JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Application Note

Cite This: J. Chem. Inf. Model. 2019, 59, 3619−3624

pubs.acs.org/jcim

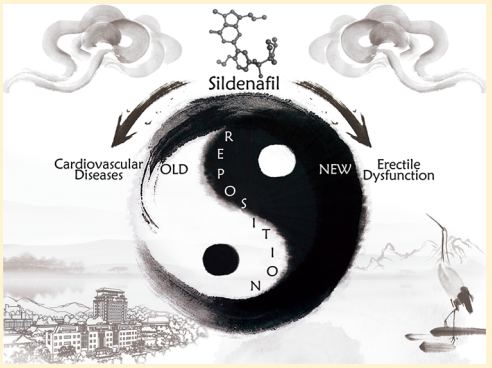
EK-DRD: A Comprehensive Database for Drug Repositioning Inspired by Experimental Knowledge

Chongze Zhao,[†] Xi Dai,[†] Yecheng Li,[†] Qingqing Guo, Jianhua Zhang, Xiaotong Zhang, and Ling Wang*

Joint International Research Laboratory of Synthetic Biology and Medicine, Guangdong Provincial Engineering and Technology Research Center of Biopharmaceuticals, School of Biology and Biological Engineering, South China University of Technology, Guangzhou 510006, China

Supporting Information

ABSTRACT: Drug repositioning, or the identification of new indications for approved therapeutic drugs, has gained substantial traction with both academics and pharmaceutical companies because it reduces the cost and duration of the drug development pipeline and the likelihood of unforeseen adverse events. To date there has not been a systematic effort to identify such opportunities, in part because of the lack of a comprehensive resource for an enormous amount of unsystematic drug repositioning information to support scientists who could benefit from this endeavor. To address this challenge, we developed a new database, the Experimental Knowledge-Based Drug Repositioning Database (EK-DRD), by using text and data mining as well as manual curation. EK-DRD contains experimentally validated drug repositioning annotation for 1861 FDA-approved and 102 withdrawn small-molecule drugs. Annotation was done at four levels using 30 944 target assay records, 3999 cell assay records, 585 organism assay records, and 8889 clinical trial records. Additionally, approximately 1799 repositioning protein or target sequences coupled with 856 related diseases and 1332 pathways are linked to the drug entries. Our web-based software displays a network of integrative relationships between drugs, their repositioning targets, and related diseases. The database is fully searchable and supports extensive text, sequence, chemical structure, and relational query searches. EK-DRD is freely accessible at http://www.idruglab.com/drd/index.php.



INTRODUCTION

The research and development of new drugs is an arduous, time-consuming, and costly task with a high rate of failure, such that the number of new drugs approved by the U.S. Food and Drug Administration (FDA) each year has shown little or no increase in successful projects despite an increasing commitment of resources.[1,2] A recent survey of 106 randomly selected approved new drugs estimated that it takes an average of 10 to 15 years and $1395 million (in 2013) to bring a new drug to market.[3] The majority of failures of drug development programs are due to the lack of efficacy of therapeutic hypotheses, with unexpected clinical side effects and tolerability being crucial issues.[4,5] Finding new uses outside the scope of the original medical indication for existing drugs, known as drug repurposing or repositioning, is one solution to achieve efficiency. As the pharmacologist and Nobel laureate James Black said, "the most fruitful basis for the discovery of a new drug is to start with an old drug." Existing drugs have already been tested in humans, have demonstrated an acceptable level of safety and tolerability, and are often approved by regulatory agencies for human use.[6,7] This could potentially increase the success rate of drug development and reduce the cost in terms of time.

Given the time and expense of developing drugs de novo, more pharmaceutical companies and academics are now scanning the existing pharmacopoeia for repositioning candidates, and the number of repositioning success stories is increasing. A classic example is sildenafil (Viagra), which was originally developed for the treatment of angina but has been repurposed for the treatment of erectile dysfunction and pulmonary arterial hypertension since the identification of an erectile response derived from its interaction with phosphodiesterase-5.[8] Many drugs have enormous potential for new drug indications in terms of polypharmacology.[1,7,9−11] Drug repurposing indications can arise from many sources, as follows: (1) clinical observations, including serendipitous or educated guesses, such as erection in the case of sildenafil[8] and reduction in erythema nodosum leprosum symptoms in the case of thalidomide;[12] (2) identification of repositioning opportunities from in vitro (phenotype- and target-based assays) and in vivo assays;[13−15] (3) epidemiological and post hoc analyses (e.g., the drug for alcohol abuse, disulfiram, exhibits activity against diverse cancer types);[16] and (4) in silico approaches, including cheminformatics, molecular

Received: May 5, 2019
Published: August 21, 2019

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624

modeling, machine learning, bioinformatics, and network-based data- or knowledge-driven mining approaches.[17−24] Many different in silico approaches to repurposing are available and have been reviewed elsewhere.[25] In silico approaches have the advantage of systematic screening of multiple candidates and are the subject of widespread interest. The usefulness of in silico algorithms in the study of drug repositioning can further be improved if they can be optimized by using experimental knowledge-based drug repositioning data.

Because of the intensifying research and accumulation of data on drug repositioning, databases relevant to drug repositioning have emerged in recent years. PROMISCUOUS was the first database that enabled users to establish and analyze networks responsible for multipharmacology by connecting the measures of structural similarity for drugs and known side effects to protein−protein interactions.[22] FDA-approved, withdrawn, or experimental drugs stored in PROMISCUOUS—25 000 in all—are included on the basis of inferred relationships through structural similarity. repDB is another database that contains approved and failed drugs and their clinical indications.[26] Although the above databases have aided research into drug repositioning, there is still no specific resource that provides comprehensive data for experimental determination of drug repositioning and further data analysis.

To cater to this need and to facilitate the scientific community's use of the experimentally determined resources for drug repositioning, we have developed the Experimental Knowledge-Based Drug Repositioning Database (EK-DRD) to host data on all aspects of experimentally validated repositioning information. EK-DRD stores repositioning records of about 30 944 target assays, 3999 cell assays, 585 organism assays, and 8889 clinical trials for 1963 drugs as well as other associated information for drugs, targets, pathways, and diseases (Table S1). To the best of our knowledge, EK-DRD is the largest database for drug repositioning with greatly improved information integration. In addition, we have developed a web-based tool for displaying a network of integrative relationships among drugs, their repositioning targets, and related diseases.

METHODS

Data Collection and Processing. The data used in EK-DRD (Figure 1) have as the main components information on drugs with FDA approval and experimental information on drug repositioning at four levels (target, cell, organism, and clinical trial). First, a total of 1963 small-molecule drugs and their corresponding information (chemical structure, name and synonyms, FDA-approved target, and indications) were retrieved from DrugBank version 4.0[27] and checked by mapping the FDA drug approval documents. Second, the drugs with available experimentally determined target assay data were searched from the public databases of ChEMBL (version 22),[28] BindingDB,[29] PubChem BioAssay,[30] and PDSP K_i (https://pdsp.unc.edu/databases/kidb.php, accessed Sept 11, 2016). The target assay data were refined with the following criteria: (1) only target-based assay data with detailed assay values (e.g., inhibition rate, K_i, or IC_{50}) were kept; (2) ADMET assay data were excluded; (3) the target assay data were filtered and obtained by mapping the FDA-approved target with an in-house script and by manual curation. Third, the cell-based assay data for repositioning were obtained by searching the ChEMBL database and refined by mapping the corresponding approved disease-related cell assay

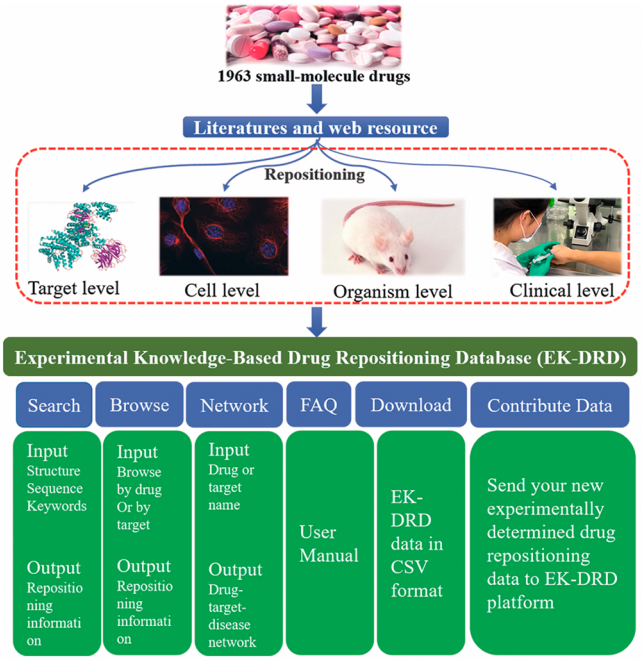


Figure 1. Overall design, construction, and contents of EK-DRD.

models. Fourth, the organism-based assay data were retrieved from PubMed (https://www.ncbi.nlm.nih.gov/pubmed/, accessed Sept 11, 2016) using combinations of the keywords "drug name and synonyms", "in vivo", "organism", and "animal". The searched literature was evaluated to find drugs with repositioning data sets by mapping the corresponding approved disease-related organism (animal) models. Finally, the repositioning data for clinical trial indications (excluding original approved indications) were obtained from the American Association of Clinical Trials Database (ClinicalTrials.gov: https://clinicaltrials.gov/ct2/, accessed Sept 11, 2016) using an in-house script. All of these repositioning data for 1963 drugs from different sources were checked by manual curation. Related information for drug repositioning, such as repositioning targets (gene, function, sequences, structures, etc.), signal transduction pathways, and diseases, were retrieved from the UniProt,[31] PDB,[32] KEGG,[33] and TTD[34] databases.

All of the drug structures are stored in EK-DRD in multiple formats (SDF, MOL, and SMILES). The structures were optimized with MOE software (version 2010.10) using the MMFF94 force field to generate three-dimensional (3D) structures. Conformational ensembles (maximum size: 200) were also generated for each drug in the database through the CAESAR algorithm[35] in the Discovery Studio software package (version 3.5; Biovia, San Diego, CA, USA).

Search and Network Display Tools. EK-DRD provides three retrieval methods for quickly searching and displaying the drug repositioning data, namely, text mining, chemical structure search, and protein sequence search. For chemical structure search, five algorithms, namely, substructure search, Markush search, two-dimensional (2D) and 3D similarity calculations, and hybrid structure-similarity calculations, are used in EK-DRD. 2D similarity calculations are based on the FP2 fingerprint and performed using OpenBabel.[36] 3D similarity adopts the weighted Gaussian algorithm (WEGA) for molecular-shape-similarity calculations,[37] which provides shape-, feature-, and coefficient-based shape-feature combo-

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624

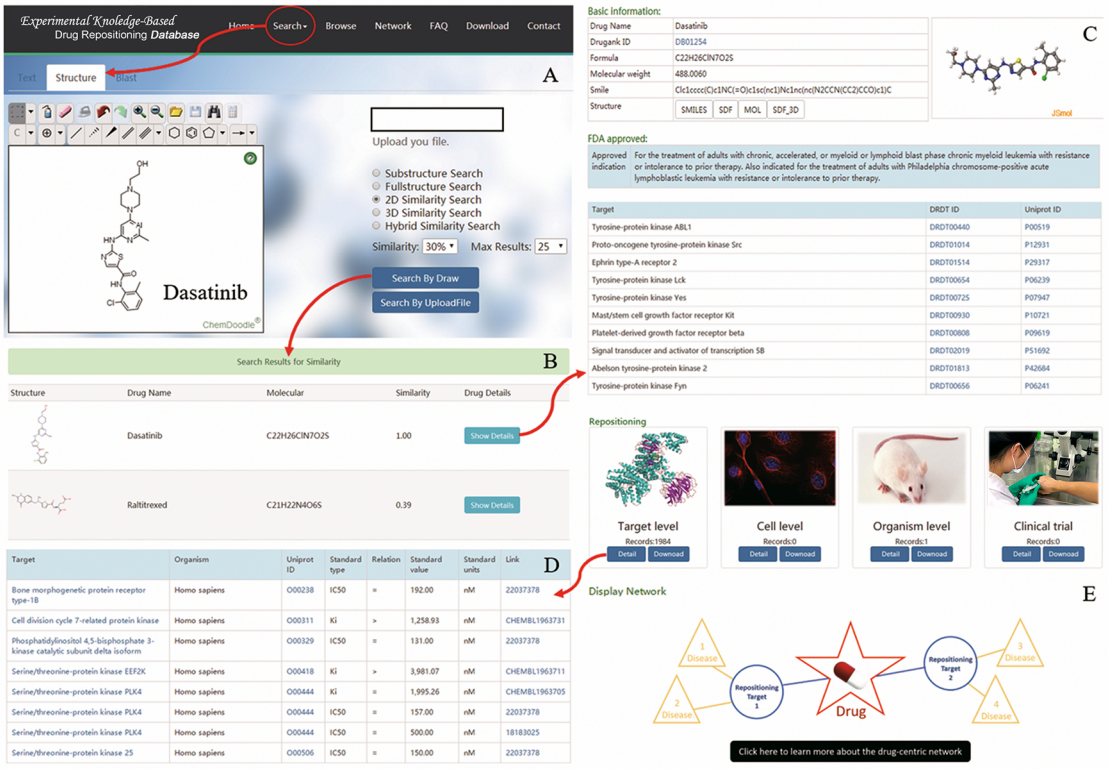


Figure 2. Schematic workflow of the chemical structure search interface in EK-DRD. (A) 2D chemical similarity for dasatinib drawn using the online ChemDoodle sketcher. (B) Snapshot of search results for dasatinib obtained using the 2D similarity search mode. (C) Snapshot of basic, FDA approval, and repositioning information on dasatinib. (D) Detailed target-based repositioning information for dasatinib presented as a table. (E) Connection concept network of dasatinib-repositioning target-related diseases.

scoring functions for user selection. Our group also encoded in EK-DRD a new hybrid similarity metric for calculating compound similarity that combines 2D fingerprint and 3D shape, called HybridSim, which was developed and validated to outperform the popular 2D FP2- and MACCS-based and 3D WEGA-based similarity methods.[38] All of the similarity methods use the Tanimoto coefficient as a similarity function to quantify the similarity between two molecules. The BLAST algorithm is used for protein sequence similarity search.[39] We also developed an online network display tool to virtually display the relationships among drugs, repositioning putative protein targets, and related diseases in the form of an interactive network.

Database and Web Interface Implementation. All of the metadata in EK-DRD are stored and managed in a MySQL database. The database query, data browser, network display, web interfaces, and access were implemented in HTML, CSS, JavaScript, PHP, and Apache HTTP server. EK-DRD allows users to input a 2D or 3D chemical structure or to draw a structure online using ChemDoodle (https://www.chemdoodle.com/) as the query structure for identifying desired drugs from EK-DRD. The input 2D structure is automatically converted into a single 3D conformer for 3D similarity calculations using the OpenBabel toolbox. The methodologies used to create EK-DRD are summarized in Table S2.

RESULTS AND DISCUSSION

Database Content. As shown in Figure 1, EK-DRD contains 1963 small-molecule drugs with four different levels of repositioning bioassay data: target, cell, organism, and clinical trials. For the target level, there are 30 944 assay data points for 1799 different repositioning targets from 123 different species. As shown in Figure S1A, approximately 89.22% of the repositioning targets come from the top nine species (*Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Cavia porcellus*, *Bos taurus*, *Bacillus subtilis*, *Equus caballus*, *Trypanosoma cruzi*, and *Escherichia coli*). It is worth noting that approximately 70% of the targets are redirected from *H. sapiens*; this suggests that the current research into drug repositioning mainly focuses on the development of drugs for human disease. For the cell level, 3999 bioassay data points were collected and stored in EK-DRD. In contrast to target- or cell-based in vitro screening assays, extensive in vivo screening of drugs for animal models is currently not possible, and as a result, only 585 organism (animal) assay records are contained in EK-DRD. For the clinical trial level, 8889 clinical trials for 666 drugs were annotated for drug repositioning by excluding the original FDA approval indications. Approximately 293 diseases are involved in these 8889 clinical trials according to classification of diseases in ICD-10 (version 2016; https://icd.who.int/browse10/2016/en#/), including general categories and subcategories (see the FAQ page in EK-DRD and Figure S2). The proportions for repositioning for different stages of clinical research (Figure S1B) are 1.46% for early phase, 11.28% for phase I, 6.24% for phases I/II, 31.09% for phase II, 4.09% for phases II/III, 13.57% for phase III, and 17.15% for phase IV.

EK-DRD contains many associated data, such as information related to 1799 repositioning targets (gene, function, sequences, structures, etc.), 1332 signal transduction pathways, and 856 related diseases involved in these repositioning targets,

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624

comprising 3762 drug-repositioning target−disease networks that are drug-centric and repositioning-target-centric. Many data fields in EK-DRD are hyperlinked to other databases (ChEMBL, KEGG, PDB, UniProt, PubMed, DrugBank, etc.).

Web Interfaces and Their Usage. EK-DRD provides fast, versatile, and user-friendly web interfaces that enable users to search, browse, display, and download all of the experimentally obtained drug-repositioning data in the database. Moreover, the Contribute data module in EK-DRD can be used to add new drug repositioning data from public users and researchers in the field.

Search. EK-DRD provides three modes to query the database, i.e., keywords (drug name, drug CAS number, target name, and UniProt ID), chemical similarity to the EK-DRD drug entries, and sequence similarity to EK-DRD target entries. Here we present a 2D similarity chemical search as an example to show how to utilize EK-DRD through the search function (Figure 2A). We sought drug repositioning information on dasatinib (a cancer drug). Using the ChemDoodle sketcher, a user can build a molecular structure of dasatinib and click the "Search By Draw" button to perform the 2D similarity search with other drugs in the EK-DRD database. The 2D similarity search results are ranked by similarity score, as shown in Figure 2B. The first record is dasatinib, which has the highest similarity score of 1; the user can click the "Show Details" button to enter the repositioning information page for dasatinib (Figure 2C), which displays basic information, the FDA-approved indication and target, and repositioning data at the target, cell, organism, and clinical trial levels. The user may check and browse the detailed data for target-level assays (Figure 2D). In addition, the connection concept network of dasatinib-repositioning target-related diseases (Figure 2E) can be found on the repositioning information page.

Browse. Repositioning information for drugs and targets can be browsed in two ways: (1) users can directly find the detailed repositioning data of the desired drugs according to drug name in alphabetical order (e.g., abacavir; Figure S3A) and (2) according to alphabetical order of the repositioning target name, the repositioning target page displays basic information on the desired target (target name, gene name, PDB ID, KEGG ID, pathway ID, and repositioning drugs) and associated descriptions of functions, related diseases, and pathways as well as the hyperlink for the repositioning-target-centric network (e.g., A7 nicotinic acetylcholine receptor; Figure S3B).

Network. This page displays drug−target−disease networks that are drug-centric and repositioning-target-centric. For the drug-centric network (e.g., as shown for dasatinib in Figure S4A), which is based on experimentally determined repositioning-target−drug interactions, the repositioning target may be involved in specific physiological functions for the treatment of certain diseases. Therefore, linking the drug and repositioning targets to their treated diseases is highly useful; this suggests that the repositioning-target-related diseases may be potentially treated by the desired drug. On the basis of this view, we built a drug-centric-based drug−target−disease network. The repositioning-target-centric-based network (e.g., as shown for A7 nicotinic acetylcholine receptor in Figure S4B) links the target and related diseases to all repositioning drugs, thus helping users to check for potential combination therapy. The user can browse the drug−target−disease network in terms of drug or target. In the drug−target−disease network, the user can double click on the identifier of the desired target or drug to browse the detailed information. In addition, the Search module (input drug, target name, and UniProt ID) on the Network page enables users to find the drug−target−disease network of the desired drug or repositioning target.

Contribute Data. If public users and researchers know of or have new experimentally determined drug repositioning data that they would like us to add, they can download the template table (CSV format) for the target, cell, organism, and clinical trial from the "Contribute Data" page, fill in the table, and send it to us using the Submit module on the "Contribute Data" page.

Download and FAQ. All of the data in the EK-DRD database can be freely downloaded from the "Download" page, and a detailed introduction and tutorial on the EK-DRD database are available on the "FAQ" page.

CONCLUSIONS

With the growing number of drug-repositioning studies, there is need for an integrated database that facilitates the exploration of data from these studies. To the best of our knowledge, EK-DRD is the first publicly available comprehensive resource for hosting and analyzing experimental knowledge-based drug-repositioning data sets. The main functions of EK-DRD enable users to search repositioning studies of a drug of interest at the levels of target, cell, organism, and clinical trial (if possible), to compare and browse the FDA-approved and repositioning targets and indications for a given drug, and to explore drug-repositioning target− disease networks. The expanded coverage of experimentally validated drug-repositioning data, together with the knowledge of the mechanisms, chemical structures, and properties of drugs as well as drug-repositioning target−disease networks, can facilitate repositioning-based drug discovery and related development, optimization, or both of in silico tools.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.9b00365.

Figures S1−S4 and Tables S1 and S2 (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: lingwang@scut.edu.cn.

ORCID

Ling Wang: 0000-0001-5116-7749

Author Contributions

†C.Z., X.D., and Y.L. are equal contributors to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the Science and Technology Program of Guangzhou (201707010063), the Natural Science Foundation of Guangdong Province (2016A030310421), the Medical Scientific Research Foundation of Guangdong Province (A2018114 and A2019021), and the Fundamental Research Funds for the Central Universities (2018ZD37).

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624

REFERENCES

(1) Strittmatter, S. M. Overcoming Drug Development Bottlenecks With Repurposing: Old Drugs Learn New Tricks. *Nat. Med.* **2014**, *20*, 590−591.

(2) DiMasi, J. A.; Feldman, L.; Seckler, A.; Wilson, A. Trends in Risks Associated with New Drug Development: Success Rates for Investigational Drugs. *Clin. Pharmacol. Ther.* **2010**, *87*, 272−277.

(3) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs. *J. Health. Econ.* **2016**, *47*, 20−33.

(4) Scannell, J. W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the Decline in Pharmaceutical R&D Efficiency. *Nat. Rev. Drug Discovery* **2012**, *11*, 191−200.

(5) Arrowsmith, J.; Miller, P. Trial Watch: Phase II and Phase III Attrition Rates 2011−2012. *Nat. Rev. Drug Discovery* **2013**, *12*, 569.

(6) Sachs, R. E.; Ginsburg, P. B.; Goldman, D. P. Encouraging New Uses for Old Drugs. *JAMA* **2017**, *318*, 2421−2422.

(7) Chong, C. R.; Sullivan, D. J., Jr. New Uses for Old Drugs. *Nature* **2007**, *448*, 645−646.

(8) Ghofrani, H. A.; Osterloh, I. H.; Grimminger, F. Sildenafil: from Angina to Erectile Dysfunction to Pulmonary Hypertension and Beyond. *Nat. Rev. Drug Discovery* **2006**, *5*, 689−702.

(9) Bertolini, F.; Sukhatme, V. P.; Bouche, G. Drug Repurposing in Oncology–Patient and Health Systems Opportunities. *Nat. Rev. Clin. Oncol.* **2015**, *12*, 732−742.

(10) Ashburn, T. T.; Thor, K. B. Drug Repositioning: Identifying and Developing New Uses for Existing Drugs. *Nat. Rev. Drug Discovery* **2004**, *3*, 673−683.

(11) 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.; Hromas, R.; Hudson, L.; Mackenzie, D. A.; Muller, C. Y.; Reed, J. C.; Simons, P. C.; Smagley, Y.; Strouse, J.; Surviladze, Z.; Thompson, T.; Ursu, O.; Waller, A.; Wandinger-Ness, A.; Winter, S. S.; Wu, Y.; Young, S. M.; Larson, R. S.; Willman, C.; Sklar, L. A. Drug Repurposing from an Academic Perspective. *Drug Discovery Today: Ther. Strategies* **2011**, *8*, 61−69.

(12) Rehman, W.; Arfons, L. M.; Lazarus, H. M. The Rise, Fall and Subsequent Triumph of Thalidomide: Lessons Learned in Drug Development. *Ther. Adv. Hematol.* **2011**, *2*, 291−308.

(13) Xu, M.; Lee, E. M.; Wen, Z. X.; Cheng, Y. C.; Huang, W. K.; Qian, X. Y.; TCW, J.; Kouznetsova, J.; Ogden, S. C.; Hammack, C.; Jacob, F.; Nguyen, H. N.; Itkin, M.; Hanna, C.; Shinn, P.; Allen, C.; Michael, S. G.; Simeonov, A.; Huang, W. W.; Christian, K. M.; Goate, A.; Brennand, K. J.; Huang, R. L.; Xia, M. H.; Ming, G. L.; Zheng, W.; Song, H. J.; Tang, H. L. Identification of Small-Molecule Inhibitors of Zika Virus Infection and Induced Neural Cell Death via a Drug Repurposing Screen. *Nat. Med.* **2016**, *22*, 1101−1107.

(14) Dittmar, A. J.; Drozda, A. A.; Blader, I. J. Drug Repurposing Screening Identifies Novel Compounds That Effectively Inhibit Toxoplasma gondii Growth. *mSphere* **2016**, *1*, No. e00042-15.

(15) Kuenzi, B. M.; Remsing Rix, L. L.; Stewart, P. A.; Fang, B.; Kinose, F.; Bryant, A. T.; Boyle, T. A.; Koomen, J. M.; Haura, E. B.; Rix, U. Polypharmacology-Based Ceritinib Repurposing Using Integrated Functional Proteomics. *Nat. Chem. Biol.* **2017**, *13*, 1222−1231.

(16) Skrott, Z.; Mistrik, M.; Andersen, K. K.; Friis, S.; Majera, D.; Gursky, J.; Ozdian, T.; Bartkova, J.; Turi, Z.; Moudry, P.; Kraus, M.; Michalova, M.; Vaclavkova, J.; Dzubak, P.; Vrobel, I.; Pouckova, P.; Sedlacek, J.; Miklovicova, A.; Kutt, A.; Li, J.; Mattova, J.; Driessen, C.; Dou, Q. P.; Olsen, J.; Hajduch, M.; Cvek, B.; Deshaies, R. J.; Bartek, J. Alcohol-Abuse Drug Disulfiram Targets Cancer via p97 Segregase Adaptor NPL4. *Nature* **2017**, *552*, 194−199.

(17) Nagaraj, A. B.; Wang, Q. Q.; Joseph, P.; Zheng, C.; Chen, Y.; Kovalenko, O.; Singh, S.; Armstrong, A.; Resnick, K.; Zanotti, K.; Waggoner, S.; Xu, R.; DiFeo, A. Using a Novel Computational Drug-repositioning Approach (DrugPredict) to Rapidly Identify Potent Drug Candidates for Cancer Treatment. *Oncogene* **2018**, *37*, 403−414.

(18) Korbee, C. J.; Heemskerk, M. T.; Kocev, D.; van Strijen, E.; Rabiee, O.; Franken, K. L. M. C.; Wilson, L.; Savage, N. D. L.; Dzeroski, S.; Haks, M. C.; Ottenhoff, T. H. M. Combined Chemical Genetics and Data-Driven Bioinformatics Approach Identifies Receptor Tyrosine Kinase Inhibitors as Host-directed Antimicrobials. *Nat. Commun.* **2018**, *9*, 358.

(19) Karatzas, E.; Bourdakou, M. M.; Kolios, G.; Spyrou, G. M. Drug Repurposing in Idiopathic Pulmonary Fibrosis Filtered by a Bioinformatics-Derived Composite Score. *Sci. Rep.* **2017**, *7*, 12569.

(20) Sawada, R.; Iwata, H.; Mizutani, S.; Yamanishi, Y. Target-Based Drug Repositioning Using Large-Scale Chemical-Protein Interactome Data. *J. Chem. Inf. Model.* **2015**, *55*, 2717−2730.

(21) Keiser, M. J.; Setola, V.; Irwin, J. J.; Laggner, C.; Abbas, A. I.; Hufeisen, S. J.; Jensen, N. H.; Kuijer, M. B.; Matos, R. C.; Tran, T. B.; Whaley, R.; Glennon, R. A.; Hert, J.; Thomas, K. L. H.; Edwards, D. D.; Shoichet, B. K.; Roth, B. L. Predicting New Molecular Targets for Known Drugs. *Nature* **2009**, *462*, 175−181.

(22) Von Eichborn, J.; Murgueitio, M. S.; Dunkel, M.; Koerner, S.; Bourne, P. E.; Preissner, R. PROMISCUOUS: A Database for Network-Based Drug-Repositioning. *Nucleic Acids Res.* **2011**, *39*, D1060−D1066.

(23) Cheng, F.; Liu, C.; Jiang, J.; Lu, W.; Li, W.; Liu, G.; Zhou, W.; Huang, J.; Tang, Y. Prediction of Drug-Target Interactions and Drug Repositioning via Network-Based Inference. *PLoS Comput. Biol.* **2012**, *8*, No. e1002503.

(24) Cheng, F.; Desai, R. J.; Handy, D. E.; Wang, R.; Schneeweiss, S.; Barabasi, A. L.; Loscalzo, J. Network-Based Approach to Prediction and Population-Based Validation of in Silico Drug Repurposing. *Nat. Commun.* **2018**, *9*, 2691.

(25) Villoutreix, B. O.; Lagorce, D.; Labbe, C. M.; Sperandio, O.; Miteva, M. A. One Hundred Thousand Mouse Clicks Down the Road: Selected Online Resources Supporting Drug Discovery Collected Over a Decade. *Drug Discovery Today* **2013**, *18*, 1081−1089.

(26) Brown, A. S.; Patel, C. J. A Standard Database for Drug Repositioning. *Sci. Data* **2017**, *4*, 170029.

(27) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; Tang, A.; Gabriel, G.; Ly, C.; Adamjee, S.; Dame, Z. T.; Han, B.; Zhou, Y.; Wishart, D. S. DrugBank 4.0: Shedding New Light on Drug Metabolism. *Nucleic Acids Res.* **2014**, *42*, D1091−D1097.

(28) Bento, A. P.; Gaulton, A.; Hersey, A.; Bellis, L. J.; Chambers, J.; Davies, M.; Kruger, F. A.; Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J. P. The ChEMBL Bioactivity Database: An Update. *Nucleic Acids Res.* **2014**, *42*, D1083−D1090.

(29) Gilson, M. K.; Liu, T. Q.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A Public Database for Medicinal Chemistry, Computational Chemistry and Systems Pharmacology. *Nucleic Acids Res.* **2016**, *44*, D1045−D1053.

(30) Wang, Y. L.; Suzek, T.; Zhang, J.; Wang, J. Y.; He, S. Q.; Cheng, T. J.; Shoemaker, B. A.; Gindulyte, A.; Bryant, S. H. Pub Chem BioAssay: 2014 Update. *Nucleic Acids Res.* **2014**, *42*, D1075−D1082.

(31) Bateman, A.; Martin, M. J.; O'Donovan, C.; Magrane, M.; Apweiler, R.; Alpi, E.; Antunes, R.; Ar-Ganiska, J.; Bely, B.; Bingley, M.; Bonilla, C.; Britto, R.; Bursteinas, B.; Chavali, G.; Cibrian-Uhalte, E.; Da Silva, A.; De Giorgi, M.; Dogan, T.; Fazzini, F.; Gane, P.; Castro, L. G.; Garmiri, P.; Hatton-Ellis, E.; Hieta, R.; Huntley, R.; Legge, D.; Liu, W.; Luo, J.; MacDougall, A.; Mutowo, P.; Nightin-Gale, A.; Orchard, S.; Pichler, K.; Poggioli, D.; Pundir, S.; Pureza, L.; Qi, G. Y.; Rosanoff, S.; Saidi, R.; Sawford, T.; Shypitsyna, A.; Turner, E.; Volynkin, V.; Wardell, T.; Watkins, X.; Zellner, H.; Cowley, A.; Figueira, L.; Li, W.; McWilliam, H.; Lopez, R.; Xenarios, I.; Bougueleret, L.; Bridge, A.; Poux, S.; Redaschi, N.; Aimo, L.; Argoud-Puy, G.; Auchincloss, A.; Axelsen, K.; Bansal, P.; Baratin, D.; Blatter, M. C.; Boeckmann, B.; Bolleman, J.; Boutet, E.; Breuza, L.; Casal-Casas, C.; De Castro, E.; Coudert, E.; Cuche, B.; Doche, M.; Dornevil, D.; Duvaud, S.; Estreicher, A.; Famiglietti, L.; Feuermann, M.; Gasteiger, E.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz-

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624

Gumowski, N.; Hinz, U.; Hulo, C.; Jungo, F.; Keller, G.; Lara, V.; Lemercier, P.; Lieberherr, D.; Lombardot, T.; Martin, X.; Masson, P.; Morgat, A.; Neto, T.; Nouspikel, N.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pozzato, M.; Pruess, M.; Rivoire, C.; Roechert, B.; Schneider, M.; Sigrist, C.; Sonesson, K.; Staehli, S.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Veuthey, A. L.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C. M.; Chen, Y. X.; Garavelli, J. S.; Huang, H. Z.; Laiho, K. T.; McGarvey, P.; Natale, D. A.; Suzek, B. E.; Vinayaka, C. R.; Wang, Q. H.; Wang, Y. Q.; Yeh, L. S.; Yerramalla, M. S.; Zhang, J. UniProt: A Hub for Protein Information. *Nucleic Acids Res.* **2015**, *43*, D204−D212.

(32) Rose, P. W.; Prlić, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Green, R. K.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. The RCSB Protein Data Bank: Integrative View of Protein, Gene and 3D Structural Information. *Nucleic Acids Res.* **2017**, *45*, D271−D281.

(33) Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* **2017**, *45*, D353−D361.

(34) Yang, H.; Qin, C.; Li, Y. H.; Tao, L.; Zhou, J.; Yu, C. Y.; Xu, F.; Chen, Z.; Zhu, F.; Chen, Y. Z. Therapeutic Target Database Update 2016: Enriched Resource for Bench to Clinical Drug Target and Targeted Pathway Information. *Nucleic Acids Res.* **2016**, *44*, D1069−D1074.

(35) Li, J.; Ehlers, T.; Sutter, J.; Varma-O'Brien, S.; Kirchmair, J. CAESAR: A New Conformer Generation Algorithm Based on Recursive Buildup and Local Rotational Symmetry Consideration. *J. Chem. Inf. Model.* **2007**, *47*, 1923−1932.

(36) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3*, 33.

(37) Yan, X.; Li, J. B.; Liu, Z. H.; Zheng, M. H.; Ge, H.; Xu, J. Enhancing Molecular Shape Comparison by Weighted Gaussian Functions. *J. Chem. Inf. Model.* **2013**, *53*, 1967−1978.

(38) Shang, J.; Dai, X.; Li, Y.; Pistolozzi, M.; Wang, L. HybridSim-VS: A Web Server for Large-Scale Ligand-Based Virtual Screening Using Hybrid Similarity Recognition Techniques. *Bioinformatics* **2017**, *33*, 3480−3481.

(39) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403−410.

DOI: 10.1021/acs.jcim.9b00365
J. Chem. Inf. Model. 2019, 59, 3619−3624