JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Prediction of Polypharmacological Profiles of Drugs by the Integration of Chemical, Side Effect, and Therapeutic Space

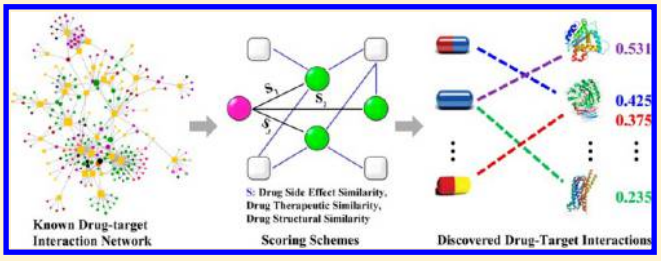
Feixiong Cheng,[†] Weihua Li,[†] Zengrui Wu,[†] Xichuan Wang,[‡] Chen Zhang,[†] Jie Li,[†] Guixia Liu,[†] and Yun Tang*[,†]

[†]Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China

[‡]Department of Surgery, Shanghai MCC Hospital, 456 Chunlei Road, Shanghai 200941, China

Supporting Information

ABSTRACT: Prediction of polypharmacological profiles of drugs enables us to investigate drug side effects and further find their new indications, i.e. drug repositioning, which could reduce the costs while increase the productivity of drug discovery. Here we describe a new computational framework to predict polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. On the basis of our previous developed drug side effects database, named MetaADEDB, a drug side effect similarity inference (DSESI) method was developed for drug−target interaction (DTI) prediction on a known DTI network connecting 621 approved drugs and 893 target proteins. The area under the receiver operating characteristic curve was 0.882 ± 0.011 averaged from 100 simulated tests of 10-fold cross-validation for the DSESI method, which is comparative with drug structural similarity inference and drug therapeutic similarity inference methods. Seven new predicted candidate target proteins for seven approved drugs were confirmed by published experiments, with the successful hit rate more than 15.9%. Moreover, network visualization of drug−target interactions and off-target side effect associations provide new mechanism-of-action of three approved antipsychotic drugs in a case study. The results indicated that the proposed methods could be helpful for prediction of polypharmacological profiles of drugs.



INTRODUCTION

Over the past decade, the "one gene, one drug, one disease" paradigm has become outdated in many cases and the concept of network pharmacology or polypharmacology was hence proposed for those drugs acting on multiple targets rather than one target.[1,2] Several published works demonstrated that a single drug target could be therapeutically insufficient, especially for some complex diseases.[3] For example, antipsychiatric drugs indicate multiple difference activities against four potential dopamine pathways in the human central nervous systems, such as the nigrostriatal dopamine pathway, mesolimbic dopamine pathway, mesocortical dopamine pathway, and tuberoinfundibular dopamine pathway.[4] However, multitarget drugs or drugs with polypharmacological profiles could have high risk of toxicity, such as undesirable profiles of off-target activity. Elucidation of undesirable profiles of off-target activity could hinder or halt the development of candidate drugs or even lead to market withdrawal by drug side effects.[2]

Drug side effects, also known as adverse drug events (ADEs) or adverse drug reactions, measure the harms associated with uses of given medications at normal dosages. Every year, thousands of people are reported to die from serious ADEs around the world. For example, there are about two million serious ADEs reported per year in the United States, resulting in about 100 000 deaths.[5] In many case, ADEs are caused by unintended activity at off-targets or drug promiscuity.[6,7] Several examples off-target toxicity included the drug−drug interactions caused by CYP450 inhibition,[8,9] cardiotoxicity caused by inhibition of the human *ether-a-go-go-related* gene potassium channel (hERG).[10] Systemically identifying polypharmacological profiles of drugs enable us to investigate drug side effects,[7] provide new strategy for multitarget drug design,[3] and further find their new indications, i.e. drug repositioning,[11] which could reduce the costs effectively while increase the productivity of drug discovery. However, how to systemically identify polypharmacological profiles of drugs is a big challenge which needs the development of a new methodology.

In the past decade, several computational methods have been published to address the issues of predicting polypharmacological profiles of drugs and drug−target interactions (DTIs).[12−20] Those methods were categorized into ligand-based, receptor-based, chemogenomics-based, and biological network-based ones. Keiser et al. predicted new molecular targets for known drugs using chemical two-dimensional (2D) structural similarity method, namely similarity ensemble approach (SEA).[21,22] Twenty-three new DTI were confirmed,

Received: January 4, 2013
Published: March 25, 2013

ACS Publications © 2013 American Chemical Society

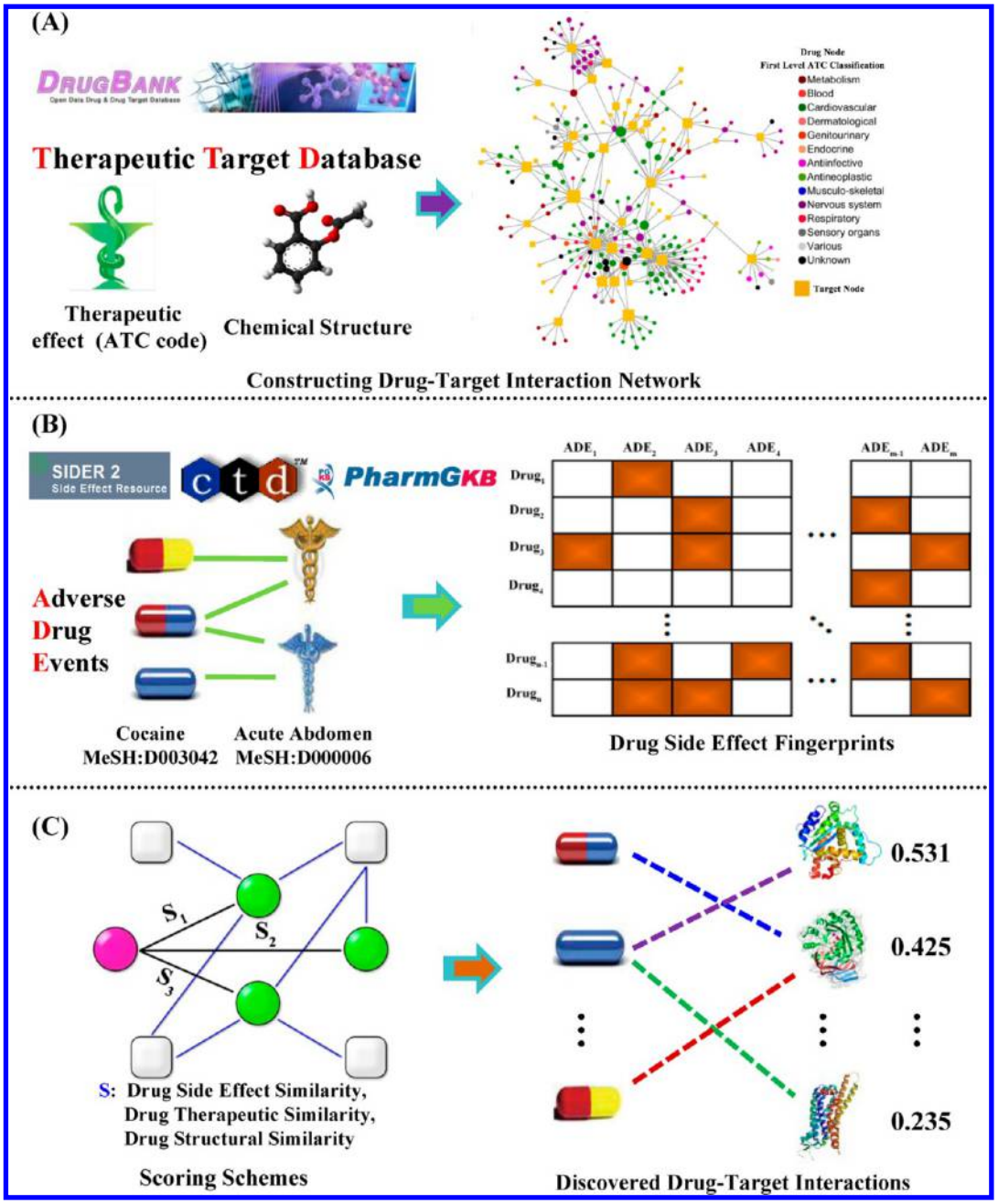


Figure 1. Schematic diagram of computational framework in this study. (**A**) Construction of drug−target interaction network from DrugBank and therapeutic target database. (**B**) Development of drug side effects similarity inference (DSESI) method. (C) Application of DSESI, drug structural similarity inference, and drug therapeutic similarity inference methods for predicting new potential drug−target interactions.

five of which were potent (K_i < 100 nM). Lounkine et al. recently used the SEA approach to predict the activity of 656 marketed drugs on 73 unintended side effect targets. Approximately half of the predictions were confirmed, either from proprietary databases unknown to the method or by new experimental assays. Affinities for these new off-targets ranged from 1 nM to 30 μM.[7] Wang et al. developed a computational chemogenomic method and used it to discovery several new ligands for four targets (i.e., GPR40, SIRT1, p38, and GSK-3β) by experimental assays.[23] Recently, our group systemically evaluated the performance of the multitarget quantitative structure−activity relationships (mt-QSAR) and computational chemogenomic methods.[19] The results indicated that there is a high false positive rate for the external validation set when using chemogenomic methods. Meanwhile, our group reported three supervised biological network inference methods, including drug-based similarity inference, target-based similarity inference, and network-based inference (NBI) methods for DTI prediction and drug repositioning. With the methods, five approved drugs were predicted and experimentally validated to have novel polypharmacological profiles on estrogen receptors and dipeptidyl peptidase-IV.[17] And two weighted biological network inference methods, namely node-weighted NBI and edge-weighted NBI, were further presented for chemical−protein interaction prediction. High performance were yielded and a "weak interaction hypothesis" was found based on two comprehensive chemical−protein interaction databases targeting G-protein-coupled receptors (GPCRs) and kinases.[18]

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

Recently, Besnard et al. described a new approach for the automated design of ligands against profiles of multiple drug targets.[3] The method is demonstrated by the evolution of an approved acetylcholinesterase inhibitor drug into brain-penetrable ligands with either specific polypharmacology or exquisite selectivity profiles for GPCRs. Several reverse docking strategies have been used for drug target binding affinity prediction, off-target prediction.[24−26] However, those methods may not be suitable for targets whose three-dimensional (3D) structures are unknown, such as most of GPCRs.

Interestingly, drug side effects can be seen as a clinical marker or an important phenotypic resource, and we can take advantage of this knowledge that, if two drugs are similar in the side effects they elicit, they are more likely to share a common drug target.[27,28] Therefore, drug side effects provide new dimensional space for predicting polypharmacological profiles of drugs recently.[29] Campillos et al. successfully predicted new DTI using drug side-effect similarity.[27] They tested 20 of those unexpected DTI and validated 13 implied DTI by in vitro binding assays, of which 11 revealed inhibition constants equal or less than 10 μM. Yang and Agarwal developed naive Bayes models to predict indications for 145 diseases using the drug side effects as features.[30] They extracted 3175 side effect-disease relationships by combining the side effect−drug relationships from drug labels and the drug−disease relationships from PharmGKB for naive Bayes model building. The area under the receiver operating characteristic curve (AUC) was above 0.8 in 92% of their models. A bottleneck in Yang's work is the lack of the benchmark negative side effect−disease relationship pairs. In our recent work,[35] we built a comprehensive drug−ADE association database so far, entitled MetaADEDB [http://www.lmmd.org/online_services/metaadedb/], which provides a comprehensive drug phenotypic resource for development of new computational methods to predict DTI and drug side effects.

In this study, we developed a new computational framework for predicting polypharmacological profiles of drugs by the integration of chemical, side effects, and therapeutic spaces (Figure 1). First, a DTI network connecting 621 approved drugs and 856 target proteins extracted from DrugBank database[31] and Therapeutic Target Database (TTD)[32] was built. The side effect similarity of each drug−drug pair was calculated from MetaADEDB, and the drug side−effect similarity inference (DSESI) method with high predictive performance was developed, which is comparative with drug structural similarity inference (DSSI) and drug therapeutic similarity inference (DTSI) methods.

METHODS AND MATERIALS

Data Set of Drug Phenotypic Source. A comprehensive drug−ADE association network, collected from our recently developed drug adverse events database, namely MetaADEDB: http://www.lmmd.org/online_services/metaadedb/. MetaADEDB was built using data integration from three drug side effects associated databases, CTD,[33] SIDER (version 2),[34] and OFFSIDES,[28] and a text mining method.[35] Only reported side effect data with experimental evidence were used. All drugs and ADE items in MetaADEDB were annotated using Medical Subject Headings (MeSH) and Unified Medical Language System (UMLS) vocabularies. And, duplicated drug−ADE associations were excluded.

Data Set of Drug−Target Interactions. Drug−target interaction data were collected from DrugBank[31] and TTD.[32] Each drug was grouped based on the Anatomical Therapeutic Chemical Classification System (ATC) code. The structures in SMILES format of each drug were collected from DrugBank.

Prediction of New Targets for Known Drugs. Denoting the target protein set as $T = \{t_1, t_2, ..., t_m\}$ and the drug set as $D = \{d_1, d_2, ..., d_n\}$, the drug−target (DT) binary interactions can be described as a bipartite DT graph $G(D, T, E)$, where $E = \{e_{ij}: d_i \in D, t_j \in T\}$. A link is drawn between d_i and t_j only if the drug d_i is interacted with the target t_j. The DT bipartite network can be presented by an $n \times m$ adjacent matrix $\{a_{ij}\}$, where $a_{ij} = 1$ when d_i and t_j interact with each other, otherwise $a_{ij} = 0$. For a DT pair d_i−t_j, a linkage between d_i and t_j is determined using the collaborative filtering algorithm[36,37] by the following predicted score:

$$v_{ij} = \frac{\sum_{l=1,l\neq i}^{n} S(d_i, d_l)a_{lj}}{\sum_{l=1,l\neq i}^{n} S(d_i, d_l)} \quad (1)$$

For different methods, including DSESI, DSSI, and DTSI, $S(d_i, d_l)$ represents different similarities of drug−drug pairs as follows:

Drug Side Effect Similarity. As shown in Figure 1B, each drug was coded by 13 060 ADE bit vectors in MetaADEDB. Each bit represents one ADE. If an ADE is associated with a given drug in MetaADEDB, the corresponding bit is set to "1"; conversely, it is set to "0". Then, the side effect similarity $S_{SE}(d_i, d_l)$ between drugs d_i and d_l was calculated by Tanimoto coefficient using the drug's ADE bit vectors.

Drug Structural Similarity. The 2D structural similarity $S_D(d_i, d_l)$ between drugs d_i and d_l were calculated via the Tanimoto similarity metric using MACCS keys, freely available from OpenBabel v2.3.1.[38]

Drug Therapeutic Similarity. The ATC codes for all drugs were collected from DrugBank database.[31] The kth level drug therapeutic similarity (S_k) between two drugs is defined via the ATC codes as

$$S_k(d_i, d_l) = \frac{\mathrm{ATC}_k(d_i) \cap \mathrm{ATC}_k(d_l)}{\mathrm{ATC}_k(d_i) \cup \mathrm{ATC}_k(d_l)} \quad (2)$$

where $\mathrm{ATC}_k(d)$ represents all ATC codes at the kth level of drug d. A score $TS(d_1, d_2)$ is used to define the therapeutic similarity between two drugs:

$$S_T(d_i, d_l) = \frac{\sum_{k=1}^{n} S_k(d_i, d_l)}{n} \quad (3)$$

where n represents the five level ATC codes (ranges from 1 to 5).[39] Note that some drugs have more than one ATC code, for example, the classic drug of nicotine has four different ATC codes (e.g., N07BA01, A11HA01, C04AC01, C10AD02). For a drug with multiple ATC codes, the therapeutic similarity was calculated for each ATC code, and then, the average therapeutic similarity was used.

Consensus Similarity. In addition, we also investigated three different consensus similarity inference methods: mean, maximum, and geometric mean as follows:

(i) Mean

$$S_M(d_i, d_l) = \frac{S_{SE}(d_i, d_l) + S_D(d_i, d_l) + S_T(d_i, d_l)}{3} \quad (4)$$

(ii) Maximum

$$S_{max}(d_i, d_l) = \max\{S_{SE}(d_i, d_l), S_D(d_i, d_l), S_T(d_i, d_l)\} \quad (5)$$

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

(iii) Geometric Mean

$$S_{GM}(d_i, d_l) = \sqrt[3]{S_{SE}(d_i, d_l)S_D(d_i, d_l)S_T(d_i, d_l)} \quad (6)$$

Where $S_M(d_i, d_l)$, $S_{max}(d_i, d_l)$, and $S_{GM}(d_i, d_l)$ represent three different combinational consensus similarities of a given drug−drug pair, $S_{SE}(d_i, d_l)$, $S_D(d_i, d_l)$, and $S_T(d_i, d_l)$ represent side effect similarity, structural similarity, and therapeutic similarity, respectively, for a given drug−drug pair. The original definition of consensus similarity is a similarity fusion method in the ligand-based virtual screening, where a user-defined reference structure is searched against a database using several different similarity measures. Detailed descriptions about consensus similarity were given in several related references.[40,41]

Performance Assessment. To test the performance of the methods, a 10-fold cross-validation technique was used and each result was repeated by 100 independent simulation tests. In the process of 10-fold cross-validation, the entire links in the DTI network were equally divided into 10 cross splits. In each step of cross-validation, the model was trained on a set of nine cross-validation splits together. The tenth subsample set was used as an internal validation set (test set). At last, two important metrics, namely the area under the receiver operating characteristic curve (AUC) and recall enhancement (RE), were calculated. The detailed descriptions of AUC and RE were given in our previous published works.[17,18]

RESULTS AND DISCUSSION

Performance of Proposed Methods. The 3195 drug-target interaction pairs (Table S1) connected 621 approved drugs and 893 target proteins were downloaded from DrugBank[31] and TTD.[32] If a target connected with only one drug (the degree of drug or target node is one) is just in the test set and the corresponding links could not be predicted by the collaborative filtering algorithm, because of no any information for those target nodes or drug nodes in the training set. Therefore, only drugs linked with two or more than two target proteins were used. First, the DTI network was constructed as a bipartite graph (Figure 1A). Then, the DSESI method was developed for prediction of new candidate DTI (Figure 1). The basic hypothesis of the DSESI method is that if two drugs are similar in the side effects they elicit, they are more likely to share a common drug target.[27,28] As given in Table 1, the high performance of the AUC value of 0.882 ± 0.011 was yielded for DSESI method measured by 100 independent simulation tests of 10-fold cross-validation.

In addition, the DTSI and DSSI methods were also developed for new DTI prediction. The DSSI method was based on the hypothesis that two drugs with high structural similarity may exhibit similar biological profiles, which was described in our previously published works.[17,18] The DTSI was based on the hypothesis that two drugs with high therapeutic similarities may exhibit similar biological profiles.[39] As given in Table 1, the high performance of 0.900 ± 0.011 and 0.912 ± 0.010 were yielded for DTSI and DSSI methods measured by 100 independent simulation tests of 10-fold cross-validation, respectively. The detailed performance of 100 independent simulation tests of 10-fold cross-validation prediction are given in Table S2 of the Supporting Information. Comparing the performance of three different methods, DTSI is marginally better than the DSSI and DSESI methods. High diversity of molecules in the training set is an important issue for the development and application of computational models.

Table 1. Performance of Prioritizing New Targets for Approved Drugs Using Different Methods[a]

method	RE ± SD	AUC ± SD
DSESI	25.4 ± 3.7	0.882 ± 0.011
DSSI	32.6 ± 4.2	0.900 ± 0.010
DTSI	41.9 ± 4.5	0.912 ± 0.010
mean	35.5 ± 4.4	0.911 ± 0.009
maximum	35.4 ± 4.2	0.908 ± 0.010
geometric mean	40.0 ± 4.6	0.912 ± 0.010

[a]Recall enhancement (RE) was evaluated based on the top five scored lists. AUC: the area under the receiver operating characteristic curve. DSESI: drug side effects similarity inference. DSSI: drug structural similarity inference. DTSI: drug therapeutic similarity inference. The standard deviation (SD) of the performance measured by 100 times independent simulated test of 10-fold cross-validation. The detailed performance is given in Table S2 of the Supporting Information.

Herein, the heat maps of side effect similarity, structural similarity, therapeutic similarity, and three different consensus similarities of drug−drug pairs for 621 approved drugs were shown in Figure 2. The mean structural similarity of 621 approved drugs was 0.318, which indicated the structural diversity of the training set. The overall therapeutic similarity of drug−drug pairs is higher than side effects and chemical structural similarities (Figure 2), which could explain that DTSI is marginally better than DSSI and DSESI methods.

Our recently works demonstrated that consensus models improve the performance of computational prediction.[8,42] In this work, three classic consensus similarity approaches, including mean, maximum, and geometric mean, were evaluated. As shown in Table 1, high performance was yielded for three different consensus similarity-based inference methods. And the geometric mean approach gives the best performance, comparing with mean and maximum methods. When comparing consensus similarity-based inference with single similarity-based inference, consensus similarity-based inference marginally outperforms single similarity-based inference, which could be caused by the redundancy in the similarity reported by Keiser et al.[21]

Discovery New Drug Targets. The 621 known drugs were computationally screened for their top 10 ranking scores to bind with 893 target proteins using DSESI, DTSI, and DSSI methods, respectively. Thousands of potential targets with high ranking scores were prioritized for 621 drugs. In order to test the feasibility of our methods, the top 10 predicted lists of each drug with high ranking scores prioritized by three different methods including DSESI, DTSI, and DSSI were overlapped. We found that about 3000 new predicted DTIs were overlapped (Supporting Information Table S3). Recently, Lounkine et al. reported a large scale prediction and testing of drug activity on side effect targets.[7] The half-maximum effective or inhibitory concentration (EC_{50} or IC_{50}) values of 151 new predicted DTIs were validated with less than 30 μM in concentration−response curves.[7] We mapped our predicted results with 151 new validated DTI in Lounkine's work. The new prioritized DTI pairs of seven approved drugs were mapped.[7] We extracted all 44 new predicted DTI pairs generated by DSESI, DTSI, and DSSI methods for those seven approved drugs (Supporting Information Table S4). Seven new predicted DTI pairs for seven approved drugs were confirmed by published experiments.[7] The range of IC_{50} or EC_{50} value is from 0.21 to 23 μM in concentration−response

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

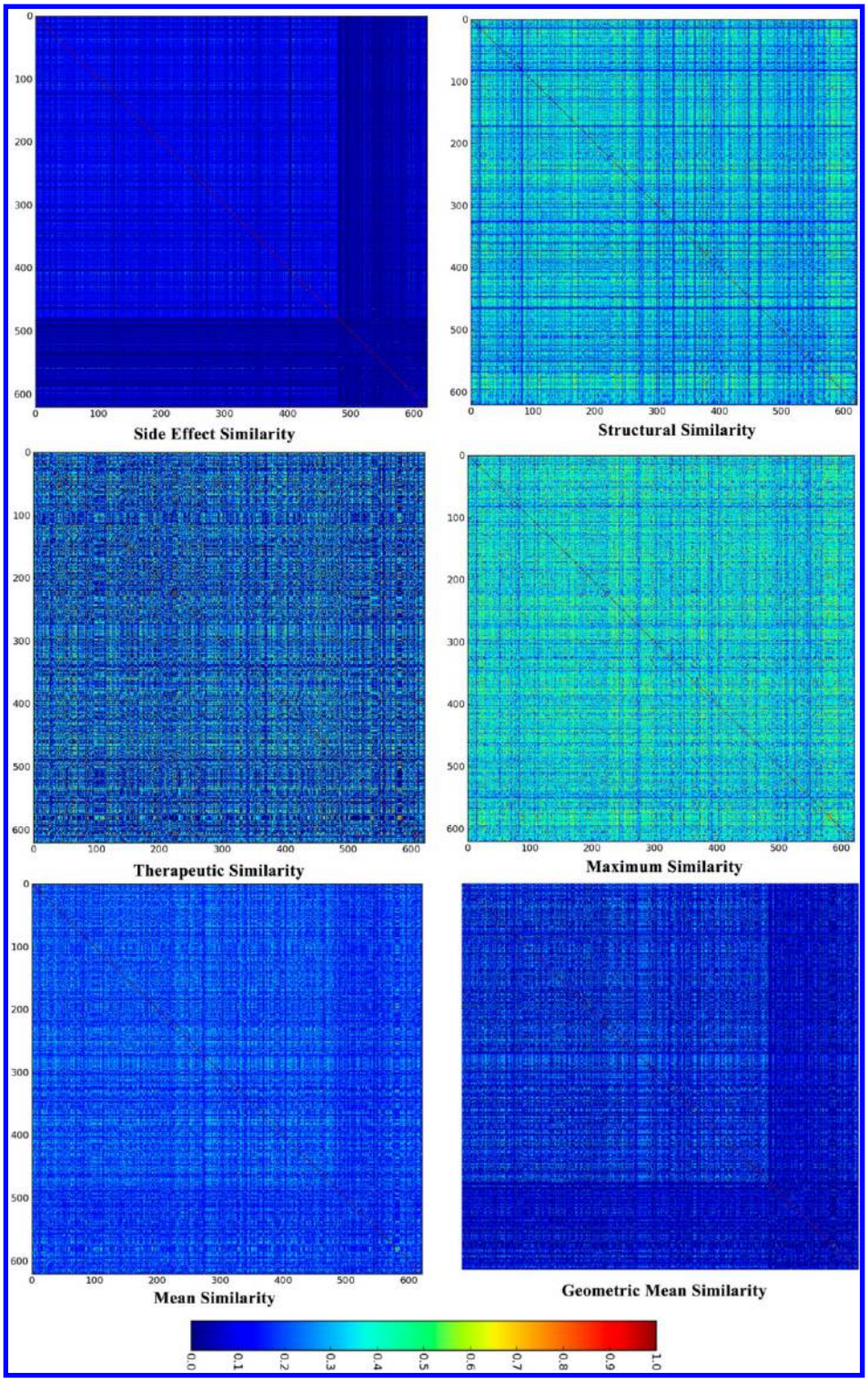


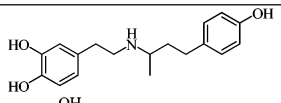
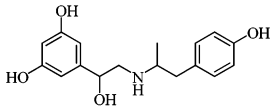
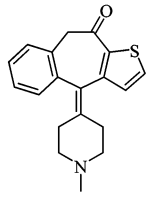
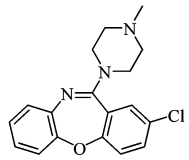
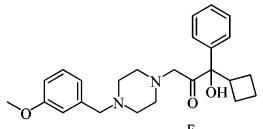
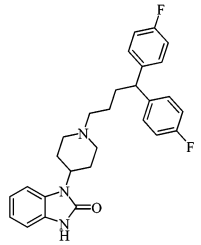
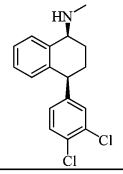
Figure 2. Heat maps of side effects similarity, structural similarity, therapeutic similarity, three different consensus (maximum, mean, and geometric mean) similarities of drug−drug pairs for 621 approved drugs.

curves (Table 2) with the successful hit rate more than 15.9%. In this study, 44 new predicted targets for seven known were mapped with the published experimental data in Lounkine's work. Only seven new predicted targets were overlapped in the published experimental data. If 44 new predicted targets were experimentally assayed, there may be more targets validated. Therefore, only a small part overlaps between our predicted targets and the published experimental data, which could explain why the success hit rate is much lower than the high accurate rate in the cross-validation prediction.

The network graph analysis could help to visualize the new and known DTI on a network level. The DTI network in Figure 3 indicated that there are obvious polypharmacological profiles of several approved drugs. By searching in DrugBank[31] and STITCH[43] databases, the orally active antipsychotic drug of pimozide binds with D_2 dopamine receptor (DRD2), D_3

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

Table 2. Predicted and Published Bioassay Results of New Identified Indications for Seven Example Approved Drugs

Drug name	Structure	[a] Primary pharmacological target	New predicted target	[b] IC_{50} or EC_{50} (μM)
Dobutamine		beta-1 Adrenergic Receptors	Adrenoceptor alpha 2A (ADRA2A)	10.83
Fenoterol		beta-2 Adrenergic Receptors	5-hydroxytryptamine receptor 2A (HTR2A)	3.40
Ketotifen		Histamine H1 Receptors	Adrenoceptor alpha 1A (ADRA1A)	10.40
Loxapine		Dopamine Receptor D2, Serotonin Receptor 2c and 7.	Muscarinic acetylcholine receptor M2 (CHRM2)	1.12
Tramadol		Opiate receptors and alpha(2)-adrenergic receptors	Muscarinic acetylcholine receptor M1 (CHRM1)	11.08
Pimozide		Dopamine D2 receptor	Adrenoceptor alpha 1A (ADRA1A)	0.21
Sertraline		Serotonin Transporter	Histamine H1 Receptors (HRH1)	23.00

[a]Original primary pharmacological target information was extracted from DrugBank[31] and STITCH[43] databases. [b]The IC_{50} or EC_{50} values were extracted from Lounkine's work.[7] IC_{50}: half maximal inhibitory concentration. EC_{50}: half maximal effective concentration.

dopamine receptor (DRD3), potassium voltage-gated channel subfamily H member 2 (KCNH2), and calmodulin (CALM1) with high biological activities. As shown in Figure 3 and Table 2, the pimozide was predicted and confirmed to bind with adrenoceptor alpha 1A (ADRA1A) with the IC_{50} value of 0.21 μM. The antiasthma drug of fenoterol binds with beta-1, 2, and 3 adrenergic receptors with high agonist activities.[31] In this study, fenoterol was predicted and confirmed to bind with 5-hydroxytryptamine receptor 2A (HTR2A) with the IC_{50} value of 3.4 μM (Table 2). The data demonstrated that our methods can successfully identify polypharmacological profiles of approved drug in actually screening work.

Polypharmacological Profiles of Approved Drugs. Network visualization of drug−target and target−ADE associations could provide a new mechanism-of-action (MOA) for marketed drugs. To assess the potential clinical relevance of predicted and validated targets systemically, we built the off-target ADE association networks for several example approved drugs. The 882 potential off-target ADE pairs among 30 off-targets and 189 ADEs for seven approved drugs (Supporting Information Table S5) were extracted from Lounkine's work.[7] The clinical validation and predicted ADEs for seven approved drugs were extracted from our previous developed MetaADEDB.[35] For an antiasthma drug of fenoterol, the validated off-target of adrenoceptor alpha 1A (ADRA1A) and two predicted off-targets of dopamine receptor D2 (DRD2) and adrenoceptor alpha 2A (ADRA2A) were linked with several ADEs, e.g. parkinsonism, parkinsonism-like phenotype (extrapyramidal disorder and hyperprolactinaemia),[44] orthostatic hypotension, hypercholesterolaemia, etc. (Supporting Information Table S6 and Figure 4A).

Pimozide is a conventional orally active antipsychotic drug which shares with other antipsychotics the ability to blockade dopaminergic receptors on neurons in the central nervous system. Like most conventional antipsychotic drugs, pimozide often generated propensity for or lack of motor side effects such as extrapyramidal side effects (EPS), tardive dyskinesia, endocrine disorder, galactorrhea, and amenorrhea (Figure 4B

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

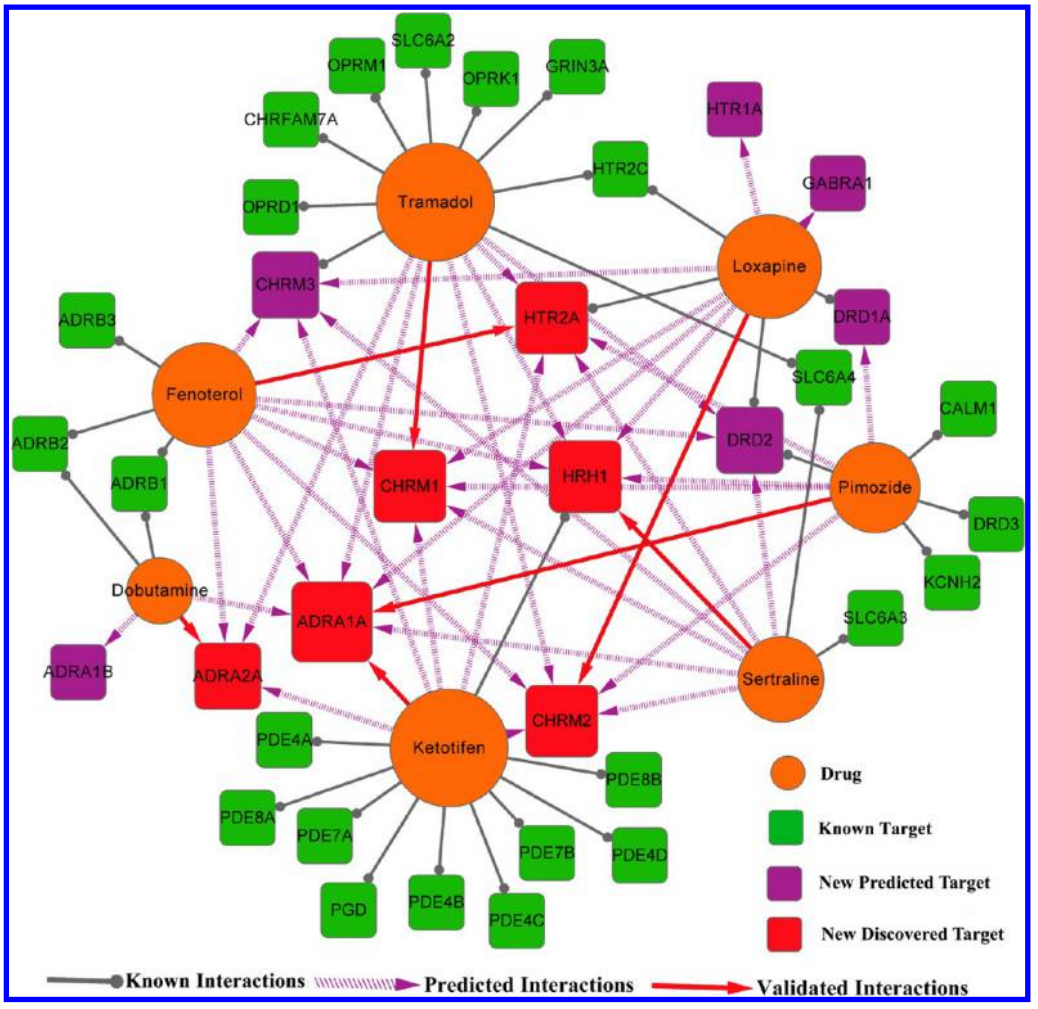


Figure 3. Discovered drug−target interactions network produced with the software package Cytoscape (http://www.cytoscape.org/). The size of the drug node is the fraction of the number of targets that the drug linked. The size of the target node is the fraction of the number of drugs that the target linked. The Cytoscape format files of Figures 3 and 4 are available in the Supporting Information as Figure 3.cys, Figure 4A.cys, Figure 4B.cys, and Figure 4C.cys.

and Supporting Information Table S6). The molecular hypothesizes of pimozide involved in ADEs is that there are four potential different dopamine pathways in human central nervous systems, such as nigrostriatal dopamine pathway, mesolimbic dopamine pathway, mesocortical dopamine pathway, and tuberoinfundibular dopamine pathway.[4] For the mesolimbic dopamine pathway, hyperactivity is thought to cause psychosis and the positive symptoms of schizophrenia, and blocking its hyperactivity could reduce psychosis or eliminate positive symptoms. However, the nigrostriatal dopamine pathway, as part of the extrapyramidal nervous system, controls movements. When blocking the dopamine D_2 receptor in this pathway by pimozide, it could lead to EPS and tardive dyskinesia. The tuberoinfundibular dopamine pathway inhibits prolactin release. If the normal function of tuberoinfundibular dopamine pathway was blocked by pimozide, several side effects, such as endocrine, galactorrhea, and amenorrhea will be generated. In addition to common ADEs of antipsychotic drugs, several other ADEs such as hypertension, electrocardiogram QT prolonged,[45] and dry mouth were also investigated. The off-target ADE network of pimozide was given in Figure 4B. The ADE of electrocardiogram QT prolonged was associated with the high inhibitive activity of IC_{50} value of 18 nM to potassium voltage-gated channel subfamily H member 2 (KCNH2).[46] The new predicted and validated target of ADRA1A with IC_{50} value of 0.21 μM was linked with the side effect of hypertension. Interestingly, the dry mouth side effect of pimozide could be explained by several new predicted associated targets of muscarinic acetylcholine receptor M1 and M2 (CHRM1, EF = 2.46, and CHRM2, EF = 2.42)[7] in Supporting Information Table S5 and Figure 4B.

Comparing with conventional antipsychotic drugs, the second generation atypical antipsychotic drug of loxapine not only binds with D_2 receptors but also binds with serotonin 5-HT_2 receptor (HTR2A) with high activities.[47] As shown in the off-target ADE network of loxapine (Figure 4C), several common ADEs, such as EPS, tardive dyskinesia, galactorrhea, and amenorrhea were also investigated for loxapien in clinical use (Supporting Information Table S6). The molecular hypothesis of ADEs was involved in the blockage of loxapine on the nigrostriatal and tuberoinfundibular dopamine pathways. The ADEs of tachycardia, urinary retention, and dry month could be associated with the new validated target of CHRM2. The known ADE of hypertension and new predicted anxiety and bradycardia could be linked with new predicted target of ADR1A1. Prediction of potential targets for old drugs could help to explore the ADEs or the molecular mechanism of known ADEs and postmarketing surveillance. Several common

dx.doi.org/10.1021/ci400010x | J. Chem. Inf. Model. 2013, 53, 753−762

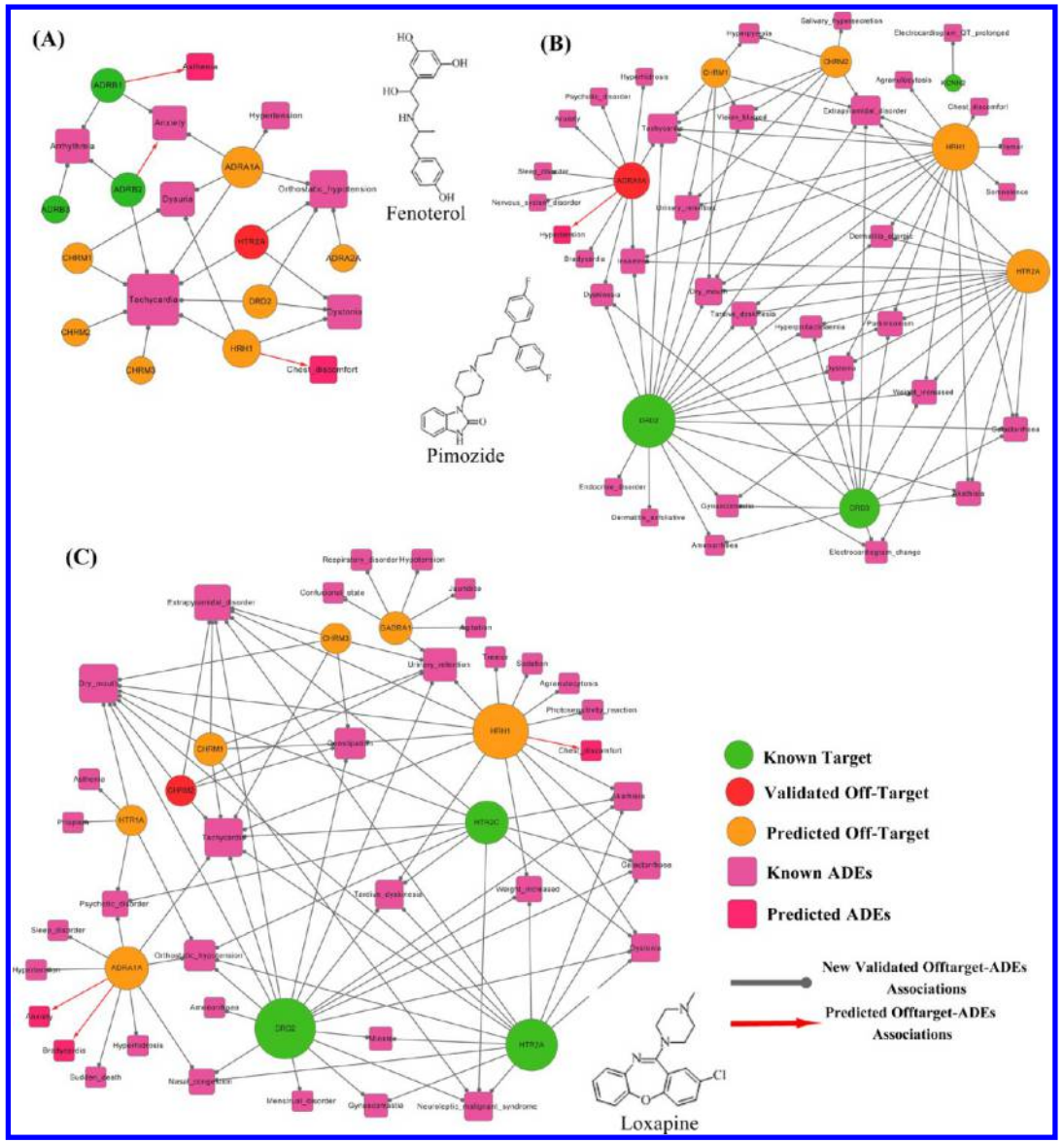


Figure 4. Drug off-target ADE network prepared with the software package Cytoscape (http://www.cytoscape.org/). Known targets (green circle) were collected from DrugBank and Therapeutic Target Database. Predicted off-target (gold circle) represents the overlap predicted results by drug side effects similarity inference, drug structural similarity inference, and drug therapeutic similarity inference methods. Known adverse drug events (ADEs) (purple square) represent the clinical validated data in MetaADEDB. Predicted ADEs (pink square) represent the new predicted data by predictive phenotypic network inference model in our previous work.[35] The size of the target node is the fraction of the number of ADEs that the target linked. The size of the ADE node is the fraction of the number of targets that the ADE linked.

ADEs with fuzzy molecular mechanisms were explained in this work.

CONCLUSION

In this study, we systemically developed several different drug-based similarity inference methods, including drug side effect similarity, drug structural similarity, drug therapeutic similarity, and consensus similarity inference methods for predicting new target proteins for approved drugs. High performance was yielded when using our methodologies by cross-validation prediction. Several new predicted target proteins for several approved drugs were confirmed by published data. Moreover, network visualization of drug−target interactions and off-target side effect associations provided a new mechanism-of-action for three antipsychotic drugs in a case study. The results will be very helpful in prediction of drug polypharmacological profiles and provide a new perspective for network pharmacology.

ASSOCIATED CONTENT

Supporting Information

Tables S1−S6. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +86-21-6425-1052. Fax: +86-21-6425-3651. E-mail: ytang234@ecust.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the 863 Project (Grant 2012AA020308), the National Natural Science Foundation of China (Grant 21072059), the Fundamental Research Funds for the Central Universities (WY1113007), and the Shanghai Committee of Science and Technology (11DZ2260600).

■ REFERENCES

(1) Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **2008**, *4*, 682−690.
(2) Bowes, J.; Brown, A. J.; Hamon, J.; Jarolimek, W.; Sridhar, A.; Waldron, G.; Whitebread, S. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat. Rev. Drug Discovery* **2012**, *11*, 909−922.
(3) Besnard, J.; Ruda, G. F.; Setola, V.; Abecassis, K.; Rodriguiz, R. M.; Huang, X. P.; Norval, S.; Sassano, M. F.; Shin, A. I.; Webster, L. A.; Simeons, F. R.; Stojanovski, L.; Prat, A.; Seidah, N. G.; Constam, D. B.; Bickerton, G. R.; Read, K. D.; Wetsel, W. C.; Gilbert, I. H.; Roth, B. L.; Hopkins, A. L. Automated design of ligands to polypharmacological profiles. *Nature* **2012**, *492*, 215−220.
(4) Stephen, M. S. Describing an Atypical Antipsychotic: Receptor Binding and Its Role in Pathophysiology. *J. Clin. Psychiat.* **2003**, *5*, 9−13.
(5) Lazarou, J.; Pomeranz, B. H.; Corey, P. N. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA, J. Am. Med. Assoc.* **1998**, *279*, 1200−1205.
(6) Cheng, F.; Yu, Y.; Zhou, Y.; Shen, Z.; Xiao, W.; Liu, G.; Li, W.; Lee, P. W.; Tang, Y. Insights into molecular basis of cytochrome p450 inhibitory promiscuity of compounds. *J. Chem. Inf. Model.* **2011**, *51*, 2482−2495.
(7) Lounkine, E.; Keiser, M. J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J. L.; Lavan, P.; Weber, E.; Doak, A. K.; Cote, S.; Shoichet, B. K.; Urban, L. Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **2012**, *486*, 361−367.
(8) Cheng, F.; Yu, Y.; Shen, J.; Yang, L.; Li, W.; Liu, G.; Lee, P. W.; Tang, Y. Classification of Cytochrome P450 Inhibitors and non-Inhibitors using Combined Classifiers. *J. Chem. Inf. Model.* **2011**, *51*, 996−1011.
(9) Cheng, F.; Li, W.; Zhou, Y.; Shen, J.; Wu, Z.; Liu, G.; Lee, P. W.; Tang, Y. admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *J. Chem. Inf. Model.* **2012**, *52*, 3099−3105.
(10) Curran, M. E.; Splawski, I.; Timothy, K. W.; Vincent, G. M.; Green, E. D.; Keating, M. T. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **1995**, *80*, 795−803.
(11) Ashburn, T. T.; Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discovery* **2004**, *3*, 673−683.
(12) Yamanishi, Y.; Araki, M.; Gutteridge, A.; Honda, W.; Kanehisa, M. Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* **2008**, *24*, i232−240.
(13) Yabuuchi, H.; Niijima, S.; Takematsu, H.; Ida, T.; Hirokawa, T.; Hara, T.; Ogawa, T.; Minowa, Y.; Tsujimoto, G.; Okuno, Y. Analysis of multiple compound-protein interactions reveals novel bioactive molecules. *Mol. Syst. Biol.* **2011**, *7*, 472.
(14) Gonzalez-Diaz, H.; Prado-Prado, F.; Garcia-Mera, X.; Alonso, N.; Abeijon, P.; Caamano, O.; Yanez, M.; Munteanu, C. R.; Pazos, A.; Dea-Ayuela, M. A.; Gomez-Munoz, M. T.; Garijo, M. M.; Sansano, J.; Ubeira, F. M. MIND-BEST: Web Server for Drugs and Target Discovery; Design, Synthesis, and Assay of MAO-B Inhibitors and Theoretical-Experimental Study of G3PDH Protein from Trichomonas gallinae. *J. Proteome Res.* **2011**, *10*, 1698−1718.
(15) Bleakley, K.; Yamanishi, Y. Supervised prediction of drug-target interactions using bipartite local models. *Bioinformatics* **2009**, *25*, 2397−2403.
(16) Yamanishi, Y.; Kotera, M.; Kanehisa, M.; Goto, S. Drug-target interaction prediction from chemical, genomic and pharmacological data in an integrated framework. *Bioinformatics* **2010**, *26*, i246−254.
(17) 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*, e1002503.
(18) Cheng, F.; Zhou, Y.; Li, W.; Liu, G.; Tang, Y. Prediction of chemical-protein interactions network with weighted network-based inference method. *PLoS One* **2012**, *7*, e41064.
(19) Cheng, F.; Zhou, Y.; Li, J.; Li, W.; Liu, G.; Tang, Y. Prediction of chemical-protein interactions: multitarget-QSAR versus computational chemogenomic methods. *Mol. Biosyst.* **2012**, *8*, 2373−2384.
(20) Cheng, F.; Li, W.; Zhou, Y.; Li, J.; Shen, J.; Lee, P. W.; Tang, Y. Prediction of human genes and diseases targeted by xenobiotics using predictive toxicogenomic-derived models (PTDMs). *Mol. Biosyst.* **2013**, DOI: 10.1039/c3mb25309k.
(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.; Edwards, D. D.; Shoichet, B. K.; Roth, B. L. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175−181.
(22) Keiser, M. J.; Roth, B. L.; Armbruster, B. N.; Ernsberger, P.; Irwin, J. J.; Shoichet, B. K. Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **2007**, *25*, 197−206.
(23) Wang, F.; Liu, D.; Wang, H.; Luo, C.; Zheng, M.; Liu, H.; Zhu, W.; Luo, X.; Zhang, J.; Jiang, H. Computational screening for active compounds targeting protein sequences: methodology and experimental validation. *J. Chem. Inf. Model.* **2011**, *51*, 2821−2828.
(24) Li, H.; Gao, Z.; Kang, L.; Zhang, H.; Yang, K.; Yu, K.; Luo, X.; Zhu, W.; Chen, K.; Shen, J.; Wang, X.; Jiang, H. TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res.* **2006**, *34*, W219−224.
(25) Yang, L.; Wang, K.; Chen, J.; Jegga, A. G.; Luo, H.; Shi, L.; Wan, C.; Guo, X.; Qin, S.; He, G.; Feng, G.; He, L. Exploring off-targets and off-systems for adverse drug reactions via chemical-protein interactome−clozapine-induced agranulocytosis as a case study. *PLoS Comput. Biol.* **2011**, *7*, e1002016.
(26) Xie, L.; Li, J.; Bourne, P. E. 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*, e1000387.
(27) Campillos, M.; Kuhn, M.; Gavin, A. C.; Jensen, L. J.; Bork, P. Drug target identification using side-effect similarity. *Science* **2008**, *321*, 263−266.
(28) Tatonetti, N. P.; Ye, P. P.; Daneshjou, R.; Altman, R. B. Data-driven prediction of drug effects and interactions. *Sci. Transl. Med.* **2012**, *4*, 125ra131.
(29) Tatonetti, N. P.; Liu, T.; Altman, R. B. Predicting drug side-effects by chemical systems biology. *Genome Biol.* **2009**, *10*, 238.
(30) Yang, L.; Agarwal, P. Systematic drug repositioning based on clinical side-effects. *PLoS One* **2011**, *6*, e28025.
(31) Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; Djoumbou, Y.; Eisner, R.; Guo, A. C.; Wishart, D. S. DrugBank 3.0: a comprehensive resource for "omics" research on drugs. *Nucleic Acids Res.* **2011**, *39*, D1035−D1041.
(32) Zhu, F.; Shi, Z.; Qin, C.; Tao, L.; Liu, X.; Xu, F.; Zhang, L.; Song, Y.; Zhang, J.; Han, B.; Zhang, P.; Chen, Y. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1128−1136.
(33) Davis, A. P.; King, B. L.; Mockus, S.; Murphy, C. G.; Saraceni-Richards, C.; Rosenstein, M.; Wiegers, T.; Mattingly, C. J. The Comparative Toxicogenomics Database: update 2011. *Nucleic Acids Res.* **2011**, *39*, D1067−1072.
(34) Kuhn, M.; Campillos, M.; Letunic, I.; Jensen, L. J.; Bork, P. A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* **2010**, *6*, 343.
(35) Cheng, F.; Li, W.; Wang, X.; Zhou, Y.; Wu, Z.; Shen, J.; Tang, Y. Adverse Drug Events: Database Construction and in Silico Prediction. *J. Chem. Inf. Model.* **2013**, DOI: 10.1021/ci4000079.
(36) Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Item-Based Collaborative Filtering Recommendation Algorithms. In *Proceedings of*

the World Wide Web Conference, Hong Kong, May 1−5, 2001; pp 285−295.
(37) Herlocker, J. L.; Konstan, J. A.; Terveen, K.; Riedl, J. T. Evaluating collaborative filtering recommender systems. *ACM T. Inf. Syst.* **2004**, *22*, 5−53.
(38) O'Boyle, N. M; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: an open chemical toolbox. *J. Cheminf.* [Online] **2011**, *3*, Article 33; http://www.jcheminf.com/content/3/1/33 (accessed Nov 19, 2011).
(39) Xu, K. J.; Song, J.; Zhao, X. M. The drug cocktail network. *BMC Syst. Biol.* **2012**, *6* (Suppl 1), S5.
(40) Willett, P. Similarity-based virtual screening using 2D fingerprints. *Drug Discovery Today* **2006**, *11*, 1046−1053.
(41) Medina-Franco, J. L.; Yongye, A. B.; Perez-Villanueva, J.; Houghten, R. A.; Martinez-Mayorga, K. Multitarget structure-activity relationships characterized by activity-difference maps and consensus similarity measure. *J. Chem. Inf. Model.* **2011**, *51*, 2427−2439.
(42) Cheng, F.; Ikenaga, Y.; Zhou, Y.; Yu, Y.; Li, W.; Shen, J.; Du, Z.; Chen, L.; Xu, C.; Liu, G.; Lee, P. W.; Tang, Y. In silico assessment of chemical biodegradability. *J. Chem. Inf. Model.* **2012**, *52*, 655−669.
(43) Kuhn, M.; Szklarczyk, D.; Franceschini, A.; Campillos, M.; von Mering, C.; Jensen, L. J.; Beyer, A.; Bork, P. STITCH 2: an interaction network database for small molecules and proteins. *Nucleic Acids Res.* **2010**, *38*, D552−D556.
(44) Farde, L.; Nordstrom, A. L.; Wiesel, F. A.; Pauli, S.; Halldin, C.; Sedvall, G. Positron emission tomographic analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. *Arch. Gen. Psychiat.* **1992**, *49*, 538−544.
(45) Aberg, K.; Adkins, D. E.; Liu, Y.; McClay, J. L.; Bukszar, J.; Jia, P.; Zhao, Z.; Perkins, D.; Stroup, T. S.; Lieberman, J. A.; Sullivan, P. F.; van den Oord, E. J. Genome-wide association study of antipsychotic-induced QTc interval prolongation. *Pharmacogenomics J.* **2012**, *12*, 165−172.
(46) Kang, J.; Wang, L.; Cai, F.; Rampe, D. High affinity blockade of the HERG cardiac K(+) channel by the neuroleptic pimozide. *Eur. J. Pharmacol.* **2000**, *392*, 137−140.
(47) Kapur, S.; Zipursky, R.; Remington, G.; Jones, C.; McKay, G.; Houle, S. PET evidence that loxapine is an equipotent blocker of 5-HT2 and D2 receptors: implications for the therapeutics of schizophrenia. *Am. J. Psychiat.* **1997**, *154*, 1525−1529.