W326–W335 *Nucleic Acids Research, 2021, Vol. 49, Web Server issue* *Published online 22 May 2021*
https://doi.org/10.1093/nar/gkab385

LigAdvisor: a versatile and user-friendly web-platform for drug design

Luca Pinzi[1], Annachiara Tinivella[1,2], Luca Gagliardelli[3], Domenico Beneventano[3] and Giulio Rastelli[1,*]

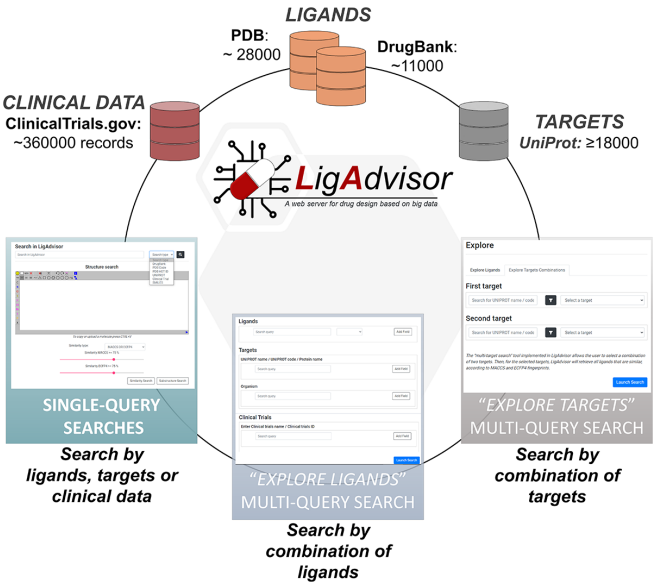
[1]Department of Life Sciences, University of Modena and Reggio Emilia, Modena 41125, Italy, [2]Clinical and Experimental Medicine, PhD Program, University of Modena and Reggio Emilia, Modena 41125, Italy and [3]Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Modena 41125, Italy

Received March 12, 2021; Revised April 19, 2021; Editorial Decision April 24, 2021; Accepted April 27, 2021

ABSTRACT

Although several tools facilitating *in silico* drug design are available, their results are usually difficult to integrate with publicly available information or require further processing to be fully exploited. The rational design of multi-target ligands (polypharmacology) and the repositioning of known drugs towards unmet therapeutic needs (drug repurposing) have raised increasing attention in drug discovery, although they usually require careful planning of tailored drug design strategies. Computational tools and data-driven approaches can help to reveal novel valuable opportunities in these contexts, as they enable to efficiently mine publicly available chemical, biological, clinical, and disease-related data. Based on these premises, we developed LigAdvisor, a data-driven webserver which integrates information reported in DrugBank, Protein Data Bank, UniProt, Clinical Trials and Therapeutic Target Database into an intuitive platform, to facilitate drug discovery tasks as drug repurposing, polypharmacology, target fishing and profiling. As designed, LigAdvisor enables easy integration of similarity estimation results with clinical data, thereby allowing a more efficient exploitation of information in different drug discovery contexts. Users can also develop customizable drug design tasks on their own molecules, by means of ligand- and target-based search modes, and download their results. LigAdvisor is publicly available at https://ligadvisor.unimore.it/.

GRAPHICAL ABSTRACT



INTRODUCTION

Drug discovery traditionally relied on the identification of compounds that selectively hit a single biological target of interest, while limiting adverse reactions. However, it is currently established that drugs very often exert their effects through the simultaneous modulation of complex networks of targets. The possibility of modulating selected combinations of targets involved in the physiopathology of multifactorial diseases to achieve improved therapeutic effects paved the way to polypharmacology (1), which nowadays is primarily pursued through *in silico* approaches (2). Similarly, repositioning approved drugs and clinically safe candidates that did not meet their primary expectations towards unmet therapeutic needs represents an opportunity that has recently gained considerable interest in drug discovery (3–5). Indeed, such an approach makes it possible to

*To whom correspondence should be addressed. Tel: +39 0592058564; Email: giulio.rastelli@unimore.it

sensibly reduce risks and costs that can usually be associated to the drug discovery process, although repurposing opportunities have mostly been discovered through retrospective analyses (3). In this context, computational tools and web platforms facilitating these drug discovery tasks have flourished in recent times, by taking advantage of data-driven approaches and big data (6–12). Common drug repositioning strategies include the identification of novel potential targets by means of their similarity with drugs already in use in therapeutic regimens for a specific disease. Strategies based on *in silico* structure-based approaches (*e.g.*, docking (13)), have also being pursued for the same purposes, although the integration of different methods is expected to provide improved drug repurposing predictions (14).

In a recent study (15), we performed extensive *in silico* investigations within the Protein Data Bank (PDB) (16), with the aim of identifying promising target combinations for polypharmacology. The results of this study allowed us to conclude that ligands reported within the PDB represent a reliable source of structural and biological information for multi-target drug design and repurposing (15). Moreover, we also discussed on how the chemical space covered by crystallographic ligands in the PDB is significantly less populated by pan-assay interference (PAINS) and potential false positive compounds, even if it is lower in size with respect to other bioactivity-oriented databases (*e.g.*, PubChem (17) and ChEMBL (18), these aspects being of crucial interest in target-driven drug discovery.

Based on these premises, we developed LigAdvisor (https://ligadvisor.unimore.it/), a webserver that allows users to perform 2D fingerprint-based similarity analyses on databases of compounds curated from DrugBank (19) and PDB (16), and to integrate the results with data reported in clinical trials (https://clinicaltrials.gov/), UniProt (20) and TTD (Therapeutic Target Database) (21). Indeed, LigAdvisor is primarily designed to assist in the identification of novel potential targets for already known drugs or clinical candidates, by means of their similarity with reported crystallographic ligands, and *vice versa*. This information is expected to help in the identification of potential biological activities for a (set of) ligand(s) of interest on a wider range of targets, compared to other drug repurposing or multi-target ligand webservers (*e.g.*, 8,9,12), thereby significantly improving the repositioning opportunities. Moreover, it provides integration of disease information and involvement in shared biological pathways for a given pair of targets, which is an essential aspect for polypharmacology design. A comparison of the potentialities offered by LigAdvisor with respect to other currently available drug repurposing and polypharmacology related websites is provided in the 'LigAdvisor and current webservers for drug repurposing and polypharmacology' section (see below).

Another important strength of LigAdvisor is the possibility to link similarity data with information available from PDB crystallographic compounds and their respective protein-ligand complexes. Indeed, this information can facilitate the interpretation of the mode of binding and structure-activity relationships across one or more targets for a query molecule, which is a key advantage in polypharmacology and drug discovery in general. Remarkably, LigAdvisor allows the integration of results of clinical trial information on the same or similar compounds, if present. Moreover, LigAdvisor can be used to analyse fully customizable molecular structures to identify potential molecular targets of already synthesized libraries of chemical compounds, which can be easily uploaded by users for the analyses. Indeed, these compounds are of great value for drug discovery, as they help to circumvent potential feasibility issues often encountered when synthesizing novel chemical scaffolds (22).

LigAdvisor is designed to also help non-expert in the field of *in silico* drug discovery in the identification of potential multi-target activities and repurposing opportunities of their compounds of interest. Indeed, our website features a fast and intuitive interface that does not require any programming knowledge, and allows users to download search results as easy-to-interpret graphs and tables. At the same time, LigAdvisor results can be post-processed by high-end users *via* more sophisticated refinement and rescoring algorithms, for example, through further *in silico* analyses (*e.g.*, docking or free energy evaluations), since the crystallographic compounds in the database have already been reported in complex with their primary targets. For example, these analyses may help in the identification of suitable combination of targets and their respective crystallographic conformations for further structure-based investigations, in both multi-target drug design and integrated *in silico* drug repurposing campaigns (13–15,23).

Another key strength of LigAdvisor is the possibility to identify targets potentially involved in common and/or similar therapeutic indications by means of the similarity between their ligands, as well as structurally similar compounds for a selected target pair of interests, through unique multi-query search runs.

LigAdvisor can be queried through different search types, which can be either ligand- or target-oriented, to enable a more effective exploitation of the advantages of both approaches. Importantly, another strength is the implementation of additional query functionality that allows researchers to combine LigAdvisor results with searches on clinical trials data. To the best of our knowledge, these types of database explorations have never been implemented together into drug design webservers. The possibility to interrogate LigAdvisor through such multi-query searches is expected to significantly facilitate and speed-up drug repurposing, target fishing and polypharmacology tasks.

In this manuscript, we will describe the curated datasets, the information employed in the LigAdvisor webserver, and the construction of the related user-friendly interfaces. We will also illustrate the different types of search implemented into the webserver, and how they can be used in multiple drug design tasks, with particular focus on polypharmacology, target fishing and drug repurposing tasks. Finally, we will discuss selected case studies to better highlight the potential offered by the LigAdvisor webserver in different drug design and discovery contexts.

MATERIALS AND METHODS

Molecular datasets preparation

Two different datasets of compounds curated from Protein Data Bank (16) and DrugBank (19) were prepared and

used as the basis for the searches implemented in LigAdvisor. In particular, ligands reported in crystallographic complexes were first downloaded from the Protein Data Bank (33 391 compounds, accessed on 25 November 2020). Then, compounds were filtered to retain ligands with 5–55 heavy atoms. Moreover, compounds corresponding to small peptides and crystallographic solvents were also removed, leading to a final database of 28 415 unique small molecule entries, crystallized across 237 374 different protein component chains. Chain-specific information was retained in this phase of dataset preparation, as ligands can potentially adopt different conformation, protonation and tautomerization states across different chains, also within the same protein complex. At the same time, ligands can also bind only to specific component chains in multi-protein crystallographic complexes. For each PDB ligand, information about crystal complexes were downloaded from PDB and matched with their respective molecular targets through the UniProt ID (20).

The overall publicly available dataset of DrugBank was downloaded and directly employed in the analyses, as it contains highly characterized ligands (10 767 compounds, accessed on 25 November 2020).

The Pandas Python library (version 1.1.3) available in Anaconda (https://anaconda.com/, version 2-2-4.0) was used to cross-reference PDB-related information to data on targets, and their associated diseases and pathways reported in the UniProt database (accessed on 25 November 2020) and TTD (accessed on 14 April 2021) (21), respectively. Moreover, the PubChemPy library (https://github.com/mcs07/PubChemPy, version 1.0.4, accessed on 25 November 2020) was used to programmatically retrieve from the PubChem database (17) molecules synonyms and identifiers across several established online platforms and databases (*e.g.*, PubChem, PDB, DrugBank, ChEMBL and others), to be associated with clinical data (16–19). Additional brand names of drugs and candidates, as well as their associations, were extracted from the FDA website (https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm, accessed on 25 November 2020), and DrugBank (https://go.drugbank.com/releases/latest#open-data, accessed on 25 November 2020) to further facilitate their association with clinical data.

Clinical data preparation

Data related to clinical trials was first downloaded from (https://clinicaltrials.gov/, accessed on 25 November 2020) and processed by using the Pandas Python library (version 1.1.3) available in Anaconda. This allowed to extract studies identifiers and information about their status, phase, type, assessed therapeutic condition and outcome description. Then, a KNIME (version 4.3.1) (24) workflow was devised to match drugs and candidate synonyms, with information reported in the pre-processed clinical data. This procedure allowed us to associate each clinical trial with the drug identifiers and candidate compounds on other databases, both alone and in combination, for a total of 359 341 records.

Similarity estimations

Similarities between ligands are evaluated by means of MACCS and Morgan (ECFP) fingerprints, implemented in the RDKit Python libraries (https://www.rdkit.org, release 2020-9, accessed on 22 February 2020). Default parameters were used for the generation of MACCS fingerprints, while a bit length of 1024 and a radius equal to 2 were employed for the Morgan fingerprints (ECFP4). The 2D similarity between ligands was evaluated in terms of the Tanimoto score (25). Pairs of compounds with similarity scores below the commonly accepted thresholds of 0.8 (MACCS) or 0.3 (Morgan) were discarded (26). Similarities between DrugBank and PDB ligands, which currently cover >2.5 million records, are pre-computed to enable fast retrieving of results from the webserver. Moreover, the obtained similarities allowed to identify 13 million target associations, which can be investigated for multi-target drug design. Statistics related to pre-calculated similarities and target combinations are reported in Figure 1. Further database updates are planned throughout the year, as more data becomes available, to increase usability of LigAdvisor for drug design.

Web implementation

The database used in the webserver was generated using PostgreSQL (https://www.postgresql.org/, version 12.6), while all web interfaces were developed using Angular (https://angular.io/, version 9.1.0).

The server-side application was developed using Python (https://www.python.org/download/releases/2.6/, version 2.6.12) with Flask (https://flask.palletsprojects.com/, version 1.1.2). The LigAdvisor server is hosted on a Linux server (Ubuntu 20.04) equipped with Intel Xeon processors (model E5-2603). To handle the requests on LigAdvisor server Nginx (https://nginx.org/en/, version 1.18) with Gunicorn (https://gunicorn.org/) is used. The LigAdvisor web application was successfully tested on various platforms, including Windows and Mac OS.

THE LIGADVISOR WEBSERVER

The LigAdvisor web platform relies on manually curated databases of ligands extracted from DrugBank, Protein Data Bank, and data related to clinical trials and target-disease associations. Together, these data are expected to assist users in drug design operations and suggest, *e.g.* potential targets for known molecules, *ad hoc* multi-target activities, and drug repurposing.

As designed, LigAdvisor allows users to perform single-query or multi-query searches, as described below. Key examples and step-by-step instructions on how to perform these analyses in LigAdvisor are provided in the downloadable tutorial in the '*Help and tutorial*' section of the website (available at: https://ligadvisor.unimore.it/tutorial), and in the Supporting Material of this manuscript.

Single-query searches implemented in LigAdvisor

Single-query searches can be performed in LigAdvisor through the dedicated '*Search in LigAdvisor*' panel in the

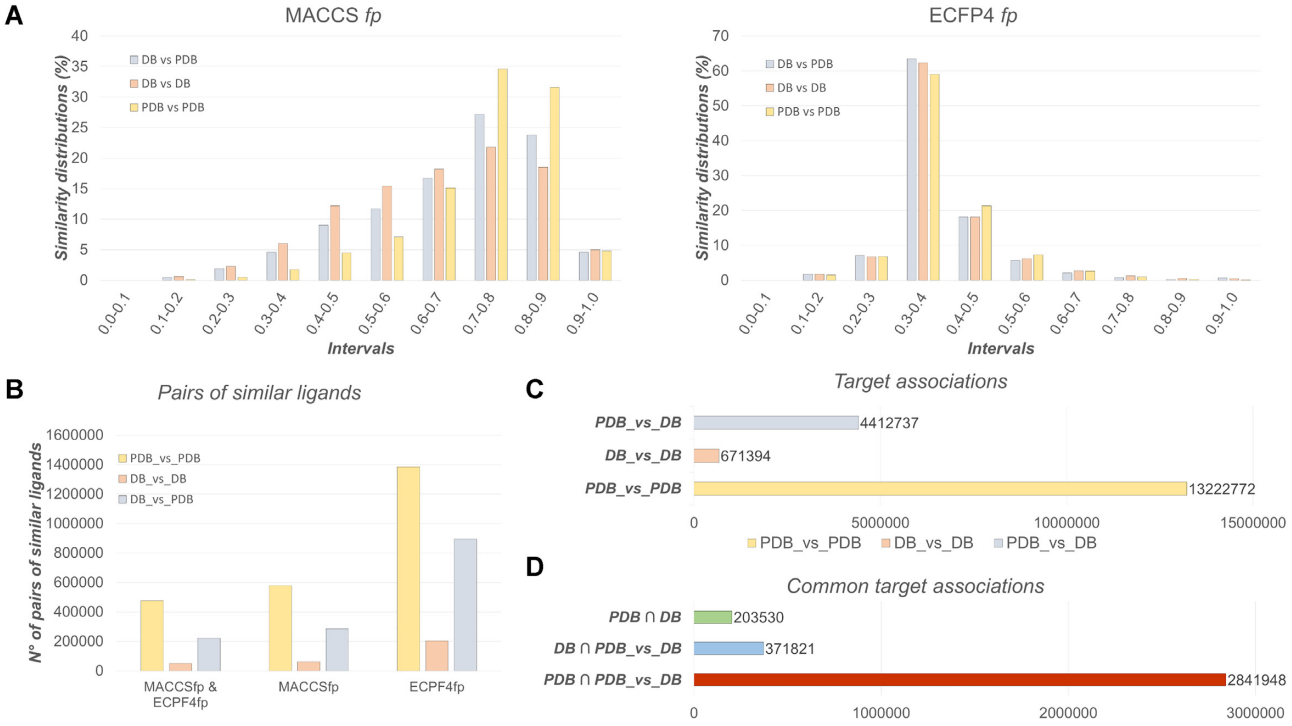


Figure 1. Statistics of ligand and target associations identified by similarity calculations on the LigAdvisor datasets. LigAdvisor includes data related to 28 415 and 10 767 unique ligands, extracted and curated from PDB and DrugBank databases, respectively. Similarity calculations performed on these curated datasets collectively provided more than 2.5 million records (*i.e.* >1.5 mln similarities from the comparison of ligands curated from the PDB; >0.2 mln similarities by ligands curated from the DrugBank, and; >0.9 from the comparison of PDB-curated ligands versus DrugBank compounds). Only similarity records with MACCS and ECFP4 (Morgan) Tanimoto scores higher than 0.8 and 0.3 were retained, respectively. Interestingly, similarities computed across the different databases provided similar trends, both in terms of MACCS (panel **A**) and ECFP4 (panel **B**). Moreover, as expected, the comparison of ligands curated from the PDB provided higher numbers, both in terms of similarity and target associations records (panels **C** and **D**, respectively), compared to the similarity estimations with those from DrugBank-related ligands. In LigAdvisor, targets are identified by means of their UniProt identifiers (>18 000 targets, mapped by PDB- and DrugBank-related compounds). Of note, several target associations resulted from the comparison of different datasets, *i.e.* the same pair of targets emerged from both the comparison of PDB ligands and DrugBank compounds (panel **D**). Indeed, these target combinations potentially represent the most valuable source for the development of polypharmacology ligands, as they are associated with crystallographic information and related biological information from literature.

homepage, which allows users to retrieve information about a specific drug, clinical or pre-clinical candidate, crystallographic ligand, target or therapeutic indication already included in the curated database. Moreover, searches on custom molecular structures can also be performed through this panel, by manually sketching them into a dedicated input box. These types of searches are particularly useful for users interested in target fishing, polypharmacology, drug repurposing and off-target predictions.

Single-query similarity analyses can be performed within the curated datasets by selecting different combinations of MACCS and ECFP4 (Morgan) score thresholds (expressed as percentage of Tanimoto similarity) to enable fine tuning of the results. In addition, searches based on the presence of substructures of interests are also implemented. LigAdvisor accepts several types of ligand queries, including SMILES identifiers and molecular structures drawn into a dedicated input box implemented in the JMSE tool (27) (Figure 2A).

The '*Search in LigAdvisor*' panel is designed to also allow users to query the database for a specific target or clinical data of interest. In this case, target and clinical data queries should be supplied to the webserver as target names, UniProt IDs, or ClinicalTrials.gov identifiers (NCT number), into the dedicated input text box (Figure 2A).

The output of single-query searches provides a summary page with statistics (*i.e.*, the number of reported and similar compounds for target and ligand queries, respectively), and identifiers with hyperlinks to internal LigAdvisor pages, if the queried compound is already included in the database (Figure 3A). Similar ligands and confirmed targets are also reported for compounds that are curated in the databases, along with their similarity scores and statistics. For custom ligand searches, a list of structurally close compounds is reported, along with their similarity scores. Selecting a record from the resulting list returns information on the related ligand in the database, as previously described. Target-based searches return information on the query target, together with a list of its associated crystallographic ligands, clinically or pre-clinically investigated compounds and approved drugs. Finally, searches based on NCT identifiers report information on clinical trials data, such as associated drugs identifiers and status of approval. This type of search is particularly helpful for drug repositioning and target fishing, and to analyse target- and clinical-related data. Remarkably, the single-query search implemented in LigAdvisor enables to query a large, high-quality dataset, comprising >30 000 unique ligands, 18 000 experimentally confirmed targets and 350 000 clinical data records. To the best

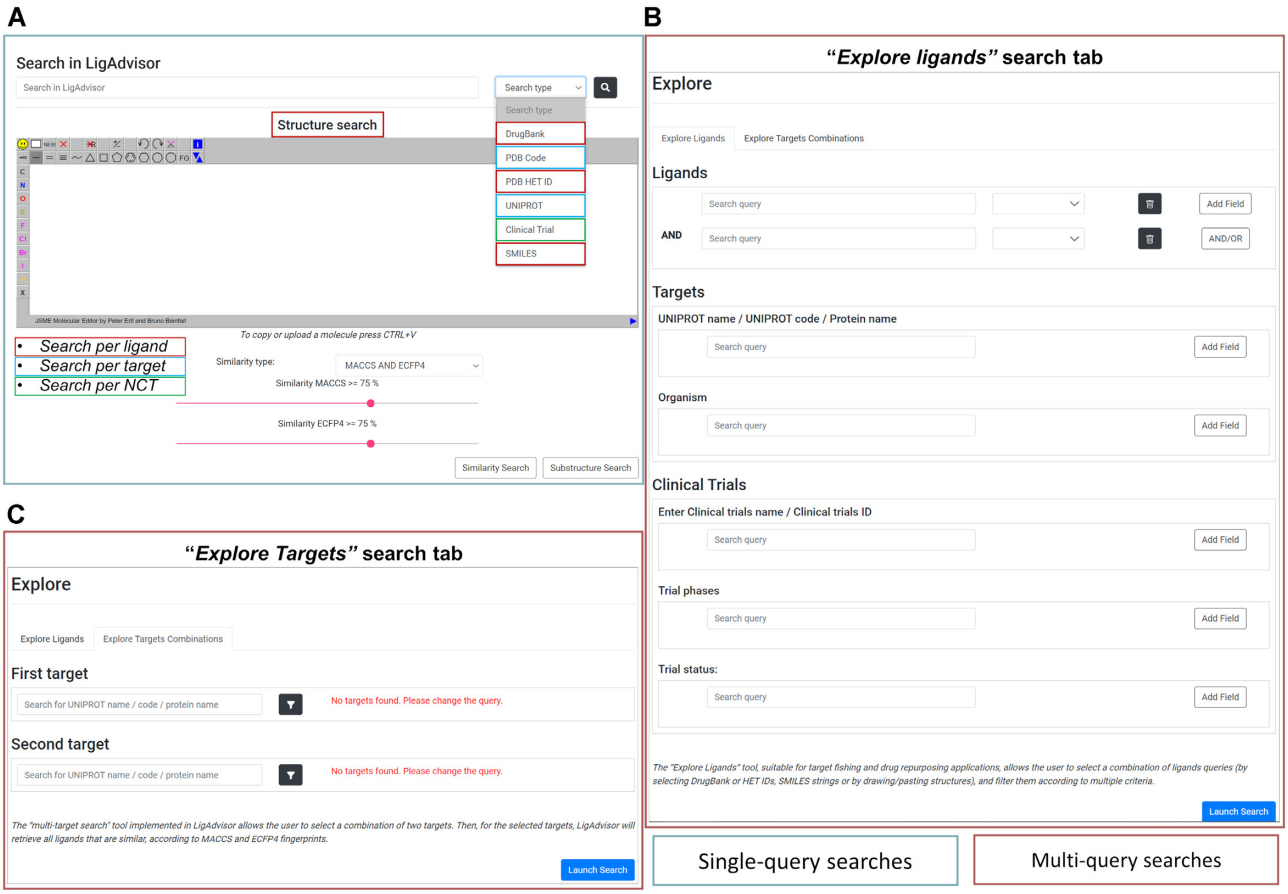


Figure 2. The three main LigAdvisor input search interfaces. Panel (**A**) reports the input interface of the 'Single-query searches implemented in LigAdvisor', which is accessible through the web site homepage. Through this window, users can perform ligand- and target-based searches. In particular, ligand-based searches can be performed by means of PDB and DrugBank identifiers, or SMILES. Moreover, molecular structures can also be drawn through a dedicated input panel. Target-related information and clinical data can be queried by means of UniProt and NCT identifiers, respectively. Panel (**B**) shows the 'Explore ligands' search interface, through which users can query the webserver for information related to a set of ligands. Further refinements of the multi-ligands query can be applied to only retain information related to specific targets and/or clinical data. Panel (**C**) reports an example of the '*Explore targets*' search interface. Through this window, users can search for information related to similar ligands between pairs of selected targets.

of our knowledge, the number of records available from this search is far higher than those of other databases and webservers focussing on drug repurposing and polypharmacology.

The '*Explore ligands*' search tab

The '*Explore ligands*' search tab is designed to query the webserver for data associated to multiple ligands simultaneously. This type of search is particularly useful, for example, when a user is interested in identifying targets potentially in common between a set of ligands, and retrieve their related clinical data. Indeed, ligands with similar therapeutic effects may derive from similar biological profiles (28). Moreover, this type of search is also suitable for target fishing and drug repurposing applications. The queries can be selected directly *via* drug names, DrugBank IDs and HET IDs of the investigated ligands (Figure 2B). Moreover, fully customizable ligand queries can also be performed through their SMILES identifiers, by typing them into an input text box, or by directly drawing the molecules in the dedicated toolbox, as described above. Multi-ligand queries in the '*Explore ligands*' search tab can also be refined to restrict the results on specific subsets of targets and/or clinical data, thereby allowing the user to focus on a particular therapeutic indication of interest.

The output of '*Explore ligands*' searches reports a list of molecular targets, ligand similarities and clinical data, organized in tabs along with their statistics, to enable fast and easy navigation across different data types in the database (Figure 3B). The records in each tab ('*DrugBank*' and '*PDB*') can be separately queried to analyse the information associated to targets and ligands resulting from the search. Confirmed targets and related clinical data for similar compounds are reported for searches on ligands curated from DrugBank, while ligands in their complexes and co-crystallized targets are reported for records related to PDB ligands. The list of molecular targets resulting from this type of search is ranked in descending order, according to the number of reported similarity records. Indeed, targets with the highest number of ligands in common are expected to be more likely to share a wider overlapping chemical space, and thus being more suitable for a polypharmacology ap-

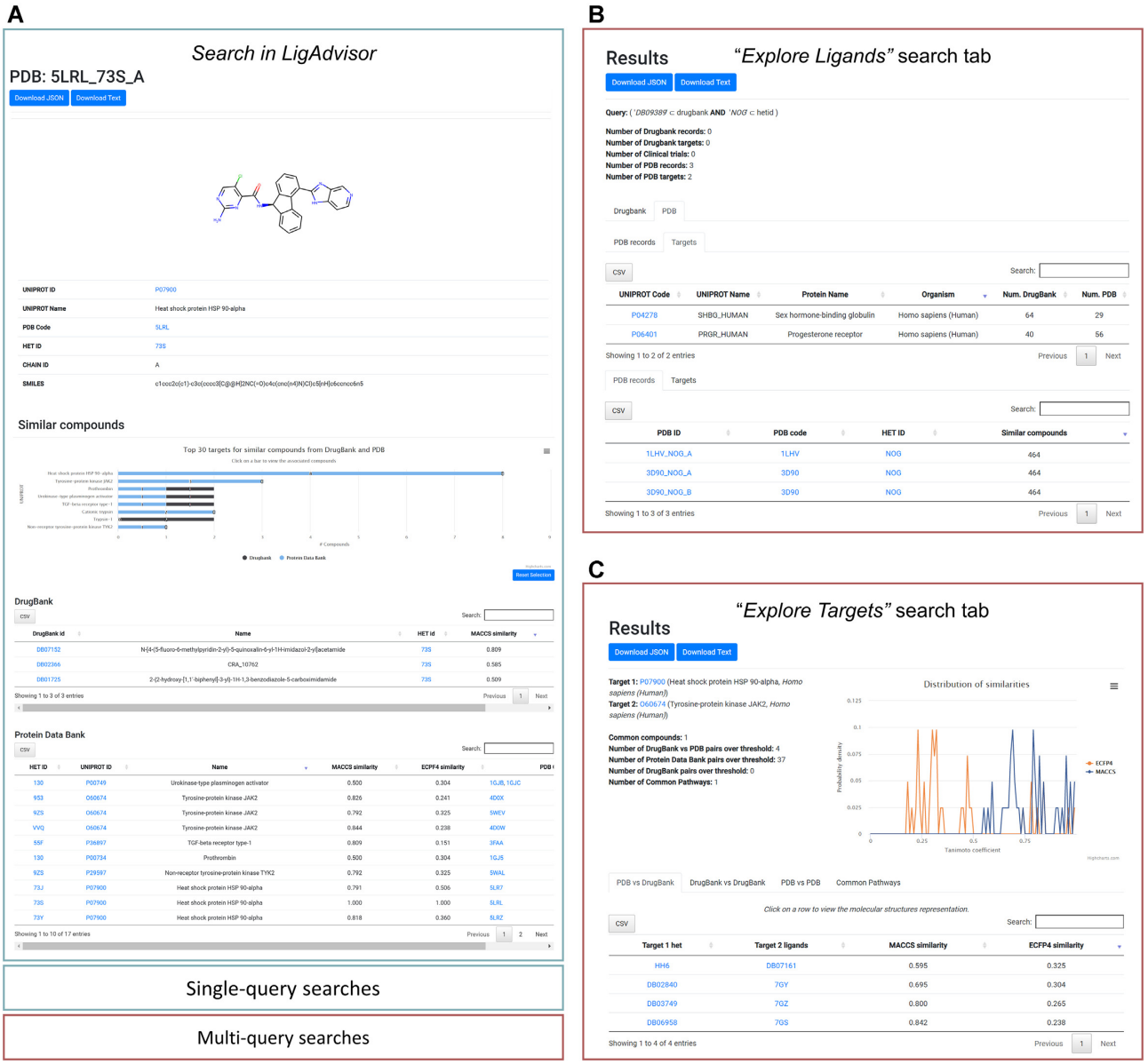


Figure 3. Examples of the results potentially obtainable by single- and multi-query searches implemented in LigAdvisor. In particular, panel (**A**) reports an example of results obtained by means of the ligand-based single-query search. As shown, the result page first provides information on targets and ligands identifiers across external databases. Clinical data is also reported, if present. Then, it reports a list of all drugs and clinical candidates ('*DrugBank*' tab), and crystallographic compounds ('*Protein Data Bank*' tab) that resulted to be similar to the query, along with their MACCS and ECFP4 scores. Panel (**B**) shows the results that can be visualized for the '*Explore ligands*' multi-query search. In this case, overall statistics related to ligands similarities are reported in the upper-left part of the page. Moreover, ligands- and targets-related records are also reported as lists at the bottom of the page, the selection of each record allowing users to access to its specific information. Panel (**C**) reports the results of the '*Explore targets*' search. In particular, the overall statistics of similarities observed between ligands of the selected targets and involvement in common pathways are reported in the upper-left part of the page, while a graphical representation of the similarity score distributions is shown in the upper-right corner. The observed ligand similarities are listed according to datasets comparison. By clicking on each similarity record in the tabs, users can visually inspect the chemical structures of related ligands. Moreover, records related to shared biological pathways between the targets of interest are also reported ('*Common pathways*' tab).

proach. Moreover, they also have the highest probability to be the most suitable targets for drug repositioning and target fishing.

The '*Explore Targets Combinations*' search tab

Searches performed through the '*Explore Targets Combinations*' tab allow users to compare the chemical spaces covered by ligands reported within the database, for pairs of targets of interest. This type of search is expected to help identify novel potential multi-target ligands, based on the concept that structurally similar compounds could have similar biological activities (29). These results could also be useful to identify the more suitable protein conformation(s) for subsequent structure-based *in silico* polypharmacology design, as previously reported (13,15,23). Targets can be

queried through their UniProt identifiers or names *via* a drop-down menu, which was designed to dynamically restrict the search upon the keyword typed in the dedicated input text box (Figure 2C).

The search returns statistics related to the targets of interests, both in terms of retrieved number of similar compounds above threshold, and involvement in shared biological pathways according to TTD data (21). Moreover, a graphical representation of the similarity distribution of compounds is also reported (Figure 3C). Results of the '*Explore Targets Combinations*' search are organized in different tabs, *i.e.* comparison of ligands curated only from PDB (reported in the '*PDB vs PDB*' tab), comparison of PDB versus DrugBank molecules ('*PDB vs DrugBank*') and comparison of DrugBank compounds ('*DrugBank vs DrugBank*'). In each tab, a list of compound pairs emerging from similarity analyses, sorted by descending ECFP4 (Morgan Fingerprint) score, is reported. By selecting (clicking on) any row, the user can visually inspect the chemical structures of the related ligand pairs. Moreover, the occurrence of any shared biological pathways for the selected target pair is also reported in the '*Common pathways*' tab. We added this latter section into the '*Explore Targets Combinations*' multi-query search to facilitate users in identifying target pairs of potentially higher therapeutic interest.

Results obtained from the different searches can be collectively downloaded as .txt and .json files to facilitate post-processing analyses by both high-qualified and non-expert users on local devices. Moreover, the results reported into different tabs can also be separately downloaded as a .csv files for more focussed investigations.

LigAdvisor and current webservers for drug repurposing and polypharmacology

In this section, we will provide the reader with an overview of available servers in the field of repurposing and polypharmacology, and how LigAdvisor fits into this context.

Currently, other databases specifically focussing on drugs, their therapeutic indications, bioactivity data and drug–target interactions have been reported to facilitate drug repurposing (8,9,11,12) (Supplementary Table S1). For example, the ReframeDB (8) and repoDB portals (12) provide information on around 12 000 compounds collected from published journals and patents with disease annotations. While these resources enable ligand repurposing through their similarity with known drugs, results can require further processing to be efficiently exploited. On the contrary, LigAdvisor provides direct annotations on more than 18 000 targets, from 28 415 PDB and 10 767 DrugBank unique ligands, resulting in a far higher number of ligand-targets associations with respect to current databases of repurposed drugs. Repurposing-related webservers entrusting on larger datasets of ligands are also reported. For example, the MuSSel website (30) facilitates the identification of novel putative targets for small molecules, based on the similarity with ChEMBL compounds. Moreover, the Promiscuous 2.0 webserver (31), which provides access to a huge reservoir of well-organised and integrated information, has also been recently helping scientists in drug repositioning and profiling tasks. The possibility to exploit such larger databases would certainly provide higher chances of repositioning for the compounds. However, the presence of potential false positive compounds on larger databases of bioactivity data should not be neglected. Indeed, results can often differ upon the adopted experimental conditions in the assays, making their mechanism of action not fully interpretable (32). One key strength of LigAdvisor in this regard is the ability to provide links to the 3D structural information available from ligand-protein complexes in the PDB. This is expected to facilitate the identification of how ligands bind to their novel putative targets. The ability to perform multi-query searches is another key asset of LigAdvisor for drug repurposing, especially when multiple compounds belonging to the same family scaffold have to be investigated. Notably, this functionality has never been embedded in a webserver focussing on drug repurposing and polypharmacology. Other repurposing-related webservers include the Drug ReposER portal (33) and PharmMapper (34), which enable the repurposing of compounds through the sub-structural similarity between the binding site interfaces of known drugs in their targets and pharmacophore mapping, respectively. Although these webtools provide valuable approaches to repurpose compounds towards novel drug targets, many of them rely on smaller datasets or do not allow a direct integration of their results with clinical data, as opposed to LigAdvisor. Programs and webservers developed to facilitate the identification of targets for ligands according to binding site alignment, or on the similarity of: (i) molecular structures; (ii) ligand-protein interactions patterns, and; (iii) residue sequences have also been recently reported (35–38). These webservers would represent valuable complementary tools to integrate for drug repurposing and polypharmacology, although their results very often require further processing to be fully exploited. Similar considerations can also be drawn for webservers specifically focussing on multi-target drug design or polypharmacology profiling (10,39). These include, for example, the Polypharma webserver (39), which allows to depict the putative polypharmacological profile of molecules by means of their similarity with crystallographic ligands, although their analyses are focused on single ligand or target queries, and do not provide clinical data information. Finally, web portals that provide drug combinations experimentally studied, such as DrugComb (40), are also published. Although these latter databases provide information of high interest for the development of multi-target ligands, they are very often focussed on cancer research. On the contrary, LigAdvisor can provide information on common pathways and mutual involvement on a wide range of diseases for any given pair of targets, thanks to its integration with TTD data. This information is of high value for both polypharmacology *de novo* design and profiling.

CASE STUDIES

Information and examples on the different types of searches that can be performed with LigAdvisor are provided in the website (https://ligadvisor.unimore.it/tutorial, accessed on 17 April 2021), and Supporting Material. The tutorial introduces users to the different types of single- and multi-

query searches implemented in LigAdvisor, through a series of dedicated examples.

In particular, the reader will be provided with representative ligand- and target-based case studies, which help in demonstrating how LigAdvisor can be used to facilitate polypharmacology and drug repositioning tasks. To this aim, validated target-disease associations were first extracted from the Therapeutic Target Database (TTD) (accessed on 6 March 2021) (21), and then integrated with information related to targets from the UniProt (20) database and similarity data from LigAdvisor. The analyses were focussed on 34 therapeutic indications, each associated with >10 different targets. A list of the investigated indications, along with their molecular targets is reported in Supplementary Table S2 (see the Supplementary Material). Data from LigAdvisor was filtered to retain only similarity records for ligands with confirmed activity on targets related to the same TTD disease indication. This allowed to identify 680 target associations across the selected 34 therapeutic indications. We are aware that not all identified target associations with pairs of structurally related ligands may have a therapeutic significance in the selected disease indications. Thus, a close analysis of the obtained data allowed the identification of target pairs that gave excellent levels of similarity. These pairs of targets represent valuable starting points for the design of multi-target ligands. Hereafter, two of the identified associations will be discussed in detail for such purpose. A first example selected among the obtained pairs comes from the association of two cancer related targets, *i.e.*, Heat shock protein 90 (Hsp90, UniProt ID: P07900) and the JAK2 Tyrosine-protein kinase (UniProt ID: O60674), which have been previously validated as a potential synergistic combination for the treatment of different tumour types (41–43). Interestingly, this pair of targets provided 31 similarity records from 27 Hsp90 ligands and 17 JAK2 compounds, belonging to different chemotypes and known binding mode (see Supplementary Table S3 in the Supporting Material). Among the identified pairs of Hsp90- and JAK2-ligands, an example of particular interest comes from the comparison of the crystallographic compounds with HET IDs 73S (in complex with Hsp90, PDB code 5LRL, DOI: 10.2210/pdb5LRL/pdb) and VVQ (crystallized with JAK2, PDB code: 4D0W) (44), which provided similarity scores of 0.884 (MACCS) and 0.238 (ECFP4). Supplementary Figure S1 highlights the 2D structure of the selected pair of ligands, and the experimentally observed binding mode in their respective targets. Interestingly, both compounds present a 2-amino-pyrimidine core scaffold through which they interact with key residues in the Hsp90 and JAK2 proteins (*i.e.*, Asp93, and Glu930 and Leu932, respectively). Indeed, these residues are often engaged in H-bond interactions with Hsp90 and JAK2 selective inhibitors, thus representing valuable common anchoring points for the development of dual ligands. The possibility to exploit common anchoring points and substructures provides advantages in the development of multi-target ligands, as they facilitate the integration of the different structural requirements of the targets of interest into a single drug-like entity. This is particularly useful for the design of compounds acting on targets that belong to different protein families, such as JAK2 and Hsp90, for which ligands with balanced dual activity have not been reported yet. Similar considerations can also be drawn for the association of progesterone receptor (UniProt ID: P06401) and the aromatase enzyme (UniProt ID: P11511), for which recent findings demonstrated that their combined inhibition provides improved efficacy in the treatment of endometriosis (45,46). To evaluate whether this association could represent a valuable starting point for the development of dual ligands, we performed target-based searches in LigAdvisor, identifying significant similarities (see Supplementary Table S4 in the Supplementary Material), especially among their known drugs or crystallographic compounds. Notably, several scaffolds of the compounds in the identified pairs (see Supplementary Figure S2 in the Supplementary Material) have already been clinically investigated and may provide clues on the structural requirements for the development of potential multi-target ligands.

Finally, we report another application of our webserver related to Alzheimer's disease (AD), for which drug repositioning campaigns are currently being pursued (47). To this aim, we first extracted a list of 22 potential repurposing candidates that are under clinical investigation for AD, as recently reported by Jordan *et al.* (48), and associated to their respective DrugBank identifiers. Then, a series of analyses were performed on each of the investigated drug, through the ligand-based single-query search approach implemented in LigAdvisor. This allowed us to obtain a subset of structurally similar ligands, with their experimentally confirmed targets. Afterwards, the information related to molecular targets identified for these drugs was integrated with disease associations reported in TTD. This allowed us to evaluate whether LigAdvisor was able to retrieve therapeutic targets associated to AD. Notably, 8 of the 22 investigated drugs were associated to at least one target related to Alzheimer's disease (see Supplementary Table S5 in the Supplementary Material), according to LigAdvisor predictions. Moreover, six of these drugs were also associated to AD-related targets, for which their activity has yet to be experimentally confirmed. These latter associations are particularly appealing and may help in unveiling new potential targets for already known drugs. An example among the identified compounds comes from DB1166 (cilostazol), which was originally approved in 1999 as a PDE3 inhibitor (49). Interestingly, cilostazol demonstrated to have positive effects in attenuating cognitive decline (50), although its biomolecular mechanism of action has not been fully elucidated yet. The analyses performed in LigAdvisor associated cilostazol to several AD-related targets reported in TTD, *i.e.*, Muscarinic acetylcholine receptor M1 (UniProt ID: P11229), D(4) dopamine receptor (UniProt ID: P21917), 5-hydroxytryptamine receptor 2A (UniProt ID: P28223), Sodium-dependent serotonin transporter (UniProt ID: P31645), D(3) dopamine receptor (UniProt ID: P35462), cAMP-specific 3',5'-cyclic phosphodiesterase 4D (UniProt ID: Q08499) and Glutamate receptor ionotropic NMDA 2A (UniProt ID: Q12879), suggesting the possibility that this drug may act through polypharmacological mechanisms.

The examples reported above represent only part of the functionalities offered by LigAdvisor. As a final note, LigAdvisor may also aid in the selection of *ad hoc* receptor con-

formation(s) for structure-based multi-target calculations, and help target fishing tasks.

CONCLUSIONS

Herein we have presented LigAdvisor, an intuitive and user-friendly web platform that combines 2D chemical similarity estimations on high quality sets of ligands curated from DrugBank and PDB, and integrates it with information on targets and clinical data reported in the UniProt, TTD and ClinicalTrials.gov databases. LigAdvisor aims at facilitating the combination of chemical, biological, structural and clinical information reported for known ligands through an intuitive interface, and to assist users in a variety of drug design tasks, including target fishing, polypharmacology design and drug repurposing. LigAdvisor is designed to perform searches on a large, high quality database of ligands with reported clinical and preclinical information, as well as crystallographic data. Indeed, the single-query and multi-query searches implemented in LigAdvisor allow users to navigate through a large, high-quality dataset including >30 000 unique ligands, 18 000 experimentally validated targets, and 350 000 clinical data record. Notably, LigAdvisor is the first webserver that enables searches *per* combination of ligands and targets, to facilitate multi-target drug design and drug repurposing. Further efforts will be addressed to expand the chemical space of LigAdvisor, for example, by including libraries of natural products, which represent a valuable source of information for drug discovery (51).

DATA AVAILABILITY

The web server is available at https://ligadvisor.unimore.it/. This website requires no login, and is free and open to all users.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Andrea Scarpino for critical reading and helpful suggestions.

FUNDING

FAR – Fondo di Ateneo per la Ricerca 2019 [166835 of 2019/30/07 to G.R.]; PhD fellowship from the Regione Emilia Romagna on Data driven technologies for drug repurposing to A.T. Funding for open access charge: FAR – Fondo di Ateneo per la Ricerca 2019 [166835 of 2019/30/07].
Conflict of interest statement. None declared.

REFERENCES

1. Anighoro,A., Bajorath,J. and Rastelli,G. (2014) Polypharmacology: challenges and opportunities in drug discovery. *J. Med. Chem.*, **57**, 7874–7887.
2. Rastelli,G. and Pinzi,L. (2015) Computational polypharmacology comes of age. *Front. Pharmacol.*, **6**, 157.
3. Pushpakom,S., Iorio,F., Eyers,P.A., Escott,K.J., Hopper,S., Wells,A., Doig,A., Guilliams,T., Latimer,J., McNamee,C. *et al.* (2019) Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov.*, **18**, 41–58.
4. Cha,Y., Erez,T., Reynolds,I.J., Kumar,D., Ross,J., Koytiger,G., Kusko,R., Zeskind,B., Risso,S., Kagan,E. *et al.* (2018) Drug repurposing from the perspective of pharmaceutical companies. *Br. J. Pharmacol.*, **175**, 168–180.
5. Oprea,T.I., Bauman,J.E., Bologa,C.G., Buranda,T., Chigaev,A., Edwards,B.S., Jarvik,J.W., Gresham,H.D., Haynes,M.K., Hjelle,B. *et al.* (2011) Drug repurposing from an academic perspective. *Drug Discov. Today Ther. Strateg.*, **8**, 61–69.
6. Reddy,A.S., Tan,Z. and Zhang,S. (2014) Curation and analysis of multitargeting agents for polypharmacological modeling. *J. Chem. Inf. Model.*, **54**, 2536–2543.
7. Costa,F.F. (2014) Big data in biomedicine. *Drug Discov. Today*, **19**, 433–440.
8. Janes,J., Young,M.E., Chen,E., Rogers,N.H., Burgstaller-Muehlbacher,S., Hughes,L.D., Love,M.S., Hull,M.V., Kuhen,K.L., Woods,A.K. *et al.* (2018) The ReFRAME library as a comprehensive drug repurposing library and its application to the treatment of cryptosporidiosis. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 10750.
9. Corsello,S.M., Bittker,J.A., Liu,Z., Gould,J., McCarren,P., Hirschman,J.E., Johnston,S.E., Vrcic,A., Wong,B., Khan,M. *et al.* (2017) The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.*, **23**, 405–408.
10. Awale,M. and Reymond,J.-L. (2017) The polypharmacology browser: a web-based multi-fingerprint target prediction tool using ChEMBL bioactivity data. *J. Cheminform.*, **9**, 11.
11. Tanoli,Z., Seemab,U., Scherer,A., Wennerberg,K., Tang,J. and Vähä-Koskela,M. (2021) Exploration of databases and methods supporting drug repurposing: a comprehensive survey. *Brief. Bioinform.* **22**, 1656–1678.
12. Brown,A.S. and Patel,C.J. (2017) A standard database for drug repositioning. *Sci. Data*, **4**, 170029.
13. Pinzi,L. and Rastelli,G. (2019) Molecular docking: shifting paradigms in drug discovery. *Int. J. Mol. Sci.*, **20**, 4331.
14. March-Vila,E., Pinzi,L., Sturm,N., Tinivella,A., Engkvist,O., Chen,H. and Rastelli,G. (2017) On the integration of in silico drug design methods for drug repurposing. *Front. Pharmacol.*, **8**, 298.
15. Pinzi,L. and Rastelli,G. (2020) Identification of target associations for polypharmacology from analysis of crystallographic ligands of the Protein Data Bank. *J. Chem. Inf. Model.*, **60**, 372–390.
16. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
17. Kim,S., Chen,J., Cheng,T., Gindulyte,A., He,J., He,S., Li,Q., Shoemaker,B.A., Thiessen,P.A., Yu,B. *et al.* (2021) PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.*, **49**, D1388–D1395.
18. Gaulton,A., Hersey,A., Nowotka,M., Bento,A.P., Chambers,J., Mendez,D., Mutowo,P., Atkinson,F., Bellis,L.J., Cibrián-Uhalte,E. *et al.* (2017) The ChEMBL database in 2017. *Nucleic Acids Res.*, **45**, D945–D954.
19. Wishart,D.S., Feunang,Y.D., Guo,A.C., Lo,E.J., Marcu,A., Grant,J.R., Sajed,T., Johnson,D., Li,C., Sayeeda,Z. *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
20. The UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
21. Wang,Y., Zhang,S., Li,F., Zhou,Y., Zhang,Y., Wang,Z., Zhang,R., Zhu,J., Ren,Y., Tan,Y. *et al.* (2020) Therapeutic Target Database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.*, **48**, D1031–D1041.
22. Jenkins,J.L., Bender,A. and Davies,J.W. (2006) In silico target fishing: Predicting biological targets from chemical structure. *Drug Discov. Today Technol.*, **3**, 413–421.
23. Pinzi,L., Caporuscio,F. and Rastelli,G. (2018) Selection of protein conformations for structure-based polypharmacology studies. *Drug Discov. Today*, **23**, 1889–1896.
24. Berthold,M.R., Cebron,N., Dill,F., Gabriel,T.R., Kötter,T., Meinl,T., Ohl,P., Sieb,C., Thiel,K. and Wiswedel,B. (2008) KNIME: The Konstanz Information Miner. In: Preisach,C., Burkhardt,H.,

Schmidt-Thieme,L. and Decker,R. (eds). *Data Analysis, Machine Learning and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 319–326.
25. Willett,P. (2005) Searching techniques for databases of two- and three-dimensional chemical structures. *J. Med. Chem.*, **48**, 4183–4199.
26. Jasial,S., Hu,Y., Vogt,M. and Bajorath,J. (2016) Activity-relevant similarity values for fingerprints and implications for similarity searching. *F1000Res*, **5**, 591.
27. Bienfait,B. and Ertl,P. (2013) JSME: a free molecule editor in JavaScript. *J. Cheminform.*, **5**, 24.
28. Xu,K.-J., Song,J. and Zhao,X.-M. (2012) The drug cocktail network. *BMC Syst. Biol.*, **6**(Suppl. 1), S5.
29. Matter,H. (1997) Selecting optimally diverse compounds from structure databases: a validation study of two-dimensional and three-dimensional molecular descriptors. *J. Med. Chem.*, **40**, 1219–1229.
30. Alberga,D., Trisciuzzi,D., Montaruli,M., Leonetti,F., Mangiatordi,G.F. and Nicolotti,O. (2019) A new approach for drug target and bioactivity prediction: the multifingerprint similarity search algorithm (MuSSeL). *J. Chem. Inf. Model*, **59**, 586–596.
31. Gallo,K., Goede,A., Eckert,A., Moahamed,B., Preissner,R. and Gohlke,B.-O. (2021) PROMISCUOUS 2.0: a resource for drug-repositioning. *Nucleic Acids Res.*, **49**, D1373–D1380.
32. Jasial,S., Hu,Y. and Bajorath,J. (2017) How frequently are pan-assay interference compounds active? Large-scale analysis of screening data reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds. *J. Med. Chem.*, **60**, 3879–3886.
33. Ab Ghani,N.S., Ramlan,E.I. and Firdaus-Raih,M. (2019) Drug ReposER: a web server for predicting similar amino acid arrangements to known drug binding interfaces for potential drug repositioning. *Nucleic Acids Res.*, **47**, W350–W356.
34. Wang,X., Shen,Y., Wang,S., Li,S., Zhang,W., Liu,X., Lai,L., Pei,J. and Li,H. (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.*, **45**, W356–W360.
35. Karasev,D., Sobolev,B., Lagunin,A., Filimonov,D. and Poroikov,V. (2019) Prediction of protein-ligand interaction based on the positional similarity scores derived from amino acid sequences. *Int. J. Mol. Sci.*, **21**, 24.
36. Konc,J., Cesnik,T., Konc,J.T., Penca,M. and Janežič,D. (2012) ProBiS-database: precalculated binding site similarities and local pairwise alignments of PDB structures. *J. Chem. Inf. Model*, **52**, 604–612.
37. Salentin,S., Schreiber,S., Haupt,V.J., Adasme,M.F. and Schroeder,M. (2015) PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.*, **43**, W443–W447.
38. Nadzirin,N., Gardiner,E.J., Willett,P., Artymiuk,P.J. and Firdaus-Raih,M. (2012) SPRITE and ASSAM: web servers for side chain 3D-motif searching in protein structures. *Nucleic Acids Res.*, **40**, W380–W386.
39. Reddy,A.S., Tan,Z. and Zhang,S. (2014) Curation and analysis of multitargeting agents for polypharmacological modeling. *J. Chem. Inf. Model*, **54**, 2536–2543.
40. Zagidullin,B., Aldahdooh,J., Zheng,S., Wang,W., Wang,Y., Saad,J., Malyutina,A., Jafari,M., Tanoli,Z., Pessia,A. *et al.* (2019) DrugComb: an integrative cancer drug combination data portal. *Nucleic Acids Res.*, **47**, W43–W51.
41. Chakraborty,S.N., Leng,X., Perazzona,B., Sun,X., Lin,Y.-H. and Arlinghaus,R.B. (2016) Combination of JAK2 and HSP90 inhibitors: an effective therapeutic option in drug-resistant chronic myelogenous leukemia. *Genes Cancer*, **7**, 201–208.
42. Bhagwat,N., Koppikar,P., Keller,M., Marubayashi,S., Shank,K., Rampal,R., Qi,J., Kleppe,M., Patel,H.J., Shah,S.K. *et al.* (2014) Improved targeting of JAK2 leads to increased therapeutic efficacy in myeloproliferative neoplasms. *Blood*, **123**, 2075–2083.
43. Fiskus,W., Verstovsek,S., Manshouri,T., Rao,R., Balusu,R., Venkannagari,S., Rao,N.N., Ha,K., Smith,J.E., Hembruff,S.L. *et al.* (2011) Heat Shock Protein 90 inhibitor is synergistic with JAK2 inhibitor and overcomes resistance to JAK2-TKI in human myeloproliferative neoplasm cells. *Clin. Cancer Res.*, **17**, 7347.
44. Brasca,M.G., Nesi,M., Avanzi,N., Ballinari,D., Bandiera,T., Bertrand,J., Bindi,S., Canevari,G., Carenzi,D., Casero,D. *et al.* (2014) Pyrrole-3-carboxamides as potent and selective JAK2 inhibitors. *Bioorg. Med. Chem.*, **22**, 4998–5012.
45. Attar,E. and Bulun,S.E. (2006) Aromatase inhibitors: the next generation of therapeutics for endometriosis? *Fertil. Steril.*, **85**, 1307–1318.
46. Garzon,S., Laganà,A.S., Barra,F., Casarin,J., Cromi,A., Raffaelli,R., Uccella,S., Franchi,M., Ghezzi,F. and Ferrero,S. (2020) Aromatase inhibitors for the treatment of endometriosis: a systematic review about efficacy, safety and early clinical development. *Expert Opin. Investig. Drugs.*, **29**, 1377–1388.
47. Ballard,C., Aarsland,D., Cummings,J., O'Brien,J., Mills,R., Molinuevo,J.L., Fladby,T., Williams,G., Doherty,P., Corbett,A. *et al.* (2020) Drug repositioning and repurposing for Alzheimer disease. *Nat. Rev. Neurol.*, **16**, 661–673.
48. Jourdan,J.-P., Bureau,R., Rochais,C. and Dallemagne,P. (2020) Drug repositioning: a brief overview. *J. Pharm. Pharmacol.*, **72**, 1145–1151.
49. Ikeda,Y. (1999) Antiplatelet therapy using cilostazol, a specific PDE3 inhibitor. *Thromb. Haemost.*, **82**, 435–438.
50. Ono,K. and Tsuji,M. (2019) Pharmacological potential of cilostazol for Alzheimer's disease. *Front. Pharmacol.*, **10**, 559.
51. Rastelli,G., Pellati,F., Pinzi,L. and Gamberini,M.C. (2020) Repositioning natural products in drug discovery. *Molecules*, **25**, 1154.