the coding variant impact prediction tools as well as the outcome variable ("neutral" vs. "non-neutral" variant) were assessed. This revealed that each score provides a different facet of information for a coding variant's impact (Figure 6A). Further, the correlogram indicated no collinearity issue. The distributions of most predictor scores for "neutral" and "non-neutral" variants were evidently different (Figure 6B).

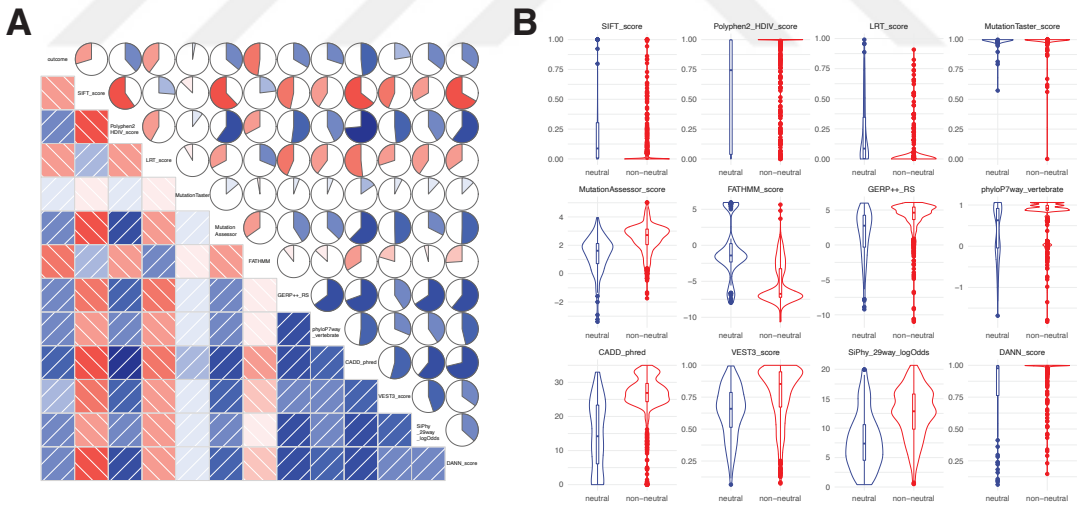


Figure 6. The overall analysis of individual variant impact predictors.

For the training and test datasets, the Receiver Operating Characteristic (ROC) curves and the AUROC values for each of the six coding variant impact metapredictor models are displayed in Figure 7.

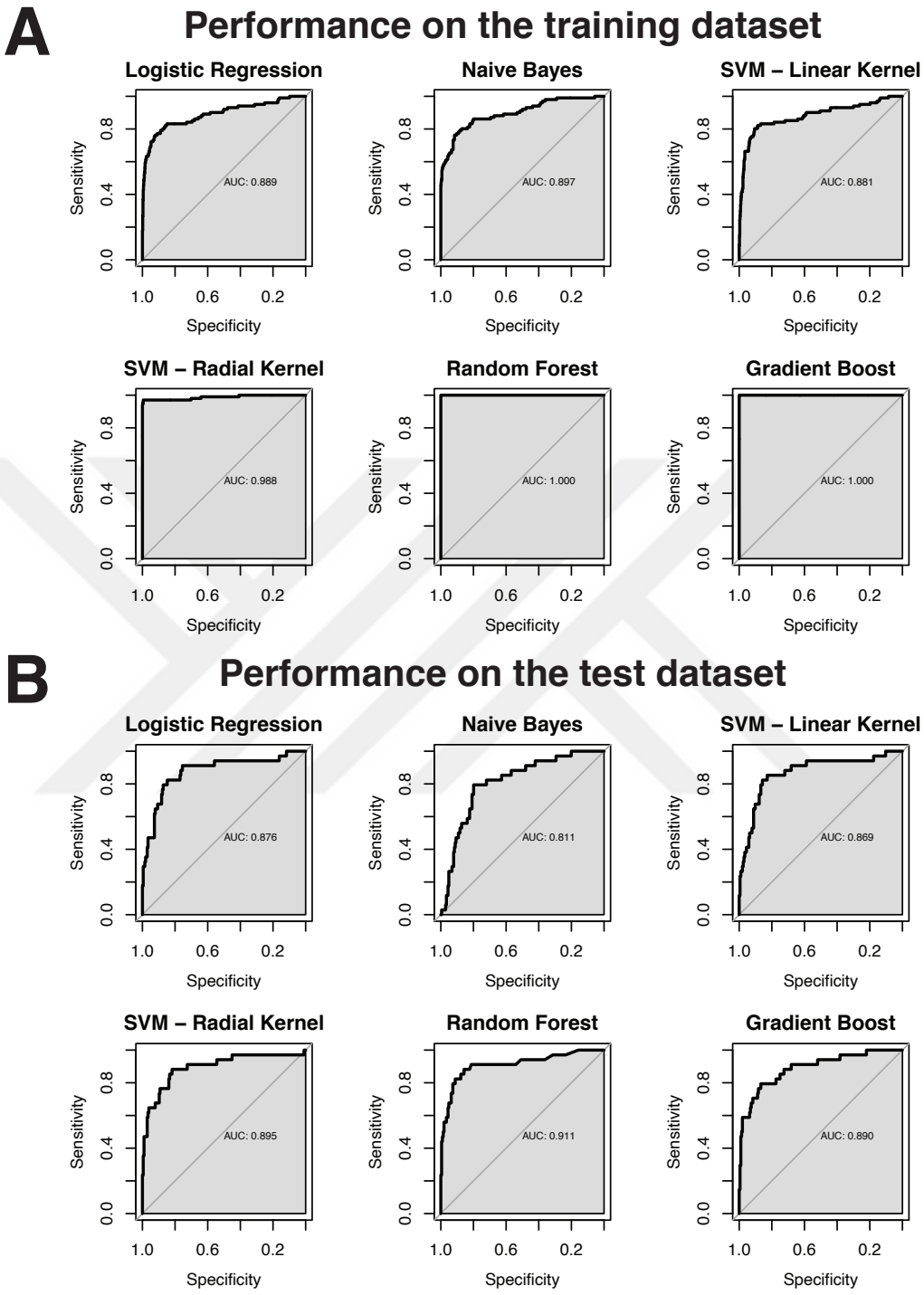


Figure 7. Performance of the coding variant impact metapredictor models on the training and test datasets

The random forest metapredictor displayed the best performance, achieving an AUROC of 0.911 (Figure 7). The random forest model also performed better than the individual impact prediction tools used to build the metapredictor models (Figure 8).

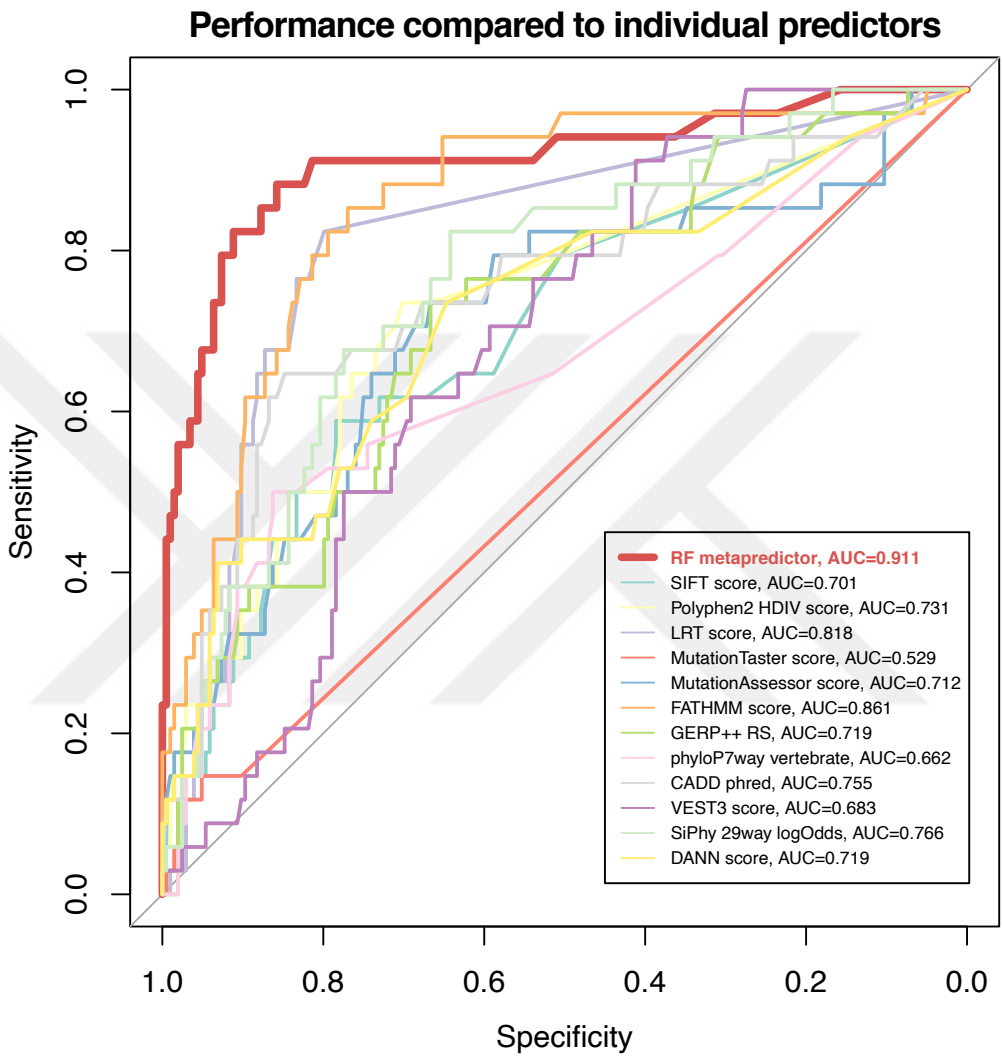


Figure 8. The performances of the coding variant impact metapredictor and individual predictors on the test dataset

4.2. Driver gene prioritization

As part of this thesis, a driver gene prioritization approach, driveR, that utilizes somatic genomics information incorporating prior biological knowledge was developed, and its performance was evaluated.

4.2.1. The MTL classification model

Based on the biclustered coefficient matrix of the MTL driver gene classification model that combines genomics data and prior biological information, the features were partitioned into 3 clusters, and the tasks (of different cancer cohorts) into 2 clusters (Figure 9).

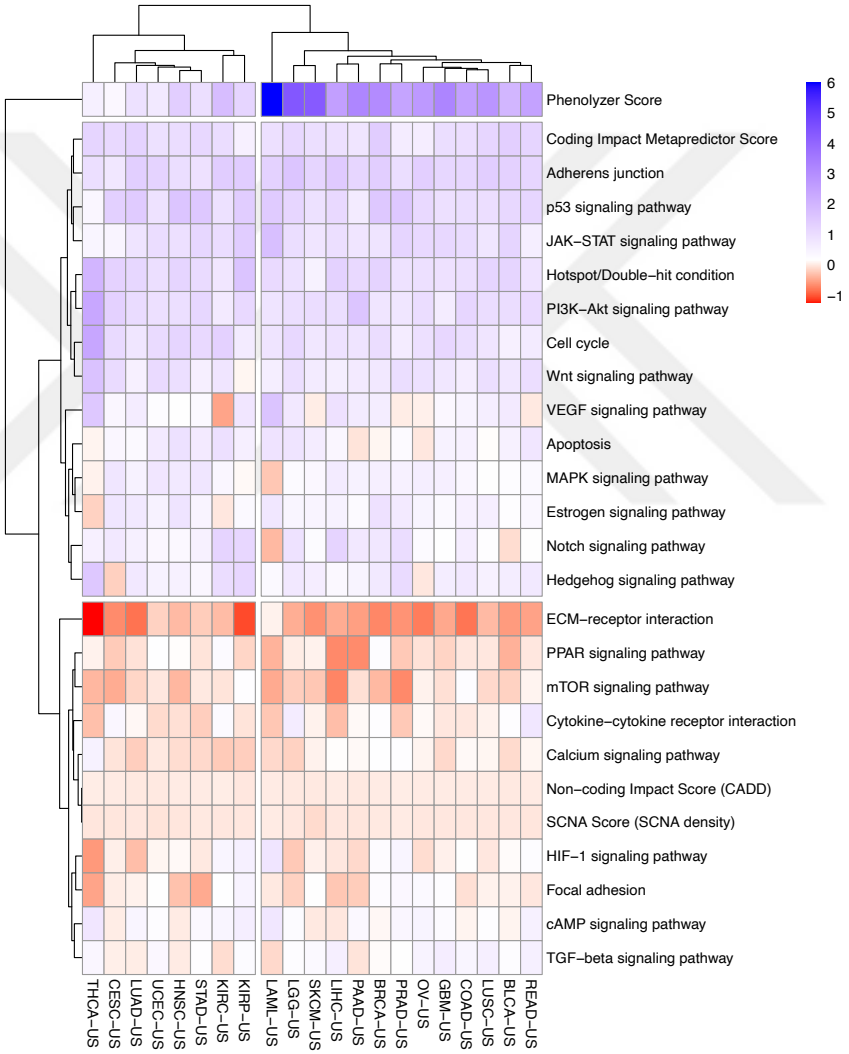


Figure 9. Biclustered heat map of the MTL model coefficient matrix. Rows are features, and columns are tasks (for specific cancer types)

As described above, based on biclustering, three clusters of features were identified: (i) a cluster containing “Phenolyzer Score”, (ii) a cluster of features with

positive coefficients, and (iii) another cluster of features with negative coefficients. The feature “Phenolyzer score” had relatively higher (and positive) coefficient estimates across all cancer types, underscoring the importance of prior biological knowledge in the MTL driver gene classification model.

The clustering of tasks (of different cancer types) was observed to be consistent as similar types of cancers (e.g., KIRC, KIRP; LGG, GBM; COAD, READ) were clustered together.

4.2.2. Performance of the MTL model

The batch analysis performance of the MTL model was assessed on the 28 test datasets (Figure 10). The median AUROC for batch analyses on the test datasets was 0.684 (range=0.651-0.861, Figure 10A). Per thresholds ranging from 0.3 to 0.7, the median number of predicted driver genes ranged from 13 to 42.5 (Figure 10B). Using cancer-type-specific thresholds (that maximized precision on the test datasets, as described in Materials and Methods), the median number of predicted driver genes was 9 (range=2-48). The median percentage percentages of true driver genes (as curated by CGC) among predicted driver genes per thresholds ranging from 0.3 to 0.7 were consistently above 60% (range = 63.84-94.59%, Figure 10C left panel). Similarly, the median percentage percentages of actionable genes (as curated by TARGET) among predicted driver genes per thresholds ranging from 0.3 to 0.7 ranged from 55.29% to 85.71% (Figure 10C right panel). Using cancer-type-specific thresholds, a median of 100% of predicted driver genes were found to be true driver genes (Figure 10C left panel), and a median of 84.2% of predicted driver genes were found to be actionable (Figure 10C right panel).

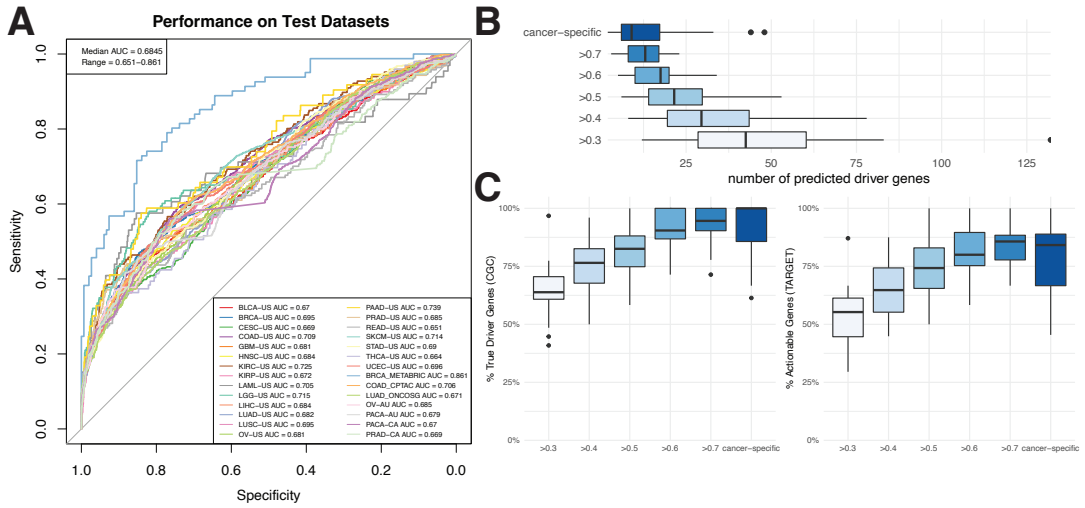


Figure 10. Batch analysis performance of driveR on 28 test datasets.

Personalized analyses of 5157 test patients using driveR resulted in an overall median AUROC of 0.773 (Figure 11A top panel, range = 0-1). Similarly, the median AUROC per test dataset was high, ranging from 0.66 to 1 (Figure 11B bottom panel). For all different choices of thresholds (values from 0.3 to 0.7, the sample-specific and the cancer-specific thresholds), the median number of predicted driver genes was 1 (Figure 11B). The median percentages of true driver genes and actionable genes among all predicted driver genes in each patient per thresholds ranging from 0.3 to 0.7, as well as per sample-specific and cancer-type-specific thresholds were 100% (Figure 11C), implying that any genes driveR predicts to be a driver gene in a tumor sample is most likely a true driver and actionable.

.

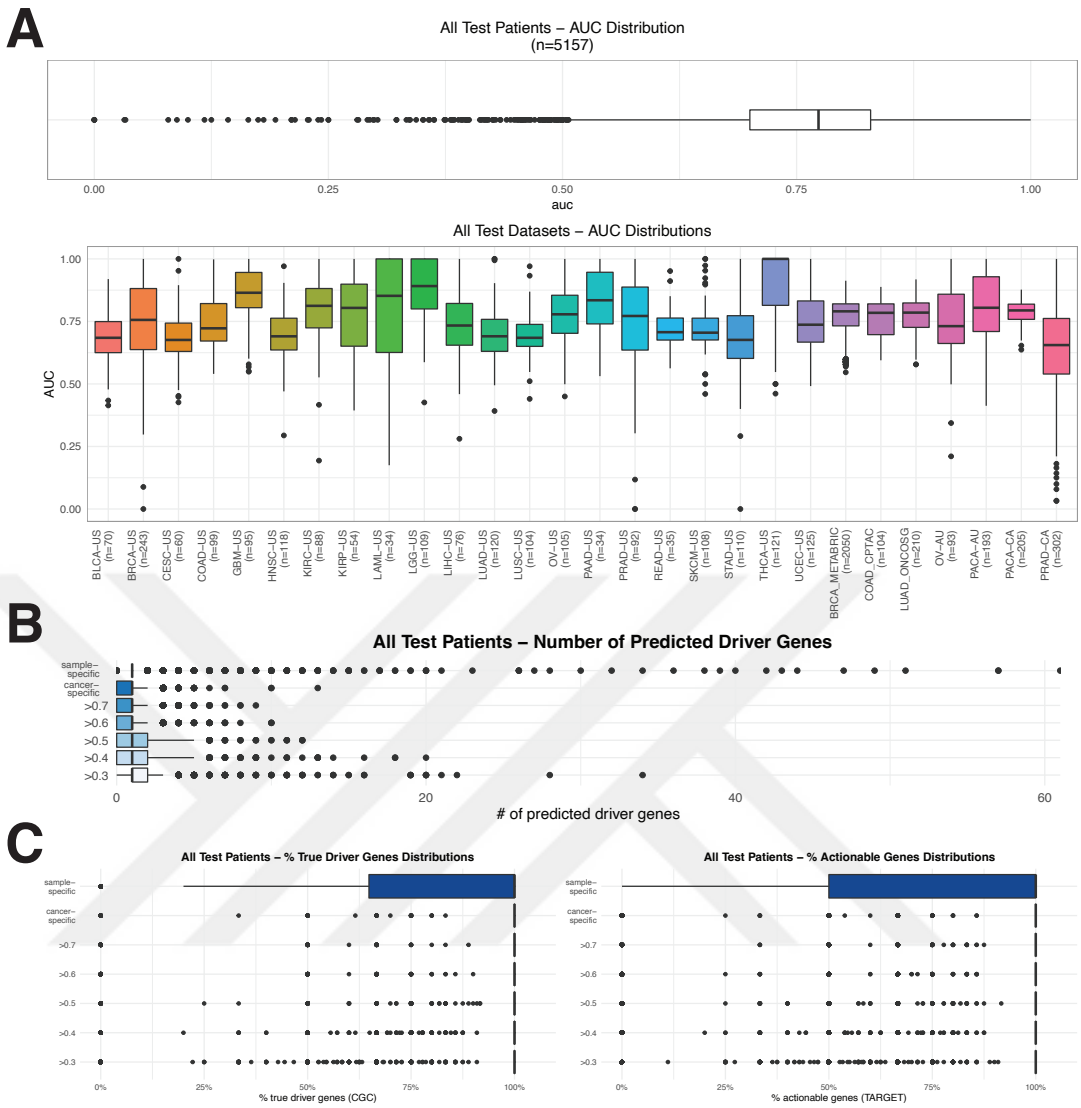


Figure 11. Personalized analysis performance of driveR on 5091 patients.

4.2.3. Comparison of performance

In the batch analysis of almost all test datasets, driveR outperformed other batch analysis methods (Figure 12).

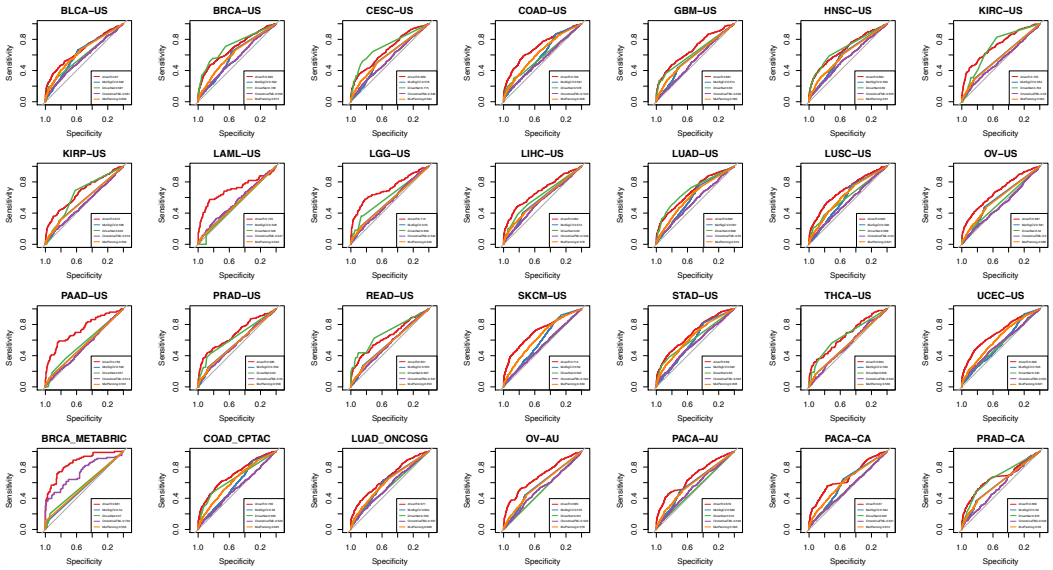


Figure 12. Comparison of performance of driveR with batch analysis approaches per test dataset.

The median AUROC of driveR across all test datasets (0.684) was significantly higher than all medians of MutSigCV (0.579, Wilcoxon rank-sum test $p < 0.001$), DriverNet (0.614, $p < 0.001$), OncodriveFML (0.546, $p < 0.001$) and MutPanning (0.591, $p < 0.001$) (Figure 13).

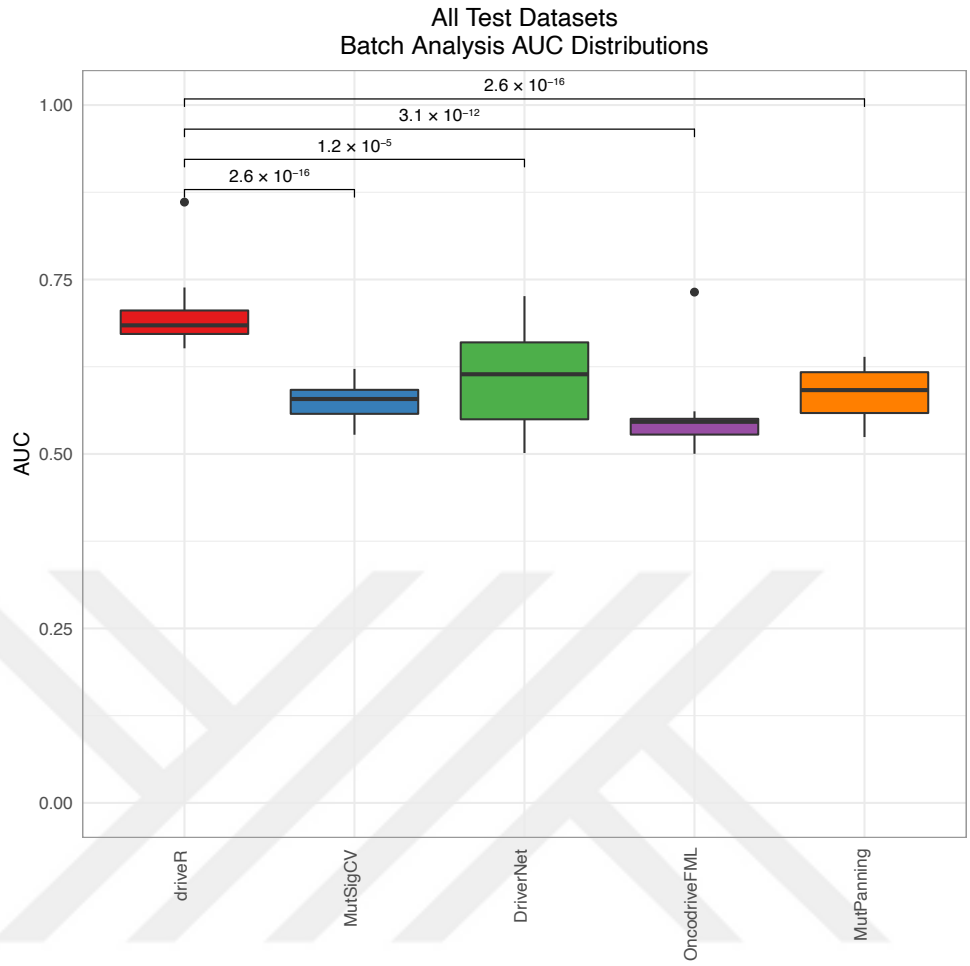


Figure 13. Overall comparison of the performance of driveR with batch analysis approaches.

Further, personalized analysis performances of driveR and other personalized analysis tools on 5157 test patients from 16 datasets were compared. When the performances were compared for all patients per dataset, driveR displayed superior performance compared to other tools (Figure 14).

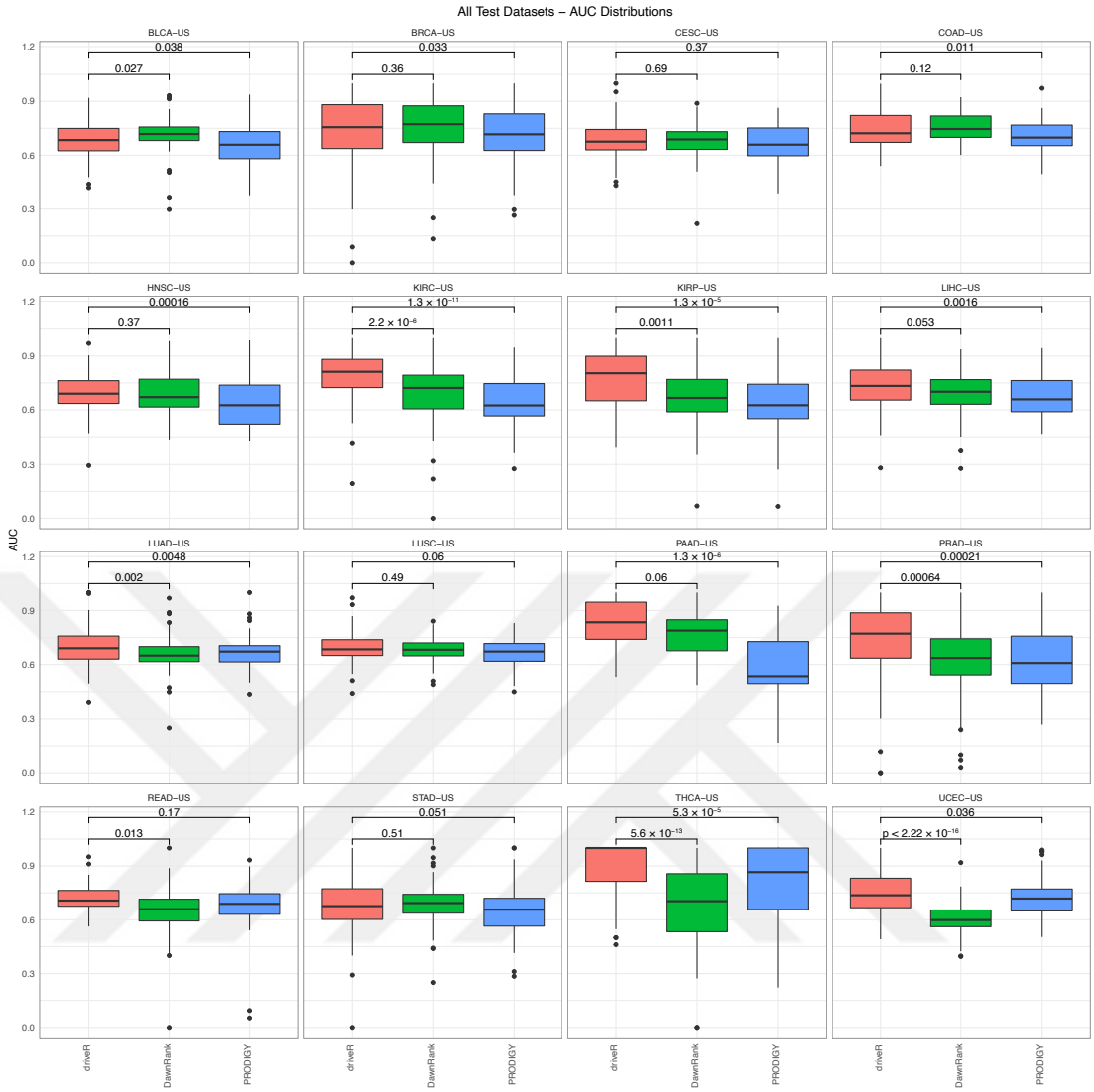


Figure 14. Comparison of performance of driveR with personalized analysis approaches per test dataset.

Overall, it was observed that driveR had higher median AUROC (0.728) compared to DawnRank (median AUROC = 0.693, $p < 0.001$) and PRODIGY (median AUROC = 0.679, $p < 0.001$) (Figure 15).

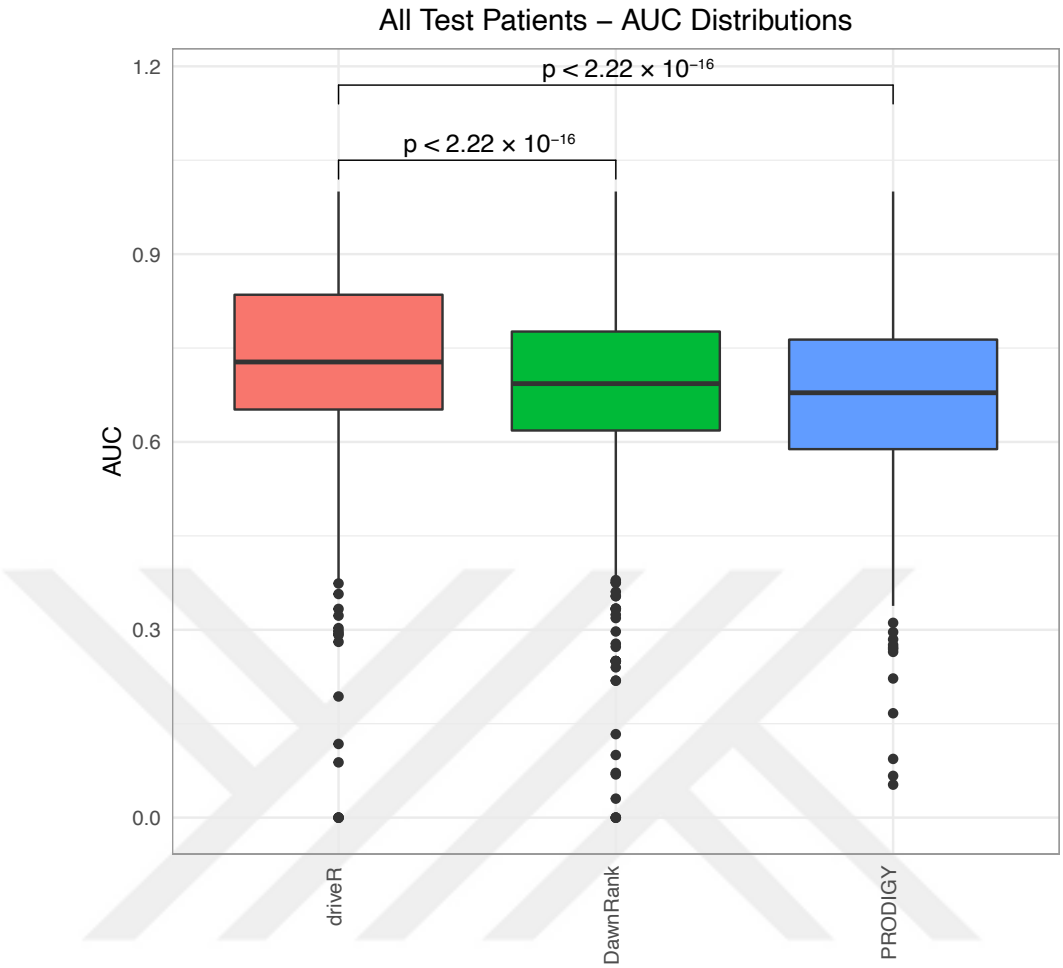


Figure 15. Comparison of performance of driveR with personalized analysis approaches on all test patients

4.3. Sample-specific prioritization of anti-cancer pharmacotherapy

4.3.1. TCGA analysis

The distance-based, the driver propagation, and driver and DEG propagation methods were performed and evaluated using the selected TCGA cohorts: LAML-US (Acute Myeloid Leukemia) was used for parameter selection of the driver propagation method, PRAD-US (Prostate Adenocarcinoma) was used for parameter selection of the driver and DEG propagation method, and all methods were assessed using THCA-US (Thyroid Cancer).

The distributions of the number of drugs for each sample in the top 5, 10, and 25 (including ties) for each method for THCA-US data are presented in Figure 16. While the numbers of selected drugs for the propagation methods showed little variability, the distance-based method displayed a large variability in the numbers of selected drugs per top n (due to drugs with tying scores).

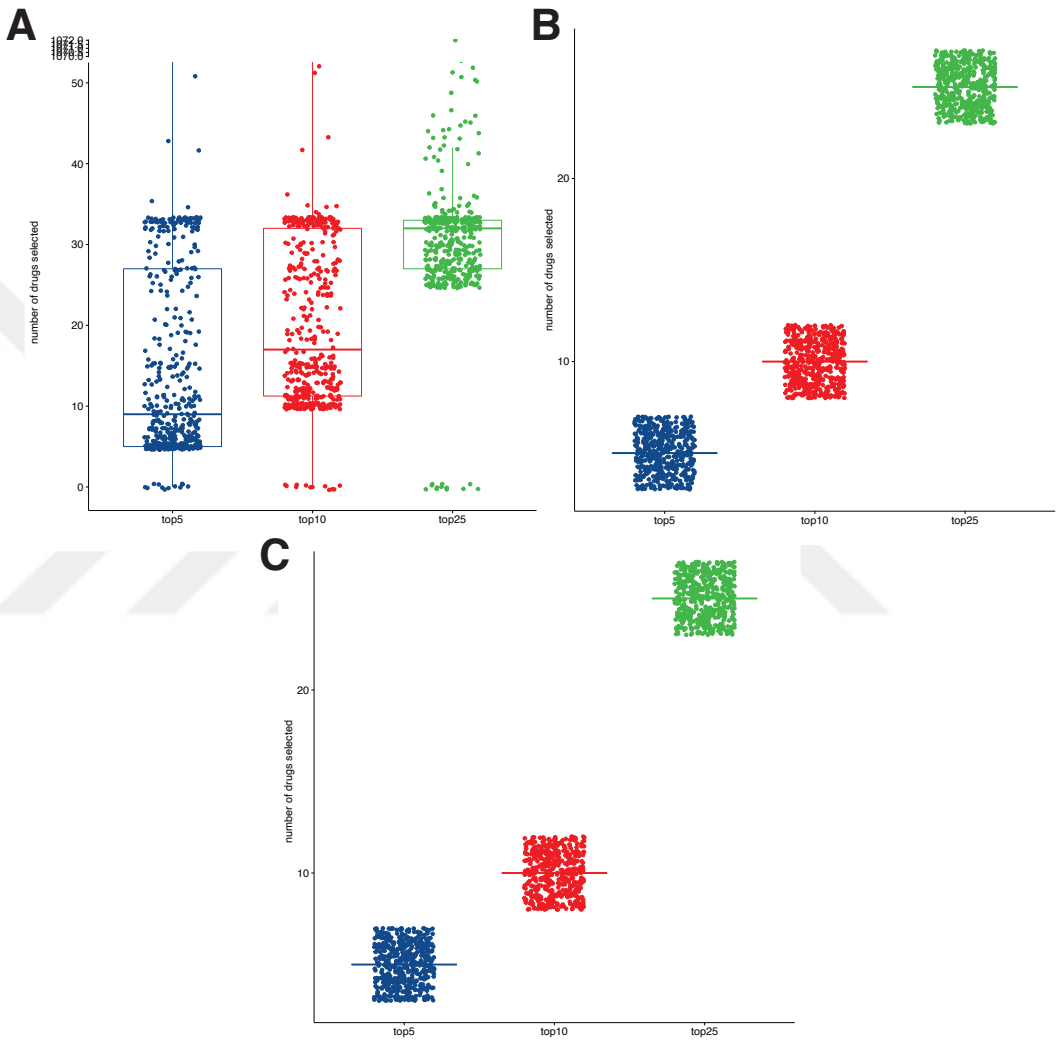


Figure 16. Numbers of drugs selected per top 5, 10, and 25 for the (A) distance-based, (B) driver propagation, and (C) driver and DEG propagation methods

Heatmaps displaying the proportions in each tier of drugs selected per the top 5, 10, and 25 per sample revealed sufficient support for the selected drugs and that the selected drugs are relevant as anti-cancer treatments (Figure 17). For all methods, almost every sample had at least one supported (Tiers 1-5) drug in the top 5. For the

distance-based method, 96.2% of samples had over 0.75 supported drugs, whereas 90.5% had over 0.9 supported drugs in the top 5. For the driver propagation method, 91.15% of samples had over 0.75 supported drugs, whereas 78.4% had over 0.9 supported drugs in the top 5. For the driver and DEG propagation method, 84.92% of samples had over 0.75 supported drugs, whereas 73.14% had over 0.9 supported drugs in the top 5.

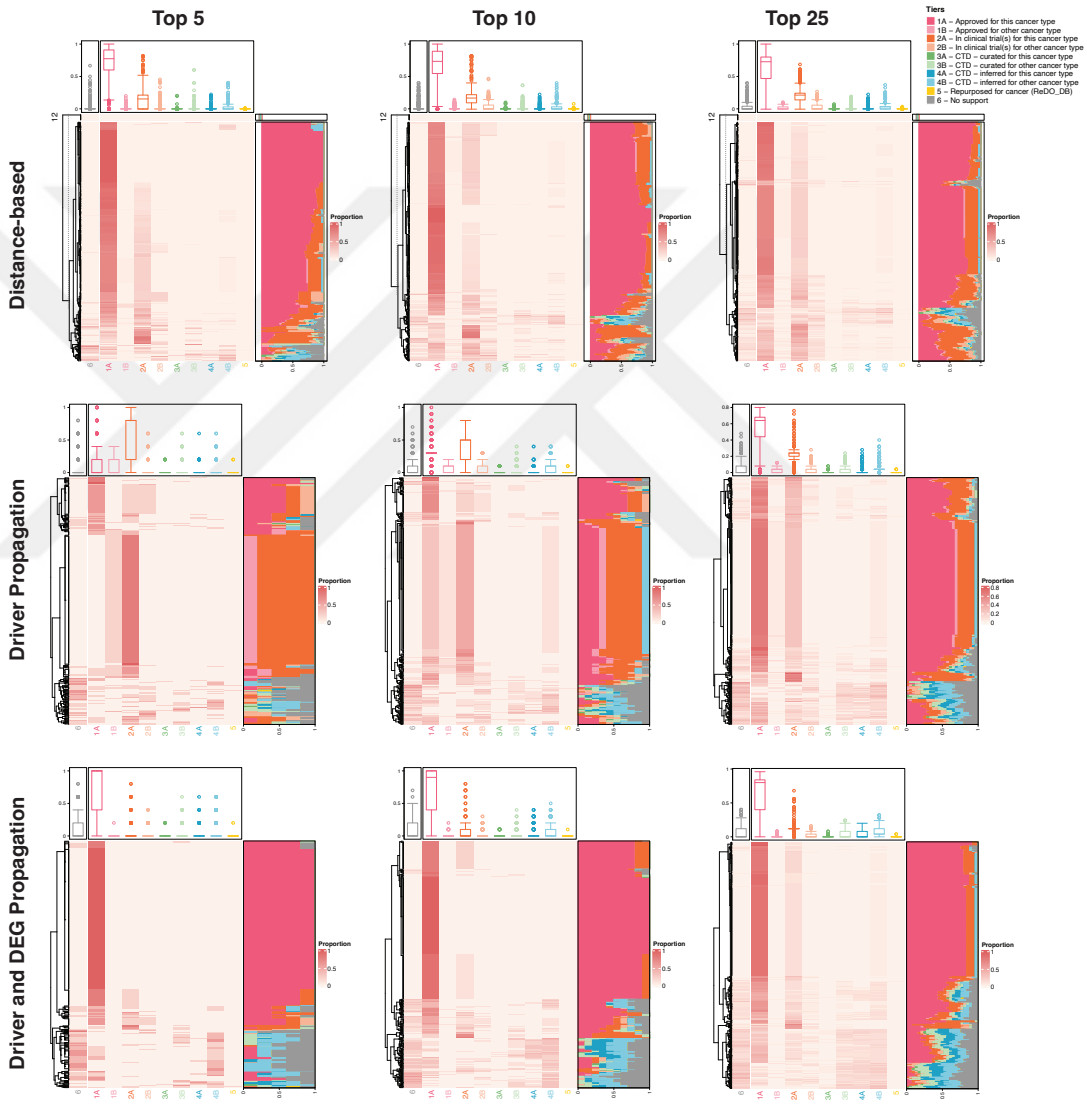


Figure 17. Heatmaps of proportions of selected drugs in each tier per top 5, 10, 25 for each method (THCA-US data, STRING PIN)

For all methods, the median proportions of tier 1-5 (all supported tiers) drugs per all top 5, 10, and 25 were higher than the expected proportion (0.53) of tier 1-5 drugs

in DGIdb (Figure 18). Similarly, the median proportions of tier 1-2 drugs (approved/clinical trial drugs) were higher than the expected proportion (0.31) of tier 1-2 drugs in DGIdb.

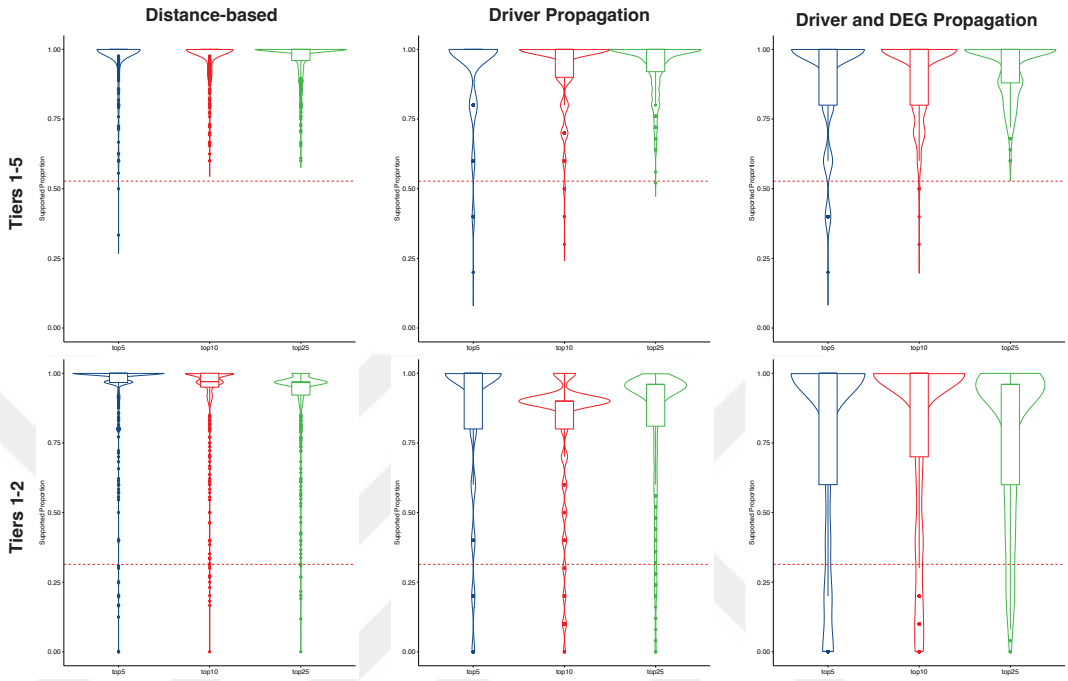


Figure 18. Distribution of proportions of supported (either Tiers 1-5 or 1-2) selected drugs per drug prioritization method (THCA-US data, STRING PIN)

Using a naïve approach and selecting all drugs that directly target the driver genes in a sample, at least one tier 1-5 drug could be recommended to 75.51% of all samples, and at least one tier 1-2 drug could be recommended to 74.69% (Table 4). The percentages of samples that were recommended at least one tier 1-5 (as well as tier 1-2) drug (as selected in the top 5) were substantially higher for the novel drug prioritization methods.

Table 4. Percentages of samples with tier 1-5 and tier 1-2 drug recommendations by different methods

	Tiers 1-5	**Tiers 1-2**
Direct Targeting	75.51%	74.69%
Distance-based Method	97.53%	96.91%
Driver Propagation Method	100.00%	96.50%

Driver and DEG Propagation Method	100.00%	90.29%

The frequencies of drugs prioritized in the top 5 per each method in each THCA-US sample had right-skewed distributions, indicating that most drugs in the top 5 were sample-specific (Appendix 1, 2, 3). The most frequently prioritized drugs in the top 5 were in cancer-associated drug tiers.

To investigate the robustness of the methods to the choice of PIN, analyses on THCA-US data were also performed using the BioGRID PIN. Appendices 4 and 5 show that the proportions of cancer-associated drugs for each method were similarly high. Hence, it can be established that the methods are robust to the choice of PIN, yielding comparable results.

4.3.2. GDSC analysis

The drug prioritization methods were further evaluated using drug-response data from the GDSC for 18 different cancer types (n = 515 cell line samples). The median number of drugs tested per cell line was 382 (inter-quartile range = 55).

The median values of maximal AUC per the top 5, 10, and 25 prioritized drugs for all methods were higher than the overall median AUC (blue dashed lines, 0.93) in the GDSC data (Figure 19A). Moreover, the median and maximum values of maximal AUC per the top 5, 10, and 25 prioritized drugs for all methods were higher than the cancer-type-specific median AUC values (blue dashed lines) (Figure 19B).

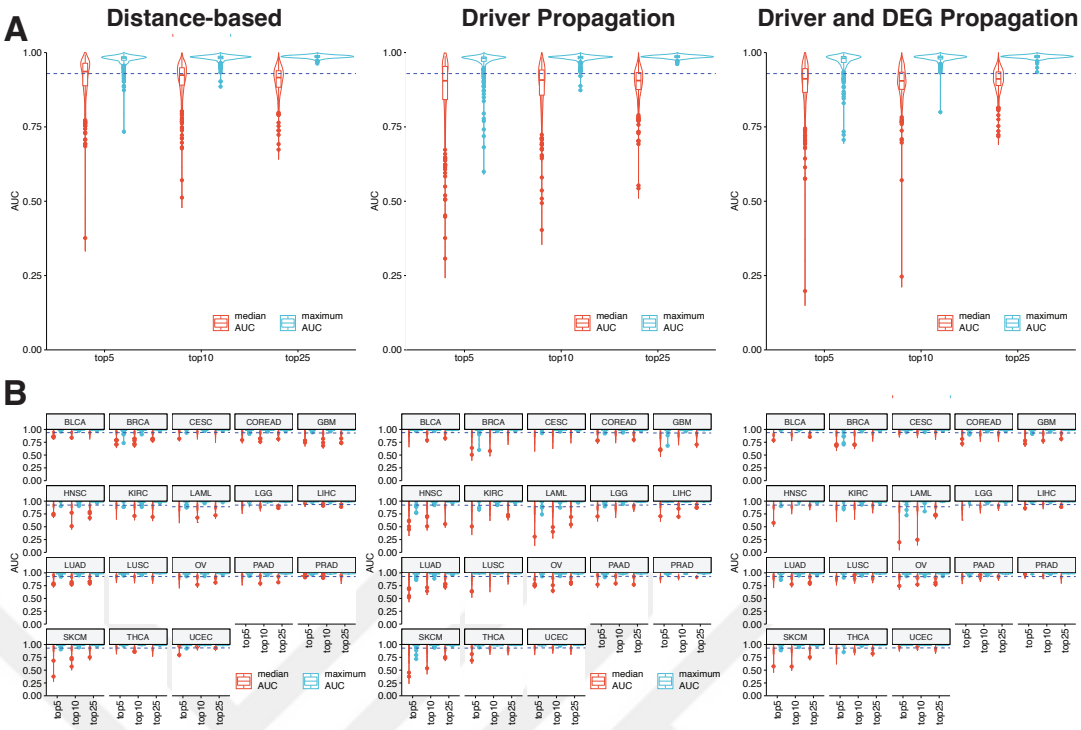


Figure 19: Distributions of drug-response AUC values for each method (A) Overall (B) Per cancer types

The proportions of drugs recommended by each method per cancer type are displayed in Figure 20. For the recommendation, the drug with the highest AUC value among the top 5 drugs for the sample was selected. Especially with the propagation methods, the relative frequencies of recommended drugs per cancer type were right-skewed, implying that most recommended drugs are sample-specific.

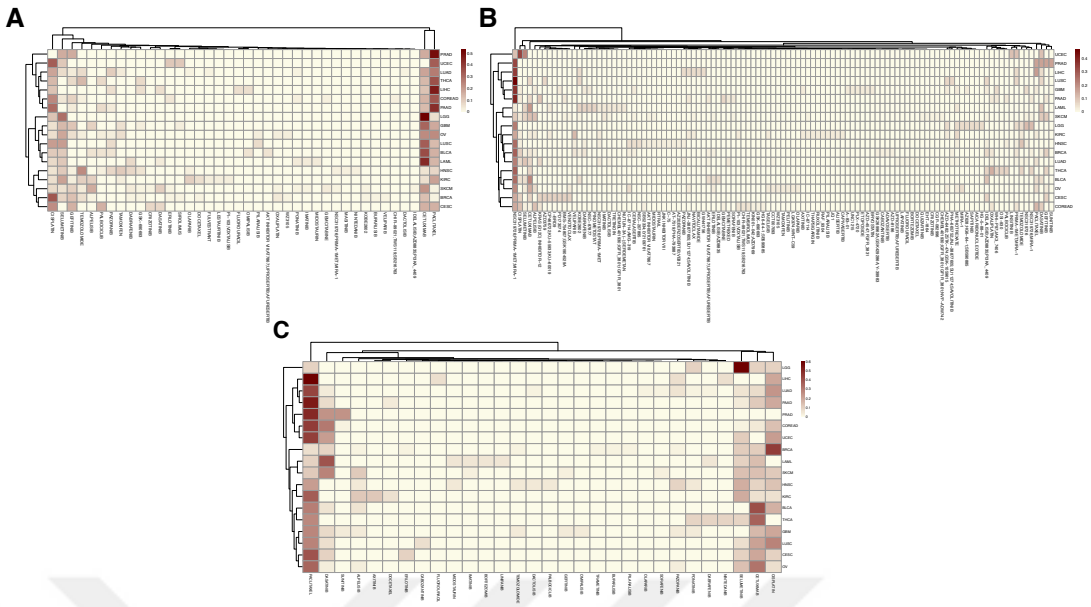


Figure 20: Heatmaps of the proportion of samples in each cancer type (rows) that were recommended each drug (columns) for GDSC analysis. (A) Distance-based, (B) Driver propagation, and (C) Driver and DEG propagation method

4.3.3. Cases with clinically demonstrated drug response

Drug prioritization analyses of the bladder cancer and lung adenocarcinoma cases were performed using the distance-based and driver propagation methods as well as iCAGES and PanDrugs (pan-cancer and cancer-type-specific) for comparison purposes. Because of the lack of expression data, the driver and DEG propagation method could not be performed in these cases.

Rankings of drugs demonstrated to be effective in these cases are presented in Table 5. For the bladder cancer case, Everolimus and Pazopanib were reported to be effective. The distance-based method ranked Pazopanib as the 1st drug and Everolimus as the 2nd drug, while the driver propagation method placed Pazopanib as the 6th and Everolimus as the 7th drug. While iCAGES did not include Pazopanib in the prioritization, it ranked Everolimus as the 3rd drug. In the PanDrugs pan-cancer analysis, Everolimus was the 9th, and Pazopanib was the 14th drug. The PanDrugs cancer-specific analyses did not include either drug for prioritization.

For the lung adenocarcinoma case, it was reported that the patient responded to Sorafenib. The distance-based method ranked Sorafenib as the 2nd drug, and the driver propagation method ranked it as the 10th drug. iCAGES placed Sorafenib as the 1st drug. PanDrugs pan-cancer analysis ranked Sorafenib as the 2nd drug while the PanDrugs cancer-specific analysis did not consider this drug.

Table 5: Ranking of effective drugs for the bladder cancer and lung adenocarcinoma cases per method

	Bladder Cancer		**Lung Adenocarcinoma**
	Everolimus	**Pazopanib**	**Sorafenib**
PANACEA - Distance-based	2	1	2
PANACEA - Driver Propagation	7	6	10
iCAGES	3	N/A	1
PanDrugs - Pan-cancer	9	14	2
PanDrugs - Cancer-specific	N/A	N/A	N/A

5. DISCUSSION

Two of the most critical challenges of personalized oncology are (i) identification of driver genes from genomics data in a tumor sample, which has great potential theragnostic benefit, and (ii) using these clinically relevant drivers to recommend rational, effective, and personalized treatments (15, 16, 139). The novel methods proposed in this dissertation aim to address these challenges.

Somatic alterations in driver genes are the foundation of tumor formation and progression (13, 14). There is a large number of experimentally validated driver genes. (31, 73, 125, 140). However, there is still a great potential benefit in personalized identification of driver genes as some unknown driver genes may act as drivers in certain cancer types and even in specific tumors.

For the purpose of driver gene prioritization, a novel model-based approach, driveR, was developed as part of this dissertation (17). For driveR, a multi-task learning classification model was trained to classify genes as “driver” or “non-driver” in a cancer-type-specific manner using publicly available somatic genomics data of 21 different cancer types. The 26 features of the MTL model were devised to incorporate the genomic landscape of the tumor and prior biological information. As input, driveR takes somatic alterations and utilizes the corresponding sub-task of this MTL model to prioritize driver genes in cancer cohorts or individual tumor samples. Through comparisons, it was illustrated that driveR performed better than the existing batch analysis approaches MutSigCV, DriverNet, OncodriveFML, and MutPanning, as well the previous personalized analysis approaches DawnRank and PRODIGY. Below, some unique aspects of driveR are discussed.

Driver genes are divergent and diverse between different types of cancers (31, 73, 125, 140). The MTL classification model at the core of driveR allows for cancer-type-specific prioritization of driver genes for 21 different cancer types. This enables driveR to prioritize driver genes specific to a specific type of cancer more adequately.

Most driver prioritization methods are developed for analyzing cancer cohorts. These batch analysis methods fail in identifying low-frequency driver genes and completely ignore sample-specific driver genes. Sample-specific driver genes are often rare or may not match the tissue-of-origin. The results indicate that driveR is an adequate option for studying driver genes not only in tumor cohorts but also in individual tumor samples.

As mentioned above, driveR incorporated prior biological knowledge in the MTL model. To score a gene’s association with cancer, Phenolyzer, a database-mining tool integrating various biological knowledge databases, was used (109). Moreover, the presence of a gene in cancer-related KEGG pathways was also utilized as a part of the features (113). Hence, the MTL driver gene classification model used was based not only on somatic alterations but also on expert biological knowledge founded on

decades of research. This incorporation of extensive biological knowledge significantly improved the accuracy of driveR.

Following the development of driveR, a collection of novel network-oriented methods for sample-specific anti-cancer pharmacotherapy prioritization, PANACEA, were established. These methods utilize driver gene prioritization results of a tumor sample and a protein interaction network extended with curated drug-gene interaction information to score anti-cancer drug candidates for the tumor sample.

Initially, PANACEA methods map the driverness probabilities of altered genes in a tumor sample onto the extended interaction network. The "distance-based" method ranks the drugs using these driverness probabilities and the distance between the drugs and the altered genes. The "driver propagation" method uses a random walk with restart framework to propagate the driverness probabilities to score the drugs. The "driver and DEG propagation" method additionally propagates the $|\log_2\text{-fold-change}|$ values of DEGs, and the score of a drug is assigned as the product of its final driverness propagation score and its final DEG score. By utilizing the network-based methods and a probabilistic definition of driver genes, the use of genes is extended beyond known drivers. Thus, each altered gene is considered while computing a drug's score. Furthermore, integration of the extended network exploits indirect interactions between drugs and altered genes, broadening the potential targets from only direct target genes. All methods were shown to be robust to the change in PIN via the analysis of TCGA data using the BioGRID Homo sapiens PIN, implying that as long as moderate to high confidence interactions from trustworthy resources are used, the prioritization results will be comparable.

The results demonstrate that the novel sample-specific anti-cancer pharmacotherapy prioritization methods adequately prioritize clinically appropriate and relevant drugs for oncology. Moreover, through analyses on drug-response data for cancer cell lines, it was shown that the prioritized drugs are highly likely to be effective. Lastly, the analyses on cancer cases illustrated that PANACEA methods can potentially be applied in the clinical setting, outperforming the previous approaches

iCAGES and PanDrugs in the bladder cancer case and yielding comparable results as the two in the lung adenocarcinoma case.

Notably, the list of potential drugs for the PANACEA methods does not necessarily need to be confined to drugs approved for cancer; drugs in clinical trials, pre-clinical drugs, or even drugs with other indications can also be included. This expands the list of potential drugs beyond only established anti-cancer drugs, enabling drug repurposing, which has gathered significant momentum in oncology in the last years (141, 142).

6. CONCLUSION

In this dissertation, the main aim was to devise an approach for prioritizing anti-cancer pharmacotherapy options based on the somatic alterations in a tumor sample. For this purpose, firstly, a novel method, driveR, for driver gene prioritization applicable to the analysis of both cancer cohorts and individual tumor samples was established. The novel driver gene prioritization method was demonstrated to perform well and outperform existing personalized and batch analysis approaches. Next, a collection of network-based methods, PANACEA, was developed and evaluated. The PANACEA methods utilize driver prioritization results from driveR and an interaction network extended with a curated drug-gene interaction resource to score and prioritize drugs. The evaluation of the PANACEA methods revealed that the top drugs selected for each method are relevant as anti-cancer agents and are also likely to be efficient.

As established through review of the current literature, methods for personalized prioritization of driver genes and personalized determination of anti-cancer pharmacotherapy are underdeveloped. Such methods are required to advance the field of personalized oncology. The novel methods proposed in this thesis, based mainly on the genomic landscape of the tumor, will hopefully provide further insight into theragnostic opportunities and help progress the research on personalized oncology.

7. REFERENCES

1. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
2. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646-74.
3. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1):31-46.
4. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719-24.
5. Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013;153(1):17-37.
6. Kamalakaran S, Varadan V, Janevski A, Banerjee N, Tuck D, McCombie WR, et al. Translating next generation sequencing to practice: opportunities and necessary steps. Mol Oncol. 2013;7(4):743-55.
7. Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011;331(6024):1553-8.
8. Kou T, Kanai M, Matsumoto S, Okuno Y, Muto M. The possibility of clinical sequencing in the management of cancer. Jpn J Clin Oncol. 2016;46(5):399-406.
9. Hyman DM, Taylor BS, Baselga J. Implementing Genome-Driven Oncology. Cell. 2017;168(4):584-99.
10. Chen HZ, Bonneville R, Roychowdhury S. Implementing precision cancer medicine in the genomic era. Semin Cancer Biol. 2019;55:16-27.
11. Berger MF, Mardis ER. The emerging clinical relevance of genomics in cancer medicine. Nat Rev Clin Oncol. 2018;15(6):353-65.
12. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007;446(7132):153-8.
13. Haber DA, Settleman J. Cancer: drivers and passengers. Nature. 2007;446(7132):145-6.
14. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. Science. 2013;339(6127):1546-58.
15. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med. 2011;17(3):297-303.
16. Schilsky RL. Personalized medicine in oncology: the future is now. Nature Reviews Drug Discovery. 2010;9(5):363-6.
17. Ülgen E, Sezerman OU. driveR: a novel method for prioritizing cancer driver genes using somatic genomics data. BMC Bioinformatics. 2021;22(1):263.
18. American Cancer Society. Global Cancer Facts & Figures 4th Edition. Atlanta: American Cancer Society2018.
19. Talbot SJ, Crawford DH. Viruses and tumours – an update. European Journal of Cancer. 2004;40(13):1998-2005.
20. Esteller M. Epigenetics in Cancer. New England Journal of Medicine. 2008;358(11):1148-59.

21. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis. 2010;31(1):27-36.
22. DeVita VT, Vincent T. DeVita JMD, Lawrence TS, Rosenberg SA. Cancer: Principles & Practice of Oncology : Primer of the Molecular Biology of Cancer. 1st ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2011.
23. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004;431(7011):931-45.
24. Mardis ER, Wilson RK. Cancer genome sequencing: a review. Hum Mol Genet. 2009;18(R2):R163-8.
25. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010;11(10):685-96.
26. Metzker ML. Sequencing technologies — the next generation. Nature Reviews Genetics. 2010;11(1):31-46.
27. Campbell PJ, Getz G, Korbel JO, Stuart JM, Jennings JL, Stein LD, et al. Pan-cancer analysis of whole genomes. Nature. 2020;578(7793):82-93.
28. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. Nature. 2020;578(7793):94-101.
29. Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM, et al. Signatures of mutation and selection in the cancer genome. Nature. 2010;463(7283):893-8.
30. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321-37.e10.
31. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173(2):371-85.e18.
32. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360(8):765-73.
33. Mardis ER, Ding L, Dooling DJ, Larson DE, McLellan MD, Chen K, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009;361(11):1058-66.
34. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363(25):2424-33.
35. Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature. 2010;463(7279):360-3.
36. Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med. 2015;372(26):2481-98.
37. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature. 2010;463(7278):191-6.

38. Gore M, Larkin J. Precision oncology: where next? Lancet Oncol. 2015;16(16):1593-5.
39. Senft D, Leiserson MDM, Ruppin E, Ronai ZA. Precision Oncology: The Road Ahead. Trends Mol Med. 2017;23(10):874-98.
40. Yang HT, Shah RH, Tegay D, Onel K. Precision oncology: lessons learned and challenges for the future. Cancer Manag Res. 2019;11:7525-36.
41. Nakagawa H, Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. Cancer Sci. 2018;109(3):513-22.
42. Rusch M, Nakitandwe J, Shurtleff S, Newman S, Zhang Z, Edmonson MN, et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. Nat Commun. 2018;9(1):3962.
43. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. Trends Genet. 2014;30(9):418-26.
44. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn. 2017;19(3):341-65.
45. Kerick M, Isau M, Timmermann B, Sültmann H, Herwig R, Krobitsch S, et al. Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. BMC Med Genomics. 2011;4:68.
46. Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J. Applications and analysis of targeted genomic sequencing in cancer studies. Comput Struct Biotechnol J. 2019;17:1348-59.
47. Allgäuer M, Budczies J, Christopoulos P, Endris V, Lier A, Rempel E, et al. Implementing tumor mutational burden (TMB) analysis in routine diagnostics-a primer for molecular pathologists and clinicians. Transl Lung Cancer Res. 2018;7(6):703-15.
48. Stenzinger A, Allen JD, Maas J, Stewart MD, Merino DM, Wempe MM, et al. Tumor mutational burden standardization initiatives: Recommendations for consistent tumor mutational burden assessment in clinical samples to guide immunotherapy treatment decisions. Genes Chromosomes Cancer. 2019;58(8):578-88.
49. Bergstrom EN, Huang MN, Mahto U, Barnes M, Stratton MR, Rozen SG, et al. SigProfilerMatrixGenerator: a tool for visualizing and exploring patterns of small mutational events. BMC Genomics. 2019;20(1):685.
50. Wang C, Liang C. MSIpred: a python package for tumor microsatellite instability classification from tumor mutation annotation data using a support vector machine. Scientific Reports. 2018;8(1):17546.
51. Ni Huang M, McPherson JR, Cutcutache I, Teh BT, Tan P, Rozen SG. MSIseq: Software for Assessing Microsatellite Instability from Catalogs of Somatic Mutations. Scientific Reports. 2015;5(1):13321.

52. Sezerman OU, Ulgen E, Seymen N, Durasi IM. Bioinformatics Workflows for Genomic Variant Discovery, Interpretation and Prioritization. Bioinformatics Tools for Detection and Clinical Interpretation of Genomic Variations: IntechOpen; 2019.
53. Van Allen EM, Wagle N, Stojanov P, Perrin DL, Cibulskis K, Marlow S, et al. Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. Nat Med. 2014;20(6):682-8.
54. Nakken S, Fournous G, Vodák D, Aasheim LB, Myklebost O, Hovig E. Personal Cancer Genome Reporter: variant interpretation report for precision oncology. Bioinformatics. 2017;34(10):1778-80.
55. Dienstmann R, Dong F, Borger D, Dias-Santagata D, Ellisen LW, Le LP, et al. Standardized decision support in next generation sequencing reports of somatic cancer variants. Molecular Oncology. 2014;8(5):859-73.
56. Berner AM, Morrissey GJ, Murugaesu N. Clinical Analysis of Whole Genome Sequencing in Cancer Patients. Current Genetic Medicine Reports. 2019;7(2):136-43.
57. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Medicine. 2018;10(1):25.
58. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn. 2017;19(1):4-23.
59. Tokheim CJ, Papadopoulos N, Kinzler KW, Vogelstein B, Karchin R. Evaluating the evaluation of cancer driver genes. Proc Natl Acad Sci U S A. 2016;113(50):14330-5.
60. Cheng F, Zhao J, Zhao Z. Advances in computational approaches for prioritizing driver mutations and significantly mutated genes in cancer genomes. Brief Bioinform. 2016;17(4):642-56.
61. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, et al. MuSiC: identifying mutational significance in cancer genomes. Genome Res. 2012;22(8):1589-98.
62. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499(7457):214-8.
63. Dietlein F, Weghorn D, Taylor-Weiner A, Richters A, Reardon B, Liu D, et al. Identification of cancer driver genes based on nucleotide context. Nature Genetics. 2020;52(2):208-18.
64. Ciriello G, Cerami E, Sander C, Schultz N. Mutual exclusivity analysis identifies oncogenic network modules. Genome Res. 2012;22(2):398-406.
65. Reyna MA, Leiserson MDM, Raphael BJ. Hierarchical HotNet: identifying hierarchies of altered subnetworks. Bioinformatics. 2018;34(17):i972-i80.
66. Paull EO, Carlin DE, Niepel M, Sorger PK, Haussler D, Stuart JM. Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE). Bioinformatics. 2013;29(21):2757-64.

67. Bashashati A, Haffari G, Ding J, Ha G, Lui K, Rosner J, et al. DriverNet: uncovering the impact of somatic driver mutations on transcriptional networks in cancer. Genome Biology. 2012;13(12):R124.
68. Kartha VK, Sebastiani P, Kern JG, Zhang L, Varelas X, Monti S. CaDrA: A Computational Framework for Performing Candidate Driver Analyses Using Genomic Features. Frontiers in Genetics. 2019;10(121).
69. Mularoni L, Sabarinathan R, Deu-Pons J, Gonzalez-Perez A, López-Bigas N. OncodriveFML: a general framework to identify coding and non-coding regions with cancer driver mutations. Genome Biology. 2016;17(1):128.
70. Collier O, Stoven V, Vert JP. LOTUS: A single- and multitask machine learning algorithm for the prediction of cancer driver genes. PLoS Comput Biol. 2019;15(9):e1007381.
71. Liu S-H, Shen P-C, Chen C-Y, Hsu A-N, Cho Y-C, Lai Y-L, et al. DriverDBv3: a multi-omics database for cancer driver gene research. Nucleic Acids Research. 2019;48(D1):D863-D70.
72. Wang T, Ruan S, Zhao X, Shi X, Teng H, Zhong J, et al. OncoVar: an integrated database and analysis platform for oncogenic driver variants in cancers. Nucleic Acids Research. 2020.
73. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A compendium of mutational cancer driver genes. Nature Reviews Cancer. 2020;20(10):555-72.
74. Kroschinsky F, Stölzel F, von Bonin S, Beutel G, Kochanek M, Kiehl M, et al. New drugs, new toxicities: severe side effects of modern targeted and immunotherapy of cancer and their management. Crit Care. 2017;21(1):89.
75. Park SR, Davis M, Doroshow JH, Kummar S. Safety and feasibility of targeted agent combinations in solid tumours. Nature Reviews Clinical Oncology. 2013;10(3):154-68.
76. Hou JP, Ma J. DawnRank: discovering personalized driver genes in cancer. Genome Med. 2014;6(7):56.
77. Guo WF, Zhang SW, Liu LL, Liu F, Shi QQ, Zhang L, et al. Discovering personalized driver mutation profiles of single samples in cancer by network control strategy. Bioinformatics. 2018;34(11):1893-903.
78. Dong C, Guo Y, Yang H, He Z, Liu X, Wang K. iCAGES: integrated CAncer GEnome Score for comprehensively prioritizing driver genes in personal cancer genomes. Genome Medicine. 2016;8(1):135.
79. Dinstag G, Shamir R. PRODIGY: personalized prioritization of driver genes. Bioinformatics. 2019;36(6):1831-9.
80. Tsimberidou AM, Fountzilas E, Nikanjam M, Kurzrock R. Review of precision cancer medicine: Evolution of the treatment paradigm. Cancer Treat Rev. 2020;86:102019.
81. Rubio-Perez C, Tamborero D, Schroeder MP, Antolín AA, Deu-Pons J, Perez-Llamas C, et al. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. Cancer Cell. 2015;27(3):382-96.
82. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, et al. The NCBI BioSystems database. Nucleic Acids Res. 2010;38(Database issue):D492-6.

83. Wang Y, Bryant SH, Cheng T, Wang J, Gindulyte A, Shoemaker BA, et al. PubChem BioAssay: 2017 update. Nucleic Acids Res. 2017;45(D1):D955-d63.
84. Piñeiro-Yáñez E, Reboiro-Jato M, Gómez-López G, Perales-Patón J, Troulé K, Rodríguez JM, et al. PanDrugs: a novel method to prioritize anticancer drug treatments according to individual genomic data. Genome Med. 2018;10(1):41.
85. Turanli B, Altay O, Borén J, Turkez H, Nielsen J, Uhlen M, et al. Systems biology based drug repositioning for development of cancer therapy. Semin Cancer Biol. 2021;68:47-58.
86. Shnaps O, Perry E, Silverbush D, Sharan R. INFERENCE OF PERSONALIZED DRUG TARGETS VIA NETWORK PROPAGATION. Pac Symp Biocomput. 2016;21:156-67.
87. Di J, Zheng B, Kong Q, Jiang Y, Liu S, Yang Y, et al. Prioritization of candidate cancer drugs based on a drug functional similarity network constructed by integrating pathway activities and drug activities. Mol Oncol. 2019;13(10):2259-77.
88. Martelotto LG, Ng CK, De Filippo MR, Zhang Y, Piscuoglio S, Lim RS, et al. Benchmarking mutation effect prediction algorithms using functionally validated cancer-related missense mutations. Genome Biol. 2014;15(10):484.
89. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Hum Mutat. 2016;37(3):235-41.
90. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 2010;38(16):e164-e.
91. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31(13):3812-4.
92. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-9.
93. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19(9):1553-61.
94. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nature Methods. 2014;11(4):361-2.
95. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Research. 2011;39(17):e118-e.
96. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. 2013;34(1):57-65.
97. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010;6(12):e1001025.
98. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010;20(1):110-21.

99. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;47(D1):D886-d94.
100. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics. 2013;14 Suppl 3(Suppl 3):S3.
101. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics. 2009;25(12):i54-i62.
102. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics. 2015;31(5):761-3.
103. Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database (Oxford). 2011;2011:bar026.
104. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-4.
105. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun. 2016;7:11479.
106. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. Cell. 2019;177(4):1035-49.e19.
107. Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East Asians. Nat Genet. 2020;52(2):177-86.
108. Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, Forbes SA. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat Rev Cancer. 2018;18(11):696-705.
109. Yang H, Robinson PN, Wang K. Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. Nat Methods. 2015;12(9):841-3.
110. Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev Cancer. 2010;10(1):59-64.
111. Kim TM, Xi R, Luquette LJ, Park RW, Johnson MD, Park PJ. Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. Genome Res. 2013;23(2):217-27.
112. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Research. 2018;47(D1):D941-D7.
113. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.
114. Cao H, Zhou J, Schwarz E. RMTL: an R library for multi-task learning. Bioinformatics. 2018;35(10):1797-8.

115. Tibshirani R. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society Series B (Methodological). 1996;58(1):267-88.
116. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet. 2006;38(5):500-1.
117. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006;34(Database issue):D535-9.
118. Oughtred R, Stark C, Breitkreutz B-J, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. Nucleic Acids Research. 2018;47(D1):D529-D41.
119. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-d13.
120. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955-61.
121. Imielinski M, Greulich H, Kaplan B, Araujo L, Amann J, Horn L, et al. Oncogenic and sorafenib-sensitive ARAF mutations in lung adenocarcinoma. The Journal of Clinical Investigation. 2014;124(4):1582-6.
122. Wagle N, Grabiner BC, Van Allen EM, Hodis E, Jacobus S, Supko JG, et al. Activating mTOR mutations in a patient with an extraordinary response on a phase I trial of everolimus and pazopanib. Cancer Discov. 2014;4(5):546-53.
123. Freshour SL, Kiwala S, Cotto KC, Coffman AC, McMichael JF, Song JJ, et al. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Research. 2020;49(D1):D1144-D51.
124. Shrager J, Tenenbaum JM, Travers M. Cancer commons: biomedicine in the internet age. Ekins/-Collaborative Computational Technologies for Biomedical Research. 2011:161-77.
125. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandoth C, Reimand J, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Scientific Reports. 2013;3(1):2650.
126. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, et al. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res. 2019;47(D1):D930-d40.
127. Griffith M, Spies NC, Krysiak K, McMichael JF, Coffman AC, Danos AM, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet. 2017;49(2):170-4.
128. Ainscough BJ, Griffith M, Coffman AC, Wagner AH, Kunisaki J, Choudhary MN, et al. DoCM: a database of curated mutations in cancer. Nat Methods. 2016;13(10):806-7.
129. Jain N, Mittendorf KF, Holt M, Lenoue-Newton M, Maurer I, Miller C, et al. The My Cancer Genome clinical trial data model and trial curation workflow. J Am Med Inform Assoc. 2020;27(7):1057-66.

130. Morgensztern D, Campo MJ, Dahlberg SE, Doebele RC, Garon E, Gerber DE, et al. Molecularly targeted therapies in non-small-cell lung cancer annual update 2014. J Thorac Oncol. 2015;10(1 Suppl 1):S1-63.
131. Rask-Andersen M, Masuram S, Schiöth HB. The druggable genome: Evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication. Annu Rev Pharmacol Toxicol. 2014;54:9-26.
132. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550.
133. Ochoa D, Hercules A, Carmona M, Suveges D, Gonzalez-Uriarte A, Malangone C, et al. Open Targets Platform: supporting systematic drug–target identification and prioritisation. Nucleic Acids Research. 2020;49(D1):D1302-D10.
134. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegers J, Wiegers TC, et al. Comparative Toxicogenomics Database (CTD): update 2021. Nucleic Acids Research. 2020;49(D1):D1138-D43.
135. Pantziarka P, Verbaanderd C, Sukhatme V, Rica Capistrano I, Crispino S, Gyawali B, et al. ReDO_DB: the repurposing drugs in oncology database. Ecancermedicalscience. 2018;12:886.
136. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, et al. A Landscape of Pharmacogenomic Interactions in Cancer. Cell. 2016;166(3):740-54.
137. Picco G, Chen ED, Alonso LG, Behan FM, Gonçalves E, Bignell G, et al. Functional linkage of gene fusions to cancer cell fitness assessed by pharmacological and CRISPR-Cas9 screening. Nature Communications. 2019;10(1):2198.
138. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7.
139. Gómez-López G, Dopazo J, Cigudosa JC, Valencia A, Al-Shahrour F. Precision medicine needs pioneering clinical bioinformaticians. Brief Bioinform. 2019;20(3):752-66.
140. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods. 2013;10(11):1081-2.
141. Schein CH. Repurposing approved drugs for cancer therapy. British Medical Bulletin. 2021;137(1):13-27.
142. Zhang Z, Zhou L, Xie N, Nice EC, Zhang T, Cui Y, et al. Overcoming cancer therapeutic bottleneck by drug repurposing. Signal Transduction and Targeted Therapy. 2020;5(1):113.

8. APPENDICES

Appendix 1

Frequencies of all drugs observed in the top 5 in each THCA-US sample via the distance-based method. Drugs that were merged (because they have the same target gene(s)) are listed separated by ";".

Drug	Freq	Tiers
EVEROLIMUS	314	1A
CETUXIMAB	282	1A
SELUMETINIB	278	1A
VEMURAFENIB	271	1A
DABRAFENIB	237	1A
PANITUMUMAB	213	1A
DACTOLISIB	210	2A
TRAMETINIB	205	1A
DASATINIB	197	1A
GEFITINIB	194	1A
SORAFENIB	189	1A
OXALIPLATIN	185	1A
ERLOTINIB	182	1A
ALPELISIB	177	1A
BINIMETINIB	174	1A
PONATINIB	168	1A
CAPECITABINE	166	1A
REGORAFENIB	164	1A
IRINOTECAN	160	1A
TEMSIROLIMUS	157	1A
CRIZOTINIB	146	1A
BEVACIZUMAB	134	1A
COBIMETINIB	129	1A
CEP-32496	122	2A
PIMASERTIB	118	2A
SORAFENIB TOSYLATE	115	2A
PLX-4720	113	4B
PERTUZUMAB	103	1A
ENCORAFENIB	102	1A

RAF-265	85	2A
LY-3009120;MLN-2480;XL-281	75	2B;2A;2A
ARQ-736;LGX-806	72	2A;6
OMIPALISIB	66	2A
CISPLATIN	58	1A
PACLITAXEL	53	1A
LIFIRAFENIB;ZELBORAF;DABRAFENIB MESYLATE;RG-7256;CHEMBL525191;PLX-8394	50	2A;6;1B;6;6;2A
PELAREOREP	45	2B
CARBOPLATIN	43	1A
RIBOCICLIB	41	1A
VANDETANIB	41	1A
DOCETAXEL	32	1A
TIPIFARNIB	32	2A
GEMCITABINE	30	1A
ABEMACICLIB	29	1A
GEDATOLISIB	29	2A
FLUOROURACIL	28	1A
OLAPARIB	28	1A
PEMBROLIZUMAB	28	1A
PF-04691502	28	2A
PILARALISIB	28	2A
QUERCETIN	28	3B
ILORASERTIB	26	2A
METFORMIN	26	2A
TEMOZOLOMIDE	26	1A
BORTEZOMIB	25	1A
COPANLISIB;BUPARLISIB;GSK-2636771	25	1A;2A;2A
GSK-690693	25	2A
IPILIMUMAB	25	1A
LAPATINIB	25	1A
CAPIVASERTIB	24	2A
ENMD-981693	24	2A
APITOLISIB	23	2A
CENISERTIB	23	2A
ENMD-2076	23	2A
IPATASERTIB	23	2A
CERALASERTIB	22	2A

DECITABINE	22	1A
DS-7423	22	2B
PALBOCICLIB	22	1A
PAZOPANIB	22	1A
DOXORUBICIN	21	1A
JNJ-26483327	21	2A
NIVOLUMAB	21	1A
TRASTUZUMAB	21	1A
INFIGRATINIB	20	2A
SIROLIMUS	20	2A
AFATINIB	19	1A
AT-9283	19	2A
PI-103;VOXTALISIB	19	3B;2A
PICTILISIB	19	2A
TAMOXIFEN	19	1A
TASELISIB	19	2B
VORINOSTAT	19	1A
BMS-690514	18	2A
IMATINIB	18	1A
XL-228	18	2B
LETROZOLE	17	1A
OCRIPLASMIN	17	6
PUQUITINIB	17	6
TALAZOPARIB	17	1A
UCN-01	17	2A
ATEZOLIZUMAB	16	1A
LY-2780301;XL-418	16	2B;2A
MK-2206	16	2A
PELITINIB;CANERTINIB DIHYDROCHLORIDE;AFATINIB DIMALEATE	16	2A;2B;1B
POZIOTINIB;OSIMERTINIB MESYLATE;DACOMITINIB;AC-480	16	2A;2B;1A;2A
SONOLISIB	16	2A
BAY-1125976	15	2A
CEP-2563	15	6
ETOPOSIDE	15	1A
GRANISETRON	15	2A
IBRUTINIB	15	1B
METHOTREXATE	15	1A
MIDOSTAURIN	15	1A

NINTEDANIB	15	1A
NIRAPARIB	15	1A
PANULISIB	15	2A
PERIFOSINE	15	2A
UPROSERTIB;TRICIRIBINE;AZD-5363;TRICIRIBINE PHOSPHATE;AFURESERTIB;ARCHEXIN;MIRANSERTIB;MK-2201	15	2A;4A;6;2A;2A;6;2B;6
ADAVOSERTIB	14	2A
BGT-226;VS-5584;SF-1126;SAMOTOLISIB;PWT-33587	14	2A;2B;2A;2A;6
CEDIRANIB	14	2A
CUDC-101	14	2A
DAUNORUBICIN	14	1A
DINACICLIB	14	2A
LIDOCAINE	14	1B
MITOMYCIN	14	1A
MSC-2363318A	14	2A
PEMETREXED	14	1A
PRAMLINTIDE	14	6
RO-5045337	14	2A
RUCAPARIB	14	1A
SUNITINIB	14	1A
VELIPARIB	14	2A
ALISERTIB	13	2A
AMG-900	13	2A
CABOZANTINIB	13	1A
FIMEPINOSTAT	13	2A
FULVESTRANT	13	1A
MASITINIB	13	2A
RG-1530	13	2A
SR-13668;NELFINAVIR	13	6;2A
TG100-801	13	6
TOPIRAMATE	13	4B
2X-121	12	2B
LENALIDOMIDE	12	1B
PD-0166285	12	4B
PREXASERTIB	12	2A
PRIMIDONE	12	4B
RONICICLIB;PHA-793887;AT-7519;AZD-5438	12	2A;2A;2A;2A

THALIDOMIDE	12	1B
ALVOCIDIB	11	2A
BOSUTINIB	11	1A
DURVALUMAB	11	1A
ENZALUTAMIDE	11	1A
HYALURONAN	11	6
LESTAURTINIB	11	2B
NERISPIRDINE	11	6
PACRITINIB	11	2B
ZOTIRACICLIB	11	2B
AEE-788	10	2A
AMUVATINIB	10	2A
CYC-116	10	2A
DALFAMPRIDINE;GUANIDINE HYDROCHLORIDE;TEDISAMIL	10	4B;6;6
DOMATINOSTAT	10	2A
DOVITINIB	10	2A
IMAGABALIN;ATAGABALIN;GABAPENTIN;PREGABALIN;BEPRIDIL HYDROCHLORIDE;GABAPENTIN ENACARBIL	10	6;6;3B;3B;6;6
KW-2449	10	2B
NERATINIB	10	1A
RG-7666;RECILISIB;WX-037;ZSTK-474;AZD-6482;PA-799;GSK-1059615	10	2B;6;2A;2A;6;2A;2A
SOTRASTAURIN	10	2B
TESEVATINIB	10	2A
ABEXINOSTAT;GIVINOSTAT	9	2A;2B
ANASTROZOLE	9	1A
AR-42	9	2B
AXITINIB	9	1A
AZM-475271;AZD-0424	9	6;2A
BARDOXOLONE METHYL	9	2B
BELINOSTAT	9	1A
BICALUTAMIDE	9	1A
BRYOSTATIN	9	3B
CARFILZOMIB	9	1A

COLLAGENASE CLOSTRIDIUM HISTOLYTICUM	9	6
CYTARABINE	9	1A
ELPETRIGINE	9	6
EXEMESTANE	9	1A
KX2-391	9	1B
LINIFANIB	9	2A
MUPARFOSTAT	9	2B
NILOTINIB	9	1A
NITRAZEPAM	9	4B
OPROZOMIB;IXAZOMIB CITRATE	9	2A;1A
PANOBINOSTAT	9	1A
PANOBINOSTAT LACTATE;TACEDINALINE;ENTINOSTAT	9	1B;2B;2A
PHENOBARBITAL	9	3A
POMALIDOMIDE	9	1B
RALFINAMIDE	9	6
RESMINOSTAT;ROMIDEPSIN;PRACINOSTAT;MOCETINOSTAT	9	2B;1A;2A;2A
RGB-286638	9	2B
ROCILETINIB;OSIMERTINIB	9	2A;1A
SARACATINIB	9	2A
SELICICLIB	9	2A
SUNITINIB MALATE	9	6
TALBUTAL;PENTOBARBITAL;BUTALBITAL;METHARBITAL	9	6;4A;6;6
AFIMOXIFENE	8	2B
AN-9;DEPAKOTE	8	4B;6
AVELUMAB	8	1A
BUTETHAL;BUTABARBITAL;HEXOBARBITAL;MEPHOBARBITAL;THIOPENTAL;SECOBARBITAL	8	6;6;4B;3B;4A;4A
CHIAURANIB	8	2A
ENFLURANE	8	4A
FORETINIB	8	2A
INDEGLITAZAR;GW501516;SODELGLITAZAR	8	6;4A;6

LENVATINIB	8	1A
MESALAMINE	8	2B
MOTESANIB	8	2A
NIMOTUZUMAB	8	2A
PEXIDARTINIB	8	1A
PROPOFOL	8	4A
QUIZARTINIB	8	2A
RIGOSERTIB	8	2A
SAFINAMIDE	8	4B
SAPANISERTIB	8	2A
SAPITINIB	8	2A
TEPROTUMUMAB	8	2A
TOZASERTIB	8	2A
VADIMEZAN	8	2A
VARLITINIB;MP-412;CANERTINIB;S-222611;ALLITINIB;LAPATINIB DITOSYLATE;PYROTINIB;TAK-285	8	2A;2A;2B;6;2B;1B;2A;2A
VINORELBINE	8	1A
XL-999	8	2A
ABITUZUMAB;INTETUMUMAB	7	2B;2B
ADINAZOLAM;CINOLAZEPAM;CLOTIAZEPAM;FLUDIAZEPAM	7	4B;6;6;6
AG-24322;RG-547	7	2B;2A
ALTIRATINIB	7	2A
AMOXAPINE	7	4B
ANLOTINIB	7	2A
AZD-4769;IMGATUZUMAB;MAB-425;RG-7160;RINDOPEPIMUT;ZALUTUMUMAB;PKI-166;NECITUMUMAB;FALNIDAMOL;DEPATUXIZUMAB;OLMUTINIB;MATUZUMAB;CHEMBL2347958;ICOTINIB;EPITINIB;THELIATINIB;SIMOTINIB;ERLOTINIB HYDROCHLORIDE	7	6;2A;2B;6;2B;2B;3B;1A;2A;2A;2A;2B;6;2A;2A;2A;2B;2A
BAFETINIB	7	2B

BRICICLIB	7	2A
BRIGATINIB	7	1A
CHEMBL3397300	7	6
CHLORDIAZEPOXIDE;CLOBAZAM;HALAZEPAM;OXAZEPAM;FLURAZEPAM;BROMAZEPAM;PRAZEPAM;QUAZEPAM;ADIPIPLON;MIDAZOLAM;ESTAZOLAM	7	4B;4B;6;3A;4B;4B;6;6;6;4B;4B
CLEVIDIPINE	7	4B
CM-082	7	2A
DANAZOL	7	3B
DESFLURANE;SEVOFLURANE	7	4A;4A
DRONEDARONE HYDROCHLORIDE	7	6
DULIGOTUZUMAB	7	2A
EMRICASAN	7	6
ERTUMAXOMAB	7	2A
FAMITINIB;SU-014813	7	2A;2A
GLPG-0187	7	2A
GLUFOSFAMIDE	7	2B
ISOFLURANE	7	3B
LITHIUM CITRATE;LITHIUM CARBONATE	7	6;2A
LUCITANIB	7	2A
LUMATEPERONE	7	6
MARIZOMIB	7	2A
MDX-447	7	6
MILCICLIB	7	2A
NAVEGLITAZAR;ALEGLITAZAR;TESAGLITAZAR;IMIGLITAZAR;MURAGLITAZAR	7	3B;6;3B;6;3B
NITROGLYCERIN	7	3B
NOSCAPINE;PX-478	7	3B;2B
OMIGAPIL	7	6
OSI-930;SEMAXANIB	7	2A;2A
PAZOPANIB HYDROCHLORIDE	7	1B
PEXMETINIB	7	2B
PROGESTERONE	7	2B
QUARFLOXIN	7	2B

REMIMAZOLAM;OCINAPLON	7	6;6
RG-4733;NIROGACESTAT	7	2A;2A
RIVICICLIB	7	6
TIDEGLUSIB;AZD-1080;LY-2090314	7	4A;4B;2A
TRIAZOLAM;CLONAZEPAM;CLORAZEPIC ACID;ALPRAZOLAM;DIAZEPAM;TEMAZEPAM;LORAZEPAM	7	4B;3B;6;4B;3B;4B;4B
UROKINASE	7	6
VATALANIB	7	2A
VERAPAMIL	7	2B
VINTAFOLIDE	7	2A
XL-820;TELATINIB	7	2A;2A
4SC-203	6	2B
ACAMPROSATE CALCIUM	6	6
ADH-1	6	2A
AFLIBERCEPT	6	1A
AMLODIPINE	6	2B
BALSALAZIDE DISODIUM;BALSALAZIDE;OLSALAZINE SODIUM	6	6;4B;6
BEPRIDIL	6	4B
BI-2536;GSK-461364;TAK-960;CAFUSERTIB;NMS-1286937;MK-1496;HMN-214	6	2A;2B;2A;2A;4B;2A;6
BMS-387032	6	2A
BMS-817378;CABOZANTINIB S-MALATE;BMS-794833;GOLVATINIB;TAS-115	6	2A;1A;6;2A;6
BPI-9016	6	2A
BRIVANIB ALANINATE;BRIVANIB	6	2A;2A
CELECOXIB	6	1A
CHEMBL1086377	6	6
CHIDAMIDE	6	6
CLOFARABINE	6	1A
CONATUMUMAB	6	2B
CONBERCEPT	6	6
CP-459632	6	6

DERSALAZINE	6	6
DEXAMETHASONE	6	1A
ENZASTAURIN	6	2A
ESTRIOL;ESTRADIOL ACETATE;ETHINYL ESTRADIOL;DIETHYLSTILBESTROL;ESTRADIOL CYPIONATE;DIENESTROL;ESTRONE;CLOMIPHENE CITRATE;RAD1901;LASOFOXIFENE;BRILANESTRANT;ARZOXIFENE;MESTRANOL;CLOMIPHENE;ESTROPIPATE;ESTRADIOL VALERATE;IODINE;POLYESTRADIOL PHOSPHATE;TAMOXIFEN CITRATE;ENDOXIFEN	6	4A;6;1A;1A;6;3B;3B;6;6;2B;2B;3B;3B;3B;3B;6;3A;1A;1B;6
ESZOPICLONE	6	6
ETARACIZUMAB	6	2A
ETOMIDATE	6	4B
FIRATEGRAST;NATALIZUMAB	6	6;2B
FP-1039	6	2A
GANCOTAMAB	6	2B
GLESATINIB;MGCD-265	6	2A;6
GTX-758	6	2B
HALOTHANE	6	4A
IBUPROFEN;SULINDAC;ASPIRIN	6	2B;2B;2A
ILOPROST	6	2B
IMATINIB MESYLATE	6	1B
INSM-18	6	3B
IRAMPANEL	6	6
K-877;BEZAFIBRATE;AVE8134;CP-778875;LY-518674;GEMFIBROZIL;FENOFIBRATE;CLOFIBRATE;FENOFIBRIC ACID;ZYH7;CHOLINE FENOFIBRATE;GW590735;DRF-10945	6	6;3B;6;6;4A;3B;3B;3B;4A;6;6;6;6
L19IL2;L19TNFA;L19SIP 131I	6	6;6;6

LEVONORGESTREL	6	3B
LOMITAPIDE MESYLATE	6	6
MEPROBAMATE;CLORAZEPATE DIPOTASSIUM;THIAMYLAL SODIUM;FLUMAZENIL;METHYPRYLON;CHLORDIAZEPOXIDE HYDROCHLORIDE;PENTOBARBITAL SODIUM;ETHCHLORVYNOL;BUTABARBITAL SODIUM;SECOBARBITAL SODIUM;METHOHEXITAL SODIUM;MIDAZOLAM HYDROCHLORIDE;THIOPENTAL SODIUM;GLUTETHIMIDE;PF-06372865;TRICLOFOS SODIUM	6	3B;6;6;4B;6;6;6;6;6;6;6;6;6;6;4B;6;6
METFORMIN HYDROCHLORIDE	6	2B
METHOXYFLURANE	6	3B
MM-111;MARGETUXIMAB	6	2B;1B
NANATINOSTAT	6	2A
NEFAZODONE HYDROCHLORIDE	6	6
NIMODIPINE	6	2B
NISOLDIPINE	6	4B
NORELGESTROMIN;DESOGESTREL;NORGESTREL;ETONOGESTREL;NORGESTIMATE;DIENOGEST;MEGESTROL	6	4B;3B;2B;4B;4B;4A;3B
ORANTINIB;NINTEDANIB ESYLATE	6	2A;6
ORPHENADRINE HYDROCHLORIDE;ORPHENADRINE CITRATE	6	6;6

OSPEMIFENE;IDOXIFENE;ACOLBIFENE;MK-6913;BAZEDOXIFENE;QUINESTROL;ESTROGENS, CONJUGATED SYNTHETIC A;SIVIFENE;TOREMIFENE;RALOXIFENE;ESTROGENS, ESTERIFIED;SR16234;ESTROGEN;FISPEMIFENE;RALOXIFEN;ESTROGENS, CONJUGATED;TOREMIFENE CITRATE;LY2245461;BAZEDOXIFENE ACETATE;DROLOXIFENE;DIETHYLSTILBESTROL DIPHOSPHATE;ESTRADIOL;SYNTHETIC CONJUGATED ESTROGENS, B;CHF4227;CHLOROTRIANISENE	6	3B;4B;6;6;2B;6;6;2B;1A;1B;2B;2B;6;6;6;6;1B;6;6;4B;1A;1B;6;6;2B
PEGPLERANIB SODIUM	6	2B
PENTOXIFYLLINE	6	2A
PF-04236921	6	6
PRILIXIMAB;CLENOLIXIMAB;TRX-1;ZANOLIMUMAB;CEDELIZUMAB;KELIXIMAB;TREGALIZUMAB;IBALIZUMAB	6	6;6;6;2B;6;6;6;6
RABEPRAZOLE SODIUM	6	1B
RAMUCIRUMAB	6	1A
RAVOXERTINIB;ULIXERTINIB;MK-8353	6	2A;2A;2A
RESEQUINIL;LOREDIPLON	6	4B;6
SILTUXIMAB;CLAZAKIZUMAB;OLOKIZUMAB;ELSILIMOMAB	6	2B;6;6;6
SIRUKUMAB	6	6
SONIDEGIB	6	1A
SULFASALAZINE	6	2B
SURUFATINIB	6	2A
TEDATIOXETINE	6	6
TOSEDOSTAT	6	2A
TRASTUZUMAB EMTANSINE	6	1A

TRAZODONE HYDROCHLORIDE	6	6
TRETINOIN	6	1A
VENETOCLAX	6	1A
VINBLASTINE	6	1A
VOLASERTIB	6	2A
X-82;TAK-593	6	2A;2A
ZONISAMIDE	6	3B
ABX-IL8;HUMAX-IL8	5	6;2A
ACETAMINOPHEN	5	2A
ALECTINIB;ALECTINIB HYDROCHLORIDE	5	1A;1B
APRINOCARSEN	5	2B
APROTININ	5	2B
AS-602868	5	4A
ASENAPINE	5	4B
AST-487	5	6
AVE0847;DB959	5	6;6
AZACITIDINE	5	1A
AZD-8055;OSI-027	5	2A;2B
AZD1305	5	6
BEGACESTAT;TARENFLURBIL	5	6;2B
C-1311	5	6
CABAZITAXEL	5	1A
CC-115	5	2B
CEP-11981	5	2A
CEP-37440	5	2A
CHEMBL247471;GSI-136;AVAGACESTAT;SEMAGACESTAT	5	6;6;4B;6

CHLORPHENTERMINE HYDROCHLORIDE;SERTRALINE HYDROCHLORIDE;PAROXETINE HYDROCHLORIDE;ESCITALOPRAM OXALATE;PAROXETINE MESYLATE;CLOMIPRAMINE HYDROCHLORIDE;FLUOXETINE HYDROCHLORIDE;FLUVOXAMINE MALEATE;CITALOPRAM HYDROBROMIDE	5	6;6;6;6;6;6;6;6;6
CILENGITIDE	5	2A
CIXUTUMUMAB	5	2A
COCAINE	5	3B
CRENOLANIB	5	2A
CROFELEMER	5	2B
CYCLOSPORINE	5	2A
DASOTRALINE;SIBUTRAMINE HYDROCHLORIDE;MAZINDOL;METHAMPHETAMINE HYDROCHLORIDE;LIAFENSINE	5	6;6;4B;6;6
DEFACTINIB	5	2A
DENILEUKIN DIFTITOX	5	1A
DESVENLAFAXINE	5	6
DIMETHYL FUMARATE	5	2B
DIPYRIDAMOLE	5	2B
ETAZOLATE	5	4B
EZOGABINE	5	4A
FAXELADOL	5	6
FELODIPINE	5	4A
FLUDARABINE PHOSPHATE;GEMCITABINE HYDROCHLORIDE	5	1A;1B
FLUNARIZINE	5	4B
FLUNITRAZEPAM	5	4A

FRUQUINTINIB;CEP-5214;L-21649;LENVATINIB MESYLATE;CEP-7055;AG-13958;TIVOZANIB;KRN-633	5	2A;6;6;1A;6;6;1A;3B
GABOXADOL	5	4B
GANAXOLONE	5	6
GANETESPIB	5	2A
GSK163090	5	6
HEMAY-022;MUBRITINIB;TUCATINIB;CP-724714	5	2B;2B;1A;2B
IBUTILIDE	5	4B
IDARUBICIN	5	1A
INDOMETHACIN	5	2B
INIPARIB	5	2A
JI-101	5	2A
LEUCOVORIN	5	1B
LEVOMILNACIPRAN HYDROCHLORIDE;VENLAFAXINE HYDROCHLORIDE;DESVENLAFAXINE SUCCINATE;PROTRIPTYLINE HYDROCHLORIDE;DULOXETINE HYDROCHLORIDE;IMIPRAMINE HYDROCHLORIDE;AMITRIPTYLINE HYDROCHLORIDE;NORTRIPTYLINE HYDROCHLORIDE	5	6;2B;4B;6;2B;6;6;6
LINSITINIB	5	2A
LOMUSTINE	5	1A
LUMACAFTOR;IVACAFTOR;QBW251;TEZACAFTOR	5	4B;4B;6;6
MDX-210	5	6
NIVOCASAN	5	6
NV-128;ME-344	5	6;2A
OBATOCLAX MESYLATE	5	2B
OLESOXIME	5	4B
ONALESPIB	5	2A
PALIFERMIN	5	2B
PEGAPTANIB SODIUM	5	2B

PF-03814735	5	2A
PF-04605412	5	2A
PONATINIB HYDROCHLORIDE	5	1B
PX-12	5	2B
RANIBIZUMAB;BEVACIZUMAB 111IN;BROLUCIZUMAB;BEVASIRANIB	5	2B;6;6;6
RETASPIMYCIN;RETASPIMYCIN HYDROCHLORIDE;TANESPIMYCIN;XL-888;ALVESPIMYCIN;CHEMBL553939	5	2A;2A;2A;2A;2A;2A
RIVOCERANIB;HENATINIB;OSI-632;PEGDINETANIB;PF-00337210;ALACIZUMAB PEGOL;ZD-4190	5	2A;2A;2B;2A;2A;6;6
RUXOLITINIB	5	1A
SU-9516;ERIBULIN	5	4A;1A
TANDUTINIB	5	2B
TAZEMETOSTAT	5	1B
TELBERMIN;ABICIPAR PEGOL	5	6;6
TREPROSTINIL	5	4B
TTP-607;SNS-314;DANUSERTIB;MK-6592	5	2B;2A;2B;2A
VIDOFLUDIMUS	5	6
VILAZODONE HYDROCHLORIDE	5	6
VINCRISTINE	5	1A
VOCLOSPORIN	5	6
VOLOCIXIMAB	5	2B
VORTIOXETINE HYDROBROMIDE	5	6
ZALTRAP	5	6
ZOLPIDEM	5	4B
ZOPICLONE	5	4B
AEW-541	4	4A
ALITRETINOIN;ETRETINATE;ADAPALENE;ACITRETIN;MOFAROTENE	4	1A;3B;3B;2B;6

AMG-548;SC-80036;R-1487;AZD-6703;TALMAPIMOD;PG-760564;TA-5493;SCIO-323;DILMAPIMOD;RALIMETINIB;RO-3201195;FX-005;DORAMAPIMOD;GSK-610677;NEFLAMAPIMOD;PS-516895;AVE-9940;SD-0006;LY-3007113;PAMAPIMOD;AZD-7624;SB-85635;TAK-715;LEO-15520;ACUMAPIMOD;PF-03715455;KC-706	4	6;6;6;6;2B;6;6;6;6;2A;6;6;3B;6;4A;6;6;6;6;6;6;6;6;6;6;6;6
AMINOCAPROIC ACID	4	4B
AMLODIPINE BESYLATE;DILTIAZEM HYDROCHLORIDE;VERAPAMIL HYDROCHLORIDE;DILTIAZEM MALATE;NICARDIPINE HYDROCHLORIDE;AMLODIPINE MALEATE	4	6;6;2B;6;6;6
ANDROGRAPHOLIDE;BARDOXOLONE	4	2B;2B
AR9281	4	6
ARHALOFENATE	4	6
ARRY-300	4	6
ARSENIC TRIOXIDE	4	1A
ASCIMINIB;NILOTINIB HYDROCHLORIDE MONOHYDRATE;FLUMATINIB	4	1B;1B;2B
BECATECARIN	4	2A
BIRABRESIB	4	2B
BMS-754807;KW-2450	4	2A;2A
BRENTUXIMAB VEDOTIN	4	1A
CARMUSTINE	4	1A
CATUMAXOMAB;SOLITOMAB	4	1A;6
CERDULATINIB	4	2B
CERITINIB	4	1A
COLCHICINE	4	2B
DASOLAMPANEL	4	6
DELCASERTIB	4	6

DIACEREIN	4	4A
DIAPEP-277	4	6
DIGOXIN;ACETYLDIGITOXIN;DESLANOSIDE;DIGITOXIN	4	2A;6;6;3B
DROSPIRENONE	4	4B
ECALLANTIDE	4	6
ENOXAPARIN	4	3B
ENTRECTINIB	4	1A
ERITORAN TETRASODIUM	4	2B
ERLOSAMIDE	4	2B
ETALOCIB	4	2A
FLINDOKALNER	4	6
FLUDARABINE	4	2B
FONTOLIZUMAB	4	6
FOSPROPOFOL	4	6
G1T28-1;RIBOCICLIB SUCCINATE	4	6;1A
GANITUMAB	4	2A
GEVOKIZUMAB;CANAKINUMAB	4	6;2B
GILTERITINIB	4	1A
HEPARIN	4	3B
IFOSFAMIDE	4	1A
LEVETIRACETAM	4	2B
LICOFELONE	4	3B
LITRONESIB;ISPINESIB	4	2A;2A
LY-2584702	4	2A
LY-2606368;AZD-7762;RABUSERTIB;PF-00477736;RG-7602;SCH-900776;RG-7741	4	2A;2A;2A;2A;2A;2B;2B
MECLOFENAMATE SODIUM	4	6
MENADIONE	4	3B
MERCAPTOPURINE	4	1A
MITOXANTRONE	4	1A

MK-0533;INT131;FARGLITAZAR;RIVOGLITAZONE;EFATUTAZONE HYDROCHLORIDE;PIOGLITAZONE;CLX-0921;ROSIGLITAZONE;MBX-2044;ATX08-001;METAGLIDASEN;BALAGLITAZONE;NETOGLITAZONE;PIOGLITAZONE HYDROCHLORIDE;EFATUTAZONE;ROSIGLITAZONE MALEATE;OMS405;MITOGLITAZONE;TROGLITAZONE	4	6;4A;4A;6;6;3B;6;3A;6;6;6;6;3B;6;2A;6;6;6;3A
MK-8033	4	2A
MLN-8054;MK-5108;TAS-119	4	2A;2A;2A
MSC-2364447;SPEBRUTINIB BESYLATE;HM-71224;TIRABRUTINIB	4	6;2B;6;2A
NAPROXCINOD	4	6
NAVITOCLAX	4	2A
NIFEDIPINE;ISRADIPINE	4	3B;4B
OLANZAPINE	4	2A
OMARIGLIPTIN;VILDAGLIPTIN;BEGELOMAB;SAXAGLIPTIN;ALOGLIPTIN;BISEGLIPTIN;LINAGLIPTIN;SITAGLIPTIN PHOSPHATE;GOSOGLIPTIN;CARMEGLIPTIN;SAXAGLIPTIN HYDROCHLORIDE;TRELAGLIPTIN SUCCINATE;SITAGLIPTIN;ALOGLIPTIN BENZOATE;DUTOGLIPTIN	4	6;4A;6;4A;6;6;2B;4A;6;6;6;6;2B;6;6
ONERCEPT;ADALIMUMAB;AZ9773;CERTOLIZUMAB PEGOL;AFELIMOMAB;NERELIMOMAB;PLACULUMAB;INFLIXIMAB;GOLIMUMAB;LENERCEPT;OZORALIZUMAB;PEGSUNERCEPT;ETANERCEPT	4	6;2A;6;6;6;6;6;2B;6;6;6;6;6;2A
OPORTUZUMAB MONATOX	4	2B

PICROTOXIN	4	4A
PIMECROLIMUS	4	3B
PKI-179;PWT-33579	4	2A;6
PROPAFENONE HYDROCHLORIDE	4	6
R-333;R-348	4	6;6
REBASTINIB	4	2A
RESATORVID;ERITORAN	4	6;2B
RG-7603;RIDAFOROLIMUS;CC-223;PALOMID-529;DS-3078A;AZD-2014	4	2B;2A;2A;3B;2B;4B
RILONACEPT	4	6
RILUZOLE	4	2B
RIMEXOLONE;BECLOMETHASONE;TRIAMCINOLONE;ALCLOMETASONE	4	4A;3B;2B;6
ROVELIZUMAB	4	6
RUBOXISTAURIN	4	4B
RUCAPARIB CAMSYLATE	4	1B
RWJ-67657;LOSMAPIMOD	4	6;6
SALINOMYCIN	4	3B
SALIRASIB	4	6
SAN-300	4	6
SAPROPTERIN	4	4A
SELURAMPANEL	4	6
SEMAPIMOD;ARRY-797;VX-702;BMS-582949;PH-797804;PIRFENIDONE	4	4A;6;6;6;6;2B
SERTINDOLE	4	4B
STIBOGLUCONATE	4	6
TALAZOPARIB TOSYLATE	4	1B
TANZISERTIB;BRIMAPITIDE, C-TERMINAL ACID;BENTAMAPIMOD;CC-401	4	6;4A;6;2B

TEPOTINIB;CAPMATINIB;ONARTUZUMAB;SAR-125844;JNJ-38877605;PF-04217903;NAQUOTINIB;AMG-337;SAVOLITINIB;BMS-698769;TIVANTINIB;AMG-208;EMD-1204831;COMETRIQ;SGX-523;TELISOTUZUMAB;MK-2461;MERESTINIB;BMS-777607	4	1A;1A;2A;2A;2A;2A;2A;2A;2A;2A;6;2A;2A;2A;6;2A;2A;2A;2A;2A
TESMILIFENE	4	3B
TEZAMPANEL	4	6
THIAMYLAL	4	4A
THIOGUANINE	4	1A
TOPSALYSIN	4	2B
TRABECTEDIN	4	1A
VINCRISTINE SULFATE;VINBLASTINE SULFATE;PACLITAXEL POLIGLUMEX;VERUBULIN;SAGOPILONE;VINORELBINE TARTRATE;INDIBULIN;VINFLUNINE;ERIBULIN MESYLATE;CROLIBULIN;FOSBRETABULIN DISODIUM;FOSBRETABULIN TROMETHAMINE;DAVUNETIDE;IXABEPILONE;LEXIBULIN;PLINABULIN	4	1B;1B;2A;2A;2A;2A;2A;1A;1B;2A;2A;2B;4B;1A;2B;2A
WX-554	4	2A
XL-844	4	2B
ZALEPLON	4	4B
ZIPRASIDONE	4	4B
ZONAMPANEL;PERAMPANEL;MIBAMPATOR;MK-8777;FARAMPATOR;BECAMPANEL;CX1739;PF-04958242;TALAMPANEL	4	6;5;6;6;6;6;6;6;6
ABCIXIMAB	3	6
ABT-510	3	2A
AMINOGLUTETHIMIDE	3	1A
AMIODARONE;DRONEDARONE	3	2B;4A

ANDECALIXIMAB	3	2B
APABETALONE	3	4B
APRINDINE	3	4B
ARSENIC	3	3B
ASM-024	3	6
ASPARAGINASE	3	1A
ASUNERCEPT	3	2B
ATAMESTANE;TESTOLACTONE;ARIMIDEX	3	2B;4B;6
AZD7325;AZD6280	3	6;6
BALICATIB;CHEMBL1085282;RELACATIB	3	6;6;6
BARASERTIB;BI-811283;TAK-901	3	2A;2A;2B
BECAPLERMIN	3	6
BELNACASAN	3	4A
BEXAROTENE;IRX-4204	3	1A;2A
BI-505;ENLIMOMAB PEGOL;ALICAFORSEN	3	2B;6;6
BIFEPRUNOX	3	4B
BMS-936559	3	6
BOTULINUM TOXIN TYPE B	3	6
BUDIODARONE	3	6
CANDOXATRIL;ILEPATRIL	3	4B;6
CARLUMAB	3	2A
CARVEDILOL PHOSPHATE;CARVEDILOL	3	6;3B
CHLOROTHIAZIDE	3	6
CLODRONIC ACID	3	3B
CORDYCEPIN	3	2B
CR 1447	3	6
CT-1578	3	6
CYT006-ANGQB	3	6
DANTROLENE SODIUM	3	6
DESMOTEPLASE;AMEDIPLASE;STREPTOKINASE; TENECTEPLASE;TRANEXAMIC ACID;DEFIBROTIDE SODIUM	3	6;6;6;6;2A;6
DEXANABINOL	3	3B
DEXTROMETHORPHAN POLISTIREX;DEXTROMETHORPHAN HYDROBROMIDE	3	6;6

DIMIRACETAM	3	6
DOXYCYCLINE	3	2A
DROXIDOPA;HYDROXYAMPHETAMINE HYDROBROMIDE;EPINEPHRINE BITARTRATE;NOREPINEPHRINE BITARTRATE;MEPHENTERMINE SULFATE;LABETALOL HYDROCHLORIDE;DIPIVEFRIN HYDROCHLORIDE	3	6;6;6;6;6;6;6
DUSIGITUMAB	3	2A
ENASIDENIB	3	1A
EPINEPHRINE	3	3B
EPIRUBICIN	3	1A
EPTIFIBATIDE;TIROFIBAN;GANTOFIBAN;TIROFIBAN HYDROCHLORIDE;TADOCIZUMAB	3	4A;6;6;6;6
ETIRINOTECAN PEGOL;IRINOTECAN HYDROCHLORIDE HYDRATE;CAMPTOTHECIN;LAPACHONE;IRINOTECAN SUCROSOFATE;LURTOTECAN;RUBITECAN;KARENITECIN	3	2A;6;3B;2A;2A;2A;2A;2B
FEDRATINIB	3	1A
FIBRINOLYSIN, HUMAN	3	6
FIGITUMUMAB;MECASERMIN;DALOTUZUMAB;PL-225B;MECASERMIN RINFABATE;ROBATUMUMAB;AVE-1642;CHEMBL401930;AXL-1717	3	2A;6;2A;2A;6;2A;2B;4B;2A
FLAVOXATE HYDROCHLORIDE	3	6
FLUOROMETHOLONE;FLURANDRENOLIDE;FLUNISOLIDE;MEDRYSONE;FLUOCINONIDE	3	6;4B;4A;6;4B

FOSTAMATINIB;R-343;ENTOSPLETINIB;R-112;R-406;HMPL-523;TAK-659;PRT-2607	3	2B;6;2A;6;6;2A;2B;6
FRESOLIMUMAB	3	2B
GALETERONE;ABIRATERONE	3	2B;1A
GAVILIMOMAB	3	6
GSK-1070916	3	2A
HALOFUGINONE	3	2A
HUHMFG1	3	2B
HYDROCORTAMATE;CLOCORTOLONE;PREDNISOLONE;DIFLORASONE;MOMETASONE;METHYLPREDNISOLONE;AMCINONIDE;PREDNICARBATE;CLOBETASOL;DESOXIMETASONE;PREDNISONE;DESONIDE;BETAMETHASONE	3	6;6;1B;4A;5;1B;6;6;4A;6;1B;6;2B
HYDROXYUREA;GALLIUM NITRATE;TEZACITABINE	3	1A;2B;2A
IDELALISIB	3	1A
IDURSULFASE;GALSULFASE	3	6;6
ILOPERIDONE	3	4B
INDANTADOL	3	6
ISOPROTERENOL HYDROCHLORIDE;CARTEOLOL HYDROCHLORIDE;LY377604	3	6;6;6
ISOSORBIDE	3	2A
IVOSIDENIB	3	1A
LABETUZUMAB	3	2B
LAPAQUISTAT ACETATE	3	6
LATREPIRDINE	3	6
LEPIRUDIN;ATECEGATRAN METOXIL;XIMELAGATRAN;BIVALIRUDIN;DESIRUDIN;ARGATROBAN;DABIGATRAN;DABIGATRAN ETEXILATE MESYLATE	3	2B;6;4B;4B;6;4A;4A;6
LUBIPROSTONE	3	4B
LUNACALCIPOL	3	6

LXR-623;CS-8080;BMS-852927;HYODEOXYCHOLIC_ACID	3	6;6;6;4B
MARIMASTAT	3	2B
MELOXICAM;CARPROFEN;ETORICOXIB;VALDECOXIB;CIMICOXIB;ROFECOXIB;APRICOXIB;PARECOXIB;LUMIRACOXIB	3	2B;4A;2B;4A;6;2B;2B;4A;4A
METELIMUMAB;LY-2382770	3	6;6
METHIMAZOLE;CARBIMAZOLE	3	2B;4B
MIFEPRISTONE	3	2B
MIPOMERSEN	3	6
MLN-0415;IMD-1041;IMD-2560;SAR-113945;AURANOFIN;IMD-0354	3	6;6;6;6;2B;4A
MYCOPHENOLATE MOFETIL;MYCOPHENOLIC ACID;AVN-944;MYCOPHENOLATE MOFETIL HYDROCHLORIDE	3	2A;2B;2B;6
NEBOGLAMINE;ESKETAMINE HYDROCHLORIDE;AV-101;AMANTADINE HYDROCHLORIDE;GW468816;MEMANTINE HYDROCHLORIDE;CNS-5161;KETAMINE HYDROCHLORIDE;APIMOSTINEL;FELBAMATE;LANICEMINE;AZD8108;DELUCEMINE;RAPASTINEL;ESKETAMINE	3	6;6;6;6;6;6;6;6;6;4B;6;6;6;6;6
NERAMEXANE MESYLATE	3	6

NKTR-171;ESLICARBAZEPINE;BENOXINATE HYDROCHLORIDE;ETHOTOIN;PHENYTOIN;ARTICAINE HYDROCHLORIDE;FOSPHENYTOIN SODIUM;RUFINAMIDE;CARBAMAZEPINE;OXCARBAZEPINE;TOCAINIDE HYDROCHLORIDE;PROPOXYCAINE HYDROCHLORIDE;ETIDOCAINE HYDROCHLORIDE;TETRACAINE;PHENACEMIDE;HEXYLCAINE HYDROCHLORIDE;INDECAINIDE HYDROCHLORIDE;TETRACAINE HYDROCHLORIDE;CHLOROPROCAINE HYDROCHLORIDE;LAMOTRIGINE;ROPIVACAINE HYDROCHLORIDE;PRILOCAINE HYDROCHLORIDE;PHENAZOPYRIDINE HYDROCHLORIDE;PROCAINAMIDE HYDROCHLORIDE;MERETHOXYLLINE PROCAINE;MEPHENYTOIN;DYCLONINE HYDROCHLORIDE;PRILOCAINE;PROPARACAINE HYDROCHLORIDE;MEPIVACAINE HYDROCHLORIDE;QUINIDINE POLYGALACTURONATE;QUINIDINE GLUCONATE;MEXILETINE HYDROCHLORIDE;MORICIZINE HYDROCHLORIDE;DISOPYRAMIDE PHOSPHATE;PHENYTOIN SODIUM;EVENAMIDE;LIDOCAINE HYDROCHLORIDE;ESLICARBAZEPINE ACETATE;PROCAINE HYDROCHLORIDE;QUINIDINE SULFATE	3	6;6;6;4B;2B;6;6;4B;3B;2A;6;6;6;4B;6;6;6;6;6;2B;2B;6;6;6;6;4B;6;2B;6;6;6;6;6;6;6;6;6;1B;6;6;6
OCTREOTIDE	3	1A
OCTREOTIDE ACETATE	3	6
ODANACATIB	3	2B
OLARATUMAB;AVAPRITINIB;TOVETUMAB	3	1A;1A;2A
OMECAMTIV MECARBIL	3	6
ORPHENADRINE	3	4A
OSI-7904;FLOXURIDINE;RALTITREXED;ARFOLITIXORIN	3	2A;1A;1A;2B
OXTRIPHYLLINE;AMINOPHYLLINE;THEOPHYLLINE SODIUM GLYCINATE	3	6;2B;6
PAGOCLONE	3	6
PARDOPRUNOX	3	4B

PASIREOTIDE DIASPARTATE;PASIREOTIDE PAMOATE;PASIREOTIDE	3	6;6;4B
PATRITUMAB	3	2A
PEVONEDISTAT	3	2A
PF-00562271;GSK-2256098;VS-4718;BI-853520	3	6;2A;2B;2A
PHENYLBUTAZONE	3	3B
PICOPLATIN	3	2A
PINOMETOSTAT	3	2B
POLMACOXIB	3	6
PRINOMASTAT	3	2B
PROBUCOL	3	3B
PROCARBAZINE	3	1A
PROPIOMAZINE;PROMAZINE	3	6;4B
PROPYLTHIOURACIL	3	3A
QUERCETIN-3'-O-PHOSPHATE	3	6
RALOXIFENE HYDROCHLORIDE;MF101;ESTRAMUSTINE PHOSPHATE SODIUM;ERTEBEREL;PRINABEREL;AUS-131	3	1B;4A;6;6;4B;3B
RETEPLASE;ALTEPLASE;ANISTREPLASE	3	6;2A;6
RILOTUMUMAB;FICLATUZUMAB	3	2B;2B
S-3304;INCYCLINIDE	3	2A;2A
SARGRAMOSTIM	3	3B
SCH-708980	3	6
SCY 635	3	6
SDX-101	3	6
SECUKINUMAB;IXEKIZUMAB	3	6;6
SELINEXOR	3	1A

SIMVASTATIN;ATORVASTATIN CALCIUM;FLUVASTATIN;ROSUVASTATIN CALCIUM;PITAVASTATIN CALCIUM;FLUVASTATIN SODIUM;LOVASTATIN;ROSUVASTATIN;PRAVASTATIN;PITAVASTATIN;CERIVASTATIN SODIUM;PRAVASTATIN SODIUM;ATORVASTATIN	3	2A;2A;2B;3A;6;2B;2B;2A;2A;2B;6;2B;2A
SOTAGLIFLOZIN	3	6
STANNSOPORFIN	3	4A
SUPROFEN;NEPAFENAC;NAPROXEN ETEMESIL;OXAPROZIN;BISMUTH SUBSALICYLATE;IBUPROFEN LYSINE;KETOPROFEN;OXYPHENBUTAZONE;PIROXICAM;OXAPROZIN POTASSIUM;BROMFENAC SODIUM;DIFLUNISAL;NAPROXEN;SALSALATE;AMINOSALICYLATE POTASSIUM;TOLMETIN;FLURBIPROFEN;NAPROXEN SODIUM;FENOPROFEN;NABUMETONE;DICLOFENAC SODIUM;ETODOLAC;BROMFENAC;DICLOFENAC EPOLAMINE;FENOPROFEN CALCIUM;TENOXICAM;VB-201;IBUPROFEN SODIUM;DICLOFENAC POTASSIUM;KETOROLAC TROMETHAMINE;TOLMETIN SODIUM;FLURBIPROFEN SODIUM;MEFENAMIC ACID;DICLOFENAC;KETOROLAC;AMINOSALICYLATE SODIUM	3	4B;6;6;4A;2B;6;4A;4B;3B;6;6;4A;2B;4A;6;4A;3B;6;4A;4A;2B;2B;4A;2B;6;6;6;6;6;4A;6;6;4A;2A;2B;4A
TAFAMIDIS;TAFAMIDIS MEGLUMINE	3	4B;6
TAMIBAROTENE	3	2B
TAREXTUMAB	3	2A

TEGAFUR	3	1A
TETRODOTOXIN	3	4A
THEOPHYLLINE	3	2B
THIORIDAZINE	3	3B
THYROTROPIN;THYROTROPIN ALFA	3	6;1A
TILARGININE;KD7040	3	2B;6
TOFACITINIB CITRATE;PEFICITINIB	3	6;6
TOFERSEN	3	6
TOPOTECAN	3	1A
TREMELIMUMAB	3	2A
TRICHLORMETHIAZIDE	3	4B
TRIMIPRAMINE MALEATE	3	6
TUCOTUZUMAB CELMOLEUKIN;ADECATUMUMAB;CITATUZUMAB BOGATOX;ING-1	3	2B;2B;2A;2A
VAPREOTIDE	3	6
VARENICLINE	3	3B
VISMODEGIB	3	1A
VISTUSERTIB	3	2A
VORUCICLIB	3	2A
ZOSUQUIDAR;TARIQUIDAR	3	2B;2B
1D09C3	2	6
ABATACEPT;BELATACEPT	2	2B;6
ACAMPROSATE	2	4B
ACARBOSE	2	3B
ACP-001;PEGVISOMANT;SOMATREM;SOMATROGON	2	6;2B;6;6
AFAMELANOTIDE	2	6
AKN-028	2	2B
ALENDRONIC ACID	2	2B
ALFIMEPRASE	2	6
ALISKIREN FUMARATE;ALISKIREN;REMIKIREN;VTP-27999	2	6;4A;6;6
AMG-108	2	6

AMRUBICIN;DAUNORUBICIN CITRATE;PODOFILOX;VALRUBICIN;ETOPOSIDE PHOSPHATE;ALDOXORUBICIN;IDARUBICIN HYDROCHLORIDE;13-DEOXYDOXORUBICIN;MITOXANTRONE HYDROCHLORIDE;AMRUBICIN HYDROCHLORIDE;DOXORUBICIN HYDROCHLORIDE;DAUNORUBICIN HYDROCHLORIDE	2	2A;2B;3B;1A;1A;2A;2B;2A;1B;2B;1A;1B
ANACETRAPIB;DALCETRAPIB;EVACETRAPIB;TORCETRAPIB	2	4A;3B;6;4A
ANG1005;ORTATAXEL;MILATAXEL;TESETAXEL	2	2A;2B;2B;2A
ANRUKINZUMAB;QAX-576;TNX-650;TRALOKINUMAB;LEBRIKIZUMAB	2	6;6;2B;2B;6
APATORSEN	2	2A
APOMAB;LEXATUMUMAB;TIGATUZUMAB;DROZITUMAB	2	6;2B;2B;2B
APOMORPHINE	2	4A
ASP-3026;LORLATINIB	2	2A;1A
AT-13148	2	2A
AZATHIOPRINE	2	2B
AZD-7451;AZD-6918;LAROTRECTINIB	2	2B;2A;1A
BAMINERCEPT	2	6
BAY-1082439;SAR-260301	2	2A;2A
BERMEKIMAB	2	2A
BERZOSERTIB	2	2A
BIMOSIAMOSE;RIVIPANSEL	2	6;6
BLEOMYCIN	2	1A
BLINATUMOMAB	2	1A
BOCOCIZUMAB;EVOLOCUMAB;ALIROCUMAB;RG-7652	2	6;6;6;6

BREMELANOTIDE	2	6
BREXPIPRAZOLE	2	4B
BRIAKINUMAB;USTEKINUMAB	2	6;2B
BURIXAFOR;MSX-122;PLERIXAFOR;MAVORIXAFOR;POL6326;CTCE-9908;BKT140	2	2B;2A;1B;2B;2B;4B;2B
CAROTUXIMAB	2	2A
CE-245677	2	6
CEBRANOPADOL	2	6
CETILISTAT	2	4B
CHLORPROTHIXENE	2	4B
CHLORZOXAZONE	2	4B
CICLESONIDE;PARAMETHASONE	2	4B;3B
CLOZAPINE	2	4A
CNV2197944;ZICONOTIDE;Z160;ZICONOTIDE ACETATE	2	6;6;6;6
COBIPROSTONE	2	6
CRISABOROLE;APREMILAST;ROFLUMILAST;AML EXANOX	2	6;6;2A;4A
CUSTIRSEN SODIUM;CUSTIRSEN	2	2B;2A
DANTROLENE	2	3B
DAPRODUSTAT;ROXADUSTAT	2	6;2B
DIAZEPINOMICIN;CHLORMEZANONE	2	2B;6
DIAZOXIDE	2	4A
DICHLORPHENAMIDE;ACETAZOLAMIDE SODIUM;ACETAZOLAMIDE	2	4B;6;2B
DOFETILIDE	2	3B
DRINABANT;SURINABANT;MARINOL;CANNABIDIOL;TARANABANT;RIMONABANT;IBIPINABANT;AZD2207;OTENABANT	2	6;4B;6;3B;4B;3B;4B;6;6
DULANERMIN	2	2B
DUTASTERIDE	2	1B
DYPHYLLINE	2	6
EFALIZUMAB;LIFITEGRAST	2	6;6

EFAPROXIRAL	2	2A
EMICIZUMAB	2	6
ENOXAPARIN SODIUM;FONDAPARINUX;TINZAPARIN SODIUM;FONDAPARINUX SODIUM;DALTEPARIN SODIUM;DANAPAROID SODIUM;HEPARIN CALCIUM;HEPARIN SODIUM;ARDEPARIN SODIUM;SEMULOPARIN SODIUM	2	2A;6;2B;6;6;6;6;6;6;6
ENSARTINIB;CHEMBL473773;LUMINESPIB	2	2B;4B;2A
ERDAFITINIB	2	1A
ERGOLOID MESYLATES	2	6
ETHOXZOLAMIDE	2	4B
EZATIOSTAT HYDROCHLORIDE;EZATIOSTAT;CANFOSFAMIDE	2	2B;2B;2B
FEBUXOSTAT;OXYPURINOL;ALLOPURINOL;ALLOPURINOL SODIUM	2	4A;4A;1A;6
FINASTERIDE	2	1B
FLANVOTUMAB	2	2B
FLUPIRTINE	2	4B
FOLLITROPIN ALFA;UROFOLLITROPIN	2	6;6

FOSINOPRIL SODIUM;ENALAPRIL MALEATE;MOEXIPRIL;BENAZEPRIL;SPIRAPRIL;RAMIPRIL;CAPTOPRIL;FOSINOPRIL;BENAZEPRIL HYDROCHLORIDE;TRANDOLAPRIL;LISINOPRIL HYDRATE;QUINAPRIL;MOEXIPRIL HYDROCHLORIDE;CILAZAPRIL;PERINDOPRIL ERBUMINE;PERINDOPRIL;ENALAPRIL;SPIRAPRIL HYDROCHLORIDE;QUINAPRIL HYDROCHLORIDE;PERINDOPRIL ARGININE	2	6;6;4B;3B;4B;2B;2A;4A;6;4A;6;4A;6;4B;6;4A;2A;6;6;6
FUTIBATINIB;PEMIGATINIB	2	2A;1A
GATAPARSEN	2	6
GIMERACIL;ENILURACIL	2	2A;2B
GLIPIZIDE;CHLORPROPAMIDE;GLYBURIDE;NATEGLINIDE;GLIMEPIRIDE;REPAGLINIDE;ACETOHEXAMIDE;TOLBUTAMIDE;TOLAZAMIDE	2	4B;4B;4A;4B;4B;2B;4B;2A;4B
GO-203-2C	2	2A
GSK2798745	2	6
GSK933776;CRENEZUMAB;PONEZUMAB;SCYLLITOL;BAN2401;SOLANEZUMAB;ADUCANUMAB;BAPINEUZUMAB;GANTENERUMAB;VANUTIDE CRIDIFICAR;CAPROSPINOL	2	6;6;2B;3B;6;6;6;6;6;6;6
IBUDILAST	2	2B
ICRUCUMAB	2	2A
IDRONOXIL	2	2A
INCOBOTULINUMTOXINA;ONABOTULINUMTOXINA;ABOBOTULINUMTOXINA	2	4A;2B;6
INDAPAMIDE	2	4B
INDIPLON;METHOHEXITAL	2	6;4A
INK-1117;SERABELISIB	2	6;2A

INSULIN ZINC SUSP RECOMBINANT HUMAN;INSULIN SUSP PROTAMINE ZINC PURIFIED PORK;INSULIN ZINC SUSP BEEF;INSULIN ASPART;INSULIN SUSP ISOPHANE RECOMBINANT HUMAN;INSULIN ZINC SUSP PURIFIED PORK;INSULIN LISPRO PROTAMINE RECOMBINANT;INSULIN;INSULIN DEGLUDEC;INSULIN LISPRO;INSULIN ZINC SUSP EXTENDED PURIFIED BEEF;INSULIN SUSP PROTAMINE ZINC BEEF/PORK;INSULIN GLARGINE;INSULIN GLULISINE;INSULIN ZINC SUSP SEMISYNTHETIC PURIFIED HUMAN;INSULIN PEGLISPRO;INSULIN ASPART PROTAMINE RECOMBINANT;INSULIN ZINC SUSP EXTENDED BEEF;INSULIN SUSP ISOPHANE SEMISYNTHETIC PURIFIED HUMAN;INSULIN ZINC SUSP PROMPT PURIFIED PORK;INSULIN SUSP ISOPHANE BEEF/PORK;INSULIN PURIFIED BEEF;INSULIN ZINC SUSP PURIFIED BEEF;INSULIN SUSP PROTAMINE ZINC PURIFIED BEEF;INSULIN SUSP ISOPHANE BEEF;INSULIN SUSP ISOPHANE PURIFIED BEEF;INSULIN DETEMIR;INSULIN PURIFIED PORK;INSULIN ZINC SUSP PURIFIED BEEF/PORK;INSULIN ZINC SUSP PROMPT BEEF;INSULIN PORK;INSULIN ZINC SUSP EXTENDED RECOMBINANT HUMAN;INSULIN SUSP ISOPHANE PURIFIED PORK	2	6;6;6;6;6;6;6;6;6;6;6;4B;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6
IROSUSTAT	2	2B
ISOTRETINOIN;FENRETINIDE;IRX4310	2	2A;2A;6
ISTAROXIME	2	6
ITOPRIDE	2	4B
IXAZOMIB	2	1A
LABETALOL	2	3B
LEUPROLIDE	2	1A
LEVOLEUCOVORIN;BHQ-880	2	6;2B
LEVOMEPROMAZINE	2	4B

LEVOTHYROXINE;LIOTHYRONINE;DEXTROTHYROXINE SODIUM;LEVOTHYROXINE SODIUM;LIOTHYRONINE SODIUM;DEXTROTHYROXINE	2	3B;5;6;6;4A;6
LISINOPRIL	2	2B
LOXAPINE	2	4B
LOXAPINE SUCCINATE;THIORIDAZINE HYDROCHLORIDE;OLANZAPINE PAMOATE;TRIFLUOPERAZINE HYDROCHLORIDE;LOXAPINE HYDROCHLORIDE	2	6;6;6;6;6
LURASIDONE	2	6
LYM-1	2	2B
LYSINE	2	4B
MAVRILIMUMAB	2	6
MEDROXYPROGESTERONE ACETATE;TELAPRISTONE ACETATE;ETHYNODIOL DIACETATE;MEGESTROL ACETATE;ONAPRISTONE;ULIPRISTAL;ULIPRISTAL ACETATE;HYDROXYPROGESTERONE CAPROATE;GESTODENE;PROGESTOGEN;LONAPRISAN;NORETHINDRONE;NORETHINDRONE ACETATE;DYDROGESTERONE;PF-05019702;ASOPRISNIL;NORETHYNODREL	2	1A;2B;4B;1A;2A;6;6;3B;4B;6;6;3B;3B;3B;6;6;3B
MEMANTINE	2	4A
MENOTROPINS	2	6
MENTHOL	2	4A
MEQUINOL;MONOBENZONE	2	4B;4B
METHAMPHETAMINE	2	4A
METHYSERGIDE	2	4B
METRELEPTIN	2	6
MIGLITOL	2	4A

MILATUZUMAB	2	2B
MINAPRINE	2	6
MINOXIDIL	2	4B
MITIGLINIDE;TOLBUTAMIDE SODIUM, STERILE	2	4B;6
MITOTANE	2	1A
MK-0752	2	2A
MLNM-2201;ERLIZUMAB;AME-133V	2	6;6;2B
MM-121;AV-203;AMG-888	2	2A;2A;2A
MODIMELANOTIDE	2	6
NABILONE;NONABINE;DRONABINOL;SAD448	2	4B;6;3B;6
NALTREXONE	2	3B
NAMINIDIL;SARAKALIM	2	6;6
NEFAZODONE	2	4A
NELARABINE	2	1A
NERAMEXANE	2	6
NI-0801;ELDELUMAB	2	6;6
NIACIN	2	3B
NICARDIPINE	2	2B
OBETICHOLIC ACID;PX-102;CHENODIOL;URSODIOL;TUROFEXORATE ISOPROPYL	2	4A;6;3B;3B;4B
OLOCROLIMUS;TACROLIMUS;GPI-1485	2	6;2A;6
ONCOLYSIN CD6	2	6
ONCOLYSIN S;LORVOTUZUMAB MERTANSINE	2	6;2B
ORLISTAT	2	3B
OXIGLUTATIONE	2	4A
OXYTOCIN	2	4A
PAC-14028;SB-705498;MR-1817;NGD-8243;RESINIFERATOXIN;MAVATREP;ZUCAPSAICIN;JTS-653;AMG-517;GRC-6211;DWP05195;ABT-102;CAPSAICIN;SYL-1001;PHE377;XEN-D0501;AZD1386;SAF312	2	6;6;6;6;3B;6;6;6;6;6;6;3B;6;6;6;6
PARATHYROID HORMONE	2	6

PARICALCITOL;CALCIPOTRIOL;CHOLECALCIFEROL;SEOCALCITOL;ERGOCALCIFEROL;CALCITRIOL;DIHYDROTACHYSTEROL;DOXERCALCIFEROL;DP001;CALCIPOTRIENE;BECOCALCIDIOL;PEFCALCITOL;ILX23-7553;ELOCALCITOL;CALCIFEDIOL;ELDECALCITOL;INECALCITOL;1.ALPHA.-HYDROXYVITAMIN D5	2	2A;5;3B;2B;4A;2A;6;2A;6;2B;6;6;3B;4A;3B;4A;2B;4B
PASCOLIZUMAB	2	6
PATECLIZUMAB	2	6
PEGNIVACOGIN	2	6
PENTOLINIUM TARTRATE;MECAMYLAMINE HYDROCHLORIDE;LARAZOTIDE ACETATE;TRIMETHAPHAN CAMSYLATE	2	4B;6;6;6
PERHEXILINE	2	4B
PF-03758309	2	2A
PHENFORMIN	2	4A
PHENSERINE, (+)-	2	6
PINACIDIL	2	4A
PLX-7486	2	2B
POZANICLINE	2	4B
PRADIGASTAT	2	6
PRALATREXATE	2	1A
PRATOSARTAN	2	6
PRITUMUMAB	2	6
PROBENECID	2	2A
PROGABIDE	2	6
PROPAFENONE	2	3B
PSEUDOEPHEDRINE	2	6
QUETIAPINE	2	5
R428	2	4B
RADIPRODIL;RISLENEMDAZ;EVT-101;TRAXOPRODIL	2	6;6;6;4B
RANIRESTAT	2	4B
REGRAMOSTIM	2	6

RESCINNAMINE;DESERPIDINE	2	4B;6
RG-7376	2	6
RIBAVIRIN	2	2A
RICOLINOSTAT	2	2B
RILAPLADIB;DARAPLADIB	2	6;4A
RITOBEGRON;MIRABEGRON;SOLABEGRON;VIBEGRON;AMIBEGRON;AMIBEGRON HYDROCHLORIDE;RAFABEGRON	2	6;6;6;6;4B;6;6
RIVASTIGMINE;ISOFLUROPHATE;RIVASTIGMINE TARTRATE;TACRINE HYDROCHLORIDE	2	4A;4A;6;6
RUXOLITINIB PHOSPHATE;BARICITINIB;MOMELOTINIB;AZD-1480	2	1B;4A;2B;2A
SAPROPTERIN DIHYDROCHLORIDE	2	6
SELONSERTIB	2	6
SERIBANTUMAB	2	2B
SILMITASERTIB	2	2B
SOBETIROME;RESMETIROM;EPROTIROME;MB07811	2	6;6;6;6
SOMATROPIN	2	6
SOTALOL	2	4B
SPARSENTAN	2	6
STX-100	2	6
SULFINPYRAZONE	2	3B
SULODEXIDE	2	6
TALNIFLUMATE	2	6

TASOSARTAN;AZILSARTAN KAMEDOXOMIL;VALSARTAN;OLMESARTAN;SAPRISARTAN;AZILSARTAN;EPROSARTAN;LOSARTAN;TELMISARTAN;CANDESARTAN;IRBESARTAN;FORASARTAN;EPROSARTAN MESYLATE;CANDESARTAN CILEXETIL;SARALASIN ACETATE;OLMESARTAN MEDOXOMIL;LOSARTAN POTASSIUM	2	4B;6;4A;4A;6;4B;4B;3B;4A;4A;4A;6;6;4A;6;4A;6
TAZAROTENE	2	2B
TC-6499	2	6
TECALCET HYDROCHLORIDE;CINACALCET HYDROCHLORIDE;CINACALCET;ASP7991;RONACALERET;VELCALCETIDE;CHEMBL180672	2	4B;1A;4B;6;6;6;6
TENAPANOR HYDROCHLORIDE	2	6
TERIPARATIDE;TERIPARATIDE ACETATE	2	6;6

TESTOSTERONE UNDECANOATE;NILUTAMIDE;TESTOSTERONE;GSK2881078;GSK2849466;ENOBOSARM;METHYLTESTOSTERONE;DAROLUTAMIDE;VK5211;TESTOSTERONE ENANTHATE;OXANDROLONE;CLASCOTERONE;OXYMETHOLONE;APC-100;FLUOXYMESTERONE;DROMOSTANOLONE;NANDROLONE DECANOATE;APALUTAMIDE;NANDROLONE;DROMOSTANOLONE PROPIONATE;TESTOSTERONE PROPIONATE;CYPROTERONE;LY2452473;AZD3514;FLUTAMIDE;ETHYLESTRENOL;MK-0773;NANDROLONE PHENPROPIONATE;HE3235;LGD-2941;STANOZOLOL;PRASTERONE;TESTOSTERONE CYPIONATE;GLPG0492	2	4A;1A;3B;6;6;6;3B;1A;6;2B;4B;6;3B;2B;1B;6;3B;1A;3B;6;3B;3B;6;2B;1A;4B;6;3B;2B;6;3B;3B;6;6
THIETHYLPERAZINE	2	4B
THYMALFASIN	2	2B
TILDRAKIZUMAB;GUSELKUMAB	2	2A;6
TILUDRONIC ACID;TRODUSQUEMINE	2	6;3B
TINZAPARIN	2	5
TRAFERMIN;AZD-4547;DERAZANTINIB;LY-2874455;ROGARATINIB	2	6;2A;2A;6;2A
TRAZODONE	2	3B
TREBANANIB	2	2A
TRIFLUOPERAZINE	2	4A
TT-232	2	6
VANTICTUMAB	2	2A
VATIQUINONE	2	6
VEROSUDIL;SAR-407899;AMA0076;DE-104;Y-39983	2	6;6;6;6;6

VIGABATRIN	2	3B
VORAPAXAR;RUSALATIDE	2	4B;6
VOSAROXIN;TENIPOSIDE;BERUBICIN HYDROCHLORIDE;AMONAFIDE	2	2A;1A;2A;6
ZILEUTON;ATRELEUTON	2	2B;6
ZIPRASIDONE HYDROCHLORIDE;CARPHENAZINE MALEATE;RISPERIDONE;ZIPRASIDONE MESYLATE;QUETIAPINE FUMARATE	2	6;6;3B;6;3B
ZOLPIDEM TARTRATE	2	6
ACEBUTOLOL;BEVANTOLOL;ESMOLOL;NEBIVOLOL;BETAXOLOL HYDROCHLORIDE;ATENOLOL;BETAXOLOL;ISOETHARINE;DOBUTAMINE;ARBUTAMINE;PRACTOLOL;ACEBUTOLOL HYDROCHLORIDE;LEVOBETAXOLOL HYDROCHLORIDE;DOPAMINE HYDROCHLORIDE;METOPROLOL;BISOPROLOL;BISOPROLOL FUMARATE;METOPROLOL SUCCINATE	1	3B;4B;4A;4A;6;3B;4B;4B;4A;4B;4B;6;6;6;3B;4A;6;6

ACOTIAMIDE;PHYSOSTIGMINE;GALANTAMINE;GALANTAMINE HYDROBROMIDE;DONEPEZIL;HEXAFLUORENIUM BROMIDE;PRALIDOXIME CHLORIDE;PYRIDOSTIGMINE;PYRIDOSTIGMINE BROMIDE;TACRINE;ECHOTHIOPHATE IODIDE;AMBENONIUM CHLORIDE;PHENSERINE;EDROPHONIUM;NEOSTIGMINE;EDROPHONIUM CHLORIDE;AMBENONIUM;ECHOTHIOPHATE;NEOSTIGMINE METHYLSULFATE;DEMECARIUM BROMIDE;DONEPEZIL HYDROCHLORIDE	1	6;3B;4B;6;2A;6;4B;6;3B;4A;4B;3B;6;3B;2B;6;6;6;6;6;6
ADL-5747	1	6
AFACIFENACIN	1	6
AFEGOSTAT	1	6
AFEGOSTAT TARTRATE	1	6
AGATOLIMOD;AGATOLIMOD SODIUM	1	2A;6
ALEFACEPT;SIPLIZUMAB	1	2B;2B
ALEPLASININ	1	6
ALMITRINE;OUABAIN	1	6;3B

ALPRENOLOL;LEVOBUNOLOL;CARTEOLOL;TIMOLOL;OXPRENOLOL;PENBUTOLOL SULFATE;NEBIVOLOL HYDROCHLORIDE;PROPRANOLOL;METIPRANOLOL;ISOPROTERENOL;OXPRENOLOL HYDROCHLORIDE;DOBUTAMINE HYDROCHLORIDE;PENBUTOLOL;BUCINDOLOL;PINDOLOL;PROPRANOLOL HYDROCHLORIDE;NADOLOL;LEVOBUNOLOL HYDROCHLORIDE;TIMOLOL MALEATE;METIPRANOLOL HYDROCHLORIDE;ESMOLOL HYDROCHLORIDE	1	4B;6;4B;4B;4B;6;6;3B;4B;3B;6;6;4A;4B;4B;6;4B;6;6;6;6
ALSEROXYLON	1	6
AMBRISENTAN;IRL-1620;BOSENTAN;MACITENTAN;ENRASENTAN;DARUSENTAN;TEZOSENTAN	1	4A;6;3B;4B;4B;4B;4B
AMIFOSTINE	1	2A
AMILORIDE	1	3B
AMISULPRIDE;BROMOCRIPTINE;CARIPRAZINE HYDROCHLORIDE;CARIPRAZINE	1	4B;3B;6;6
AMITRIPTYLINE	1	4A
AMPHETAMINE	1	3B
ANAGRELIDE HYDROCHLORIDE	1	1B
ANAKINRA	1	2A
ANDOLAST;BMS-223131;CROMOLYN	1	6;6;2B
ANISINDIONE	1	6
AR-12	1	2B
ARBACLOFEN;LESOGABERAN;ARBACLOFEN PLACARBIL;BACLOFEN;SODIUM OXYBATE	1	6;6;6;3B;4B
ARCITUMOMAB	1	6

ARIPIPRAZOLE;ARIPIPRAZOLE LAUROXIL	1	4A;6
ASELIZUMAB	1	6
ASENAPINE MALEATE;SB-773812;MESORIDAZINE;PALIPERIDONE;LURASIDONE HYDROCHLORIDE;THIOTHIXENE HYDROCHLORIDE;RISPERDAL;MESORIDAZINE BESYLATE;THIOTHIXENE;MOLINDONE HYDROCHLORIDE	1	6;6;4B;6;4B;6;6;6;6;6
ASK-8007	1	6
AT-406	1	2A
AVANAFIL;VARDENAFIL HYDROCHLORIDE;VARDENAFIL;SILDENAFIL;UDENAFIL;SILDENAFIL CITRATE;CP-461	1	6;6;5;2A;4A;2A;2B
AZATHIOPRINE SODIUM	1	6
AZD-8186	1	2B
AZD2423;CCX140;PLOZALIZUMAB	1	6;6;2B
AZD4818;CCX354	1	6;6
AZD7009	1	6
BATEFENTEROL	1	6
BAY-1217389;BAY-1161909	1	2A;2A
BETAINE HYDROCHLORIDE	1	6
BETHANIDINE	1	6
BIIB059	1	6
BIMAGRUMAB;RAMATERCEPT	1	6;6
BIMATOPROST	1	6
BIMOCLOMOL	1	4B
BMS-779788	1	6

BUDESONIDE;BETAMETHASONE BENZOATE;METHYLPREDNISOLONE ACETATE;FOSDAGROCORAT;BECLOMETHASONE DIPROPIONATE;HE3286;HYDROCORTISONE CYPIONATE;HYDROCORTISONE BUTYRATE;TRIAMCINOLONE HEXACETONIDE;HYDROCORTISONE SODIUM PHOSPHATE;PREDNISOLONE SODIUM PHOSPHATE;TRIAMCINOLONE DIACETATE;HYDROCORTISONE SODIUM SUCCINATE;MEPREDNISONE;PARAMETHASONE ACETATE;ALCLOMETASONE DIPROPIONATE;ORG-34517;CLOBETASOL PROPIONATE;FLUTICASONE PROPIONATE;PREDNISOLONE TEBUTATE;HYDROCORTISONE VALERATE;BETAMETHASONE SODIUM PHOSPHATE;HYDROCORTISONE ACETATE;DIFLUPREDNATE;BETAMETHASONE VALERATE;FLUOROMETHOLONE ACETATE;DEXAMETHASONE ACETATE;DEXAMETHASONE SODIUM PHOSPHATE;BETAMETHASONE DIPROPIONATE;MOMETASONE FUROATE;TRIAMCINOLONE ACETONIDE;FLUTICASONE;HYDROCORTAMATE HYDROCHLORIDE;FLUTICASONE FUROATE;PREDNISOLONE ACETATE;HYDROCORTISONE PROBUTATE;CORT 125134;FLUOCINOLONE ACETONIDE;HALOBETASOL PROPIONATE;HYDROCORTISONE;CORT 108297;CORTISONE ACETATE;DAGROCORAT;BETAMETHASONE ACETATE;CLOCORTOLONE PIVALATE;DIFLORASONE DIACETATE;METHYLPREDNISOLONE SODIUM SUCCINATE;LOTEPREDNOL ETABONATE;DEFLAZACORT;FLUPREDNISOLONE ;HALCINONIDE;FLUMETHASONE PIVALATE;MAPRACORAT	1	2B;6;3B;6;6;6;6;6;6;6;6;6;6;1B;4B;6;2B;6;6;6;6;6;6;4B;6;4A;6;4B;4A;4B;4A;1B;4A;6;6;4B;6;2A;4B;6;2A;6;1B;6;6;6;6;1B;4B;4B;6;6;6;6
BUMETANIDE	1	2B
BUSPIRONE	1	4A
BVT.115959;ISTRADEFYLLINE;TOZADENANT;REG ADENOSON;BINODENOSON;PRELADENANT	1	6;4B;6;6;6;2A
CAFFEINE	1	3A
CAFFEINE, CITRATED;ADENOSINE	1	6;4A

CANTUZUMAB RAVTANSINE;CANTUZUMAB MERTANSINE;AR-20.5;SONTUZUMAB;PEMTUMOMAB	1	2B;6;6;6;6
CARGLUMIC ACID	1	4B
CARPHENAZINE;ZUCLOPENTHIXOL	1	6;4B
CDP-860	1	6
CER-002;MBX-8025;KD3010	1	6;6;6
CG250 177LU;CG250 111IN;IODINE I 124 GIRENTUXIMAB	1	6;6;6
CHLOROQUINE;PEGARGIMINASE	1	2A;2B
CHORIOGONADOTROPIN ALFA;AVICINE;GONADOTROPIN, CHORIONIC;LH (MENOTROPINS);GOSERELIN;LUTROPIN ALFA	1	6;6;6;6;1A;6
CINPANEMAB	1	6
CINTREDEKIN BESUDOTOX	1	2B
COBICISTAT	1	2B

CODEINE;HYDROMORPHONE;NALOXONE HYDROCHLORIDE;AXOMADOL;ODELEPRAN;CODEINE POLISTIREX;BUTORPHANOL;BENZHYDROCODONE HYDROCHLORIDE;HYDROCODONE;NALMEFENE;NALOXONE;MORPHINE;OXYMORPHONE HYDROCHLORIDE;CODEINE PHOSPHATE;NALBUPHINE;BUPRENORPHINE;NALMEFENE HYDROCHLORIDE;OXYCODONE;PROPOXYPHENE;SUFENTANIL;HYDROCODONE POLISTIREX;CODEINE SULFATE;HYDROCODONE BITARTRATE;NALTREXONE HYDROCHLORIDE;NALBUPHINE HYDROCHLORIDE	1	3B;3B;6;6;6;6;4B;6;6;4B;4A;3B;6;6;4B;4B;6;4B;4B;4B;6;6;6;6;6
CONIVAPTAN HYDROCHLORIDE;CONIVAPTAN	1	6;6
CYCLOBENZAPRINE HYDROCHLORIDE	1	6
CYCLOTHIAZIDE	1	4B
D-3263	1	2A
DACARBAZINE	1	1A
DARATUMUMAB	1	1A
DEMCIZUMAB	1	2A
DESIPRAMINE	1	2A
DESOXYCORTICOSTERONE PIVALATE;FLUDROCORTISONE;EPLERENONE;FLUDROCORTISONE ACETATE;LY2623091;SPIRONOLACTONE;FINERENONE;MT-3995;XL550;DESOXYCORTICOSTERONE ACETATE;PF-03882845	1	6;4A;2B;4B;6;3B;6;6;6;4A;6

DEXLANSOPRAZOLE;ESOMEPRAZOLE MAGNESIUM;LANSOPRAZOLE;PANTOPRAZOLE SODIUM;RABEPRAZOLE;OMEPRAZOLE MAGNESIUM;OMEPRAZOLE;ESOMEPRAZOLE STRONTIUM;ESOMEPRAZOLE SODIUM	1	2B;1B;2A;1B;1B;6;2A;1B;6
DEXTROMETHORPHAN	1	4A
DIBUCAINE HYDROCHLORIDE;DIBUCAINE;HEXYLCAINE	1	6;4B;6
DINOPROSTONE	1	3B
DINUTUXIMAB	1	1A
DISULFIRAM	1	2A
DOXEPIN	1	2B
DROTRECOGIN ALFA (ACTIVATED)	1	6
DUVOGLUSTAT HYDROCHLORIDE;DUVOGLUSTAT	1	6;4B
E-6201	1	2A
E-FLUPENTIXOL;ACETOPHENAZINE;PERPHENAZINE;PERGOLIDE;FLUPHENAZINE	1	4A;4B;4B;3B;2B
ECABET	1	6
ECULIZUMAB;PEXELIZUMAB;AVACINCAPTAD PEGOL SODIUM	1	2B;6;6
EFLORNITHINE HYDROCHLORIDE;EFLORNITHINE	1	2B;2A
EGAPTIVON PEGOL;CAPLACIZUMAB	1	6;6
ELENBECESTAT	1	6

ELINOGREL;CLOPIDOGREL BISULFATE;TICLOPIDINE HYDROCHLORIDE;CANGRELOR TETRASODIUM;PRASUGREL;CANGRELOR;PRASUGREL HYDROCHLORIDE;TICLOPIDINE;TICAGRELOR;CLOPIDOGREL;REGRELOR DISODIUM	1	6;2B;6;6;6;4B;3B;4B;4B;2B;6
ELOBIXIBAT	1	6
EMIXUSTAT;EMIXUSTAT HYDROCHLORIDE	1	6;6
EMPAGLIFLOZIN;IPRAGLIFLOZIN;DAPAGLIFLOZIN;TOFOGLIFLOZIN;CANAGLIFLOZIN;REMOGLIFLOZIN ETABONATE;BEXAGLIFLOZIN;DAPAGLIFLOZIN PROPYLENE GLYCOL;ERTUGLIFLOZIN	1	4B;6;2B;6;3B;6;6;6;6
ENPROFYLLINE;OGLEMILAST	1	6;6
ENSITUXIMAB	1	2B
ENTOLIMOD	1	2A
EPTINEZUMAB	1	6
ESMIRTAZAPINE MALEATE	1	6
ETROLIZUMAB	1	6
EXENATIDE;LIRAGLUTIDE;DULAGLUTIDE;SEMAGLUTIDE;LIXISENATIDE;TASPOGLUTIDE;ALBIGLUTIDE	1	3A;3A;4A;6;3A;6;6
EZETIMIBE	1	2B
F16SIP 131I;81C6 131I;TENATUMOMAB;F16IL2	1	6;6;6;6
FARLETUZUMAB	1	2A
FASINUMAB;FULRANUMAB;TANEZUMAB	1	6;6;6
FASUDIL	1	4A
FE 202767;CARBETOCIN;BARUSIBAN;RETOSIBAN;EPELSIBAN	1	6;6;6;6;6

FEZAKINUMAB	1	6
FIBOFLAPON;FIBOFLAPON SODIUM	1	6;6
FLUOXETINE	1	3B
FOMEPIZOLE	1	4A
FOSPHENYTOIN;FLECAINIDE;DISOPYRAMIDE;MORICIZINE;BENZONATATE;ENCAINIDE;QUINIDINE;INDECAINIDE;TOCAINIDE;MEXILETINE;PROCAINAMIDE;FLECAINIDE ACETATE;LICARBAZEPINE	1	6;4B;4B;4B;6;6;4A;6;6;4A;3B;6;6
GALIXIMAB	1	2B
GALUNISERTIB	1	2A
GDC-0917	1	2A
GIRENTUXIMAB	1	2B
GLEMBATUMUMAB VEDOTIN	1	2B
GLPG-0259	1	6
GRC-15300	1	6
GRC-17536;CB-189625	1	6;6
GS 6201;LAS101057	1	6;6
GSK-1004723;BETAHISTINE	1	6;4B
GSK-249320	1	6
GSK189254;PITOLISANT;CIPRALISANT MALEATE;ABT-288;MK-0249;PF-03654764;JNJ-17216498;IRDABISANT;BAVISANT;GSK239512;PF-03654746	1	6;4B;6;6;6;6;6;6;6;6;6
HALOPERIDOL DECANOATE;CHLORPROMAZINE;HALOPERIDOL;HALOPERIDOL LACTATE;CHLORPROMAZINE HYDROCHLORIDE;PROMAZINE HYDROCHLORIDE;DROPERIDOL	1	4B;2B;4A;6;6;6;4B

HYDROCHLOROTHIAZIDE;QUINETHAZONE;BENZTHIAZIDE;BENDROFLUMETHIAZIDE;POLYTHIAZIDE;CHLOROTHIAZIDE SODIUM	1	4A;6;6;4B;6;6
HYDROFLUMETHIAZIDE;METOLAZONE;METHYCLOTHIAZIDE;CHLORTHALIDONE	1	6;4B;6;4B
HYDROXYCHLOROQUINE SULFATE;HYDROXYCHLOROQUINE	1	6;2A
IBANDRONIC ACID;RISEDRONIC ACID;ALENDRONATE SODIUM;PAMIDRONATE DISODIUM;RISEDRONATE SODIUM;IBANDRONATE SODIUM;PAMIDRONIC ACID	1	2B;2B;2B;1B;2B;2B;2B
ICM3	1	6
IDRAPARINUX SODIUM;APIXABAN;OTAMIXABAN;EDOXABAN;IDRAPARINUX;RIVAROXABAN;EDOXABAN TOSYLATE;LETAXABAN;BETRIXABAN;CHEMBL1271162	1	6;2B;6;6;6;2A;6;6;6;6
IMC-3C5	1	2A
IMETELSTAT;IMETELSTAT SODIUM;TERTOMOTIDE	1	2B;2A;2B
IMIQUIMOD;AZD-8848	1	1B;4A
INDACATEROL	1	6

ISOETHARINE HYDROCHLORIDE;ARFORMOTEROL TARTRATE;SALMETEROL;TERBUTALINE;ALBUTEROL;OLODATEROL HYDROCHLORIDE;PROTOKYLOL HYDROCHLORIDE;FORMOTEROL;ABEDITEROL;MILVETEROL;OLODATEROL;TERBUTALINE SULFATE;METAPROTERENOL;RITODRINE HYDROCHLORIDE;PF-00610355;LEVALBUTEROL HYDROCHLORIDE;AZD3199;FENOTEROL;CLENBUTEROL;BITOLTEROL MESYLATE;PROCATEROL;ALBUTEROL SULFATE;VILANTEROL TRIFENATATE;BAMOSIRAN;ARFORMOTEROL;SALMETEROL XINAFOATE;PIRBUTEROL ACETATE;METAPROTERENOL SULFATE;BEDORADRINE SULFATE;VILANTEROL;RITODRINE;GSK159802;INDACATEROL MALEATE;CARMOTEROL;ISOETHARINE MESYLATE	1	6;6;6;4A;4A;6;6;6;6;6;4B;6;4A;6;6;6;6;4A;4A;6;4A;6;6;6;4A;3B;6;6;6;6;4B;6;6;6;6
ISOSORBIDE MONONITRATE	1	4B
IVABRADINE HYDROCHLORIDE	1	6
KRN-330	1	2B
LANREOTIDE ACETATE	1	6
LAS190792	1	6
LCL-161;BIRINAPANT	1	2A;2A
LENZILUMAB;OTILIMAB;NAMILUMAB	1	2B;6;6
LERDELIMUMAB;TRABEDERSEN;BELAGENPUMATUCEL-L	1	6;2B;2A
LESINURAD	1	6
LEVOSIMENDAN	1	6
LIAROZOLE	1	3B
LISDEXAMFETAMINE	1	6

LISURIDE	1	4B
LOMETREXOL	1	2A
LOMITAPIDE;IMPLITAPIDE;SLX-4090	1	6;6;6
LONAFARNIB	1	2B
LUCATUMUMAB;TENELIXIMAB;DACETUZUMAB;PG-102	1	2B;6;2B;6
LUMILIXIMAB	1	2B
LY2590443	1	6
LY2881835;FASIGLIFAM	1	6;4A
LY2940094	1	6
LY545694;DIPRAGLURANT;RASEGLURANT;AZD2516;STX107;AZD2066;RG7342;MAVOGLURANT;BASIMGLURANT	1	6;6;6;6;6;6;6;6;6
M195;ONCOLYSIN M;LINTUZUMAB;GEMTUZUMAB OZOGAMICIN	1	6;6;2B;1A
MACIMORELIN	1	6
MARAVIROC;LERONLIMAB;INCB-9471;APLAVIROC HYDROCHLORIDE;VICRIVIROC MALEATE;AZD5672;VICRIVIROC	1	2B;2A;6;4B;6;6;2B
MB-07803	1	6
MBX-2982	1	6
MDX-070;CAPROMAB	1	6;6
MEBENDAZOLE	1	2B
MEPENZOLATE	1	6
MEPOLIZUMAB;RESLIZUMAB	1	6;6
METHAZOLAMIDE	1	2B
METHOCARBAMOL	1	6
METHOTREXATE SODIUM;TRIMETREXATE;TALOTREXIN	1	1B;2B;2A
METOCLOPRAMIDE	1	3B
METOCLOPRAMIDE HYDROCHLORIDE	1	6
METYRAPONE	1	4A
METYROSINE	1	6
MIANSERIN	1	4A

MIBEFRADIL DIHYDROCHLORIDE;ETHOSUXIMIDE;TRIMETHADIONE;PARAMETHADIONE;PHENSUXIMIDE;METHSUXIMIDE	1	6;4A;3B;6;6;6
MIFAMURTIDE	1	2B
MIGALASTAT	1	4B
MIGLUSTAT;ELIGLUSTAT	1	2B;6
MILRINONE	1	4A
MIROGABALIN BESYLATE	1	6
MIRTAZAPINE	1	4B
MISOPROSTOL	1	4A
MOCLOBEMIDE	1	4B
MOTOLIMOD	1	2B
MUROMONAB-CD3	1	2A
NEPICASTAT	1	4B
NESVACUMAB	1	2A
NICOTINE	1	3B
NITISINONE	1	4B
NITRIC OXIDE;RIOCIGUAT;SODIUM NITROPRUSSIDE;ISOSORBIDE DINITRATE	1	3B;4B;6;4B
NRX195183	1	2B
NS-018;GANDOTINIB;AC-430;BMS-911543;XL-019;LS-104	1	6;2A;6;2A;2B;6
OBATOCLAX;OBLIMERSEN	1	2B;2A
ONCOLYSIN B;TAPLITUMOMAB PAPTOX;COLTUXIMAB RAVTANSINE;MDX-1342;INEBILIZUMAB	1	6;6;2B;2B;2A
ONTUXIZUMAB	1	2A
ORE-1001	1	4A
ORTERONEL;ABIRATERONE ACETATE;ZYTIGA	1	2A;1A;6
OTELIXIZUMAB;FORALUMAB;VISILIZUMAB	1	6;6;2B
OTLERTUZUMAB	1	2B
OXYMORPHONE;FENTANYL	1	4B;4B
OZANEZUMAB	1	6

PALOSURAN	1	6
PAMREVLUMAB	1	2B
PANTOPRAZOLE;ESOMEPRAZOLE	1	1B;1B
PEGFILGRASTIM;FILGRASTIM;BALUGRASTIM;LI PEGFILGRASTIM	1	2A;2A;6;2B
PEGINTERFERON ALFA-2A;INTERFERON ALFA-N3;INTERFERON BETA-1B;INTERFERON ALFA-2B;INTERFERON ALFACON-1;PEGINTERFERON BETA-1A;ALBINTERFERON ALFA-2B;INTERFERON BETA-1A;PEGINTERFERON ALFA-2B;INTERFERON ALFA-2A	1	2A;6;4B;1B;3B;6;6;6;1B;2A
PEGINTERFERON LAMBDA-1A	1	6
PENTOSTATIN	1	1A
PEPTIDE B27PD	1	6
PERGOLIDE MESYLATE;LEVODOPA;ETILEVODOPA;MELEVODOPA	1	6;3B;6;6
PF-04634817;CENICRIVIROC	1	6;6
PHA-543613;GTS-21;ENCENICLINE;TC-6987;AZD0328;APN1125;BRADANICLINE;AVL-3288;JNJ-39393406;AQW051;ND0801;SSR180711;ABT-126;ABT-107	1	6;4B;6;6;6;6;6;6;6;6;6;4B;6;6
PHENELZINE SULFATE;PHENELZINE;ISOCARBOXAZID;PARGYLINE HYDROCHLORIDE;TRANYLCYPROMINE	1	2B;2B;4B;6;2B

PHENTOLAMINE MESYLATE;TETRAHYDROZOLINE HYDROCHLORIDE;NAPHAZOLINE HYDROCHLORIDE;OXYMETAZOLINE HYDROCHLORIDE;RAUWOLFIA SERPENTINA;PHENOXYBENZAMINE HYDROCHLORIDE	1	2B;6;4B;6;6;6
PHENYLEPHRINE BITARTRATE;TERAZOSIN HYDROCHLORIDE HYDRATE;PRAZOSIN HYDROCHLORIDE;MIDODRINE HYDROCHLORIDE;TAMSULOSIN HYDROCHLORIDE;METARAMINOL BITARTRATE;MIDODRINE;METHOXAMINE HYDROCHLORIDE;ALFUZOSIN HYDROCHLORIDE;DOXAZOSIN MESYLATE;PHENYLEPHRINE HYDROCHLORIDE;PHENDIMETRAZINE;METHOXAMINE	1	6;6;6;6;6;6;6;6;6;6;6;6;4A
PHENYLPROPANOLAMINE HYDROCHLORIDE;PHENYLPROPANOLAMINE POLISTIREX	1	6;6
PHENYLPROPANOLAMINE;OXYMETAZOLINE;BENZPHETAMINE;PHENTOLAMINE	1	4A;4B;4B;2B
PIMOZIDE	1	3B
PPI-2458;BELORANIB	1	2B;6
PROCHLORPERAZINE	1	4B
PRX-321;DUPILUMAB	1	2B;6
QAF805	1	6
RANOLAZINE	1	4B

RASAGILINE MESYLATE;RASAGILINE;SAFINAMIDE MESYLATE;SELEGILINE;SELEGILINE HYDROCHLORIDE	1	6;4B;6;2B;6
RELCOVAPTAN	1	4B
REMINERTANT	1	4B
REMOXIPRIDE;SULPIRIDE;FLUPHENAZINE DECANOATE;MOLINDONE;FLUPHENAZINE ENANTHATE;ORDOPIDINE;DOMPERIDONE;FLUSPIRILENE;THIETHYLPERAZINE MALATE;CABERGOLINE;PROCHLORPERAZINE EDISYLATE;APLINDORE FUMARATE;PROCHLORPERAZINE MALEATE;PRIDOPIDINE;ACETOPHENAZINE MALEATE;BENZQUINAMIDE HYDROCHLORIDE;APLINDORE;THIETHYLPERAZINE MALEATE;FLUPHENAZINE HYDROCHLORIDE;SERIDOPIDINE;TRIFLUPROMAZINE HYDROCHLORIDE;JNJ-37822681	1	4B;4B;6;4B;6;6;3B;5;6;3B;6;6;6;6;6;6;6;6;2B;6;6;6
RENZAPRIDE	1	6
REPOTRECTINIB	1	2A
RG3487	1	3B
RIGOSERTIB SODIUM;ATOPAXAR;VORAPAXAR SULFATE	1	2A;4B;6
RIMCAZOLE;PENTAZOCINE LACTATE	1	4B;6
RINTATOLIMOD	1	2A

ROPINIROLE;PRAMIPEXOLE;ROPINIROLE HYDROCHLORIDE;BROMOCRIPTINE MESYLATE;APOMORPHINE HYDROCHLORIDE;PRAMIPEXOLE DIHYDROCHLORIDE;DEXPRAMIPEXOLE	1	4A;4B;6;6;6;6;6
ROTIGOTINE	1	4B
RUPLIZUMAB;TORALIZUMAB;DAPIROLIZUMAB PEGOL	1	6;6;6
SACUBITRIL	1	6
SAMALIZUMAB	1	2A
SARIZOTAN HYDROCHLORIDE	1	6
SATRALIZUMAB	1	6
SENICAPOC	1	6
SENREBOTASE;DEBRISOQUIN;NOREPINEPHRINE	1	6;4B;3B
SETMELANOTIDE;PF-00446687	1	6;6
SIFALIMUMAB	1	6
SILODOSIN;MEPHENTERMINE;NICERGOLINE;METARAMINOL;EPHEDRINE;PHENOXYBENZAMINE;DAPIPRAZOLE HYDROCHLORIDE;DABUZALGRON;TOLAZOLINE	1	4B;6;6;6;4B;3B;6;6;4B
SIVELESTAT;ERDOSTEINE	1	6;4A
SONIDEGIB PHOSPHATE;PATIDEGIB;GLASDEGIB;TAK-441;LEQ506;TALADEGIB;BMS-833923	1	1B;2A;1A;2B;2B;2A;2A
SOTALOL HYDROCHLORIDE	1	6
SSR125543;ONO-2333MS;VERUCERFONT;EMICERFONT;PEXACERFONT	1	6;6;4B;6;4B
TADALAFIL	1	2A
TALABOSTAT	1	2A
TALAROZOLE	1	4B

TAMSULOSIN;DAPIPRAZOLE;PERICIAZINE;TERAZOSIN;DOXAZOSIN;PHENYLEPHRINE;ALFUZOSIN;PRAZOSIN	1	4B;6;4B;3B;3B;4A;4B;3B
TEDALINAB;PRS-211375;GW842166X;LY2828360	1	6;6;6;6
TEPLIZUMAB	1	6
THROMBIN	1	4B
TIAGABINE	1	4B
TIMOLUMAB;HYDRALAZINE;VEPALIMOMAB	1	6;4A;6
TIPIRACIL HYDROCHLORIDE	1	1B
TIRASEMTIV	1	6
TOCILIZUMAB;SARILUMAB	1	2A;2B
TOLAZOLINE HYDROCHLORIDE	1	6
TOLCAPONE;ENTACAPONE	1	2B;2B
TOLVAPTAN;LIXIVAPTAN;SATAVAPTAN	1	2A;6;6
TRIFLUPROMAZINE	1	4B
TRILOSTANE	1	2B
TXA127	1	2A
VABICASERIN HYDROCHLORIDE	1	6
VADADUSTAT	1	6
VALPROATE SODIUM;DIVALPROEX SODIUM;VALPROIC ACID	1	2B;2A;2A
VALZIFLOCEPT	1	6
VASOPRESSIN;VASOPRESSIN TANNATE;TERLIPRESSIN;DESMOPRESSIN;LYPRESSIN;DESMOPRESSIN ACETATE	1	4B;6;6;4B;6;6
VATELIZUMAB	1	6
VEDOLIZUMAB	1	2B
VELDOREOTIDE	1	6
VOGLIBOSE	1	4B
VONAPANITASE	1	6
VORTIOXETINE	1	4B

ZAFIRLUKAST;NEDOCROMIL;CR-3465;MONTELUKAST SODIUM;MONTELUKAST;CINALUKAST;PRANLUKAST	1	2B;6;6;6;2A;6;4B
ZOLBETUXIMAB	1	2A
ZOLEDRONIC ACID	1	1B
ZOLIMOMAB ARITOX	1	6

Appendix 2

Frequencies of all drugs observed in the top 5 in each THCA-US sample via the driver propagation method. Drugs that were merged (because they have the same target gene(s)) are listed separated by ";".

Drug	Freq	Tiers
LIFIRAFENIB;ZELBORAF;DABRAFENIB MESYLATE;RG-7256;CHEMBL525191;PLX-8394	281	2A;6;1B;6;6;2A
ARQ-736;LGX-806	274	2A;6
LY-3009120;MLN-2480;XL-281	268	2B;2A;2A
RAF-265	261	2A
CEP-32496	244	2A
PELAREOREP	57	2B
IPILIMUMAB	38	1A
METFORMIN	38	2A
RIBOCICLIB	38	1A
PANITUMUMAB	37	1A
BORTEZOMIB	33	1A
NERISPIRDINE	31	6
TOPIRAMATE	22	4B
OCRIPLASMIN	21	6
CARFILZOMIB	20	1A
BINIMETINIB	16	1A
DALFAMPRIDINE;GUANIDINE HYDROCHLORIDE;TEDISAMIL	16	4B;6;6
SELUMETINIB	16	1A
TIPIFARNIB	16	2A
CISPLATIN	15	1A
EVEROLIMUS	15	1A
COLLAGENASE CLOSTRIDIUM HISTOLYTICUM	13	6
METFORMIN HYDROCHLORIDE	12	2B
PACLITAXEL	12	1A
OPROZOMIB;IXAZOMIB CITRATE	10	2A;1A
DASATINIB	9	1A

IMAGABALIN;ATAGABALIN;GABAPENTIN;PREGABALIN;BEPRIDIL HYDROCHLORIDE;GABAPENTIN ENACARBIL	9	6;6;3B;3B;6;6
VANDETANIB	9	1A
NV-128;ME-344	8	6;2A
PELITINIB;CANERTINIB DIHYDROCHLORIDE;AFATINIB DIMALEATE	8	2A;2B;1B
POZIOTINIB;OSIMERTINIB MESYLATE;DACOMITINIB;AC-480	8	2A;2B;1A;2A
RONICICLIB;PHA-793887;AT-7519;AZD-5438	8	2A;2A;2A;2A
AFATINIB	7	1A
BAY-1125976	6	2A
BMS-690514	6	2A
CHEMBL1086377	6	6
IPATASERTIB	6	2A
OLANZAPINE	6	2A
STANNSOPORFIN	6	4A
2X-121	5	2B
CERALASERTIB	5	2A
ENCORAFENIB	5	1A
QUERCETIN	5	3B
SELURAMPANEL	5	6
SR-13668;NELFINAVIR	5	6;2A
TEZAMPANEL	5	6
VELIPARIB	5	2A
AMLODIPINE	4	2B
CNV2197944;ZICONOTIDE;Z160;ZICONOTIDE ACETATE	4	6;6;6;6
DABRAFENIB	4	1A
DACTOLISIB	4	2A
JNJ-26483327	4	2A
NIRAPARIB	4	1A
PROPYLTHIOURACIL	4	3A
SALIRASIB	4	6
TASELISIB	4	2B
VERAPAMIL	4	2B
ZONISAMIDE	4	3B
ABITUZUMAB;INTETUMUMAB	3	2B;2B

ACAMPROSATE CALCIUM	3	6
ADAVOSERTIB	3	2A
BAY-1082439;SAR-260301	3	2A;2A
BGT-226;VS-5584;SF-1126;SAMOTOLISIB;PWT-33587	3	2A;2B;2A;2A;6
CLEVIDIPINE	3	4B
DANTROLENE	3	3B
DANTROLENE SODIUM	3	6
DIGOXIN;ACETYLDIGITOXIN;DESLANOSIDE;DIGITOXIN	3	2A;6;6;3B
DIPYRIDAMOLE	3	2B
ENZALUTAMIDE	3	1A
ERLOSAMIDE	3	2B
GSK-690693	3	2A
IDELALISIB	3	1A
LEUCOVORIN	3	1B
LEVETIRACETAM	3	2B

NKTR-171;ESLICARBAZEPINE;BENOXINATE HYDROCHLORIDE;ETHOTOIN;PHENYTOIN;ARTICAINE HYDROCHLORIDE;FOSPHENYTOIN SODIUM;RUFINAMIDE;CARBAMAZEPINE;OXCARBAZEPINE;TOCAINIDE HYDROCHLORIDE;PROPOXYCAINE HYDROCHLORIDE;ETIDOCAINE HYDROCHLORIDE;TETRACAINE;PHENACEMIDE;HEXYLCAINE HYDROCHLORIDE;INDECAINIDE HYDROCHLORIDE;TETRACAINE HYDROCHLORIDE;CHLOROPROCAINE HYDROCHLORIDE;LAMOTRIGINE;ROPIVACAINE HYDROCHLORIDE;PRILOCAINE HYDROCHLORIDE;PHENAZOPYRIDINE HYDROCHLORIDE;PROCAINAMIDE HYDROCHLORIDE;MERETHOXYLLINE PROCAINE;MEPHENYTOIN;DYCLONINE HYDROCHLORIDE;PRILOCAINE;PROPARACAINE HYDROCHLORIDE;MEPIVACAINE HYDROCHLORIDE;QUINIDINE POLYGALACTURONATE;QUINIDINE GLUCONATE;MEXILETINE HYDROCHLORIDE;MORICIZINE HYDROCHLORIDE;DISOPYRAMIDE PHOSPHATE;PHENYTOIN SODIUM;EVENAMIDE;LIDOCAINE HYDROCHLORIDE;ESLICARBAZEPINE ACETATE;PROCAINE HYDROCHLORIDE;QUINIDINE SULFATE	3	6;6;6;4B;2B;6;6;4B;3B;2A;6;6;6;4B;6;6;6;6;6;2B;2B;6;6;6;6;4B;6;2B;6;6;6;6;6;6;6;6;6;1B;6;6;6
OLAPARIB	3	1A
PENTOXIFYLLINE	3	2A
PERIFOSINE	3	2A
PRIMIDONE	3	4B
RG-7666;RECILISIB;WX-037;ZSTK-474;AZD-6482;PA-799;GSK-1059615	3	2B;6;2A;2A;6;2A;2A
RO-5045337	3	2A
TEPROTUMUMAB	3	2A
TETRODOTOXIN	3	4A
UPROSERTIB;TRICIRIBINE;AZD-5363;TRICIRIBINE PHOSPHATE;AFURESERTIB;ARCHEXIN;MIRANSERTIB;MK-2201	3	2A;4A;6;2A;2A;6;2B;6
ABAGOVOMAB;OREGOVOMAB	2	2B;2B
ADINAZOLAM;CINOLAZEPAM;CLOTIAZEPAM;FLUDIAZEPAM	2	4B;6;6;6

ALPELISIB	2	1A
ANASTROZOLE	2	1A
APREPITANT;VOFOPITANT;CASOPITANT	2	4B;4B;6
ASENAPINE	2	4B
AZD-8186	2	2B
BELINOSTAT	2	1A
BERZOSERTIB	2	2A
BICALUTAMIDE	2	1A
BRYOSTATIN	2	3B
BUTETHAL;BUTABARBITAL;HEXOBARBITAL;MEPHOBARBITAL;THIOPENTAL;SECOBARBITAL	2	6;6;4B;3B;4A;4A
CAPIVASERTIB	2	2A
CELECOXIB	2	1A
CEP-1347	2	4A
CHLORDIAZEPOXIDE;CLOBAZAM;HALAZEPAM;OXAZEPAM;FLURAZEPAM;BROMAZEPAM;PRAZEPAM;QUAZEPAM;ADIPIPLON;MIDAZOLAM;ESTAZOLAM	2	4B;4B;6;3A;4B;4B;6;6;6;4B;4B
CILENGITIDE	2	2A
COPANLISIB;BUPARLISIB;GSK-2636771	2	1A;2A;2A
DASOLAMPANEL	2	6
DNK333;BURAPITANT;FIGOPITANT;VESTIPITANT MESYLATE;FOSAPREPITANT DIMEGLUMINE;TRADIPITANT;ORVEPITANT;EZLOPITANT;ROLAPITANT HYDROCHLORIDE;SERLOPITANT;NETUPITANT	2	6;6;6;6;6;6;6;6;2B;6;2A
DOXORUBICIN	2	1A
DS-7423	2	2B
ELPETRIGINE	2	6
ENFLURANE	2	4A
ENMD-981693	2	2A
EXEMESTANE	2	1A
FLUOROURACIL	2	1A
GEFITINIB	2	1A
GLPG-0187	2	2A
IMETELSTAT;IMETELSTAT SODIUM;TERTOMOTIDE	2	2B;2A;2B

INK-1117;SERABELISIB	2	6;2A
ISOFLURANE	2	3B
LETROZOLE	2	1A
LIDOCAINE	2	1B
METHIMAZOLE;CARBIMAZOLE	2	2B;4B
MIPOMERSEN	2	6
MK-2206	2	2A
NICOTINE	2	3B
NITRAZEPAM	2	4B
PANOBINOSTAT	2	1A
PIMASERTIB	2	2A
PKI-179;PWT-33579	2	2A;6
PROPIOMAZINE;PROMAZINE	2	6;4B
RAMUCIRUMAB	2	1A
REMIMAZOLAM;OCINAPLON	2	6;6
REMINERTANT	2	4B
RESMINOSTAT;ROMIDEPSIN;PRACINOSTAT;MOCETINOSTAT	2	2B;1A;2A;2A
RG3487	2	3B
RILUZOLE	2	2B
RIVOCERANIB;HENATINIB;OSI-632;PEGDINETANIB;PF-00337210;ALACIZUMAB PEGOL;ZD-4190	2	2A;2A;2B;2A;2A;6;6
SAPANISERTIB	2	2A
SURUFATINIB	2	2A
TALAZOPARIB	2	1A
TELBERMIN;ABICIPAR PEGOL	2	6;6
TG100-115;CAL-263;RP-6530;DUVELISIB	2	4A;6;2B;1A
TRAFERMIN;AZD-4547;DERAZANTINIB;LY-2874455;ROGARATINIB	2	6;2A;2A;6;2A
TRAMETINIB	2	1A
TRIAZOLAM;CLONAZEPAM;CLORAZEPIC ACID;ALPRAZOLAM;DIAZEPAM;TEMAZEPAM;LORAZEPAM	2	4B;3B;6;4B;3B;4B;4B
VARENICLINE	2	3B
VISMODEGIB	2	1A
ABEMACICLIB	1	1A
ABEXINOSTAT;GIVINOSTAT	1	2A;2B
ACARBOSE	1	3B

ACP-001;PEGVISOMANT;SOMATREM;SOMATROGON	1	6;2B;6;6
AFIMOXIFENE	1	2B
ALECTINIB;ALECTINIB HYDROCHLORIDE	1	1A;1B
ALEMTUZUMAB	1	1B
ALMITRINE;OUABAIN	1	6;3B
AMG-900	1	2A
AMINOGLUTETHIMIDE	1	1A
AMP579	1	6
AN-9;DEPAKOTE	1	4B;6
ANDOLAST;BMS-223131;CROMOLYN	1	6;6;2B
APTO-253	1	6
AR-42	1	2B
ARRY-300	1	6
ASM-024	1	6
ASP-3026;LORLATINIB	1	2A;1A
AT-13148	1	2A
ATAMESTANE;TESTOLACTONE;ARIMIDEX	1	2B;4B;6
ATEZOLIZUMAB	1	1A
AVANAFIL;VARDENAFIL HYDROCHLORIDE;VARDENAFIL;SILDENAFIL;UDENAFIL;SILDENAFIL CITRATE;CP-461	1	6;6;5;2A;4A;2A;2B
AZACITIDINE	1	1A
AZD-7451;AZD-6918;LAROTRECTINIB	1	2B;2A;1A
AZD-8055;OSI-027	1	2A;2B
AZD7325;AZD6280	1	6;6
BALSALAZIDE DISODIUM;BALSALAZIDE;OLSALAZINE SODIUM	1	6;4B;6
BARDOXOLONE METHYL	1	2B
BAY-1217389;BAY-1161909	1	2A;2A
BECAPLERMIN	1	6
BETAINE HYDROCHLORIDE	1	6
BIMOSIAMOSE;RIVIPANSEL	1	6;6
BMS-817378;CABOZANTINIB S-MALATE;BMS-794833;GOLVATINIB;TAS-115	1	2A;1A;6;2A;6
BUMETANIDE	1	2B

CAPECITABINE	1	1A
CARBIDOPA	1	4A
CDP-860	1	6
CEP-2563	1	6
CEP-37440	1	2A
CETUXIMAB	1	1A
CHEMBL3397300	1	6
CHLOROTHIAZIDE	1	6
CHLORPHENTERMINE HYDROCHLORIDE;SERTRALINE HYDROCHLORIDE;PAROXETINE HYDROCHLORIDE;ESCITALOPRAM OXALATE;PAROXETINE MESYLATE;CLOMIPRAMINE HYDROCHLORIDE;FLUOXETINE HYDROCHLORIDE;FLUVOXAMINE MALEATE;CITALOPRAM HYDROBROMIDE	1	6;6;6;6;6;6;6;6;6
CHLORZOXAZONE	1	4B
CHORIOGONADOTROPIN ALFA;AVICINE;GONADOTROPIN, CHORIONIC;LH (MENOTROPINS);GOSERELIN;LUTROPIN ALFA	1	6;6;6;6;1A;6
CIXUTUMUMAB	1	2A
COBICISTAT	1	2B
CONTEZOLID ACEFOSAMIL	1	6
CP-459632	1	6
CROFELEMER	1	2B
CT-1578	1	6
CUDC-101	1	2A
CYC-116	1	2A
CYCLOSPORINE	1	2A
CYTARABINE	1	1A
DACLIZUMAB;ALDESLEUKIN;BASILIXIMAB	1	2A;1B;2A
DAPRODUSTAT;ROXADUSTAT	1	6;2B
DARATUMUMAB	1	1A
DARBEPOETIN ALFA;PEGINESATIDE;EPOETIN ALFA;PEGINESATIDE ACETATE;EPOETIN BETA	1	2A;4A;6;6;2A
DAUNORUBICIN	1	1A
DECITABINE	1	1A

DEFACTINIB	1	2A
DESFLURANE;SEVOFLURANE	1	4A;4A
DIAZEPINOMICIN;CHLORMEZANONE	1	2B;6
DINACICLIB	1	2A
DOCETAXEL	1	1A
DRINABANT;SURINABANT;MARINOL;CANNABIDIOL;TARANABANT;RIMONABANT;IBIPINABANT;AZD2207;OTENABANT	1	6;4B;6;3B;4B;3B;4B;6;6
DRONEDARONE HYDROCHLORIDE	1	6
DULANERMIN	1	2B
DULIGOTUZUMAB	1	2A
ECULIZUMAB;PEXELIZUMAB;AVACINCAPTAD PEGOL SODIUM	1	2B;6;6
EMICIZUMAB	1	6
EMRICASAN	1	6
ENFORTUMAB VEDOTIN	1	1A
ENSARTINIB;CHEMBL473773;LUMINESPIB	1	2B;4B;2A
ENSITUXIMAB	1	2B
ENZASTAURIN	1	2A
EPINASTINE	1	4A
ERDAFITINIB	1	1A
ERITORAN TETRASODIUM	1	2B
ETOPOSIDE	1	1A
EZOGABINE	1	4A
FEDRATINIB	1	1A
FIBRINOLYSIN, HUMAN	1	6
FIGITUMUMAB;MECASERMIN;DALOTUZUMAB;PL-225B;MECASERMIN RINFABATE;ROBATUMUMAB;AVE-1642;CHEMBL401930;AXL-1717	1	2A;6;2A;2A;6;2A;2B;4B;2A
FIRATEGRAST;NATALIZUMAB	1	6;2B
FLINDOKALNER	1	6
FLUNARIZINE	1	4B
FOSPROPOFOL	1	6
FRUQUINTINIB;CEP-5214;L-21649;LENVATINIB MESYLATE;CEP-7055;AG-13958;TIVOZANIB;KRN-633	1	2A;6;6;1A;6;6;1A;3B

FUTIBATINIB;PEMIGATINIB	1	2A;1A
GANAXOLONE	1	6
GAVILIMOMAB	1	6
GEMCITABINE	1	1A
GLEPAGLUTIDE;TEDUGLUTIDE;ELSIGLUTIDE	1	6;6;6
GLPG-0259	1	6
GLUCAGON HYDROCHLORIDE RECOMBINANT;GLUCAGON HYDROCHLORIDE;PF-06291874;SAR425899;GLUCAGON;MK-0893	1	6;6;6;6;4A;6
GRANISETRON	1	2A
GRC-15300	1	6
GSK-2269557;INCB-40093;ME-401;UMBRALISIB TOSYLATE;AMG-319;ACALISIB	1	6;2A;6;1B;2B;2B
HALOTHANE	1	4A
IC14	1	6
IDARUBICIN	1	1A
IDRAPARINUX SODIUM;APIXABAN;OTAMIXABAN;EDOXABAN;IDRAPARINUX;RIVAROXABAN;EDOXABAN TOSYLATE;LETAXABAN;BETRIXABAN;CHEMBL1271162	1	6;2B;6;6;6;2A;6;6;6;6
ILOPROST	1	2B
INDAPAMIDE	1	4B

INSULIN ZINC SUSP RECOMBINANT HUMAN;INSULIN SUSP PROTAMINE ZINC PURIFIED PORK;INSULIN ZINC SUSP BEEF;INSULIN ASPART;INSULIN SUSP ISOPHANE RECOMBINANT HUMAN;INSULIN ZINC SUSP PURIFIED PORK;INSULIN LISPRO PROTAMINE RECOMBINANT;INSULIN;INSULIN DEGLUDEC;INSULIN LISPRO;INSULIN ZINC SUSP EXTENDED PURIFIED BEEF;INSULIN SUSP PROTAMINE ZINC BEEF/PORK;INSULIN GLARGINE;INSULIN GLULISINE;INSULIN ZINC SUSP SEMISYNTHETIC PURIFIED HUMAN;INSULIN PEGLISPRO;INSULIN ASPART PROTAMINE RECOMBINANT;INSULIN ZINC SUSP EXTENDED BEEF;INSULIN SUSP ISOPHANE SEMISYNTHETIC PURIFIED HUMAN;INSULIN ZINC SUSP PROMPT PURIFIED PORK;INSULIN SUSP ISOPHANE BEEF/PORK;INSULIN PURIFIED BEEF;INSULIN ZINC SUSP PURIFIED BEEF;INSULIN SUSP PROTAMINE ZINC PURIFIED BEEF;INSULIN SUSP ISOPHANE BEEF;INSULIN SUSP ISOPHANE PURIFIED BEEF;INSULIN DETEMIR;INSULIN PURIFIED PORK;INSULIN ZINC SUSP PURIFIED BEEF/PORK;INSULIN ZINC SUSP PROMPT BEEF;INSULIN PORK;INSULIN ZINC SUSP EXTENDED RECOMBINANT HUMAN;INSULIN SUSP ISOPHANE PURIFIED PORK	1	6;6;6;6;6;6;6;6;6;6;6;6;4B;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6;6
IRINOTECAN	1	1A
ISOSORBIDE MONONITRATE	1	4B
ISTAROXIME	1	6
K-877;BEZAFIBRATE;AVE8134;CP-778875;LY-518674;GEMFIBROZIL;FENOFIBRATE;CLOFIBRATE;FENOFIBRIC ACID;ZYH7;CHOLINE FENOFIBRATE;GW590735;DRF-10945	1	6;3B;6;6;4A;3B;3B;3B;4A;6;6;6;6
KRN-330	1	2B
LANREOTIDE ACETATE	1	6
LAPATINIB	1	1A
LEVOMEPROMAZINE	1	4B

LEVOMILNACIPRAN HYDROCHLORIDE;VENLAFAXINE HYDROCHLORIDE;DESVENLAFAXINE SUCCINATE;PROTRIPTYLINE HYDROCHLORIDE;DULOXETINE HYDROCHLORIDE;IMIPRAMINE HYDROCHLORIDE;AMITRIPTYLINE HYDROCHLORIDE;NORTRIPTYLINE HYDROCHLORIDE	1	6;2B;4B;6;2B;6;6;6
LEVOTHYROXINE;LIOTHYRONINE;DEXTROTHYROXINE SODIUM;LEVOTHYROXINE SODIUM;LIOTHYRONINE SODIUM;DEXTROTHYROXINE	1	3B;5;6;6;4A;6
LISOFYLLINE	1	6
LUMACAFTOR;IVACAFTOR;QBW251;TEZACAFTOR	1	4B;4B;6;6
MAPATUMUMAB	1	2B
MARIZOMIB	1	2A
MB-07803	1	6
MBX-2982	1	6
MDX-070;CAPROMAB	1	6;6
MECLOFENAMATE SODIUM	1	6
MENOTROPINS	1	6
MENTHOL	1	4A
METRELEPTIN	1	6
MIBEFRADIL DIHYDROCHLORIDE;ETHOSUXIMIDE;TRIMETHADIONE;PARAMETHADIONE;PHENSUXIMIDE;METHSUXIMIDE	1	6;4A;3B;6;6;6
MIGLUSTAT;ELIGLUSTAT	1	2B;6
MITOMYCIN	1	1A
MITOTANE	1	1A
MM-111;MARGETUXIMAB	1	2B;1B
MM-121;AV-203;AMG-888	1	2A;2A;2A
MSC-2363318A	1	2A
NABILONE;NONABINE;DRONABINOL;SAD448	1	4B;6;3B;6
NAMINIDIL;SARAKALIM	1	6;6

NAPTUMOMAB ESTAFENATOX;ANATUMOMAB MAFENATOX	1	2B;6
NAVEGLITAZAR;ALEGLITAZAR;TESAGLITAZAR; IMIGLITAZAR;MURAGLITAZAR	1	3B;6;3B;6;3B
NESVACUMAB	1	2A
NINTEDANIB	1	1A
NITRIC OXIDE;RIOCIGUAT;SODIUM NITROPRUSSIDE;ISOSORBIDE DINITRATE	1	3B;4B;6;4B
NITROGLYCERIN	1	3B
NIVOLUMAB	1	1A
NMS-1116354;RXDX-103;BMS-863233	1	2A;6;6
NS-018;GANDOTINIB;AC-430;BMS-911543;XL-019;LS-104	1	6;2A;6;2A;2B;6
OCTREOTIDE	1	1A
ODULIMOMAB	1	6
OLESOXIME	1	4B
OMECAMTIV MECARBIL	1	6
ORLISTAT	1	3B
OTLERTUZUMAB	1	2B
OXALIPLATIN	1	1A
OXELUMAB	1	6
PACRITINIB	1	2B
PAGOCLONE	1	6
PALBOCICLIB	1	1A
PANULISIB	1	2A
PASIREOTIDE DIASPARTATE;PASIREOTIDE PAMOATE;PASIREOTIDE	1	6;6;4B
PATRITUMAB	1	2A
PD-0166285	1	4B
PEGFILGRASTIM;FILGRASTIM;BALUGRASTIM;LI PEGFILGRASTIM	1	2A;2A;6;2B

PEGINTERFERON ALFA-2A;INTERFERON ALFA-N3;INTERFERON BETA-1B;INTERFERON ALFA-2B;INTERFERON ALFACON-1;PEGINTERFERON BETA-1A;ALBINTERFERON ALFA-2B;INTERFERON BETA-1A;PEGINTERFERON ALFA-2B;INTERFERON ALFA-2A	1	2A;6;4B;1B;3B;6;6;6;1B;2A
PEMBROLIZUMAB	1	1A
PEMETREXED	1	1A
PERTUZUMAB	1	1A
PF-04605412	1	2A
PHA-543613;GTS-21;ENCENICLINE;TC-6987;AZD0328;APN1125;BRADANICLINE;AVL-3288;JNJ-39393406;AQW051;ND0801;SSR180711;ABT-126;ABT-107	1	6;4B;6;6;6;6;6;6;6;6;6;4B;6;6
PHENOBARBITAL	1	3A
PILARALISIB	1	2A
PONATINIB	1	1A
POZANICLINE	1	4B
PRAMLINTIDE	1	6
PRAMLINTIDE ACETATE	1	6
PRATOSARTAN	1	6
PROBENECID	1	2A
PUMOSETRAG;DOLASETRON	1	6;4B
PUQUITINIB	1	6
QUARFLOXIN	1	2B
QUERCETIN-3'-O-PHOSPHATE	1	6
RALFINAMIDE	1	6
RANITIDINE;NIZATIDINE;FAMOTIDINE;CIMETIDINE;METHANTHELINE BROMIDE;BETAZOLE HYDROCHLORIDE;RANITIDINE HYDROCHLORIDE;BETAZOLE;CIMETIDINE HYDROCHLORIDE;HISTAMINE;RANITIDINE BISMUTH CITRATE	1	2A;4B;3B;2B;6;6;6;6;6;3B;4A
REBASTINIB	1	2A

REFAMETINIB;BALAMAPIMOD;AS-703988;CI-1040;RG-7304;TRAMETINIB DIMETHYL SULFOXIDE;TAK-733;RO-4987655;PD-0325901;AZD-8330;RG-7167;RG-7420;COBIMETINIB FUMARATE	1	2A;6;2A;2B;4B;1B;2B;2A;2A;2A;6;6;1B
REGORAFENIB	1	1A
REPOTRECTINIB	1	2A
RESATORVID;ERITORAN	1	6;2B
RG-7376	1	6
RIGOSERTIB	1	2A
RIVASTIGMINE;ISOFLUROPHATE;RIVASTIGMINE TARTRATE;TACRINE HYDROCHLORIDE	1	4A;4A;6;6
ROCURONIUM BROMIDE;CISATRACURIUM BESYLATE;PIPECURONIUM BROMIDE;RAPACURONIUM BROMIDE;TUBOCURARINE CHLORIDE;VECURONIUM BROMIDE;MIVACURIUM CHLORIDE;GALLAMINE TRIETHIODIDE;SUCCINYLCHOLINE CHLORIDE;DOXACURIUM CHLORIDE;METOCURINE IODIDE;DECAMETHONIUM BROMIDE;ATRACURIUM BESYLATE;PANCURONIUM BROMIDE	1	6;6;6;6;6;4B;6;4B;6;6;6;6;6;6
ROLOFYLLINE;TECADENOSON;TONAPOFYLLINE;PBF-680;FK453;DERENOFYLLINE;SELODENOSON;GS-9667;TRABODENOSON;BAY1067197;CAPADENOSON;GW493838;NAXIFYLLINE;RPR749	1	6;6;6;6;6;6;6;6;6;6;6;6;6;6
ROZROLIMUPAB;ATOROLIMUMAB;ROLEDUMAB	1	6;6;6
RUPLIZUMAB;TORALIZUMAB;DAPIROLIZUMAB PEGOL	1	6;6;6
RUXOLITINIB PHOSPHATE;BARICITINIB;MOMELOTINIB;AZD-1480	1	1B;4A;2B;2A
S-3304;INCYCLINIDE	1	2A;2A

SAFINAMIDE	1	4B
SAN-300	1	6
SARGRAMOSTIM	1	3B
SCY 635	1	6
SOMATROPIN	1	6
SONIDEGIB	1	1A
SOTATERCEPT	1	2B
SOTRASTAURIN	1	2B
STX-100	1	6
SULFINPYRAZONE	1	3B
SUNITINIB	1	1A
TADALAFIL	1	2A
TALBUTAL;PENTOBARBITAL;BUTALBITAL;METHARBITAL	1	6;4A;6;6
TASIMELTEON;MELATONIN;RAMELTEON	1	6;3B;4B
TC-6499	1	6
TEDATIOXETINE	1	6
TELOTRISTAT ETHYL;TELOTRISTAT	1	2B;2B
TEMOZOLOMIDE	1	1A
TEMSIROLIMUS	1	1A
TEPOTINIB;CAPMATINIB;ONARTUZUMAB;SAR-125844;JNJ-38877605;PF-04217903;NAQUOTINIB;AMG-337;SAVOLITINIB;BMS-698769;TIVANTINIB;AMG-208;EMD-1204831;COMETRIQ;SGX-523;TELISOTUZUMAB;MK-2461;MERESTINIB;BMS-777607	1	1A;1A;2A;2A;2A;2A;2A;2A;2A;2A;6;2A;2A;2A;6;2A;2A;2A;2A;2A
THEOPHYLLINE	1	2B
THIETHYLPERAZINE	1	4B
THROMBIN	1	4B
THYROTROPIN;THYROTROPIN ALFA	1	6;1A
TINZAPARIN	1	5
TORAPSEL;RG-1512;INCLACUMAB;CRIZANLIZUMAB	1	6;6;6;6
TOSEDOSTAT	1	2A
TOZASERTIB	1	2A
TREBANANIB	1	2A

TRICHLORMETHIAZIDE	1	4B
TTP-607;SNS-314;DANUSERTIB;MK-6592	1	2B;2A;2B;2A
TXA127	1	2A
UCN-01	1	2A
VANTICTUMAB	1	2A
VAPREOTIDE	1	6
VARESPLADIB METHYL;VARESPLADIB	1	4A;4B
VELDOREOTIDE	1	6
VEMURAFENIB	1	1A
VEROSUDIL;SAR-407899;AMA0076;DE-104;Y-39983	1	6;6;6;6;6
VIDOFLUDIMUS	1	6
VILAZODONE HYDROCHLORIDE	1	6
VOLOCIXIMAB	1	2B
WX-554	1	2A
ZOLPIDEM	1	4B
ZONAMPANEL;PERAMPANEL;MIBAMPATOR;MK-8777;FARAMPATOR;BECAMPANEL;CX1739;PF-04958242;TALAMPANEL	1	6;5;6;6;6;6;6;6;6
ZOPICLONE	1	4B

Appendix 3

Frequencies of all drugs observed in the top 5 in each THCA-US sample via the driver and DEG propagation method. Drugs that were merged (because they have the same target gene(s)) are listed separated by ";".

Drug	Freq	Tiers
TRAMETINIB	314	1A
SELUMETINIB	305	1A
REGORAFENIB	201	1A
CETUXIMAB	153	1A
VEMURAFENIB	145	1A
DASATINIB	117	1A
CRIZOTINIB	98	1A
NERISPIRDINE	89	6
TOPIRAMATE	86	4B
OCRIPLASMIN	72	6
SORAFENIB	63	1A
PANITUMUMAB	45	1A
PRIMIDONE	45	4B
VANDETANIB	38	1A
DALFAMPRIDINE;GUANIDINE HYDROCHLORIDE;TEDISAMIL	35	4B;6;6
COLLAGENASE CLOSTRIDIUM HISTOLYTICUM	34	6
PONATINIB	34	1A
EVEROLIMUS	28	1A
PERTUZUMAB	28	1A
BORTEZOMIB	26	1A
DACTOLISIB	25	2A
GEFITINIB	18	1A
IMAGABALIN;ATAGABALIN;GABAPENTIN;PREGABALIN;BEPRIDIL HYDROCHLORIDE;GABAPENTIN ENACARBIL	17	6;6;3B;3B;6;6
BINIMETINIB	16	1A
CISPLATIN	15	1A
ENFLURANE	14	4A
ALPELISIB	13	1A
PELAREOREP	13	2B
TIPIFARNIB	12	2A

CARFILZOMIB	10	1A
QUERCETIN	9	3B
LAPATINIB	8	1A
PACLITAXEL	7	1A
CEP-32496	6	2A
DABRAFENIB	6	1A
RALFINAMIDE	6	6
SORAFENIB TOSYLATE	6	2A
ZONISAMIDE	6	3B
DESFLURANE;SEVOFLURANE	5	4A;4A
ERLOTINIB	5	1A
INFIGRATINIB	5	2A
ISOFLURANE	5	3B
OLAPARIB	5	1A
PICTILISIB	5	2A
RONICICLIB;PHA-793887;AT-7519;AZD-5438	5	2A;2A;2A;2A
ACAMPROSATE CALCIUM	4	6
AFATINIB	4	1A
DOVITINIB	4	2A
GSK-690693	4	2A
OLANZAPINE	4	2A
OPROZOMIB;IXAZOMIB CITRATE	4	2A;1A
POZIOTINIB;OSIMERTINIB MESYLATE;DACOMITINIB;AC-480	4	2A;2B;1A;2A
SAFINAMIDE	4	4B
TALBUTAL;PENTOBARBITAL;BUTALBITAL;METHARBITAL	4	6;4A;6;6
DANTROLENE	3	3B
DANTROLENE SODIUM	3	6
DOCETAXEL	3	1A
DOXORUBICIN	3	1A
ENMD-2076	3	2A
GEDATOLISIB	3	2A
IRINOTECAN	3	1A
METFORMIN HYDROCHLORIDE	3	2B
NIRAPARIB	3	1A
NITRAZEPAM	3	4B
NV-128;ME-344	3	6;2A
OMIPALISIB	3	2A
PENTOXIFYLLINE	3	2A
TEMOZOLOMIDE	3	1A

BENTIROMIDE	2	6
BRYOSTATIN	2	3B
CHLORDIAZEPOXIDE;CLOBAZAM;HALAZEPAM;OXAZEPAM;FLURAZEPAM;BROMAZEPAM;PRAZEPAM;QUAZEPAM;ADIPIPLON;MIDAZOLAM;ESTAZOLAM	2	4B;4B;6;3A;4B;4B;6;6;6;4B;4B
DIPYRIDAMOLE	2	2B
ENMD-981693	2	2A
FLUOROURACIL	2	1A
GEMCITABINE	2	1A
JNJ-26483327	2	2A
LY-2780301;XL-418	2	2B;2A
PANOBINOSTAT	2	1A
PEMBROLIZUMAB	2	1A
PILARALISIB	2	2A
PROPYLTHIOURACIL	2	3A
RUCAPARIB	2	1A
RUXOLITINIB	2	1A
TEPROTUMUMAB	2	2A
TEZAMPANEL	2	6
TRASTUZUMAB	2	1A
TRIAZOLAM;CLONAZEPAM;CLORAZEPIC ACID;ALPRAZOLAM;DIAZEPAM;TEMAZEPAM;LORAZEPAM	2	4B;3B;6;4B;3B;4B;4B
VORINOSTAT	2	1A
ABITUZUMAB;INTETUMUMAB	1	2B;2B
ACARBOSE	1	3B
ACP-001;PEGVISOMANT;SOMATREM;SOMATROGON	1	6;2B;6;6
AMINOGLUTETHIMIDE	1	1A
AMLODIPINE	1	2B
APREPITANT;VOFOPITANT;CASOPITANT	1	4B;4B;6
AR-42	1	2B
ASENAPINE	1	4B
ASM-024	1	6
AT-9283	1	2A
ATAMESTANE;TESTOLACTONE;ARIMIDEX	1	2B;4B;6

AVANAFIL;VARDENAFIL HYDROCHLORIDE;VARDENAFIL;SILDENAFIL;UDENAFIL;SILDENAFIL CITRATE;CP-461	1	6;6;5;2A;4A;2A;2B
BELINOSTAT	1	1A
BEPRIDIL	1	4B
BEVACIZUMAB	1	1A
BICALUTAMIDE	1	1A
BMS-690514	1	2A
BMS-817378;CABOZANTINIB S-MALATE;BMS-794833;GOLVATINIB;TAS-115	1	2A;1A;6;2A;6
BRIGATINIB	1	1A
BUMETANIDE	1	2B
BUTETHAL;BUTABARBITAL;HEXOBARBITAL;MEPHOBARBITAL;THIOPENTAL;SECOBARBITAL	1	6;6;4B;3B;4A;4A
CARBOPLATIN	1	1A
CEDIRANIB	1	2A
CEP-2563	1	6
CERITINIB	1	1A
CHLOROTHIAZIDE	1	6
CHLORZOXAZONE	1	4B
CIXUTUMUMAB	1	2A
CP-459632	1	6
CRENOLANIB	1	2A
CUDC-101	1	2A
CYCLOSPORINE	1	2A
CYTARABINE	1	1A
DAUNORUBICIN	1	1A
DIGOXIN;ACETYLDIGITOXIN;DESLANOSIDE;DIGITOXIN	1	2A;6;6;3B
DNK333;BURAPITANT;FIGOPITANT;VESTIPITANT MESYLATE;FOSAPREPITANT DIMEGLUMINE;TRADIPITANT;ORVEPITANT;EZLOPITANT;ROLAPITANT HYDROCHLORIDE;SERLOPITANT;NETUPITANT	1	6;6;6;6;6;6;6;6;2B;6;2A
DRINABANT;SURINABANT;MARINOL;CANNABIDIOL;TARANABANT;RIMONABANT;IBIPINABANT;AZD2207;OTENABANT	1	6;4B;6;3B;4B;3B;4B;6;6
ELPETRIGINE	1	6

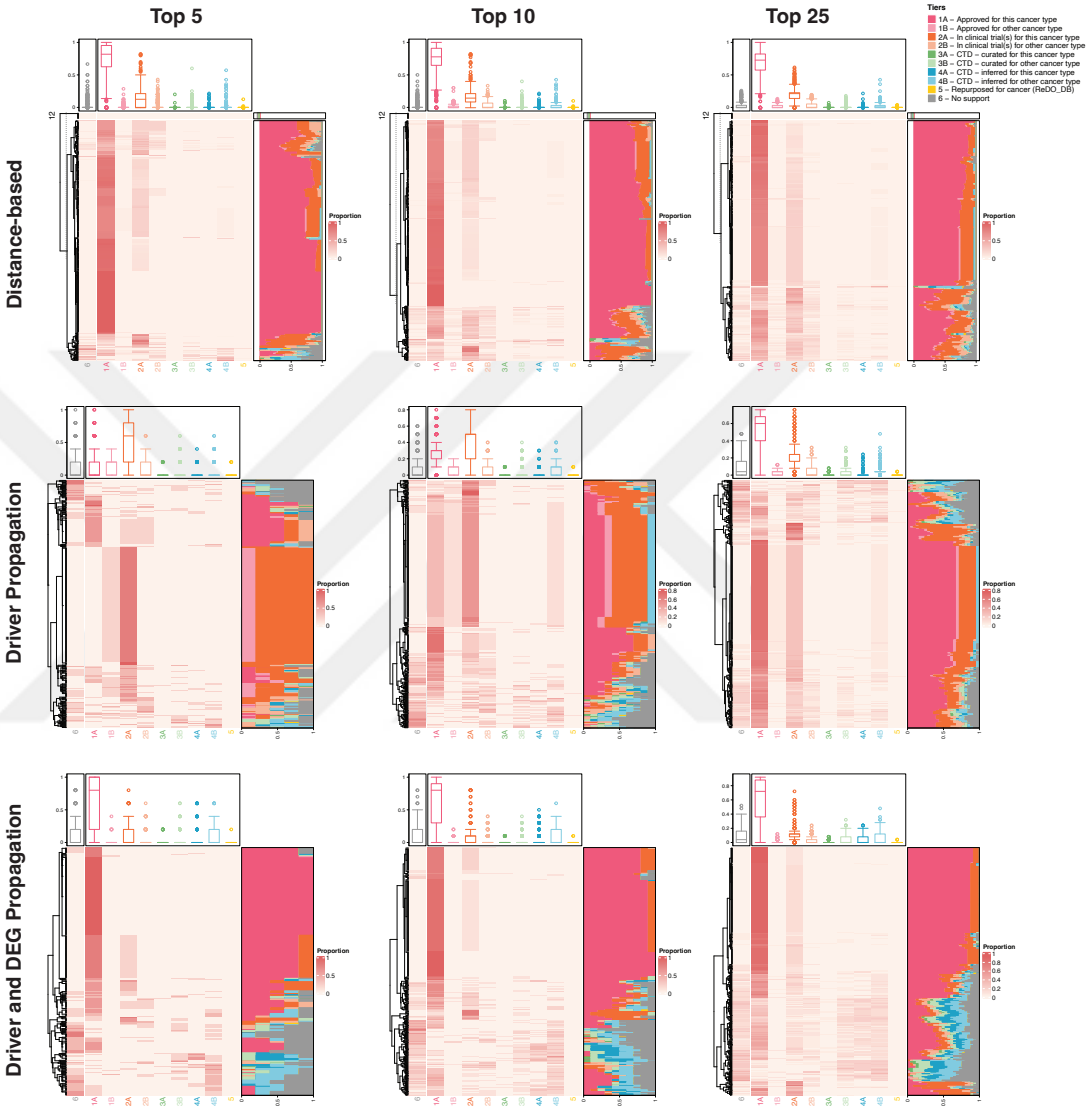
EMICIZUMAB	1	6
ENFORTUMAB VEDOTIN	1	1A
ENTRECTINIB	1	1A
EPINASTINE	1	4A
ERITORAN TETRASODIUM	1	2B
ERLOSAMIDE	1	2B
ETARACIZUMAB	1	2A
FLINDOKALNER	1	6
FLUNARIZINE	1	4B
FORETINIB	1	2A
GLEPAGLUTIDE;TEDUGLUTIDE;ELSIGLUTIDE	1	6;6;6
GLPG-0187	1	2A
GLUFOSFAMIDE	1	2B
HALOTHANE	1	4A
IDRAPARINUX SODIUM;APIXABAN;OTAMIXABAN;EDOXABAN;IDRAPARINUX;RIVAROXABAN;EDOXABAN TOSYLATE;LETAXABAN;BETRIXABAN;CHEMBL1271162	1	6;2B;6;6;6;2A;6;6;6;6
IMETELSTAT;IMETELSTAT SODIUM;TERTOMOTIDE	1	2B;2A;2B
INDEGLITAZAR;GW501516;SODELGLITAZAR	1	6;4A;6
ISOSORBIDE MONONITRATE	1	4B
ISTAROXIME	1	6
LANREOTIDE ACETATE	1	6
LESTAURTINIB	1	2B
LEVETIRACETAM	1	2B
MARIZOMIB	1	2A
MDX-070;CAPROMAB	1	6;6
MENADIONE	1	3B
METRELEPTIN	1	6
MIBEFRADIL DIHYDROCHLORIDE;ETHOSUXIMIDE;TRIMETHADIONE;PARAMETHADIONE;PHENSUXIMIDE;METHSUXIMIDE	1	6;4A;3B;6;6;6
MIDOSTAURIN	1	1A
NABILONE;NONABINE;DRONABINOL;SAD448	1	4B;6;3B;6

NAVEGLITAZAR;ALEGLITAZAR;TESAGLITAZAR; IMIGLITAZAR;MURAGLITAZAR	1	3B;6;3B;6;3B
NITRIC OXIDE;RIOCIGUAT;SODIUM NITROPRUSSIDE;ISOSORBIDE DINITRATE	1	3B;4B;6;4B
NITROGLYCERIN	1	3B
NKTR-171;ESLICARBAZEPINE;BENOXINATE HYDROCHLORIDE;ETHOTOIN;PHENYTOIN;ARTI CAINE HYDROCHLORIDE;FOSPHENYTOIN SODIUM;RUFINAMIDE;CARBAMAZEPINE;OXCAR BAZEPINE;TOCAINIDE HYDROCHLORIDE;PROPOXYCAINE HYDROCHLORIDE;ETIDOCAINE HYDROCHLORIDE;TETRACAINE;PHENACEMIDE; HEXYLCAINE HYDROCHLORIDE;INDECAINIDE HYDROCHLORIDE;TETRACAINE HYDROCHLORIDE;CHLOROPROCAINE HYDROCHLORIDE;LAMOTRIGINE;ROPIVACAINE HYDROCHLORIDE;PRILOCAINE HYDROCHLORIDE;PHENAZOPYRIDINE HYDROCHLORIDE;PROCAINAMIDE HYDROCHLORIDE;MERETHOXYLLINE PROCAINE;MEPHENYTOIN;DYCLONINE HYDROCHLORIDE;PRILOCAINE;PROPARACAINE HYDROCHLORIDE;MEPIVACAINE HYDROCHLORIDE;QUINIDINE POLYGALACTURONATE;QUINIDINE GLUCONATE;MEXILETINE HYDROCHLORIDE;MORICIZINE HYDROCHLORIDE;DISOPYRAMIDE PHOSPHATE;PHENYTOIN SODIUM;EVENAMIDE;LIDOCAINE HYDROCHLORIDE;ESLICARBAZEPINE ACETATE;PROCAINE HYDROCHLORIDE;QUINIDINE SULFATE	1	6;6;6;4B;2B;6;6;4B;3B;2A;6;6 ;6;4B;6;6;6;6;6;2B;2B;6;6;6;6; 4B;6;2B;6;6;6;6;6;6;6;6;6;1B; 6;6;6
OCTREOTIDE	1	1A
OCTREOTIDE ACETATE	1	6
OMECAMTIV MECARBIL	1	6
ORPHENADRINE HYDROCHLORIDE;ORPHENADRINE CITRATE	1	6;6
PASIREOTIDE DIASPARTATE;PASIREOTIDE PAMOATE;PASIREOTIDE	1	6;6;4B
PF-04691502	1	2A

PHA-543613;GTS-21;ENCENICLINE;TC-6987;AZD0328;APN1125;BRADANICLINE;AVL-3288;JNJ-39393406;AQW051;ND0801;SSR180711;ABT-126;ABT-107	1	6;4B;6;6;6;6;6;6;6;6;6;4B;6;6
PHENOBARBITAL	1	3A
PUQUITINIB	1	6
QUERCETIN-3'-O-PHOSPHATE	1	6
RESATORVID;ERITORAN	1	6;2B
RG3487	1	3B
RIBOCICLIB	1	1A
RIVASTIGMINE;ISOFLUROPHATE;RIVASTIGMINE TARTRATE;TACRINE HYDROCHLORIDE	1	4A;4A;6;6
SELURAMPANEL	1	6
SOMATROPIN	1	6
SOTATERCEPT	1	2B
STANNSOPORFIN	1	4A
SUNITINIB	1	1A
TADALAFIL	1	2A
TALAZOPARIB	1	1A
TEDATIOXETINE	1	6
THEOPHYLLINE	1	2B
THROMBIN	1	4B
THYROTROPIN;THYROTROPIN ALFA	1	6;1A
TINZAPARIN	1	5
TOLAZOLINE HYDROCHLORIDE	1	6
TRICHLORMETHIAZIDE	1	4B
VARENICLINE	1	3B
VARESPLADIB METHYL;VARESPLADIB	1	4A;4B
VELDOREOTIDE	1	6
VELIPARIB	1	2A
VERAPAMIL	1	2B

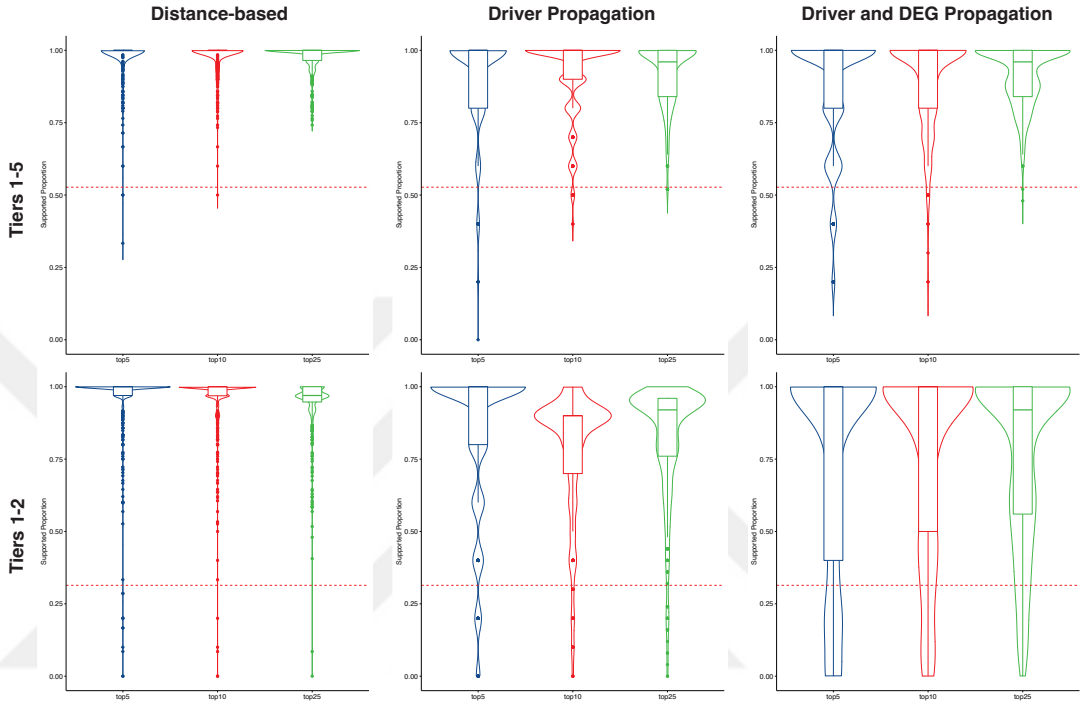
Appendix 4

Heatmaps of proportions of selected drugs in each tier per top 5, 10, 25 for each method (THCA-US data, BioGRID PIN)



Appendix 5

Distribution of proportions of supported (either Tiers 1-5 or 1-2) selected drugs per drug prioritization method (THCA-US data, BioGRID PIN)



9. CURRICULUM VITAE