Briefings in Bioinformatics, 00(00), 2018, 1–10

doi: 10.1093/bib/bby119
Advance Access Publication Date: 18 December 2018
Review article

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

Interactive visual analysis of drug–target interaction networks using Drug Target Profiler, with applications to precision medicine and drug repurposing

Ziaurrehman Tanoli, Zaid Alam*, Aleksandr Ianevski*, Krister Wennerberg, Markus Vähä-Koskela and Tero Aittokallio

Corresponding authors: Ziaurrehman Tanoli, Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland.
E-mail: zia.rehman@helsinki.fi; Tero Aittokallio, Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Helsinki Institute for Information Technology (HIIT), Aalto University, Espoo and Department of Mathematics and Statistics, University of Turku, Turku, Finland.
Tel.: +318-503182426; Fax: +358 2941 25737; E-mail: tero.aittokallio@helsinki.fi
*These authors contributed equally to this work.

Abstract

Knowledge of the full target space of drugs (or drug-like compounds) provides important insights into the potential therapeutic use of the agents to modulate or avoid their various on- and off-targets in drug discovery and precision medicine. However, there is a lack of consolidated databases and associated data exploration tools that allow for systematic profiling of drug target-binding potencies of both approved and investigational agents using a network-centric approach. We recently initiated a community-driven platform, Drug Target Commons (DTC), which is an open-data crowdsourcing platform designed to improve the management, reproducibility and extended use of compound–target bioactivity data for drug discovery and repurposing, as well as target identification applications. In this work, we demonstrate an integrated use of the rich bioactivity data from DTC and related drug databases using Drug Target Profiler (DTP), an open-source software and web tool for interactive exploration of drug–target interaction networks. DTP was designed for network-centric modeling of mode-of-action of multi-targeting anticancer compounds, especially for precision oncology applications. DTP enables users to construct an interaction network based on integrated bioactivity data across selected chemical compounds and their protein targets, further customizable using various visualization and filtering options, as well as cross-links to several drug and protein databases to provide comprehensive information of the network nodes and interactions. We demonstrate here the operation of the DTP tool and its unique features by several use cases related to both drug discovery and drug repurposing applications, using examples of anticancer drugs with shared target profiles. DTP is freely accessible at http://drugtargetprofiler.fimm.fi/.

Key words: network visualization; drug–target interactions; drug mode-of-action; drug repurposing; precision oncology

Ziaurrehman Tanoli is a postdoctoral researcher at the Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland.
Zaid Alam is a research assistant at the FIMM, University of Helsinki, Finland.
Aleksandr Ianevski is a PhD student at the FIMM, University of Helsinki, Finland.
Krister Wennerberg is a professor and group leader at the Biotech Research & Innovation Centre, University of Copenhagen, Denmark.
Markus Vähä-Koskela is a senior researcher at the FIMM, University of Helsinki, Finland.
Tero Aittokallio is a professor and group leader at FIMM, University of Helsinki, Finland.
Submitted: 11 September 2018; **Received (in revised form):** 1 November 2018

© The Author(s) 2018. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

The effect of small molecules on various protein targets is critical for understanding the compound's mode-of-action (MoA), therapeutic potential and potential side effects prior to clinical trials. In recent years, multiple resources have been developed based on diverse compound collections to define primary targets of small molecules [1] and to identify potent molecular probes for specific molecular targets [2]. Toward mapping a wider spectrum of compound–target interactions, including both intended 'on-targets' as well as secondary 'off-targets', a number of high-throughput target profiling studies have published a large amount of compound–target bioactivity data, which are available through open-access databases, such as ChEMBL [3], BindingDB [4] and PubChem [5], and recently integrated and annotated in our open-access Drug Target Commons (DTC) platform [6, 7]. Similarly, compound sensitivity data have been made available across hundreds of human cancer cell lines in open-access databases, including Cancer Cell Line Encyclopedia (CCLE) [8], Genomics of Drug Sensitivity in Cancer (GDSC) [9] and Cancer Therapeutic Response Portal (CTRP) [10], and recently integrated and standardized in an open-access PharmacoDB platform [11]. While these resources have been useful, the compound bioactivity data are mostly provided in a tabular format and without explicit connections across multiple protein targets and cell lines, which has made it difficult to understand the compound's context-dependent MoA on a systems level, especially for multi-targeting compounds with various on- and off-target interactions.

To enable a network-centric approach to interactive analysis and visual exploration of compound–target interaction patterns in a systematic manner, we have combined the standardized drug–target potency and drug sensitivity data from both the DTC and PharmacoDB platforms, respectively. Based on these existing data resources, we then implemented an open-source web application, named Drug Target Profiler (DTP), which enables a systems-level profiling of the full target space and potencies of approved and investigational drugs to provide insights into their therapeutic and biological potentials. DTP implements a number of unique features, including protein family classification, novel interaction scoring, compound response and gene expression profiles, as well as disease and mutational information for protein targets, which are critical for finding selective inhibitors. We demonstrate here the wide applicability of DTP using three use cases related to both anticancer drug discovery and drug repurposing. In the first use case, we show how linking the on- and off-target potencies with compound sensitivity and target expression profiles across cancer cell lines enables one to better understand the polypharmacological effects of multi-targeting agents such as imatinib and bosutinib. The second use case explores interactions among clinically approved drugs with mutated targets to facilitate drug discovery efforts for precision oncology applications. In the third use case, we extend to drug combinations by exploiting relationships among drug pairs based on their DTP interaction scores.

Drug–target interaction prediction methods

A number of recent review articles have surveyed the applications of bioinformatic methods for drug–target interaction (DTI) prediction and demonstrated their value in multiple downstream analyses. For instance, Wang and Kurgan [12] reviewed 35 methods for DTI prediction using similarity-based approaches. They grouped these methods based on three types of similarities and then compared their key properties such as source databases and predictive models. Rifaioglu *et al.* [13] reviewed the recent applications of machine-learning techniques (especially deep learning) to computational drug discovery and virtual screening. They summarized the main components of virtual screening methods, including various compound and protein features, as well as toolkits, bioactivity databases and gold standard data sets for systems training and benchmarking. Ezzat *et al.* [14] provided an overview and empirical evaluation of several DTI prediction techniques. They discussed the pros and cons of each method and highlighted potential avenues for further enhancement of DTI prediction performance. To extend accessibility of the methods to scientists without a computational background, Sam *et al.* [15] provided a survey on web tools that support drug repositioning. Fang *et al.* [16] provided a review on polypharmacological profiles of five natural products that are currently being considered as cancer therapies. They also highlighted combination therapies that target tumor ecosystems by exploiting the immunological and inflammatory side effects of natural products. Lotfi Shahreza *et al.* [17] reviewed especially network-based methods for predicting drug targets for drug repositioning. Chen *et al.* [18] provided a general review on computational models for DTI prediction as well as offered several future directions based on network-based drug discovery, tumor clone-based networks and cancer hallmark-based methods.

Several bioinformatics tools are currently available for predicting DTIs and helping in drug repurposing and precision medicine applications. For instance, substructure-drug-target network-based inference [19] is a cheminformatics tool for systematic prediction of DTIs and drug repositioning. RFDT [20] is another tool, based on random forest model, which predicts potential DTIs by using evolutionary information of proteins and structural properties of the drugs. DrugTargetInspector [21] is an interactive assistance tool for patient treatment stratification. It analyzes genomic, transcriptomic and proteomic data sets and provides information on deregulated drug targets, enriched biological pathways and deregulated subnetworks, as well as mutations and their potential effects on putative drug targets and genes of interest. RepurposeDB [22] is a reference database for drug repurposing investigations. It currently combines information on 253 drugs and 1125 diseases and identifies pharmacological and epidemiological factors mediating drug repositioning. These bioinformatics tools can provide comprehensive DTI information based on databases and computational prediction methods. However, there are not many tools or case studies that provide visual insights into the complex DTI networks. Drug Target Explorer [23] is a recently developed web tool, which enables network-based visualization of the compound–target interaction spaces to assist identification of structurally similar molecules and their targets. Similarly, STITCH [24] is a useful and widely used protein–chemical interactions visualization tool that integrates experimental data with text-mining information and interaction predictions. DTP provides novel and complementary options for interactive DTI network analysis and visualization, as well as large-scale and manually-annotated data for training more accurate and comprehensive DTI prediction methods.

Unique features of DTP

The DTP software is built on the bioactivity data and annotations from our recent community platform, DTC, an open-data crowdsourcing portal designed to improve the management,

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

consensus and extended use of compound–target bioactivity data for drug discovery, target identification and repurposing [6]. DTP web software is linked with the DTC database and it enables users to visually explore and interactively analyze the curated data from one of the most comprehensive bioactivity knowledgebase available. The common aims of DTC and DTP include speeding up the experimental efforts and de-risking the clinical trial designs by (i) using computational tools as systematic and cost-effective means for guiding the compound–target mappings, (ii) prioritizing most potential compound–target interactions for further experimental or pre-clinical evaluation and (iii) avoiding the modulation of 'antitargets' related to toxic side effects.

DTP additionally combines standardized compound sensitivity bioactivity data from the PharmacoDB platform [11] and gene expression profiles across hundreds of cancer cell lines from CCLE [8], GDSC [9] and Genentech Cell Line Screening Initiative (gCSI) resources [25]. PharmacoDB harmonizes the compound sensitivity and genomic profiles from multiple studies that have used different experimental protocols and aims to reduce the technical and biological variation in the standardized *in vitro* pharmacological profiles, enabling rigorous comparison and integrative analysis of large-scale drug screening data sets. PharmacoDB provides the largest cancer pharmacogenomic database published to date, which was made possible by manual curation of both the cell line and chemical compound identifiers that maximizes the overlap between studies, and therefore improves the reproducibility of the pharmacogenomics data that are otherwise notoriously difficult to integrate and cross-compare [26].

Compared to the STITCH interaction score, in which only the strongest reported or predicted binding affinity between any compound and protein is considered in the construction of compound–protein interaction networks for the given compounds or proteins, the interaction strength score in DTP is based on comprehensive and manually curated bioactivity data across various experimental assays and studies, stored in the open-access DTC database, resulting in a robust and standardized source of evidence for constructing compound–target interaction networks (see Implementation details). Compared to the existing compound/target data resources, DTC implements several novel features for data curation, annotation and intra-resource integration of quantitative compound–target bioactivity profiles [6]. In addition, the DTP platform makes use of the novel features of DTC version 2.0 [7], such as protein family classification and disease information for protein targets, to enable its key unique features:

- Systems-wide and harmonized target profiles include not only the primary targets but also off-targets and disease-related mutant targets, making available the full spectrum of differential target potencies in DTP, which is important for drug repurposing and precision oncology applications.
- Customized bioactivity assay annotations provide reusable and reproducible target profiles, making it possible to calculate an integrated interaction score for compound–target pairs in DTP (see below), based on multiple target profiling studies that use differing assay types and end points.
- By linking the on- and off-target potencies with compound sensitivity and target expression profiles across hundreds of cancer cell lines, DTP provides the users with an improved understanding of the context-specific mechanisms behind polypharmacological effects of the multi-targeting agents.

Implementation details

DTP back-end was implemented using Python (Django 1.9), the bioactivity database using PostgreSQL (v. 9.0) and the network visualization front-end using JavaScript library (D3.4). The web software is freely available without login requirements at http://drugtargetprofiler.fimm.fi/. We have also provided comprehensive step-by-step user instructions with example network visualization outputs on the DTP help page (http://drugtargetprofiler.fimm.fi/help), along with a step-by-step video tutorial to help the end-users with various applications. The source code has been made publicly available at GitHub, https://github.com/zrehman/DrugTargetProfiler, released under the GNU General Public License 3.0.

For the compound–target network construction, bioactivity data from dose-response measurements (IC_{50}, EC_{50}, XC_{50}, AC_{50}, K_d, K_i, and potency) were extracted from DTC [6]. DTP's current version includes 937 269 compounds, 5077 protein targets (192 mutant targets) and 4 429 858 bioactivities among the compounds and targets (Table 1). Bioactivity data values were further converted into nM to have a standard unit. Our novel interaction potency scores for compound–target pairs were computed based on aggregated bioactivity measurements and by considering the protein family and assay format information (Table S2) using the following steps for each compound–target pair:

(i) Map the multiple bioactivity values (nM) across various bioassays and bioactivity readouts into a summary interaction score (0–1), where higher values indicate more potency (see Supplementary Data for details).
(ii) Take median value of the interaction scores across various assays computed in step (i). Median interaction score over the replicate measurements provides a robust way of summarizing interaction strength.

All the compounds are cross-referenced in the current DTP version to >15 databases using UniChem. Clinical information for a subset of ~3500 investigational compounds and approved drugs are extracted from ClinicalTrials resource. The proteins are cross-referenced to >20 databases using UniProt, and the protein–disease associations are currently extracted from DisGeNet [27]. The purpose of cross-referencing proteins and compounds with external resources is to provide the end-users with further information on the targets or agents present in multiple resources. The full list of names, URLs and description of the cross-referencing resources are provided in Supplementary Data Table S1.

To better understand the context-specific MoA of the compounds in various cancer cell backgrounds, standardized drug sensitivity scores (DSSs) for a panel of >1000 cancer cell lines across 7 different data sources were extracted from PharmacoDB [11]. To investigate the target expression, normalized gene expression data for the protein targets across the cancer cell lines were extracted from two data portals, CCLE and GDSC, and from the gCSI [25]. More details of how to use these data sources are available in the DTP user guide (http://drugtargetprofiler.fimm.fi/help).

Based on several in-house test rounds, bugs were initially identified and fixed. Based on the chemical biology applications in-house and by our collaborators, the software usability and visualization options have been improved and a variety of useful data resources have been integrated with the DTP software. In addition to personal contacts, feedback forum is provided at the DTP website in order to have additional suggestions from

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

Table 1. Data and information integrated into DTP from publicly available key external resources

Data resource	Data type	Description and current statistics in DTP
DTC [6] (https://drugtargetcommons.fimm.fi/)	Bioactivity data (IC_{50}, EC_{50}, XC_{50}, AC_{50}, K_d, K_i and potency)	Dose-response and single-dose bioactivity data among compounds and protein targets that quantify the strength of their interactions. Bioactivities, 4 429 858; compounds, 937 269; proteins, 5077; mutant proteins, 192.
UniChem [28] (https://www.ebi.ac.uk/unichem/)	Compound cross-referencing	Using standard InChiKeys, DTP compounds were cross-linked to >15 drug databases to provide further insights into the compound's properties (see Supplementary Data Table S1 for details).
UniProt [29] (http://www.uniprot.org/)	Protein target cross-referencing	Using UniProt protein ID, protein targets in DTP were cross-linked to >20 protein/genomic databases (e.g. UniProt, Ensembl, ENA, PDBe, HGNC and Uniref; see Supplementary Data Table S1 for details).
PubChem [5] (https://pubchem.ncbi.nlm.nih.gov/)	Compound 2D structures	Using standard InChiKeys, 2D structures for ~91 000 DTP compounds were obtained from PubChem, and these can be visualized by clicking compound nodes in DTP.
PDB [30] (https://www.rcsb.org)	Protein 3D structures	Structural images for 1871 DTP protein targets were obtained from PDB, and these are available as pop-up windows by clicking protein target nodes in DTP.
DisGeNET [27] (http://www.disgenet.org)	Disease–gene associations	Collection of human disease phenotypes associated with the genes coding for the DTP protein targets. Genes, 1573; associated diseases, 4123; total number of disease–gene associations, 331 514.
ChEMBL [31] (https://www.ebi.ac.uk/chembl/), Panther [32] (http://pfam.xfam.org) and Pfam [33] (http://www.pantherdb.org/)	Superfamily classification of proteins	Protein targets in DTP are classified into 7 super families: kinases, 541; enzymes, 3029; GPCRs, 744; nuclear receptors, 151; ion channels, 301; epigenetic receptors, 118; transporters, 216; other, 1331. The total number of classified proteins in DTP is 6431.
ClinicalTrials (https://clinicaltrials.gov/)	Clinical trial information for compounds	Clinical information for 3532 DTP compounds having 292 218 indications, including study details, development phases, symptoms, MeSH terms, adverse effects, participant information, eligibility criteria and reference publications.
CCLE [8] (https://portals.broadinstitute.org/ccle)	Gene expression intensity values	Microarray measurements across 1036 cell lines for a total of 18 898 genes. Among these, 2726 were matched with genes in DTP.
GDSC [9] (http://www.cancerrxgene.org/)	Gene expression intensity values	Microarray measurements across 967 cell lines for 17 417 genes. Among these, 2688 genes were matched with those in DTP.
gCSI [25]	Gene expression intensity values	RNA-seq measurements across 675 cell lines for 19 042 genes. Among these, 2837 genes were matched with those in DTP.
PharmacoDB [11] (https://pharmacodb.pmgenomics.ca/)	Standardized drug sensitivity profiles	Standardized DSS for 757 compounds across >1000 cancer cell lines combined from 7 data sources (FIMM, CCLE, GDSC, gCSI, GRAY, CTRPv2 and UHNBreast).

the end-users. We are continuously improving the software implementation according to the user suggestions and feedback. With increased number of users, we hope to get diverse suggestions, both from computational and translational researchers, which will make DTP a widely adopted and highly useful tool for the broad community of end-users.

DTP workflow

After entering the names of compounds or proteins of interest, DTP generates an interactive network layout with multiple visualization options for the user (Figure 1). A combined network visualization of compounds and their various targets (including

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

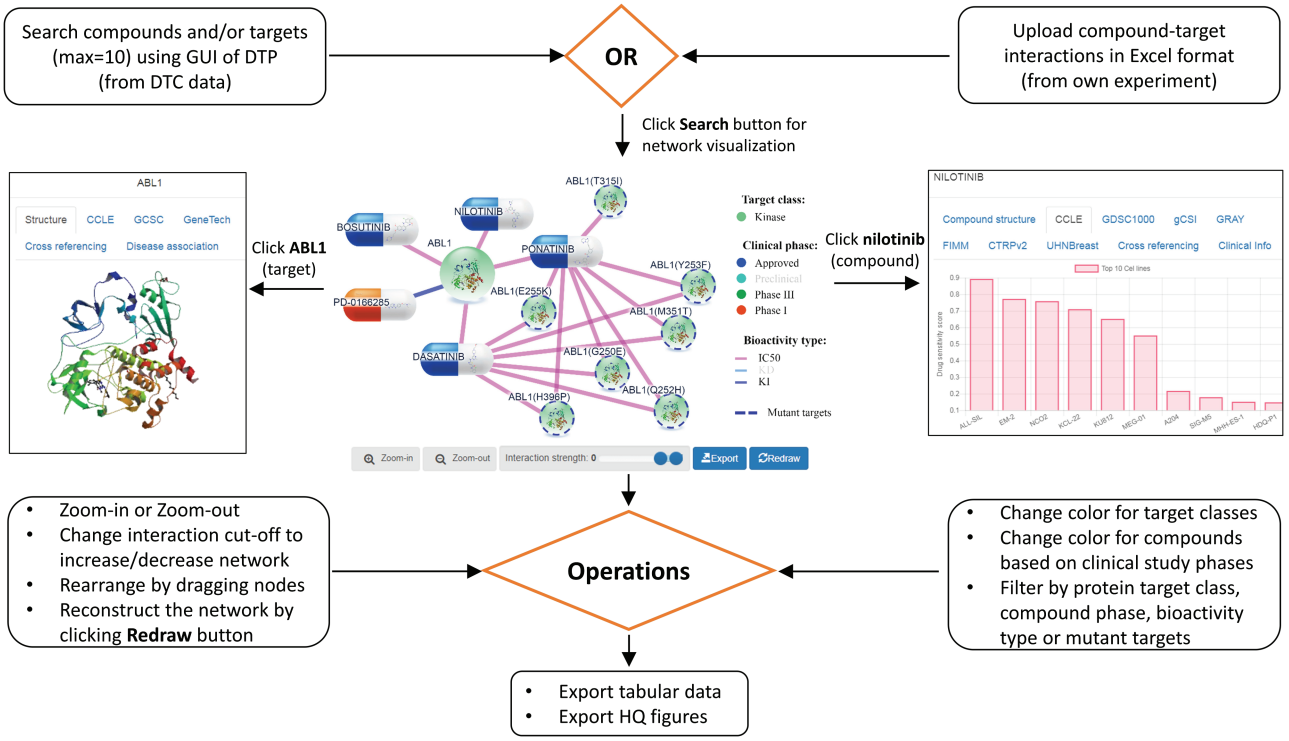

Figure 1. A schematic DTP workflow through various interactive input/output operations and data visualization options.

on- and off-targets) can be made for a maximum of 10 search entries. Users may also upload their own interaction data and generate customized visualizations for new compounds, targets and their interactions beyond those that are available in the DTP database. The user-uploaded data may contain new and existing compounds or targets. A data input template for uploading compound–target interactions via an Excel file is provided at the software's landing page (http://drugtargetprofiler.fimm.fi/). DTP features a variety of user options, including filter by protein family, bioactivity type, compound development phase or mutant versus wild-type protein targets, which are all visible as network figure legends (see http://drugtargetprofiler.fimm.fi/help).

After an initial network layout (constructed using the Barnes–Hut algorithm), the node sizes, colors and positions can be set manually. The network size can be increased or decreased gradually by adding or removing protein targets or by adjusting the interaction strength cut-off (0–1). A visual workspace can be shared with the collaborators through web link. Upon clicking a protein target, the Graphical User Interface (GUI) displays top-expressed cell lines, 3D structure for the protein, protein–disease associations and cross-links to ~20 public protein and genomic databases (such as UniProt, PDB and Ensembl). For compounds, GUI shows the DSS for top-sensitive cell lines, drug indications and cross-links to ~15 compound databases (such as ChEMBL, PubChem and DrugBank). For more details of these cross-referenced resources, see Supplementary Data Table S1. User-customized compound–target interaction networks can be exported in Portable Network Graphics (PNG) format. The selected data can be downloaded in Comma Separated Values (CSV) format.

Application use cases

Polypharmacological effects of multi-targeting drugs

We first demonstrated the operation of the DTP by exploring the target spaces of two well-studied ABL1 kinase-inhibiting drugs—imatinib and bosutinib—which are both approved for the treatment of patients with chronic myeloid leukemia (CML) [34], with bosutinib showing efficacy also in imatinib-resistant cases due to its greater ability to inhibit common mutants of ABL1. As can be seen from Figure 2A, imatinib and bosutinib not only have both overlapping and unique targets, confirming the shared MoA, but also gives clues about their distinct therapeutic and toxicity profiles. As imatinib and bosutinib have different target spaces, they elicit different toxicities in patients suffering from distinct cancers, or even subtypes of the same cancer, driving continued development of ABL1-multi-target inhibitors. DSSs for the top 10 cancer cell lines that have been tested for bosutinib and imatinib show relatively consistent drug sensitivities (Figure 2B) but more variability in ABL expression profiles across eight CML cell lines and two BCR-ABL1 acute lymphoblastic leukemia (ALL) cell lines, suggesting context-specific on-target MoA in cancer cells with a high dependence on ABL1.

To further demonstrate the utility of the DTI network analysis, we searched for drug-repurposing cases of drug efficacy in off-label indications. For example, imatinib targets CSFR1 (Figure 2A), and it has been found to be effective also in Philadelphia chromosome negative myeloproliferative patient cells due to the CSF1R signaling abnormalities [35]. Similarly, imatinib, but not bosutinib, targets DDR1 and DDR2, which play important roles in several human disorders, not just in cancers [36]. Bosutinib was originally developed to inhibit ABL1 and SRC more

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

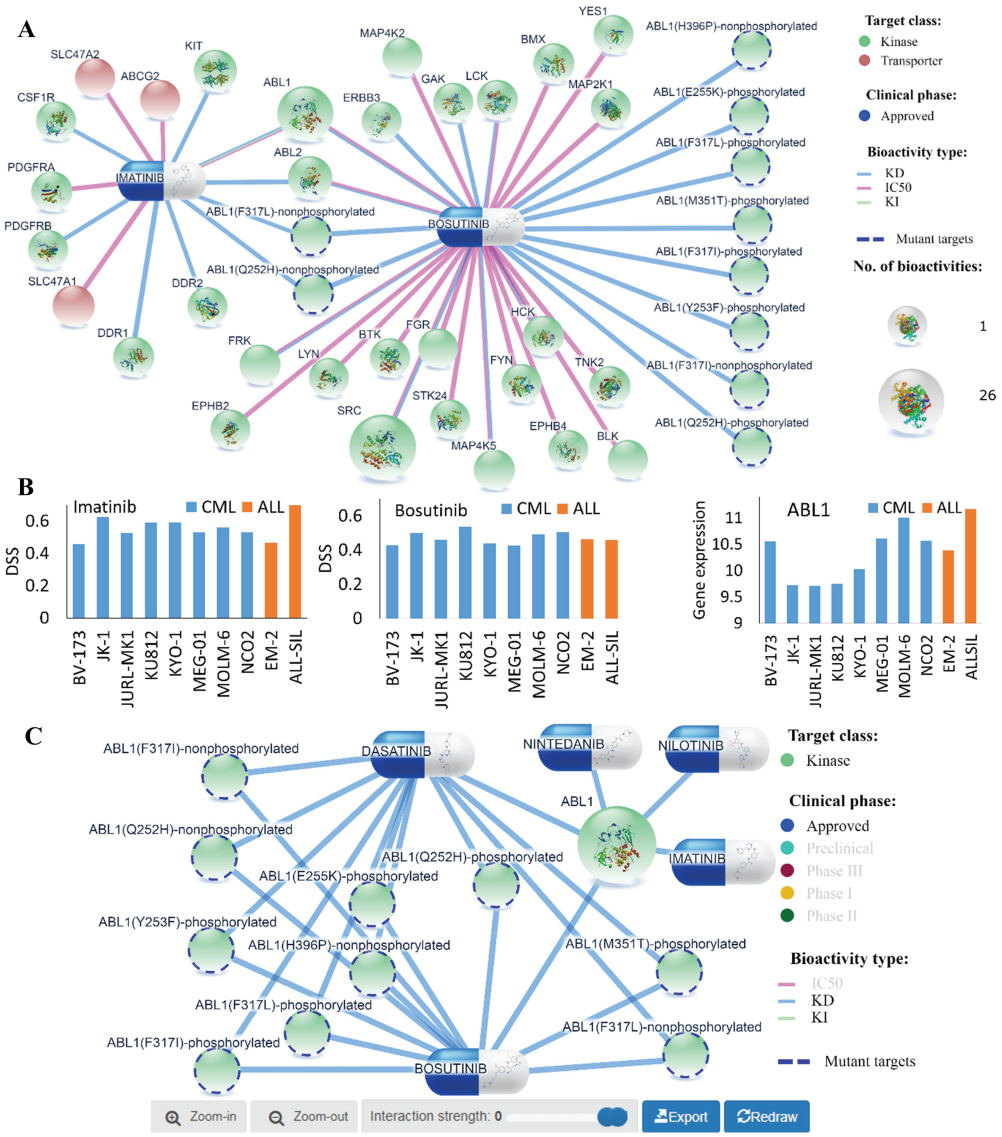

Figure 2. **(A)** Integrated network visualization generated for bosutinib and imatinib using an interaction score threshold of ≥0.9 in DTP. The thickness of the line indicates the interaction score between the compound and target (ranging between 0 and 1, where larger values indicate higher target potency). **(B)** Normalized DSS for the top 10 cancer cell lines that have the highest sensitivity to both bosutinib and imatinib (integrated data from PharmacoDB). The right-most plot shows expression levels of ABL1 across eight CML and two BCR-ABL1 (Philadelphia chromosome) positive ALL cell lines (data integrated from CCLE, GDSC and gCSI). **(C)** Target-centric visualization for ABL1 kinase comprising of interacting compounds with an interaction score of ≥0.9 (ABL1 target network with interaction score cut-off of 0.6 and with multiple bioactivity types is shown in Supplementary Data Figure S2 for comparison).

effectively than imatinib, yet avoid modulating PDGFR, as imatinib does [37]. These interactions are also evident from the network of Figure 2A. Furthermore, comprehensive kinome-wide profiling and target-space comparison of eight ABL1 inhibitors underscores the greater capacity of bosutinib to inhibit several receptor tyrosine kinases that trigger MAPK pathway signaling more potently than imatinib, also evident in Figure 2A. This at least partly contributes to the ability of bosutinib to interfere with MAPK signaling-driven diseases [38], leading to further extension in its potential application space.

In addition to providing compound-centric views for multi-targeted compounds, DTP can also be used from a target-centric perspective. As an example, we generated a target-based network visualization for ABL1 kinase, which is the intended primary target for both imatinib and bosutinib (Figure 2A). At a relatively stringent interaction strength cut-off of 0.9, the network visualization reveals already a highly interconnected system of compounds and their various off-targets (Figure 2C), which can be useful in designing pre-clinical experiments that explore the roles of compounds in complex multi-target systems, for instance, in living cells. Decreasing the interaction cut-off to 0.6 increases the size and complexity of the ABL1 network and includes also the activities of imatinib, axitinib, crizotinib and nilotinib across ABL1 mutations (Supplementary Data Figure S2). As an additional example of target-centric visualization, we generated a combined interaction network for two kinases, ABL1 and KIT, with maximal interaction cut-off (Supplementary Data Figure S3), highlighting broad polypharmacological effects of two multi-targeting drugs (dasatinib and nintedanib) on these and other kinase targets.

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

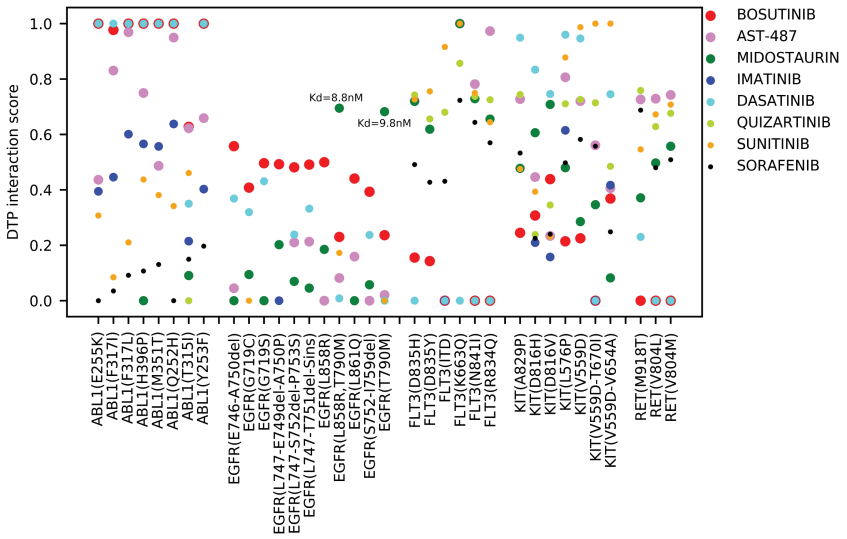

Figure 3. Activity of select clinically relevant well-annotated multi-kinase inhibitors across unrelated, but frequently, mutated targets in hematological and solid cancers (the *x*-axis). The interaction score (*y*-axis) was calculated as median across various assays (midostaurin K_d bioactivities shown as examples) and exported using the data download options in DTP. The interaction overview allows an easy and user-friendly perusal and summarization of a vast number of bioactivity data points that would be difficult to parse manually.

Exploration of compound interactions with oncogenic mutations

A core utility of DTP is to assist users in parsing drugs and targets that are actively being pursed in clinical trials and oncological practice. By systematically comparing multiple such drugs across mutated targets, experimental observations may be easier to explain and new hypotheses can become more refined. Based on its integrated features, DTP facilitates the discovery and summarization of multiple findings that would be otherwise easily missed when focusing on individual studies at a time. To highlight these features, we used the DTP search options, interaction summary score and output data export to compile a summary of the activity of several clinically relevant and well-annotated multi-kinase inhibitors on different but frequently encountered oncogenic mutations in both solid and hematological cancers (Figure 3). This application showcases the utility of DTP in providing an overview of both drugs and targets to guide interpretation of the target profiles and potential follow-up experimental design.

For example, midostaurin, which was approved in 2017 for FLT3-mutated AML [39], displays a measurable activity against a gatekeeper mutation T790M in EGFR (K_d values highlighted in Figure 3), hence supporting the testing of midostaurin for EGFR-gatekeeper-mutated cancers such as Non-Small Cell Lung Cancer (NSCLC) [40]. In contrast, another multi-kinase/FLT3-inhibitor, AST-487, does not show similar activity, arguing against development of this compound for similar indications, rather supporting its use as an FLT3/multi-kinase but EGFR(T790M)-non-targeting negative control during follow-up experimental testing. While the DTP interaction score calculated as median across various assays do not alone guarantee activity in patient samples or animal models, the score provides a useful comparison of compound potencies to facilitate a more careful scrutiny on a case-by-case basis, i.e. hypothesis generation for precision oncology drug development. The adverse effect profiles available from the clinical trial data further support these developments.

Correlation of compound combinations targeting oncogenic mutations

An extended utility of data in DTP is to focus on compounds used in combinations or in comparison to new agents in clinical trials and then to compare their relative activity against mutant targets. As case examples, we used four pairs of targeted kinase inhibitors that have been explored in clinical trials and compared their DTP interaction scores with a set of annotated oncogenic mutations across four kinases (Figure 4). In these clinical trials, imatinib in combination with nilotinib was tested in clinical study for CML, in which ABL1 mutations act as drivers in most cases (NCT01819389); EGFR-inhibitor erlotinib was combined with sunitinib for both NSCLC (NCT00581789) and kidney cancer (NCT00425386); and erlotinib was used sequentially with gefitinib for NSCLC (NCT02747953).

Overall, the high correlation between the interaction scores of these agents with the target mutations known to confer resistance to the treatments indicates that none of these clinical trials were designed to overcome mutations that typically accrue during treatment. For example, the resistance-conferring T790M mutation of EGFR almost always appears in cells already harboring activating EGFR mutations, such as L858R, but neither erlotinib nor gefitinib are potent against EGFR(T790M), and do not significantly differ from each other in potency against other EGFR mutations either. Instead, the combination studies highlighted in Figure 4 have typically been designed to give drugs in sequence or in lower individual doses to minimize side effects, as well as to target other potentially synergizing oncogenic processes.

For the nilotinib and imatinib combination (the top-left panel in Figure 4), unsurprisingly the overall correlation was strong, indicating that these two drugs are more or less able to bind to and inhibit the same targets (Supplementary Data Figure S2). However, even if the correlation is strong, there remains unique selectivity for each of the compounds. For example, nilotinib is not as potent as imatinib in targeting the double-mutated KIT receptor (V559D and V654A) [41], and, vice versa, imatinib

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

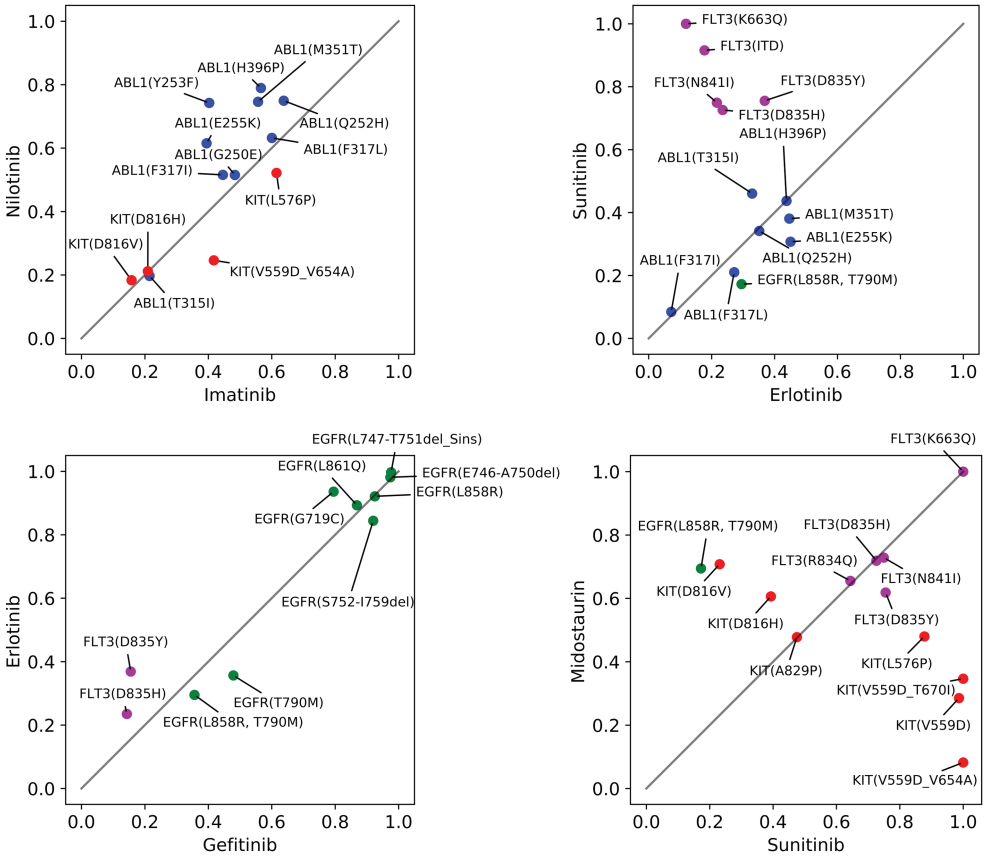

Figure 4. Correlations of the DTP interaction scores between drug pairs used in combination (in completed clinical trials) for target mutations across four kinases (color-coded separately). The individual points in the scatterplot correspond to median DTP interaction score calculated across both on- and off-mutant targets between the drug pairs (x- and y-axes). High correlation of DTP scores across certain kinase mutations conferring resistance to one of the agents shows that the mutation is also protective against the other kinase inhibitors.

does not inhibit ABL1 (Y253F) as potently as nilotinib [42]. The sunitinib and erlotinib combination (the top-right panel) shows less correlation since erlotinib is not as strong inhibitor of FLT3 as sunitinib, instead erlotinib was developed as an EGFR inhibitor. The same applies also for gefitinib (the lower-left panel of Figure 4).

Comparing the two top examples focusing on ABL1, neither sunitinib nor erlotinib appears as potent as nilotinib or imatinib in inhibiting ABL1 mutants. Lastly, the two broad kinase inhibitors, midostaurin and sunitinib (the lower-right panel of Figure 4), differ decisively from each other in targeting particular KIT mutations [41]. This is translationally relevant, as KIT mutations are found in FLT3-mutated AML, for which midostaurin was recently approved. Overall, this case study underscores the utility of DTP, and especially its interaction score, to provide convenient access and visualization of compound activity against driver and select resistance-conferring mutations, which may help to design combinatorial therapies targeting one or more oncogenic pathways that may lead to unexpected synergistic effects [43].

Discussion

Compound–target interaction networks enable systematic analysis of the compound's inhibition potency across both its intended primary 'on-targets' as well as secondary 'off-targets'. By further comparing the interaction potencies with the compound sensitivity profiles as well as gene expression and mutation patterns across multiple cell lines, one can reveal mechanistic and predictive insights into context-specific MoA of molecularly-targeted agents [44]. However, software tools to help in such integrated analyses using network-centric approach have been lacking. To bridge this gap, we implemented DTP for interactive exploration of DTI networks, built on top of the harmonized drug target and response profiling data from DTC and PharmacoDB, respectively, and with cross-references to multiple compound and target databases (Table 1). We demonstrated in the selected use cases how DTP can help the end-users to better understand the polypharmacological effects of the multi-targeting drugs in precision oncology and drug repurposing applications.

Given its wide range of features supporting multiple use cases (Figure 1), we believe that DTP will become a broadly used tool for many exciting applications. For instance, DTP implements tools for both (i) target-based drug repurposing applications (e.g. identification of candidate compounds that selectively inhibit a particular disease- or resistance-related target) and (ii) phenotype-based drug discovery applications (i.e. mapping the efficacy target space of a given drug molecule or probe). The web tool is designed so that it is usable also by researchers without bioinformatics skills, hence supporting its wide adoption by broad user communities, including chemical biologists and translational researchers. Finally, the visual analytic tools based on comprehensive bioactivity data should prove useful

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

also for the computational biologists in the process of extending the target space of compounds, for instance, by implementing machine-learning models for guiding compound–target mapping efforts [45–47]. The rich information in DTP and its visualization options should help in model training and follow-up prioritization of most potent compound–target interactions for further experimental evaluation.

Current limitations

The current DTP version is especially focused on kinase inhibitors, due to their clinical importance [48], but the next versions will be extended to a wider spectrum of drug and target classes. We are also planning to extend the DTP web platform with additional features to support the exploration of massive search space of potential drug combinations for personalized combinatorial treatment predictions [49, 50]. In the next release of DTP, we will integrate also target pathway information to give further insights into the MoA of multi-targeting agents, for instance, the biological role of gene expression and signal transmissions of the therapeutics targets. Finally, as the DTP database and source code are freely available, we encourage the community to further modify and extend the current implementation and potential application areas, such as considering also interactions between small molecules and non-coding RNAs or microRNAs [51, 52].

As with any computational interaction summarization approach, we note that there may be cases that result in an apparent discrepancy between a compound's interaction score and its actual biological effects. As seen in Figure 2A, for instance, while bosutinib binds ABL1 with greater potency than imatinib, it also binds to a greater number of other targets at concentrations where ABL1 is inhibited compared to imatinib. Correspondingly, bosutinib elicits stronger side effects than imatinib in clinical use, which at least partly can be attributed to these off-target activities. It is therefore important to recognize that DTB interaction scores, similar to binding affinities, do not reflect clinical effectiveness or toxicities and should instead be used for exploring polypharmacological effects of compounds, as guidelines for systematic positioning of the compounds of interest into concentration categories relevant for the majority of targeted agents annotated in the DTP database.

Key Points

- Integration of multiple open-data drug and target resources are required to provide comprehensive, network-centric information for the modeling of the mode-of-action of multi-targeting drugs and for predicting their therapeutic and side effect profiles computationally.
- Systems-wide and harmonized target profiles across the drug's on- and off-targets as well as among disease- or response-related mutant targets capture the full spectrum of target potencies, which is critical for multiple drug repurposing and precision oncology applications.
- Minimal yet sufficient bioactivity assay annotation provides reproducible target profiles, making it possible to calculate an integrated interaction score for compound–target pairs in DTP that can be used for exploring relationships between compounds with shared target profiles.
- Linking the on- and off-target potencies with compound sensitivity and target expression profiles across cancer cell lines provides researchers with an improved understanding of the context-specific mechanisms behind selective polypharmacological effects of the multi-targeting agents.

Supplementary Data

Supplementary data are available online at https://academic.oup.com/bib.

Funding

European Union's Horizon 2020 Research and Innovation Program 2014–2020 (634143); Cancer Society of Finland, Sigrid Jusélius Foundation and Academy of Finland (277293, 292611, 295504, 310507 and 313267).

References

1. Santos R, Ursu O, Gaulton A, *et al*. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* 2017;**16**:19–34.
2. Arrowsmith CH, Audia JE, Austin C, *et al*. The promise and peril of chemical probes. *Nat Chem Biol* 2015;**11**:536–41.
3. Bento AP, Gaulton A, Hersey A, *et al*. The ChEMBL bioactivity database: an update. *Nucleic Acids Res* 2014;**42**:D1083–90.
4. Gilson MK, Liu T, Baitaluk M, *et al*. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res* 2016;**44**:D1045–53.
5. Wang Y, Bryant SH, Cheng T, *et al*. PubChem BioAssay: 2017 update. *Nucleic Acids Res* 2016;**45**:D955–63.
6. Tang J, Tanoli ZU, Ravikumar B, *et al*. Drug Target Commons: a community effort to build a consensus knowledge base for drug–target interactions. *Cell Chem Biol* 2018;**25**:224–9.
7. Tanoli Z, Alam Z, Vähä-Koskela M, *et al*. Drug Target Commons 2.0: a community platform for systematic analysis of drug–target interaction profiles. *Database (Oxford)* 2018;**2018**:1–13.
8. Barretina J, Caponigro G, Stransky N, *et al*. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;**483**:603.
9. Iorio F, Knijnenburg TA, Vis DJ, *et al*. A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;**166**:740–54.
10. Seashore-Ludlow B, Rees MG, Cheah JH, *et al*. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discov* 2015;**5**:1210–23.
11. Smirnov P, Kofia V, Maru A, *et al*. PharmacoDB: an integrative database for mining *in vitro* anticancer drug screening studies. *Nucleic Acids Res* 2018;**46**:D994–D1002.
12. Wang C, Kurgan L. Review and comparative assessment of similarity-based methods for prediction of drug–protein interactions in the druggable human proteome. *Brief Bioinform* 2018.
13. Rifaioglu AS, Atas H, Martin MJ, *et al*. Recent applications of deep learning and machine intelligence on *in silico* drug discovery: methods, tools and databases. *Brief Bioinform* 2018.

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019

14. Ezzat A, Wu M, Li XL, *et al*. Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey. *Brief Bioinform* 2018.
15. Sam E, Athri P. Web-based drug repurposing tools: a survey. *Brief Bioinform* 2017.
16. Fang J, Liu C, Wang Q, *et al*. *In silico* polypharmacology of natural products. *Brief Bioinform* 2017.
17. Lotfi Shahreza M, Ghadiri N, Mousavi SR, *et al*. A review of network-based approaches to drug repositioning. *Brief Bioinform* 2017;**19**:878–92.
18. Chen X, Yan CC, Zhang X, *et al*. Drug–target interaction prediction: databases, web servers and computational models. *Brief Bioinform* 2015;**17**:696–712.
19. Wu Z, Cheng F, Li J, *et al*. SDTNBI: an integrated network and chemoinformatics tool for systematic prediction of drug–target interactions and drug repositioning. *Brief Bioinform* 2016;**18**:333–47.
20. Wang L, You ZH, Chen X, *et al*. RFDT: a rotation forest-based predictor for predicting drug–target interactions using drug structure and protein sequence information. *Curr Protein Pept Sci* 2018;**19**:445–54.
21. Schneider L, Stöckel D, Kehl T, *et al*. DrugTargetInspector: an assistance tool for patient treatment stratification. *Int J Cancer* 2016;**138**:1765–76.
22. Shameer K, Glicksberg BS, Hodos R, *et al*. Systematic analyses of drugs and disease indications in RepurposeDB reveal pharmacological, biological and epidemiological factors influencing drug repositioning. *Brief Bioinform* 2017;**19**:656–78.
23. Allaway RJ, La Rosa S, Guinney J, *et al*. Probing the chemical–biological relationship space with the Drug Target Explorer. *Journal of Cheminformatics* 2018;**10**:41.
24. Szklarczyk D, Santos A, von Mering C, *et al*. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* 2016;**44**:D380–4.
25. Klijn C, Durinck S, Stawiski EW, *et al*. A comprehensive transcriptional portrait of human cancer cell lines. *Nat Biotechnol* 2015;**33**:306–12.
26. Haibe-Kains B, El-Hachem N, Birkbak NJ, *et al*. Inconsistency in large pharmacogenomic studies. *Nature* 2013;**504**:389–93.
27. Piñero J, Bravo À, Queralt-Rosinach N, *et al*. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res* 2017;**45**:D833–9.
28. Chambers J, Davies M, Gaulton A, *et al*. UniChem: a unified chemical structure cross-referencing and identifier tracking system. *J Cheminform* 2013;**5**:5.
29. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res* 2014;**43**:D204–12.
30. Gutmanas A, Alhroub Y, Battle GM, *et al*. PDBe: protein data bank in Europe. *Nucleic Acids Res* 2014;**42**:D285–91.
31. Gaulton A, Hersey A, Nowotka M, *et al*. The ChEMBL database in 2017. *Nucleic Acids Res* 2016;**45**:D945–54.
32. Mi H, Huang X, Muruganujan A, *et al*. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* 2016;**45**:D183–9.
33. Finn RD, Coggill P, Eberhardt RY, *et al*. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 2015;**44**:D279–85.
34. Rix LLR, Rix U, Colinge J, *et al*. Global target profile of the kinase inhibitor bosutinib in primary chronic myeloid leukemia cells. *Leukemia* 2009;**23**:477–85.
35. Chase A, Schultheis B, Kreil S, *et al*. Imatinib sensitivity as a consequence of a CSF1R-Y571D mutation and CSF1/CSF1R signaling abnormalities in the cell line GDM1. *Leukemia* 2009;**23**:358.
36. Day E, Waters B, Spiegel K, *et al*. Inhibition of collagen-induced discoidin domain receptor 1 and 2 activation by imatinib, nilotinib and dasatinib. *Eur J Pharmacol* 2008;**599**:44–53.
37. Konig H, Holyoake TL, Bhatia R. Effective and selective inhibition of chronic myeloid leukemia primitive hematopoietic progenitors by the dual Src/Abl kinase inhibitor SKI-606. *Blood* 2008;**111**:2329–38.
38. Winter GE, Rix U, Carlson SM, *et al*. Systems-pharmacology dissection of a drug synergy in imatinib-resistant CML. *Nat Chem Biol* 2012;**8**:905.
39. Levis M. Midostaurin approved for FLT3-mutated AML. *Blood* 2017;**129**:3403–6.
40. Lee HJ, Schaefer G, Heffron TP, *et al*. Noncovalent wild-type-sparing inhibitors of EGFR T790M. *Cancer Discov* 2013;**3**:168–81.
41. Davis MI, Hunt JP, Herrgard S, *et al*. Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotechnol* 2011;**29**:1046–51.
42. Ren X, Pan X, Zhang Z, *et al*. Identification of GZD824 as an orally bioavailable inhibitor that targets phosphorylated and nonphosphorylated breakpoint cluster region–Abelson (Bcr–Abl) kinase and overcomes clinically acquired mutation-induced resistance against imatinib. *J Med Chem* 2013;**56**:879–94.
43. Tang J, Karhinen L, Xu T, *et al*. Target inhibition networks: predicting selective combinations of druggable targets to block cancer survival pathways. *PLoS Comput Biol* 2013;**9**:e1003226.
44. Rees MG, Seashore-Ludlow B, Cheah JH, *et al*. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* 2016;**12**:109.
45. Pahikkala T, Airola A, Pietilä S, *et al*. Toward more realistic drug–target interaction predictions. *Brief Bioinform* 2014;**16**:325–37.
46. Cichonska A, Ravikumar B, Parri E, *et al*. Computational–experimental approach to drug–target interaction mapping: a case study on kinase inhibitors. *PLoS Comput Biol* 2017;**13**:e1005678.
47. Yamanishi Y, Araki M, Gutteridge A, *et al*. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* 2008;**24**:i232–40.
48. Elkins JM, Fedele V, Szklarz M, *et al*. Comprehensive characterization of the published kinase inhibitor set. *Nat Biotechnol* 2016;**34**:95.
49. He L, Tang J, Andersson EI, *et al*. Patient-customized drug combination prediction and testing for T-cell prolymphocytic leukemia patients. *Cancer Res* 2018;**78**:2407–18.
50. Madani Tonekaboni SA, Soltan Ghoraie L, Manem VSK, *et al*. Predictive approaches for drug combination discovery in cancer. *Brief Bioinform* 2016;**19**:263–76.
51. Qu J, Chen X, Sun Y-Z, *et al*. Inferring potential small molecule–miRNA association based on triple layer heterogeneous network. *J Cheminform* 2018;**10**:30.
52. Chen X, Sun Y-Z, Zhang D-H, *et al*. NRDTD: a database for clinically or experimentally supported non-coding RNAs and drug targets associations. *Database (Oxford)* 2017;**2017**:bax057.

Downloaded from https://academic.oup.com/bib/advance-article-abstract/doi/10.1093/bib/bby119/5232987 by Turun Yliopiston Kirjasto user on 13 February 2019