Improving the efficacy-safety balance of polypharmacology in multi-target drug discovery

Abstract

Introduction: Polypharmacology has emerged as an essential paradigm for modern drug discovery process. Multiple lines of evidence suggest that agents capable of modulating multiple targets in a selective manner may offer also improved balance between therapeutic efficacy and safety compared to single-targeted agents. We review here the recent progress made in experimental and computational strategies for addressing the critical challenges with rational discovery of selective multi-targeted agents within the context of polypharmacological modelling.

Areas covered: Specific focus is placed on multi-targeted mono-therapies, although examples of combinatorial polytherapies are also covered as an important part of the current polypharmacology paradigm. We focus mainly on anti-cancer treatment applications, where polypharmacology is playing a key role in determining the efficacy-toxicity trade-off of multi-targeting strategies. Specific focus areas of recent developments cover both experimental *in-vitro* profiling strategies and screening libraries, as well as computational *in-silico* approaches (target-centric, compound-centric and network-centric), along with related databases and web-applications that support polypharmacological modelling in the multi-target drug discovery or drug repurposing applications.

Expert opinion Even though it is widely appreciated that complex polypharmacological interactions can contribute both to therapeutic and adverse side-effects, systematic approaches for improving this balance by means of integrated experimental-computational strategies are still lacking. Future developments will be needed for comprehensive collection and harmonization of systems-wide target selectivity data, enabling better utilization and control for multi-targeted activities in the drug development process. Additional areas of future developments include model-based strategies for drug combination screening, and improved pre-clinical validation options for the model predictions, before entering into costly clinical development phases.

Article highlights box

- Polypharmacological approaches are revolutionizing the drug discovery process, but successful applications of such multi-targeting approaches lie in the selective trade-off between the efficacy-toxicity ratio exhibited by polypharmacological agents.
- Novel screening strategies based on chemogenomic and combinatorial drug libraries are critical for rational drug discovery and repurposing approaches, challenging the notion of serendipity in identifying polypharmacological agents.
- Data-driven computational frameworks spanning a wide array of approaches (compound-centric, target-centric and network-centric) are imperative in predicting the various efficacy and toxicity-related outcomes of multi-targeted drugs.
- To best support the scientific community, the computational models and current and emerging knowledge of compound-target interactions should be made freely available and implemented as easy-to-use web-applications and comprehensive databases.
- Although anti-cancer treatment applications are currently leading the way in the use of polypharmacology approaches, also other complex disease with genetic and epigenetic heterogeneity should benefits from multi-targeting drug treatment strategies.

Keywords: polypharmacology, efficacy-toxicity ratio, drug repositioning, multi-target drug design, profiling strategies, screening libraries, drug combination therapies, computational models, data resources, web-tools.

1. Introduction

As evidenced by the increasing number of publications in the recent years, current trends in drug discovery are witnessing a steady paradigm shift from the ideology of selective agents (‘magic bullets’) toward selectively promiscuous multi-targeting drugs (MTDs, or so-called

'magic shotguns') [1, 2]. Polypharmacology-based 'one drug, multiple targets' paradigm has surpassed the time and cost constrains encountered when following the conventional *de novo* approaches for drug discovery [3, 4]. As detailed by Peters [5], polypharmacological effect of a compound is determined by the trade-off between the compound's selective activity toward therapeutic targets and its promiscuous activity toward non-therapeutic targets, which dictates factors related to the efficacy, safety and adverse toxic effects of the agent. Formally, polypharmacology is the ability of a single drug molecule to exhibit an acceptable degree of specificity towards multiple protein targets, present either in a single or multiple disease pathways. In cases of complex disease mechanisms, such as cancers, the formal definition of polypharmacology has also been extended to drug combinatorial strategies, wherein multiple selective drugs act on distinct targets with cumulative effects on physiological processes such as proliferation, angiogenesis and cell-cell interactions [5, 6]. In addition to providing synergistic inhibition of distinct targets in complex disease networks, combinatorial therapies can also reduce the toxic effects of individual drug molecules by decreasing the therapeutic dosage required for the treatment effect [7].

Investigations of the bioactivity profiles of compound collections in publicly available data resources have shown that it is not at all uncommon for drug molecules to interact with multiple targets [8, 9]. Relatively often such promiscuous interactions can be attributed to unintended targets, or 'anti-targets', which results in adverse drug reactions (ADRs) [10]. One such example is fenfluramine, a widely-used anorexigen, which exhibit unintended interactions with non-therapeutic target, serotonin 5-HT_{2B} receptor, leading to toxic side-effects. Pre-clinical testing for such inadvertent interactions and ADRs is often performed with safety screening panels, which facilitate the screening of candidate drug compounds across numerous safety-relevant targets, hence serving as initial filter to disregard compounds with undesirable side-

effects. Computational tools can also help to identify non-selective compounds, for instance, Pan Assay Interference Compounds (PAINS) includes a number of substructure features for promiscuity obtained from various high-throughput screening experiments, which thereby can be used for the removal of cross-interference compounds prior to biochemical screening experiments [11].

Besides leading to ADRs, the intriguing feature of polypharmacological interactions is the possibility of the unintended targets to be sufficiently selective and therapeutically beneficial [12]. Polypharmacological treatment approaches thereby explore such therapeutic effects of agents that lead to improved balance between efficacy and safety compared to single-targeted agents. A particularly exciting possibility of multi-targeted agents is to rationally discover selective off-targets for an already approved drug molecule, resulting in the development of drug repurposing strategies [13, 14]. When compared to the traditional *de novo* designing of drug molecules, repurposing or repositioning strategies are much faster and reasonably risk-free. It has also been argued that polypharmacological drugs with high enough specificity towards multiple therapeutic targets exhibit selectively synergy and are less liable for toxic side-effects. One famous example of polypharmacological drugs includes multi-targeting kinase inhibitors [15]. Most of the currently-approved kinase inhibitors for cancer treatment target multiple protein kinases at a varying degree of specificity. Although initially designed to be selective, most ATP-competitive kinase inhibitors have a broad spectrum of off-target activities owing to the conserved ATP-binding pocket across the kinase superfamily. Currently, such multi-targeting treatments are facilitated either through single-drug mono-therapy (for example, sunitinib is a multi-targeted tyrosine kinase inhibitor potent towards VEGFR1, VEGFR2, VEGFR3, KIT, PDGFRa and PDGFRb resulting in both antitumor and antiangiogenic activities [16]) or through selective drug combination polytherapy (e.g., one surprising combination

strategy combines an anti-fungal agent itraconazole that inhibits the activation of VEGFR2 and FGFR3 with an approved chemotherapy drug, pemetrexed, for the treatment of non-squamous non-small-cell lung cancer [17]).

Until recently, endeavors to identify potential polypharmacological drugs have traditionally required a profound understanding of the compounds' bioactivity target profiles and mechanism of action (MoA). Examples of drugs such as sildenafil and aspirin suggests that many novel indications of drugs were mainly discovered by the virtue of serendipity during a biochemical screening process, rather than by rational discovery process. However, the recent advancements in high-throughput screening (HTS) and multi-phenotypic screening strategies, in concordance with the development of novel computational and bioinformatics approaches, have challenged the previous notion of accidental discoveries. These emerging efforts are augmented with the availability of copious data resources of compounds bioactivity profiles, combination therapies, novel compound screening libraries and drug response-related -omics data, which has resulted in the development and implementations of streamlined approaches to systematically identify polypharmacological drugs. We argue that these pursuits are revolutionizing the drug discovery process and will become a growing area of interest for the drug discovery community.

The nature of efficacy-toxicity duality engraved in polypharmacological drug discovery and the practical benefits gained through drug-repurposing have been extensively described in previous reviews [18, 19]. In this review, we instead focus on the recent advancements made in experimental *in-vitro* profiling strategies and screening libraries, as well as in computational *in-silico* approaches (target-centric, compound-centric and network-centric), along with related databases and web-applications that can be used for addressing the selectivity vs. non-selectivity balance of polypharmacology in multi-target drug discovery applications (**Figure 1**).

2. Experimental screening strategies for polypharmacology

In-vitro profiling of compounds for efficacy and ADR's has become an indispensable part of any drug discovery process. Profiling for ADR's typically consists of targets spanning across various protein families including G-protein couple receptors (GPCR's), nuclear receptors, enzymes and ion channels. Such broad-spectrum targets are prone to have a high hit rate in screening studies, and the collaborative efforts between AstraZeneca, Pfizer, Novartis and GlaxoSmithKline, led to standardization of 44 such targets that are deemed as the 'minimal safety screening panel' [20]. Although such safety screening panels aid in excluding non-specific or cross-inference compounds in the early stages of drug discovery process, the limited target profile hinders more systematic approaches to deconvoluting the compound's MoA or ascertaining its effect on the druggable proteome. In the case of promiscuous kinase inhibitors, for instance, more comprehensive *in-vitro* profiling of kinase inhibitors has shown that the cytotoxic effects of certain inhibitors can be attributed to their cross-reactive binding to non-kinase targets [21]. As an example, tivantinib is a non-ATP competitive tyrosine kinase MET inhibitor, currently in phase III clinical trial for the treatment of hepatocellular carcinoma (HCC). Cell-based assay to deconvolute tivantinib's cellular activity has shown that tivantinib elicits cytotoxic effects even in cells independent of MET, and that such effects are likely explained by its activity towards impairing microtubule dynamics [22].

In complex diseases like cancer, treatment outcomes are often linked to point mutations resulting in either therapeutic efficacy or drug resistance. In particular, identifying potent inhibitors to confront acquired drug resistance still remains a huge clinical challenge. An illustrative case study that highlights both the adverse and therapeutic polypharmacological effects encountered when evading such resistance mechanisms involve the drugs ponatinib and

axitinib in the treatment of BCR-ABL(T315I) driven chronic myeloid leukemia (CML). This particular point mutation confers resistance to all currently-approved ABL1 inhibitors except ponatinib. Ponatinib is a potent BCR-ABL(T315I) small molecule inhibitor that was temporarily withdrawn from the market owing to its adverse off-target vascular effects [23]. On the other hand, axitinib was designed as a highly selective VEGFR inhibitor, but through a recent cell-based profiling study, it was found to have a therapeutic off-target potency also towards ABL1(T315I) mutant cells, compared to their wild-type counterpart [24]. Axitinib is currently undergoing a phase II drug repurposing clinical trial for T315I-driven CML (NCT02782403).

In-vivo profiling of compounds' activity is not as amenable as *in-vitro* phenotypic screening, as screening of extensive compound libraries in *in-vivo* models is often economically unfeasible and extremely time consuming. Circumventing these pre-clinical limitations has led to the emergence and development of efficient chemogenomic libraries. Apart from elucidating the MoA of compounds, a hit from a chemogenomic screening sets can also aid in identifying targets for pharmacological modulations [25, 26]. Chemogenomic libraries such as Mechanism Interrogation PlatE library (MIPE), Sigma-Aldrich Library of Pharmacologically Active Compounds (LOPAC^{1280}) and Mixture-Based Synthetic Combinatorial library have been meticulously designed to address specific requirements during the HTS process. For example, compound collections in the MIPE library include compounds that are either advanced clinical or pre-clinical agents, making them suitable candidates for drug-repurposing studies [27]. Similarly, Mixture-Based Synthetic Combinatorial libraries are generated through the systematic arrangements of a mixture of synthetic compounds, and are shown to be ideally suited for identifying synergistic effects among compounds. *In-vivo* testing after such chemogenomic screening has shown to reduce the currently high drug attrition rates [28]. A

summary of the various chemogenomic libraries useful for polypharmacological studies are provided in **Table 1**.

A decade ago, Wermuth proposed an innovative screening strategy, called Selective Optimization of Side Activities (SOSA), as an alternative to HTS for drug repositioning [29]. SOSA capitalizes on the side-effects of compounds tested by screening a significantly smaller panel of structurally and therapeutically diverse compounds across wide spectrum of pharmacological targets. Hits from the screens are then later rationally optimized such that the observed side effects of the compounds are transformed to the intended primary effect and by eliminating other pharmacological activity. Varbanov et al. recently implemented a similar screening strategy to repurpose drugs approved for other diseases in the treatment of lung and pancreatic cancers [30].

Drug combination therapies offer the potential for more effective and sustainable clinical outcomes, and thereby combinatorial modalities have emerged as an exciting alternative to monotherapies especially for complex diseases. Although monotherapy profiling has streamlined the identification of distinct therapeutic (or toxicity) targets (so-called target deconvolution), improved responses in combination therapy can be attained through screening of a set of drug combinations selected either rationally by considering individual drugs' MoA, or as in majority of cases traditionally, through trial-and-error [31]. In comparison to the effect observed when using individual drug molecules alone, the polypharmacological outcomes of drug combination screens can be categorized as antagonistic, additive or synergistic effects [32]. Dose-adjusted combination screens have become an imperative approach especially for cancer treatment in overcoming issues of drug resistance encountered during therapy [7]. For example, trametinib, an FDA approved MEK inhibitor is being used in combination with navitoclax, a

BCL2-family inhibitor, for the treatment of advanced or metastatic solid-tumors [33]. This combination is currently in phase Ib/II clinical trial for patients with *KRAS* mutation-positive advanced lung adenocarcinomas. Recently, combination therapies have been extended from the conventional pairwise combinations to higher-order drug combinations. For instance, Horn et al. introduced a systematic screening approach and identified novel higher-order drug combinations effective in killing drug-resistant colorectal cancer cells [34].

The above-mentioned pharmacological profiling strategies have led to a plethora of high-throughput and low-throughput screening initiatives, including both target-based and phenotypic screening efforts. Examples of cancer-related efforts include the kinase-specific profiling studies, carried out by Davis et al. [35] and Metz et al. [35, 36], and the extensive cell-based pharmacogenetic screens, such as Cancer Cell Line Encyclopedia (CCLE), Genomics of Drug Sensitivity in Cancer (GDSC) and Cancer Therapeutics Response Portal (CTRP) [37-39]. These pharmacological profiling datasets, especially when coupled with other -omics related datasets, provide an invaluable data resource for various in-silico models developed for polypharmacological investigations.

3. Computational *in-silico* approaches for polypharmacology

A large array of computational methodologies has been designed and implemented to address both the adverse and therapeutic aspects of polypharmacology. These methodologies are predominantly data-driven and stem from the prior knowledge of compounds, targets and their interactions. The existing information comprises both the 2D or 3D structure of compounds and targets, the protein-protein interaction networks among targets, the available activity profiles among drugs and targets, along with the gene expression, copy number variation and other -omics datasets related to the drugs' MoA. These data resources have assisted in the

development of various *in-silico* approaches spanning a wide array of areas such as unique virtual screening strategies for polypharmacology, innovative chemoinformatic and statistical approaches for adverse and therapeutic polypharmacological effects, supervised machine-learning algorithms for predicting drug-target interactions, and graph theory models aiding in network pharmacology (reviewed below).

These computational methodologies can be broadly categorized as target-centric, compound/ligand-centric or network pharmacological approaches. Such approaches have been successfully employed in elucidating ADRs by predicting compounds' off-targets and ADME properties [40]. Machine-learning models, both binary classification and regression algorithms, have been deployed to predict novel compound-target interactions [41]. Network models enable analysis of complex interaction networks between drugs and their targets, with the aim to aid in identifying functional modules, elucidating compounds MoA and target deconvolution phases [42, 43]. Statistical models that make use of data integration framework have been effectively applied both to predict novel drug indications and to identify potent drug combination strategies [44]. Computational–experimental frameworks have also been implemented for systematical design of multi-targeted drugs [45]. Excellent reviews detailing the application of the various computational approaches in drug repositioning are available [19, 46, 47].

3.1 Target-centric approaches

In-silico target-centric approaches are based on the underlying hypothesis that proteins that are structurally similar are likely to exhibit similar selectivity properties and hence are expected to bind to similar compounds. Despite their wide and successful applications in drug discovery, target-centric protocols are drastically limited by the availability of the X-ray crystal or nuclear magnetic resonance (NMR) spectroscopy structure of targets, thereby making the Protein Data

Bank (PDB) as an invaluable data repository for target-centric studies. Established methods for target-centric approaches for polypharmacological modelling can be broadly categorized into two categories; the first one is based on algorithms that estimate binding pocket similarity of multiple targets [14], and the other uses molecular docking strategies that facilitate *in-silico* screening of drugs to multiple targets [48].

Binding pocket similarity estimation between two targets is often carried out by local structural alignment of targets and calculating the geometrical distance between residues (C-alpha or centers of functional groups) of the superimposed proteins. These binding pockets are either predefined or identified through sequence alignment and structure refinement strategies, and computational tools that enables binding site predictions are available including COFACTOR, FINDSITE and Concavity [49]. Structural alignment algorithms provide a discrete representation of the targets' binding site as geometric patterns or numerical physiochemical fingerprints [50], which in-turn are subjected to weighted scoring functions for the similarity estimation. Examples of estimation algorithms include Sequence-Order Independent Profile-Profile Alignment (SOIPAA) [51], a graph-based similarity algorithm was successfully used to identify off-targets for cholesteryl ester transfer protein inhibitors [52], and SiteAlign [53], a similarity estimation algorithm that is based on numerical fingerprint was used to identify novel targets for staurosporine [54].

Traditional molecular docking approaches in drug discovery involve virtual screening compound libraries against a given target to identify potential lead molecules. Recent advancements in computational efficiency have facilitated the optimization of standard docking algorithms to address specific demands when predicting off-targets or identifying drug repositioning opportunities. Such docking strategies are termed as inverse or reverse docking,

wherein a given compound can be virtually screened across multiple targets. There are currently only few docking algorithms that have successfully implemented inverse docking strategy, including idTarget, INVDOCK and DRAR-CPI [55-57]. For instance, a recent study by Kumar et al. revealed seven novel targets for kinetin using idTraget algorithm, four of which were experimentally validated [58]. Similarly, Ye et al. used PharmMapper [59] in combination with DRAR-CPI to identify a potentially novel target for capsaicin [60]. These inverse docking algorithms are often followed by a target ranking schema based on comparing the docking score, affinity profile or target-ligand ligand profile to identify the most potent targets.

A recent focus in target-centric approaches for drug repositioning and off-target effect prediction involves integrating protein-ligand interaction fingerprints (IFS) to rescore docking results. Approaches using IFS are best suited for cases of remote similarities, i.e., in which target proteins lack both sequence or structural similarities when aligning proteins globally (whole protein) or locally (protein binding site). IFS in general ignores the identity of amino acid involved in binding and abstracts the interactions that exist between the receptor and ligand, thereby providing a better similarity estimation of distantly related proteins. For instance, Desaphy et al. showed that such interaction fingerprint similarity strongly correlates with binding site similarity and is computationally more efficient to calculate when handling large dataset [61]. A recent review by Salentin et al. comprehensively documents the various types of interaction profiles, established methods for calculating IFS and their applications in drug discovery [62]. As a more recent work, BioGPS implements pocket similarity workflow that combines IFS with Molecular Interaction Fields (MIFs), and was used to identify potential off-target effects of estrogen modulators [63].

3.2 Compound-centric approaches

Compound-centric (chemo-centric or ligand-based) approaches, in contrast to the target-centric methods, rely on the compound's quantitative properties and their existing bioactivity profiles. The quantitative properties consist of 2D and 3D description of the compounds, represented as fingerprints, or the physiochemical properties, represented as descriptors. The key principle behind the compound-centric approaches is that compounds with shared structural or chemical similarity are inclined to bind to the same targets. These approaches are mostly data-driven and require the prior knowledge of compounds' activity towards multiple targets, thereby often limited to targets whose ligands are well-documented. A number of compound-centric computational methodologies have been developed to address both the adverse and therapeutic aspects of polypharmacology.

The most prominent compound-centric method that has gained popularity in the recent years is the Similarity Ensemble Approach (SEA) [44]. SEA is a statistical model that relates protein targets based on the chemical similarity of ligand sets that actively bind to the particular target. For instance, Lin et al. used SEA to relate non-GPCR's to GPCR protein family by showing how proteins that lack sequence similarity can be related in the chemical space, and predicted four compounds from various non-GPCR's that were experimentally validated to bind to GPCR's [64]. SEA has also been used to discover the off-targets of compounds to predict ADR's among 656 drugs and 73 off-targets by Lounkine et al. [65]. Another ligand-based target fishing approach is Rapid Overlay of Chemical Structures (ROCS), which works by querying the 3D pharmacophore profiles of drug molecules generated for individual protein targets [66]. Similarly, Gaussian Ensemble Screening (GSE) approach is based on Compound Polypharmacological Fingerprints (CFPs) that encode the ligand promiscuity information by comparing the 3D shape and chemistry of the ligands that bind to the targets [67].

Supervised machine learning (ML) models have also been developed for compound-centric target prediction, implemented through a wide variety of classification or regression algorithms such as support vector machines (SVM), decision trees (DT), k-nearest neighbor (k-NN), naïve Bayesian models and artificial neural networks (ANN) [68]. For instance, Cheng et al. developed a combined classifier that integrates many of the above ML algorithms to classify potent inhibitors and non-inhibitors of cytochrome P450; interestingly, the prediction accuracy of the combined classifier was significantly improved when compared to the independent classifiers [69]. A recent review by Lavecchia comprehensively describes these ML models in a great detail and with relevant examples, and also explains the recent variations incorporated into these models that have led to improved prediction accuracy, providing an excellent overview of the benefits and limitation of these ML methods [41]. A novel addition to the compound-centric toolbox for drug-target prediction is the deep learning model developed by Unterthiner et al. [70]. In comparison to other ML models, deep network (DN) architecture can improve multi-task learning and aid in constructing complex features. Accordingly, DN model was shown to have a better performance both in terms of target prediction and toxicity prediction [71].

The robustness of predictions from the ML models have been improved by combining the precise molecular features from both compounds and targets that attribute to a drug-target interaction, thus paving a way to model polypharmacological effects. Recently, Cichonska et al. implemented a kernel-based Kronecker's Regularized Least Square (KronRLS) ML model, where kernels encoding both target features (protein sequence and structure similarities) and compound features (2D and 3D shape and fingerprint similarities) were used to predict off-targets of tivozanib, an investigational VEGFR inhibitor [72]. In cases where specific disease models are available, the predictions from these ML methods can be enhanced by integrating

the drug response profiles with molecular and genomics profiling of the disease models, using e.g. copy number variation, proteomic, transcriptomic, methylation, and exome and RNA-Seq datasets. As a part of the NCI-DREAM challenge project, Costello et al. implemented a Bayesian multitask multiple kernel learning (MKL) method, a non-linear regression algorithm that integrates genomic, epigenomic and molecular profiling datasets to predict drug responses in breast cancer cell line models [73].

It is only recently that *in-silico* approaches have been successfully extended for predicting also drug combination effects. Parallel to the ML models highlighted above, other types of computational algorithms to predict drug combination synergies work by integrating dose-response profiles of single drugs with various -omics datasets, and in some cases, includes also the chemical and structural description of the compounds. For instance, Drug-Induced Genomic Residual Effect (DIGRE) is one such method that integrates the similarity of genome-wide gene expression responses to single drugs with their dose-response measurements to estimate the residual effect induced by two drug combinations [74]. DIGRE was found most effective in ranking 91 compound-pairs in the order of their synergistic or antagonistic behavior in a recent DREAM Challenge [75]. Recently, Korkut and colleagues proposed a quantitative model that links drug perturbations, proteomic changes and phenotypic outcomes to predict cellular responses to untested combinatorial interventions for RAF-inhibitor resistant melanoma cells [76]. To avoid the use of molecular response profiles that often are not available in translational applications, Gayvert et al. in their recent study used merely single-dose drug response measurements and limited number of combination testing to predict synergistic effects of pairwise drug combinations for the treatment of advanced melanoma [77].

In comparison with polytherapies, multi-targeted drugs (MTDs) provide multi-targeting activity through a monotherapy, and thereby can alleviate some pharmacokinetic problems of combination therapies, yet shown to have high synergistic effects [45]. Although screening approaches, such as SOSA, could help identifying molecular scaffolds necessary for designing MTDs with intended and beneficial off-target effects [29], computational methodologies to rationally design MTDs with desired polypharmacological profile are still in exploratory phase. Established *in-silico* approaches for designing MTDs are fragment-based methods that rely on fragment-growing strategies. For instance, Besnard introduced an automated, adaptive drug design method involving a Bayesian probabilistic model, where ligands optimally designed/evolved against a desired polypharmacological profile were used to make 800 novel ligand-target predictions, of which 75% were experimentally confirmed [78]. Recently, novel *de novo* design strategies have also been proposed for designing MTDs, in terms of both to formulate a polypharmacological drug and to design effective combinatorial screening libraries, currently limited for a dual-target profile only [79, 80].

3.3 Network polypharmacology approaches

A seminal work that instigated the concept of network polypharmacology in drug discovery was conducted in 2006 by Paolini et al., where they comprehensive catalogued and linked various protein targets, disease indications and chemical structures to generate a pharmacological mapping of interactions between ligands and protein targets [81]. A re-illustration of the global pharmacological space by Paolini et al. is depicted in **Figure 2**. This polypharmacological map was used in a probabilistic Bayesian framework to predict novel drug-target interactions (DTI). Building on the concept of network pharmacology [82], Yildirim et al. constructed a DTI bipartite graphs between FDA approved drug and their protein targets to quantitatively represent the overabundance of follow-on drugs [83]. Later, Yamanishi et al. posed a supervised learning

problem on the bipartite graph model by integrating the chemical structure and genomic sequence information to predict unknown DTIs [84]. Chiang and Butte developed a network based, guilt-by-association method, that links diseases based on shared similar therapies to suggests novel drug repurposing strategies [85]. After these seminal works, several network-based inference (NBI) algorithms that integrated gene regulatory, metabolic networks, protein-protein and drug-protein interaction networks have been developed and implemented to identify novel drug-target interactions and drug repositioning opportunities. Recent literature reviews that comprehensive detail these approaches are given elsewhere [46, 47].

A systematic study that evaluated the performance of NBI algorithms against standard target-based inference and compound-based inference algorithms was carried out by Cheng et al. [42]. Their implementation of the NBI algorithm outperformed both the target and compound-centric based methods, and was successful in repositioning drugs for human estrogen receptor (ER). Another network polypharmacology approach, Target Inhibition using Maximization and Minimization Averaging (TIMMA), uses only monotherapy responses in a given cell sample to identify selective target combination through target inhibition networks. TIMMA model makes use of the target space of promiscuous kinase inhibitors to predict efficacies for novel drug-target combinations [86]. As a case study, TIMMA was used to identify selective target combinations for breast cancer and pancreatic cell-lines, and the prediction was later experimentally validated using siRNA-mediated silencing. Recently, a gene regulatory network-based approach was developed to elucidate compounds MoA using algorithm called Detecting Mechanism of Action by Network Dysregulation (DeMAND) [43]. DeMAND interrogates tissue-specific gene regulatory network changes following compound perturbation, and identifies novel proteins that are accompanied with the established MoA of the given compound. This method identified glutathione peroxidase 4 (GPX4) as a novel MoA effector

for altretamine, which was later experimentally validated using liquid chromatography-mass spectrometry (LC-MS)-based GPX4 assay.

There exist also NBI-based methods for modelling MoA and suggesting drug repositioning using drug-drug interactions networks. For instance, Mode of Action by NeTwoRk Analysis (MANTRA) is one such approach that constructs drug-drug similarity networks using comparative gene expression profiles across cell-lines after drug treatment. Using network topological properties, MANTRA identifies compound hubs, which were shown to be enriched with compounds having similar MoA, or tend to inhibit similar biological pathway [87]. MANTRA was used to predict MoA of nine anticancer compounds and it also discovered repositioning possibilities of a Rho-inhibitor, fasudil, to enhance cellular autophagy. A semi-supervised based on similar concepts was recently developed by Iorio et al. for refining transcriptional drug response signatures and making drug repositioning predictions [88]. Gottileb et al. developed a NBI-based platform for predicting drug indications (PREDICT). This platform combines several drug-drug and disease-disease similarity measures in a logistic regression model to infer drug-disease associations, thereby finding novel indications not only for existing and approved drugs, but also for investigational agents [89]. The authors validated the predictions by means of an overlap analysis with drug indications that were under clinical trials, and by their agreement with tissue-specific expression information on the drug targets based on public resources.

4. Data resources and web-tools for polypharmacology

In addition to widely-used chemical databases, such as ChEMBL, PubChem, SuperTargets, BindingDB and DrugBank, which compiles various bioactivity profiling information along with the structure of compounds and targets, dedicated data resources have recently been developed

to address specifically the polypharmacological modelling needs (**Table 2**). For instance, recent efforts by Corsello and colleagues led to the development of Drug Repositioning Hub (DRH), a drug-library resource containing a comprehensive collection of 5627 compounds that are either approved and marketed for clinical use or previously approved and withdrawn, or are currently in different stages of clinical trials [90]. This library of drugs was initially collected and curated from existing public or proprietary databases, then subjected to chemical-structure analysis, annotated for functional and developmental status, and later experimentally checked for their identity and purity. The DRH library is proposed as a go-to resource for repositioning applications. Similarly, RepoDB is a gold-standard database consisting of both approved and failed drugs with their clinical indications [91]. RepoDB was developed as a standard benchmark dataset to evaluate computational models addressing repurposing strategies. Our group has also developed Drug Target Commons (DTC) platform, a community-driven web-based database for comprehensive resource of drug-target bioactivity data [92]. The main focus of DTC is to address the underlying heterogeneity in the existing bioactivity data, thereby providing a reliable platform for data-driven polypharmacology applications.

Either associated with the above-mentioned databases or as stand-alone software, various computational tools based on either machine learning or statistical algorithms have been implemented to support polypharmacological studies, many of which also come with an easy-to-use web-application (**Table 3**). These web-applications are based on a computational framework, be it target-centric, compound-centric, or network-centric (as reviewed in Section 3), enabling to study both the adverse and therapeutic aspects of polypharmacology. For instance, SPiDER web-tool uses self-organization maps (SOMs) and their consensus to identify molecular targets for known and *de novo* designed drug molecules [93]. TarPred is a recently developed web-tool that aids in identifying and ranking the top interacting proteins for a given

drug molecule by elucidating the compounds' MoA or its toxic side-effects, and it also highlights the possible disease indications [94]. CSNAP and systemsDock are recent network-centric web-tools to predict drug-target interactions, where the former uses a graph-similarity estimation and the latter permits also efficient molecular docking simulation [95, 96]. C-SPADE is a compound-centric web-tool recently developed to interactively analyze and explore drug screening data. C-SPADE generates compound similarity clusters using various chemical fingerprint measurements and overlays the compound bioactivity data (either against protein targets or cell lines), thereby aiding users to explore the chemical diversity of the screening panel [97].

Only a handful of data resources and web-applications are currently available for modelling or analyzing drug combinations (see **Table 2 and Table 3**). Drug Combination Database (DCDB) and Antifungal Synergistic Drug Combination Database (ASDCD) are two examples of drug combination databases [98, 99]. ASDCD is solely dedicated for synergistic drug combination information for fungal infection, whereas DCDB consists of curated, both successful and unsuccessful drug combinations, compiled from multiple resources such as ClinicalTrials.gov, FDA orange book and PubMed database for various human diseases. As for drug combination web-applications, DT-Web is a recently-implement NBI-based tool that combines domain-specific information of drug and target similarity to predict drug-target interactions and drug combination effects [100]. Drug combination effects quantified as Combination Index (CI) can be calculated in DT-Web using both dose-response-based and effect-based scoring approaches, by comparing the combination responses to those of the single agents when used alone. Other popular references models for the synergy quantification include the Highest Single Agent (HSA), Bliss independence model, Loewe additivity model and Zero Interaction Potency (ZIP) models. SynergyFinder is a recently developed web-application that aids in calculating synergy

scores using these different reference models, and it provides the users with an interactive visualization of the drug combination responses, both as 2D and 3D dose-response synergy and efficacy matrices and landscapes [101].

5. Conclusion

Most drug discovery processes encounter molecules that are promiscuous in their nature. Rather than trying to avoid multi-targeting properties of compounds, recent efforts comprising both screening strategies and computational methodologies have been streamlined to better exploit and control for such polypharmacological effects. Unless meticulously designed, polypharmacological activities of a promiscuous compound behave in a two-edged manner, giving rise to both adverse toxic effects and therapeutically beneficiary responses. Examples from the past show that such therapeutic responses were mostly discovered through mere serendipity during the screening process, but recent advancements in the drug development pipelines, encompassing novel experimental profiling, library designing and computational strategies, have paved way to explore the polypharmacological behavior in a more rationalized manner. Compared to the conventional approach of *de novo* designing of drug molecules for specific molecular targets, these polypharmacological approaches serve as an amenable alternative to the traditional drug discovery process. For instance, the design of efficient combinatorial mixture-based libraries has already aided in identifying compounds and MTDs with the desired polypharmacological effects. Furthermore, the advent of rational drug combination screening strategies has facilitated the identification of clinically relevant synergistic drug combinations to combat drug resistance problems in complex diseases like cancers. Data-driven models based on polypharmacological information of compounds have also been successfully applied to predicting toxic side-effects and ADME properties of compounds. To best support the scientific community, the computational models and current

knowledge on compound-target interactions are being implemented as easy-to-use web-applications and comprehensive databases, respectively.

As revolutionary these polypharmacological applications are to the modern drug discovery, it would be unwise failing to address their current challenges and limitations. A major challenge attributed to any phenotype-based polypharmacology study is related to the target deconvolution; our current understanding of the molecular disease pathways and compounds' MoA is still limited, thereby making it increasingly difficult to assess the global polypharmacological behavior of the multi-targeting agents. Likewise, all the aforementioned computational approaches are developed independently to resolve specific biological question (e.g., predicting novel targets, drug indications, interactions, toxicity, etc.). As most of these approaches are inherently data-driven, they are strongly dependent on the availability of the specific data type, including the structure of proteins and compounds, or the prior compounds' bioactivity profiling information. Further, the accuracy and specificity of these ML approaches depends drastically on the size and diversity of the training data used to estimate the model prior to predictions. Another significant limitation is the availability of tools and databases for studying drug combination effects, which are currently fairly few compared to those developed for predicting monotherapy indications or side effects. The recent AstraZeneca-Sanger Drug Combination Prediction DREAM Challenge (https://www.synapse.org/#!Synapse:syn4231880) will soon release a large number of experimentally tested drug combinations over 118 drugs and 85 cancer cell lines, together with the monotherapy responses and baseline genomic data, which will provide rich data resource to explore molecular mechanisms that underlie effective combination treatments and synergistic drug behavior (manuscript submitted). Adding and harmonizing these and other emerging data resources for the needs of polypharmacological modelling requires a careful annotation of the multiple targets of the combinations, in order to

map the combination effects into the existing pharmacogenetic mappings between drugs and targets and their effects in different diseases and cell contexts.

Despite these obstacles, novel systematic frameworks utilizing polypharmacology have already been successful in speeding-up and de-risking the currently costly and lengthy drug discovery process. We are confident that with the ongoing and emerging improvements both in the experimental and computation strategies, polypharmacology will remain as an essential paradigm for the next generation of selective, multi-target drug discovery.

6. Expert Opinion

A critical challenge in translating polypharmacological responses into clinical applications is how to better control for the adverse effects of multi-targeted activities. Even after excluding compounds with non-selective polypharmacological activities, using e.g. the safety screening panels or PAINS and other filters, the compounds may still have selective activities against a number of proteins in distinct biological pathways. It is not enough to merely map the potential target space of the candidate compound, but also to have quantitative information about the potency of the compound to inhibit (or activate) each individual target. This is important for adjusting the concentration levels of the multi-targeting compounds and for providing insights into the potential therapeutic window in clinical applications. Even if the exact doses are hard to predict based on the biochemical or cell-based bioactivity data, these can at least provide some clues about the MoA of the compound and the order in which the multiple-targets are being modulated. The intended primary targets are often modulated at lower concentrations, whereas activity against the secondary targets often comes only with higher doses, many times leading to toxic side effects. However, there are also many examples where the off-targets are actually

modulated with similar or even lower doses than the intended target. Many of such cases are being discovered after carrying out more comprehensive target selectivity profiling (or getting access to the historic target selectivity data from pharma companies). The most exciting cases are naturally those where the off-target effect is therapeutically beneficial, rather than toxic, as was the case with the BCR-ABL(T315I) example in CLL treatment (see Section 2). This is why we argue that comprehensive target selectivity profiling is so important for the success of the polypharmacological modelling, not only for better characterizing novel lead molecules but also for new drug repurposing opportunities of existing drugs. For the promiscuous kinase inhibitors, there are already several chemogenomic screening efforts ongoing to comprehensively map the "druggable" kinome, but similar community efforts are needed also for other important compound and target families. This should not be only an academic endeavor, but we also encourage industrial partners to join these efforts. In particular, pharma companies will benefit from having more extensive target selectivity profiles for their molecules, both approved and investigational ones, and even for abandoned leads that might be still re-activated or directed toward academy-driven drug development. Collecting and integrating the heterogeneous target profiling data from various sources and bioactivity assays is a massive IT, statistical and curation challenge, and this is where community-based platforms, such as our DrugTargetCommons (https://drugtargetcommons.fimm.fi/), will become important in standardizing and annotating the heterogeneous bioactivity data for downstream modelling.

We also believe that the use computational modelling is highly beneficial also when designing the compound or target selectivity experiments, not only in the analysis of their results. Data driven computational models should prove useful especially for experiments where the chemical or target spaces are too high-dimensional to be mapped solely by experimental approaches. For instance, for comprehensive mapping of binding interactions between kinase inhibitors and the

whole human kinome, the dimensionality of both the compound and target spaces is in hundreds, not to mention the number of potential compound-target interactions. We have already shown in a pilot study that ML-based predictive models can guide such compound-kinase mapping efforts by prioritizing the most potent interactions for further experimental validation [72]. Such model-guided experimental designs are expected to provide more effective alternative to the systematic experimentation also in other high-dimensional applications (Azencott et al., Nature Methods, in press). One such application is testing of drug combinations, which poses significant experimental challenges; even if we consider only the approved drugs and exclude agents with redundant mechanisms, the number of pairwise drug combinations remains in thousands. With investigational compounds and higher-order combinations, the size of the combination space grows exponentially. The recent Drug Combination DREAM Challenge should provide useful data for developing multi-task predictive models that can leverage the information across compounds and cell lines, using their similarities and differences, with the aim to suggest what are the most informative areas of further combinatorial experiments in the similar cell contexts. These massive experimental datasets will be also useful for systematic evaluation of the model predictions using cross-validation or hold-out datasets. So far, many of the ML models for drug-target interaction or drug repurposing predictions have been validated using inconsistent evaluation strategies, published data or 'gold standard' databases that may not be optimal for the particular application [102]. We argue that although the use of publicly available data is important for the initial 'sanity check', the practical performance of the prediction models cannot be evaluated without new experiments, generated based on the model predictions, both for the positive and negative predictions; otherwise, there is always the risk of information leakage if the validation data are available at the time of making the predictions. For validation of polypharmacological models, biochemical or cell lines may not be realistic enough, but we should go more toward

pre-clinical animal models to study both the therapeutic efficacy and safety of the multi-targeted treatments in more disease-informative models.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Roth B L, Sheffler D J, and Kroeze W K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat Rev Drug Discov 2004;3(4):353-9.

2. Strebhardt K and Ullrich A. Paul Ehrlich's magic bullet concept: 100 years of progress. Nat Rev Cancer 2008;8(6):473-80.

3. Dickson M and Gagnon J P. Key factors in the rising cost of new drug discovery and development. Nat Rev Drug Discov 2004;3(5):417-29.

4. Dimasi J A. Risks in new drug development: approval success rates for investigational drugs. Clin Pharmacol Ther 2001;69(5):297-307.

5. Peters J U. Polypharmacology - foe or friend? J Med Chem 2013;56(22):8955-71.

•• A detailed review documenting numerous examples of adverse and therapeutic effects of polypharmacology.

6. Petrelli A, Polypharmacological Kinase Inhibitors: New Hopes for Cancer Therapy, in Polypharmacology in Drug Discovery, First Edition. Edited by Jens-Uwe Peters., J.W. Sons, Editor. 2012, John Wiley & Sons. p. 149-165.

7. Bayat Mokhtari R, Homayouni T S, Baluch N, et al. Combination therapy in combating cancer. Oncotarget 2017;8(23):38022-38043.

• An excellent review comprising of numerous examples of effective drug combination treatments that were proved to be fruitful in cancer therapy.

8. Hu Y and Bajorath J. Compound promiscuity: what can we learn from current data? Drug Discov Today 2013;18(13-14):644-50.

9. Santos R, Ursu O, Gaulton A, et al. A comprehensive map of molecular drug targets. Nat Rev Drug Discov 2017;16(1):19-34.

• Describes a comprehensive curation process focused towards deconvolution of molecular targets for approved drugs to aid in polypharmacology and drug efficacy studies.

10. Dar A C, Das T K, Shokat K M, et al. Chemical genetic discovery of targets and anti-targets for cancer polypharmacology. Nature 2012;486(7401):80-4.

11. Baell J B and Holloway G A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 2010;53(7):2719-40.

• A novel filtering approach that compiles numerous key sub structures of highly promiscuous compounds that accounts for compound side effects and aids in molecular screening studies.

12. Reddy A S and Zhang S. Polypharmacology: drug discovery for the future. Expert Rev Clin Pharmacol 2013;6(1):41-7.

13. Chong C R and Sullivan D J, Jr. New uses for old drugs. Nature 2007;448(7154):645-6.

14. Haupt V J and Schroeder M. Old friends in new guise: repositioning of known drugs with structural bioinformatics. Brief Bioinform 2011;12(4):312-26.

15. Apsel B, Blair J A, Gonzalez B, et al. Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases. Nat Chem Biol 2008;4(11):691-9.

16. Papaetis G S and Syrigos K N. Sunitinib: a multitargeted receptor tyrosine kinase inhibitor in the era of molecular cancer therapies. BioDrugs 2009;23(6):377-89.

17. Tsubamoto H, Sonoda T, Ikuta S, et al. Combination Chemotherapy with Itraconazole for Treating Metastatic Pancreatic Cancer in the Second-line or Additional Setting. Anticancer Res 2015;35(7):4191-6.

18. Li J, Zheng S, Chen B, et al. A survey of current trends in computational drug repositioning. Brief Bioinform 2016;17(1):2-12.

19. Lavecchia A and Cerchia C. In silico methods to address polypharmacology: current status, applications and future perspectives. Drug Discov Today 2016;21(2):288-98.

20. Bowes J, Brown A J, Hamon J, et al. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. Nat Rev Drug Discov 2012;11(12):909-22.

21. Munoz L. Non-kinase targets of protein kinase inhibitors. Nat Rev Drug Discov 2017;16(6):424-440.

22. Basilico C, Pennacchietti S, Vigna E, et al. Tivantinib (ARQ197) displays cytotoxic activity that is independent of its ability to bind MET. Clin Cancer Res 2013;19(9):2381-92.

23. Senior M. FDA halts then allows sales of Ariad's leukemia medication. Nat Biotechnol 2014;32(1):9-11.

24. Pemovska T, Johnson E, Kontro M, et al. Axitinib effectively inhibits BCR-ABL1(T315I) with a distinct binding conformation. Nature 2015;519(7541):102-5.

•• An elegant drug repurposing example through phenotypic screening that led to the identification of novel drug-target interaction efficient in overcoming drug resistance in leukemia patients.

25. Bunnage M E, Chekler E L, and Jones L H. Target validation using chemical probes. Nat Chem Biol 2013;9(4):195-9.

26. Jones L H and Bunnage M E. Applications of chemogenomic library screening in drug discovery. Nat Rev Drug Discov 2017;16(4):285-296.

27. Austin C P, Brady L S, Insel T R, et al. NIH Molecular Libraries Initiative. Science 2004;306(5699):1138-9.

28. Houghten R A, Pinilla C, Giulianotti M A, et al. Strategies for the use of mixture-based synthetic combinatorial libraries: scaffold ranking, direct testing in vivo, and enhanced deconvolution by computational methods. J Comb Chem 2008;10(1):3-19.

29. Wermuth C G. Selective optimization of side activities: the SOSA approach. Drug Discov Today 2006;11(3-4):160-4.

•• A seminal study that highlighted a novel screening approach that capitalizes on the compounds off-target effects to facilitate drug repositioning studies.

30. Varbanov H P, Kuttler F, Banfi D, et al. Repositioning approved drugs for the treatment of problematic cancers using a screening approach. PLoS One 2017;12(2):e0171052.

31. Kummar S, Chen H X, Wright J, et al. Utilizing targeted cancer therapeutic agents in combination: novel approaches and urgent requirements. Nat Rev Drug Discov 2010;9(11):843-56.

32. Foucquier J and Guedj M. Analysis of drug combinations: current methodological landscape. Pharmacol Res Perspect 2015;3(3):e00149.

• A comprehensive review detailing the various drug combination methodologies and scoring functions currently in use for combination therapy investigations.

33. Manchado E, Weissmueller S, Morris J P t, et al. A combinatorial strategy for treating KRAS-mutant lung cancer. Nature 2016;534(7609):647-51.

34. Horn T, Ferretti S, Ebel N, et al. High-Order Drug Combinations Are Required to Effectively Kill Colorectal Cancer Cells. Cancer Res 2016;76(23):6950-6963.

35. Davis M I, Hunt J P, Herrgard S, et al. Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol 2011;29(11):1046-51.

36. Metz J T, Johnson E F, Soni N B, et al. Navigating the kinome. Nat Chem Biol 2011;7(4):200-2.

37. Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483(7391):603-7.

38. Yang W, Soares J, Greninger P, 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.

39. Basu A, Bodycombe N E, Cheah J H, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell 2013;154(5):1151-61.

40. PREDICTIVE ADMET: Integrative Approaches in Drug Discovery and Development, ed. J. Wang and L. Urban. 2014, Hoboken, New Jersey: John Wiley & Sons.

41. Lavecchia A. Machine-learning approaches in drug discovery: methods and applications. Drug Discov Today 2015;20(3):318-31.

• An excellent review of the various machine learning models used in predicting drug-target interactions with examples of their application.

42. Cheng F, Liu C, Jiang J, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput Biol 2012;8(5):e1002503.

43. Woo J H, Shimoni Y, Yang W S, et al. Elucidating Compound Mechanism of Action by Network Perturbation Analysis. Cell 2015;162(2):441-451.

44. Keiser M J, Roth B L, Armbruster B N, et al. Relating protein pharmacology by ligand chemistry. Nat Biotechnol 2007;25(2):197-206.

•• The first instance on the use of a statistical model to explain polypharmacology across a protein family with details on the use of ligand set similarity to identify novel drug-target interactions.

45. Bottegoni G, Favia A D, Recanatini M, et al. The role of fragment-based and computational methods in polypharmacology. Drug Discov Today 2012;17(1-2):23-34.

46. Cichonska A, Rousu J, and Aittokallio T. Identification of drug candidates and repurposing opportunities through compound-target interaction networks. Expert Opin Drug Discov 2015;10(12):1333-45.

47. Lotfi Shahreza M, Ghadiri N, Mousavi S R, et al. A review of network-based approaches to drug repositioning. Brief Bioinform 2017.

• An excellent review that details the various network based approaches that have been implemented for drug repositioning strategies.

48. Chen Y Z and Zhi D G. Ligand–Protein Inverse Docking and Its Potential Use in the Computer Search of Protein Targets of a Small Molecule. Proteins: Struct., Funct., Genet. 2001;43:217–226.

49. Roy A and Zhang Y. Recognizing protein-ligand binding sites by global structural alignment and local geometry refinement. Structure 2012;20(6):987-97.

50. Kellenberger E, Schalon C, and Rognan D. How to Measure the Similarity Between Protein Ligand-Binding Sites? Curr Comput Aided Drug Des 2008;4.

51. Xie L and Bourne P E. Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. Proc Natl Acad Sci U S A 2008;105(14):5441-6.

52. Xie L, Li J, Xie L, et al. Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. PLoS Comput Biol 2009;5(5):e1000387.

53. Schalon C, Surgand J S, Kellenberger E, et al. A simple and fuzzy method to align and compare druggable ligand-binding sites. Proteins 2008;71(4):1755-78.

54. Defranchi E, Schalon C, Messa M, et al. Binding of protein kinase inhibitors to synapsin I inferred from pair-wise binding site similarity measurements. PLoS One 2010;5(8):e12214.

55. Wang J C, Chu P Y, Chen C M, et al. idTarget: a web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. Nucleic Acids Res 2012;40(Web Server issue):W393-9.

56. Chen Y Z and Ung C Y. Prediction of potential toxicity and side effect protein targets of a small molecule by a ligand-protein inverse docking approach. J Mol Graph Model 2001;20(3):199-218.

57. Luo H, Chen J, Shi L, et al. DRAR-CPI: a server for identifying drug repositioning potential and adverse drug reactions via the chemical-protein interactome. Nucleic Acids Res 2011;39(Web Server issue):W492-8.

58. Kumar S P, Parmar V R, Jasrai Y T, et al. Prediction of protein targets of kinetin using in silico and in vitro methods: a case study on spinach seed germination mechanism. J Chem Biol 2015;8(3):95-105.

59. Liu X, Ouyang S, Yu B, et al. PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. Nucleic Acids Res 2010;38(Web Server issue):W609-14.

60. Ye X Y, Ling Q Z, and Chen S J. Identification of a Potential Target of Capsaicin by Computational Target Fishing. Evid Based Complement Alternat Med 2015;2015:983951.

61. Desaphy J, Raimbaud E, Ducrot P, et al. Encoding protein-ligand interaction patterns in fingerprints and graphs. J Chem Inf Model 2013;53(3):623-37.

62. Salentin S, Haupt V J, Daminelli S, et al. Polypharmacology rescored: protein-ligand interaction profiles for remote binding site similarity assessment. Prog Biophys Mol Biol 2014;116(2-3):174-86.

• A detailed documentation of various protein-ligand interaction profiles, their enumeration and usage in virtual screening and docking studies.

63. Siragusa L, Cross S, Baroni M, et al. BioGPS: navigating biological space to predict polypharmacology, off-targeting, and selectivity. Proteins 2015;83(3):517-32.

64. Lin H, Sassano M F, Roth B L, et al. A pharmacological organization of G protein-coupled receptors. Nat Methods 2013;10(2):140-6.

65. Lounkine E, Keiser M J, Whitebread S, et al. Large-scale prediction and testing of drug activity on side-effect targets. Nature 2012;486(7403):361-7.

66. AbdulHameed M D, Chaudhury S, Singh N, et al. Exploring polypharmacology using a ROCS-based target fishing approach. J Chem Inf Model 2012;52(2):492-505.

67. Perez-Nueno V I, Karaboga A S, Souchet M, et al. GES polypharmacology fingerprints: a novel approach for drug repositioning. J Chem Inf Model 2014;54(3):720-34.

68. Marchese Robinson R L, Palczewska A, Palczewski J, et al. Comparison of the Predictive Performance and Interpretability of Random Forest and Linear Models on Benchmark Data Sets. J Chem Inf Model 2017;57(8):1773-1792.

• A comparative study of various machine learning methods and their interpretability, which provides a usable benchmark dataset for other computational models.

69. Cheng F, Yu Y, Shen J, et al. Classification of cytochrome P450 inhibitors and noninhibitors using combined classifiers. J Chem Inf Model 2011;51(5):996-1011.

70. Unterthiner T, Mayr A, Klambauer G n, et al., Deep Learning as an Opportunity in Virtual Screening, in Deep Learning and Representation Learning Workshop, NIPS. 2014, MIT Press: Cambridge, MA.

•• A seminal work on the use of deep learning machine learning algorithm in predicting novel drug-target interactions, where the authors have also successfully used this approach for toxicity predictions.

71. Unterthiner T, Mayr A, Klambauer G n, et al. Toxicity Prediction using Deep Learning. arXiv preprint 2015:arXiv:1503.01445.

72. 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(8):e1005678.

73. Costello J C, Heiser L M, Georgii E, et al. A community effort to assess and improve drug sensitivity prediction algorithms. Nat Biotechnol 2014;32(12):1202-12.

74. Yang J, Tang H, Li Y, et al. DIGRE: Drug-Induced Genomic Residual Effect Model for Successful Prediction of Multidrug Effects. CPT Pharmacometrics Syst Pharmacol 2015;4(2):e1.

75. Bansal M, Yang J, Karan C, et al. A community computational challenge to predict the activity of pairs of compounds. Nat Biotechnol 2014;32(12):1213-22.

76. Korkut A, Wang W, Demir E, et al. Perturbation biology nominates upstream-downstream drug combinations in RAF inhibitor resistant melanoma cells. Elife 2015;4.

77. Gayvert K M, Aly O, Platt J, et al. A Computational Approach for Identifying Synergistic Drug Combinations. PLoS Comput Biol 2017;13(1):e1005308.

78. Besnard J, Ruda G F, Setola V, et al. Automated design of ligands to polypharmacological profiles. Nature 2012;492(7428):215-20.

•• A novel approach that integrates a computational-experimental framework to design and test multi-targeted drugs.

79. Reutlinger M, Rodrigues T, Schneider P, et al. Multi-objective molecular de novo design by adaptive fragment prioritization. Angew Chem Int Ed Engl 2014;53(16):4244-8.

80. Shang E, Yuan Y, Chen X, et al. De novo design of multitarget ligands with an iterative fragment-growing strategy. J Chem Inf Model 2014;54(4):1235-41.

81. Paolini G V, Shapland R H, van Hoorn W P, et al. Global mapping of pharmacological space. Nat Biotechnol 2006;24(7):805-15.

•• An excellent article that instigated the idea of network polypharmacology for modelling inter- and intra-relationships between protein families using compound promiscuity measure.

82. Hopkins A L. Network pharmacology. Nat Biotechnol 2007;25(10):1110.

83. Yildirim M A, Goh K I, Cusick M E, et al. Drug-target network. Nat Biotechnol 2007;25(10):1119-26.

84. 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(13):i232-40.

85. Chiang A P and Butte A J. Systematic evaluation of drug-disease relationships to identify leads for novel drug uses. Clin Pharmacol Ther 2009;86(5):507-10.

86. 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(9):e1003226.

87. Iorio F, Bosotti R, Scacheri E, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci U S A 2010;107(33):14621-6.

88. Iorio F, Shrestha R L, Levin N, et al. A Semi-Supervised Approach for Refining Transcriptional Signatures of Drug Response and Repositioning Predictions. PLoS One 2015;10(10):e0139446.

89. Gottlieb A, Stein G Y, Ruppin E, et al. PREDICT: a method for inferring novel drug indications with application to personalized medicine. Mol Syst Biol 2011;7:496.

90. Corsello S M, Bittker J A, Liu Z, et al. The Drug Repurposing Hub: a next-generation drug library and information resource. Nat Med 2017;23(4):405-408.

91. Brown A S and Patel C J. A standard database for drug repositioning. Sci Data 2017;4:170029.

92. Tang J, Tanoli Z-u-R, Ravikumar B, et al. DrugTargetCommons: a community-effort to build a consensus knowledgebase for drug-target interactions. Cell Chemical Biology. (manuscript accepted)

93. Reker D, Rodrigues T, Schneider P, et al. Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. Proc Natl Acad Sci U S A 2014;111(11):4067-72.

94. Liu X, Gao Y, Peng J, et al. TarPred: a web application for predicting therapeutic and side effect targets of chemical compounds. Bioinformatics 2015;31(12):2049-51.

95. Lo Y C, Senese S, Li C M, et al. Large-scale chemical similarity networks for target profiling of compounds identified in cell-based chemical screens. PLoS Comput Biol 2015;11(3):e1004153.

96. Hsin K Y, Matsuoka Y, Asai Y, et al. systemsDock: a web server for network pharmacology-based prediction and analysis. Nucleic Acids Res 2016;44(W1):W507-13.

97. Ravikumar B, Alam Z, Peddinti G, et al. C-SPADE: a web-tool for interactive analysis and visualization of drug screening experiments through compound-specific bioactivity dendrograms. Nucleic Acids Res 2017;45:W495–W500.

98. Liu Y, Wei Q, Yu G, et al. DCDB 2.0: a major update of the drug combination database. Database (Oxford) 2014;2014:bau124.

99. Chen X, Ren B, Chen M, et al. ASDCD: Antifungal Synergistic Drug Combination Database. PLoS One 2014;9(9):e86499.

100. Alaimo S, Bonnici V, Cancemi D, et al. DT-Web: a web-based application for drug-target interaction and drug combination prediction through domain-tuned network-based inference. BMC Syst Biol 2015;9 Suppl 3:S4.

101. Ianevski A, He L, Aittokallio T, et al. SynergyFinder: a web application for analyzing drug combination dose-response matrix data. Bioinformatics 2017;33(15):2413-2415.

102. Brown A S and Patel C J. A review of validation strategies for computational drug repositioning. Brief Bioinform 2016.

103. Drewry D H, Wells C I, Andrews D M, et al. Progress towards a public chemogenomic set for protein kinases and a call for contributions. PLoS One 2017;12(8):e0181585.

Table 1. Chemogenomic and combinatorial libraries for polypharmacological studies.

Library	Compound enumerations	Description
Protein Kinase Inhibitor Set - GlaxoSmithKline (GSK-PKIS)	367	Developed to identify lead molecules for orphan kinases and a screening study across 24 annotated primary targets was used in estimating the performance of various machine learning models. The authors have recently released a comprehensive kinase chemogenomic set called PKIS2 [103].
NIH Clinical Collection (NCC)	446	Consists entirely of compounds that are in phase I to phase III of clinical trials with availability of their bioactivity data.
Sigma Library of Pharmacologically Active Compounds (LOPAC 1280)	1280	A flexible target characterization and assay validation library, consisting of marketed and pharmaceutically relevant structures annotated with biological activities against target classes such as GPCRs and kinases.
Prestwick Chemical Library	1280	A compilation of approved drugs from FDA, EMA and other sources selected based on the compounds bioavailability, target diversity and safety.
Mechanism Interrogation PlatE (MIPE-4.0)	1912	An oncology-targeted compound collection, consisting of drugs that are either approved or in various stages of clinical trials and well suited for anticancer profiling studies.
MicroSource Spectrum	2000	A collection of compounds from commercial drug repositories that show significant diversity in their structure, bioactivity and function. The collections include drug components, natural products and other bioactive components.
National Cancer Institute Collection (NCI)	2277	A cancer drug derivative collection of compounds with 1990 compounds from NCI diversity set, 230 from natural products and 57 from their challenge set.
Specs consortium collection	30000	A combinatorial chemistry collection of compounds that exhibit structural characteristics of a biologically active compound, and meet Absorption Distribution Metabolism Excretion (ADME) requirements.
ChemBridge DIVERSet	50000	A combinatorial chemistry collection consisting of highly diverse drug-like and lead compounds with pharmacophore space representing potential interactions between compounds and biological targets.

Maybridge	57809	Compound collection consisting of diverse compounds obtained using clustering algorithm based on the Tanimoto similarity of compound fingerprints.

Table 2. A summary of representative data resources to aid polypharmacology.

Name	Description	URL
Drug Repurposing HuB	A hand-curated collection of drugs that are approved or currently used in clinical trials. Compounds purity has been checked and developed to serve as a repositioning library.	https://clue.io/repurposing-app
RepoDB	The collection consists of compounds that have either approved or failed with their intended indications. Designed to serve as a benchmark dataset for computational drug repurposing methods.	http://apps.chiragjpgroup.org/repoDB/
Repurpose DB	The database consisting of successfully repurposed drugs, drug targets and associated disease indications. Designed to serve as a centralized knowledgebase to identify pharmacological, biological and epidemiological factors influencing drug repositioning strategies.	http://repurposedb.dudleylab.org/index
PROMISCIOUS	A compendium of withdrawn and experimental drugs annotated with protein-protein and drug-protein interaction information, to aid in network-based approach for polypharmacological studies.	http://bioinformatics.charite.de/promiscuous/
Multiple Target Ligand Databases (MTLD)	MTLD was developed by mining the PDB databases to retrieve ligands that bind to multiple targets, of which 222 were approved drugs and 1334 drug-like compounds. Well-suited for developing of polypharmacological drugs.	www.mtdcadd.com
FDA's Adverse Event Reporting System (FAERS)	FAERS databases consist of a compilation of the adverse event and medication failures submitted to FDA.	https://open.fda.gov/data/faers/
DSigDB	A Drug Signature Databases that relates drugs and their target genes for gene set enrichment analysis.	http://tanlab.ucdenver.edu/DSigDB/DSigDBv1.0/
Drug Target Commons (DTC)	A crowd-sourcing platform designed to improve the consensus and use of drug-target interactions by addressing the inherent heterogeneity in target profiling data through manual annotation of assay information.	https://drugtargetcommons.fimm.fi/
Probes and Drug portal	A comprehensive database dedicated for chemical probes documenting attributes such	https://www.probes-drugs.org/home/

	as potency, selectivity and MoA. Enables library filtration and comparisons, augmented with visualization tools to understand the chemical space.	
Drug Central	Complies drug information from FDA, EMA and PMDA. Provides information on the drug MoA, pharmacological action and approved indications.	http://drugcentral.org/
ChemProt-3.0	Complies and annotates drug-protein-disease interactions. The database consists of 1.7 million compounds, 19504 targets with 7.8 million activity measurements.	http://potentia.cbs.dtu.dk/ChemProt/
LINCS	Apart from the collection of small molecules, LINCS databases catalogs the gene expression and molecular responses to the drug perturbation, thereby providing a network-based understanding for drug development.	http://lincs.hms.harvard.edu/db/
e-Drug3D	A comprehensive resource of FDA approved drugs and active metabolites with structural and bioactivity information. Also provides the possibility to virtually screen the drugs and search for pharmacophores.	http://chemoinfo.ipmc.cnrs.fr/MOLDB/index.html
Drug Combination Database (DCDB)	A collection of 1363 drug combinations among 904 drugs and 805 targets. It consists of 330 approved drugs, 1033 investigational drugs and 237 unsuccessfully combinations.	http://www.cls.zju.edu.cn/dcdb/
Antifungal Synergistic Drug Combination Database (ASDCD)	Consists of 210 antifungal synergistic drug combination and 1255 drug-target interaction involving 105 drugs.	http://ASDCD.amss.ac.cn

Table 3. A list of useful web-based tools implemented to aid polypharmacological applications.

Name	Description	URL
SEA	Implementation of statistical ensemble approach to identify novel drug-target interaction by comparing the chemical space between protein targets	http://sea.bkslab.org/
Galahad	A web server that uses drug effect analysis from gene expression profiles to elucidate the compounds MoA.	https://galahad.esat.kuleuven.be/

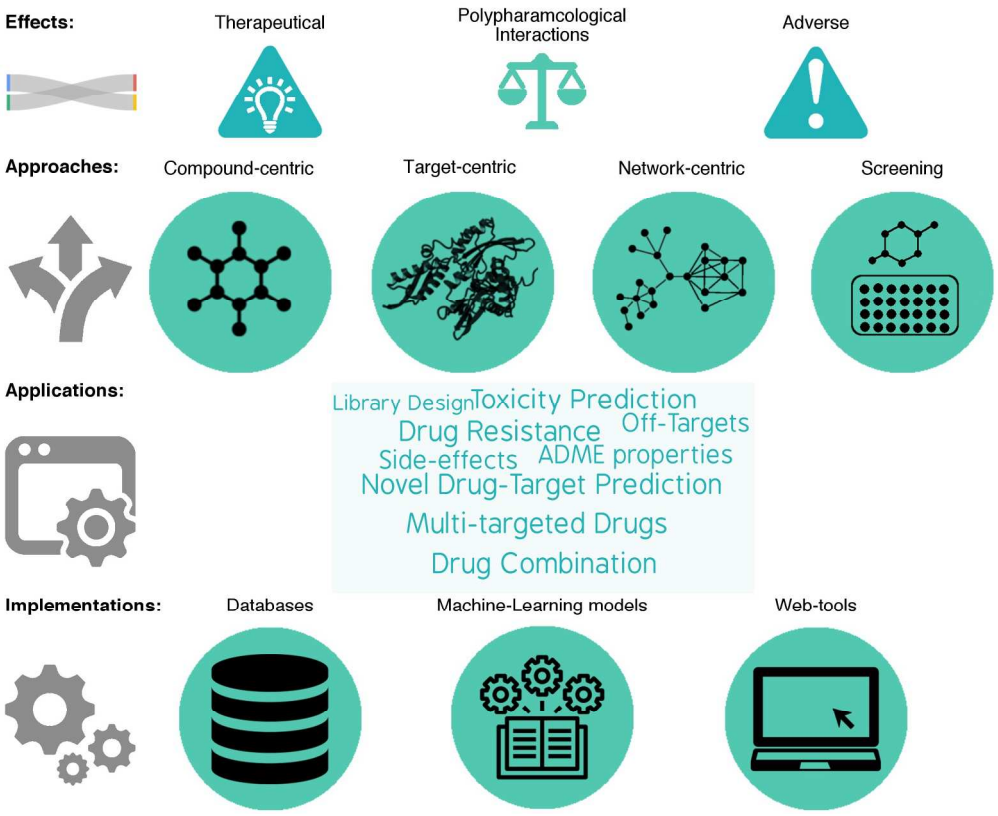
CSNAP	A network-centric approach that estimates the chemical similarly network of target profiles to identify compounds for cell-line based screening study.	http://services.mbi.ucla.edu/CSNAP/
C-SPADE	A web application to visualized compound similarity dendrogram augmented with their profiling information, suitable for target-based, cell-based and other phenotypic screening data.	https://cspade.fimm.fi/
SynergyFinder	A web-tool that estimates combination synergy in dose-response matrix studies by implementing dose-response and effect-based scoring functions for drug combination study.	https://synergyfinder.fimm.fi/
systemsDock	A network pharmacological based approach that implements novel docking strategies to predict drug-target interactions.	http://systemsdock.unit.oist.jp/iddp/home/index
DT-Web	A web-based application to predict novel drug-target interaction and drug combinations.	https://alpha.dmi.unict.it/dtweb/index.php
DSEA	Drug set Enrichment Analysis based on drug induced gene expression profiles to identify molecular pathway targeted by drugs.	http://dsea.tigem.it/
IntSide	A web-tool that integrates chemical and biological similarity to explore the molecular processes underlying drugs side-effects.	http://intside.irbbarcelona.org/
idTarget	Identifies molecular targets for small molecules using a divide-and-conquer docking approach.	http://idtarget.rcas.sinica.edu.tw/
cDRUG	A web-server that is developed from the NCI60 database to predict the anticancer efficacy of chemical compounds.	http://bsb.kiz.ac.cn/CDRUG
iDrug	A comprehensive application that aids pharmacophore searching, identifying potential molecular targets, elucidating protein binding sites and 3D similarity estimation.	http://lilab.ecust.edu.cn/idrug/doc.html
ProBis	A web-application that uses protein binding site similarities to predict novel ligand interactions.	http://probis.cmm.ki.si/ligands/
SPiDER	Uses self-organizing consensus to identify molecular targets of compounds or de novo designed molecules.	http://modlab-cadd.ethz.ch/software/spider/

TarPred	A web-tool to predict the side-effects of compounds along with the associated disease indications.	http://www.dddc.ac.cn/tarpred/

Figure legends:

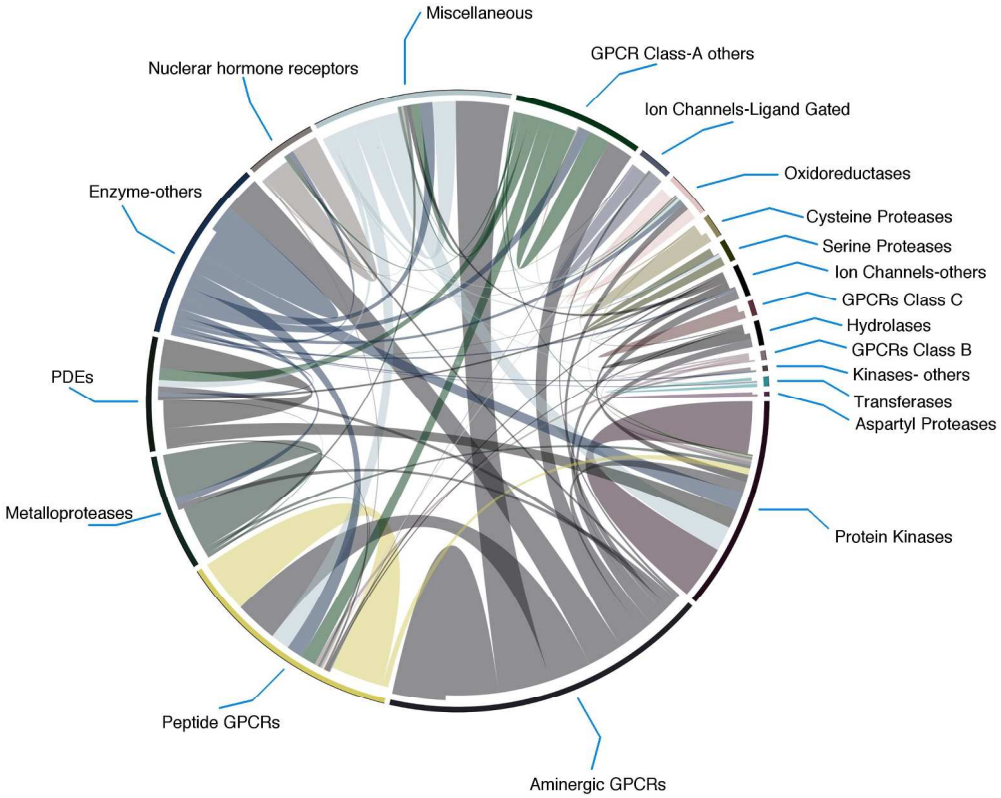
Figure 1. A schematic illustration of the various approaches to polypharmacology, along with their applications and implementations.

Figure 2. A chord diagram re-illustrating global polypharmacological interactions as depicted by Paolini et al. [81]. The degree of promiscuity is represented as the area of the chord corresponding to each family, and the links indicate the range of promiscuity shared between the intra- and inter-target families.



A schematic illustration of the various approaches to polypharmacology, along with their applications and implementations.

192x157mm (300 x 300 DPI)



A chord diagram re-illustrating global polypharmacological interactions as depicted by Paolini et al. [81]. The degree of promiscuity is represented as the area of the chord corresponding to each family, and the links indicate the range of promiscuity shared between the intra- and inter-target families.

208x166mm (300 x 300 DPI)