# Target-Based Drug Repositioning Using Large-Scale Chemical–Protein Interactome Data

Ryusuke Sawada,[†] Hiroaki Iwata,[†] Sayaka Mizutani,[‡] and Yoshihiro Yamanishi[*,†,¶]

[†]Division of System Cohort, Multi-scale Research Center for Medical Science, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

[‡]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

[¶]Institute for Advanced Study, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

Ⓢ *Supporting Information*

ABSTRACT: Drug repositioning, or the identification of new indications for known drugs, is a useful strategy for drug discovery. In this study, we developed novel computational methods to predict potential drug targets and new drug indications for systematic drug repositioning using large-scale chemical–protein interactome data. We explored the target space of drugs (including primary targets and off-targets) based on chemical structure similarity and phenotypic effect similarity by making optimal use of millions of compound–protein interactions. On the basis of the target profiles of drugs, we constructed statistical models to predict new drug indications for a wide range of diseases with various molecular features. The proposed method outperformed previous methods in terms of interpretability, applicability, and accuracy. Finally, we conducted a comprehensive prediction of the drug–target–disease association network for 8270 drugs and 1401 diseases and showed biologically meaningful examples of newly predicted drug targets and drug indications. The predictive model is useful to understand the mechanisms of the predicted drug indications.
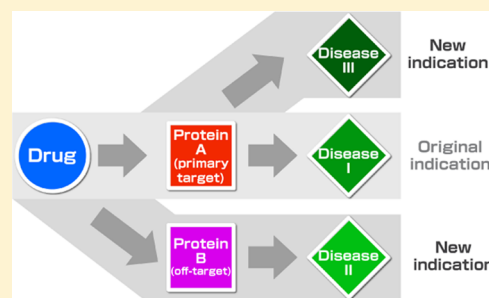
INTRODUCTION

In recent years, the development of new drugs based on conventional drug discovery processes has become increasingly difficult,[1] as reflected by a decrease in the number of newly approved drugs worldwide. Drug repositioning is the identification of new drug indications (i.e., new applicable diseases) for known drugs, and it has been recognized as a useful strategy for drug discovery.[2,3] With drug repositioning, some of the information on known drugs (e.g., pharmacokinetics, safety for humans, and manufacturing processes) can be reused. This could potentially increase the success rate of drug development and reduce the cost in terms of time and expenditure.[4]

Most drugs are chemical compounds that interact with therapeutic target proteins implicated in a disease of interest. However, drugs may interact not only with the therapeutic target proteins but also with additional proteins (hereafter referred to as off-targets), resulting in unexpected side effects.[5−7] Drug side effects derived from off-targets are undesired, but they may occasionally be beneficial for new therapeutic indications. For example, sildenafil (Viagra) was developed to treat angina, but it was redeveloped to treat erectile dysfunction owing to the discovery of an erectile response derived from its interaction with a phosphodiesterase (PDE5). Many drugs have enormous potential for new drug indications in terms of polypharmacology.[8−10] Therefore, there is a strong incentive to develop computational methods to predict both drug targets and drug indications for systematic drug repositioning.[11]

Recently, there has been an exponential growth in the volume of molecular data available on drugs, proteins, and diseases. Consequently, a variety of computational methods for systematic drug repositioning based on various molecular data have been proposed. Examples include the connectivity map-based approach,[12−15] the molecular similarity-based approach,[16−20] and the machine learning approach.[21−25] In these methods, prediction is based on drug features (e.g., drug-induced gene expression, chemical structures, side effects, target protein sequences, and biological pathways) and disease features (e.g., symptomatic state and phenotypes). However, most of the previous methods are purely predictive and the predictive models are not biologically interpretable. This makes it difficult both to understand the mechanisms of drug actions from the prediction results and to extensively evaluate the usefulness in clinical trials. Some previous studies were conducted from the viewpoints of biological interpretations. For example, known target proteins of drugs were linked to disease-susceptibility genes,[18] and the docking scores of drug–protein pairs were linked to drug indications or side effects.[20] However, the scalability of methods is limited in terms of the coverage of target proteins and the range of diseases. Therefore,

one of the most challenging issues in drug repositioning research is to identify both new drug indications and associated target proteins on a large scale.

In this study, we developed novel computational methods to predict potential drug targets and new drug indications for systematic drug repositioning using large-scale chemical-protein interactome data. We explored the target space of drugs (including primary targets and off-targets) based on chemical structure similarity and phenotypic effect similarity by making optimal use of millions of compound−protein interactions. Based on the target profiles of drugs, we constructed statistical models to predict new drug indications for a wide range of diseases with various molecular features. Finally, we conducted a comprehensive prediction of the drug−target−disease association network for 8270 drugs and 1401 diseases. The results indicate that the predictive model is useful to understand the mechanisms of the predicted drug indications.

■ MATERIALS

Chemical Structures of Compounds and Drugs. Chemical structure data of drugs and compounds were obtained from ChEMBL[26] and were represented by the following three chemical descriptors: Extended-Connectivity FingerPrints (ECFPs),[27] Chemistry Development Kit (CDK) fingerprints,[28,29] and KEGG Chemical Function and Substructures (KCF-S).[30] ECFP descriptors were calculated using ChemAxon's JChem Base (www.chemaxon.com), CDK fingerprint descriptors (Daylight-type fingerprints) were calculated using the CDK Java library,[28] and KCF-S descriptors were calculated using an in-house program. The diameter parameter for ECFP was set to 4 (ECFP4). We used default parameters for other chemical descriptors. Each drug or each compound was represented by a high-dimensional feature vector, in which the number of features (i.e., chemical substructures) of ECFPs, CDK fingerprints, and KCF-S were 1024, 1024, and 475 692, respectively. The resulting feature vector is referred to as the "chemical profile."

Phenotypic Effects of Drugs. Phenotypic effect data of drugs were obtained from Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS). FAERS has collected adverse event reports for approved drugs submitted to FDA by healthcare professionals from the first quarter of 2004 to the third quarter of 2012. We used adverse drug reactions (ADRs) and side effects as listed in the FDA database as the phenotypic effects of drugs in this study. FAERS drug labels were converted into KEGG D numbers by complete and partial name matching followed by manual inspection. In the partial name matching, FAERS drugs are mapped to the Anatomical Therapeutic Chemical (ATC) classification system by partially matching the drug labels to the ATC class description at the fifth level (chemical structure level) using KEGG's ATC classification file (http://www.kegg.jp/kegg-bin/get_htext?br08303.keg), where KEGG drugs are located to each of the ATC categories. In this way, FAERS drugs having the main structure but different salts (e.g., drugs labeled as "acebutolol", "acebutolol HCL", and "acebutolol hydrochloride" are identified as the same drug entry). The primary suspect (PS) and secondary suspect (SS) of side effects were used, but concomitants were not used. As a result, 2594 drugs were associated with a total of 16 075 phenotypic effects. For each of 2594 drugs, we constructed a 16 075-dimensional binary vector whose elements encode for the presence or absence of the phenotypic effects by 1 or 0, respectively. There is a huge bias

in the frequency of reported side effects, so the feature vector for each drug was normalized such that the total sum of all elements was one. The resulting feature vector is referred to as the "phenotypic profile."

Chemical−Protein Interactome Data. Data of compound−protein interaction pairs were obtained from seven published databases: ChEMBL,[26] MATADOR,[31] Drug Bank,[32] Psychoactive Drug Screening Program K_i (PDSP-K_i),[33] KEGG DRUG,[34] BindingDB,[35] and Therapeutic Target Database.[36] For ChEMBL, we selected only compound−protein interaction pairs that were clearly denoted as active interactions or with binding affinity less than 30 μM (e.g., at IC50). In total, 1 287 404 compound−protein interactions involving 519 061 compounds and 3736 proteins were obtained. This data set is referred to as "chemical−protein interactome data."

Drug−Target Interactions. There are 8270 drugs that are registered as approved drugs in Japan, the USA, or Europe. Data of drug−target interactions were obtained from the above chemical−protein interactome data by selecting compounds that were registered as approved drugs. In total, 21 646 drug−target interactions involving 2787 drugs and 1497 proteins were obtained. The drug−target interactions were used as gold standard data in cross-validation (CV) experiments to evaluate the performance of drug target estimation. Unknown target proteins of all 8270 drugs were predicted in the new prediction in the Results section.

Molecular Features of Diseases. Data of various molecular features of diseases were obtained from the 10th revision of the International Classification of Diseases (ICD-10)[37] and KEGG DISEASE.[34,38] In total, 1401 diseases were represented by 6342 molecular features (e.g., 525 diagnostic markers, 4373 disease-causing genes, 247 disease-related pathways, and 230 environmental factors). Each disease was represented by a 6342-dimensional binary vector whose elements encoded for the presence or absence of each molecular feature by 1 or 0, respectively.

Drug Indications. Drug indications are regarded as drug−disease associations. Data of drug−disease associations were obtained from medical books[39] and KEGG DRUG.[34] Diseases defined in ICD-10[37] were used in this study. In total, 5830 drug−disease associations involving 2271 drugs and 463 diseases were obtained. This data set was denoted as the "Full dataset." The drug−disease associations were used as gold standard data in the CV experiments to evaluate the performance of drug indication prediction.

In addition to the Full data set, we constructed another drug indication data set consisting of drugs for which chemical structures, phenotypic effects, and drug−target interactions were available. In total, we collected 1855 drug−disease associations involving 762 drugs and 223 diseases. This data set was denoted as the "Common dataset."

The Common data set consisted of drugs for which all data (target information, chemical structure information and phenotypic information) are available. In contrast, the Full data set included drugs that lacked some data; thus, the missing features of drugs in the Full data set were imputed by zeros. Drugs in the Common data set had richer information; thus, the Full data set was more difficult to reconstruct in the CV experiment compared with the Common data set.

Disease−Target Association Template. The information on therapeutic target proteins for each disease was obtained from KEGG DRUG.[34] Drugs can activate or inhibit therapeutic target proteins known to be useful for the treatment of each

disease. Note that target proteins that are not known to be associated with diseases are not taken into consideration. In total, 2062 disease-target associations involving 250 diseases and 462 therapeutic target proteins were obtained. This data set is referred to as "disease−target association template."

■ METHODS

Overview of Target-Based Drug Repositioning. Drug phenotypic effects are basically derived from the interactions between drug molecules and target proteins, including the primary target and off-targets. Thus, we propose a new systematic drug repositioning approach based on all potential target proteins of a given query drug, which we call "target-based drug repositioning." Figure 1 shows a conceptual diagram
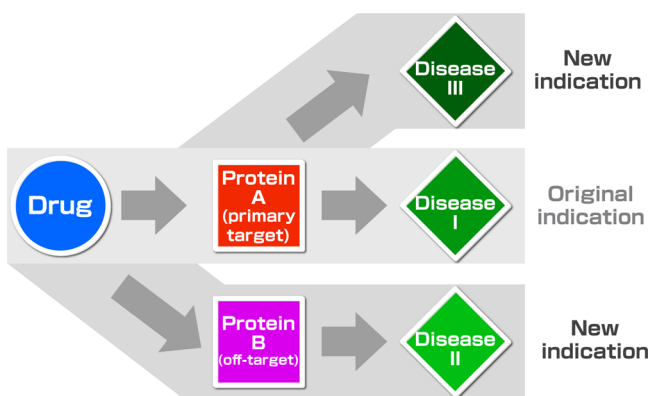

Figure 1. Conceptual diagram of target-based drug repositioning. The prediction of new drug indications is performed by linking drugs to diseases via drug target proteins (including the primary target or off-targets).

of target-based drug repositioning. New indications (i.e., new applicable diseases) for a given query drug are predicted based not only on the primary target protein but also on additional off-target proteins.

Figure 2 shows the procedure of our proposed method to predict new drug indications. The prediction procedure consists of two parts. Figure 2A shows the first part, in which all potential target proteins of a given query drug are estimated by performing the target estimation with similarity search (TESS). Figure 2B shows a process of the second part, in which new indications of a given query drug are predicted by performing indication prediction with template matching (IPTM). Figure 2C shows another process of the second part, in which new indications of a given query drug are predicted by performing indication prediction with supervised classification (IPSC). The details of each process are described below.

Target Estimation with Similarity Search (TESS). To predict all potential target proteins (including primary target and off-targets) of the query drug, we propose to use large-scale chemical−protein interactome data and various molecular features of drugs to reveal the target space of drugs as completely as possible. Recall that our chemical−protein interactome data consist of millions of compound−protein interactions. Compounds and proteins in the chemical−protein interactome data are referred to as interactome compounds and interactome proteins, respectively.

We propose TESS as a new method to estimate drug targets using a similarity search (see Figure 2A). A query drug is represented by a feature vector as $X^{(\text{query})} = (x'_1, x'_2, ..., x'_d)^{\text{T}}$ and

each interactome compound is also represented by a feature vector as $X^{(\text{interactome})} = (x_1, x_2, ..., x_d)^{\text{T}}$, where d is the number of features. Each interactome compound is also represented by a target protein interaction profile as $Y^{(\text{interactome})} = (y_1, y_2, ..., y_p)^{\text{T}}$, where y_k indicates the presence or absence of the interaction with the kth protein by 1 or 0 ($k = 1, 2, ..., p$), and p is the number of interactome proteins (3736 proteins in this study). We consider predicting the target protein interaction profile of the query drug, which is represented as $Y^{(\text{query})}$.

First, we compute similarity scores for pairs between a given query drug and all interactome compounds in our chemical-protein interactome data. Second, from interactome compounds that are known to interact with the kth protein ($k = 1, 2, ..., p$), we select an interactome compound that has the highest similarity with the query drug and use the corresponding similarity score as a prediction score $\hat{y}_k$ for the possibility that the query drug interacts with the kth protein. Third, we repeat this procedure for all p interactome proteins ($k = 1, 2, ..., p$) and assign the prediction scores to query drug−protein pairs (pairs between the query drug and all interactome proteins) as $\hat{y}_k$ ($k = 1, 2, ..., p$). Finally, the predicted target protein interaction profile for the query drug is represented as $Y^{(\text{query})} = (\hat{y}_1, \hat{y}_2, ..., \hat{y}_p)^{\text{T}}$.

To compute the similarity between $X^{(\text{query})}$ and $X^{(\text{interactome})}$, we propose a generalized Jaccard correlation coefficient for real-valued feature vectors, which is defined as follows:

$$\text{corr}(X^{(\text{query})}, X^{(\text{interactome})}) = \frac{\sum_{i=1}^{d} \min(x'_i, x_i)}{\sum_{i=1}^{d} \max(x'_i, x_i)}$$

where $\min(\cdot, \cdot)$ is an operation to take a minimum value and $\max(\cdot, \cdot)$ is an operation to take a maximum value for two given values. Using the generalized Jaccard correlation coefficient, we compute chemical similarity scores based on chemical profiles and phenotypic similarity scores based on phenotypic profiles. Note that the original Jaccard coefficient (Tanimoto coefficient) can only treat binary vectors (called fingerprints). The distributions of highest similarity from both chemical structures and phenotypic effects of drugs are shown in Figure S5 in the Supporting Information.

Indication Prediction by Template Matching (IPTM). Suppose that we are given a query drug and we want to predict novel drug indications (i.e., applicable diseases) for the query drug. We consider predicting drug indications based on target proteins of the query drug (including the primary target and off-targets) and the disease−target association template.

Here we propose IPTM, a new method to predict drug indications using template matching (see Figure 2B). We look for target proteins that appear in both the query drug's target profile and the disease−target association template. In this process, we take into account not only known drug−target interactions but also drug−target interactions newly predicted by TESS (see previous subsection). Our proposed IPTM method is described in detail below.

First, we take one target protein of the query drug out of its target proteins. Second, we look for the same target protein in the disease−target association template, select diseases associated with the matched target protein, and link the query drug to the selected diseases via the matched target protein. Third, we repeat this procedure for all target proteins of the query drug, and merge the selected diseases into a disease list for the query drug. Fourth, we assign prediction scores to the newly linked drug−disease pairs. The prediction
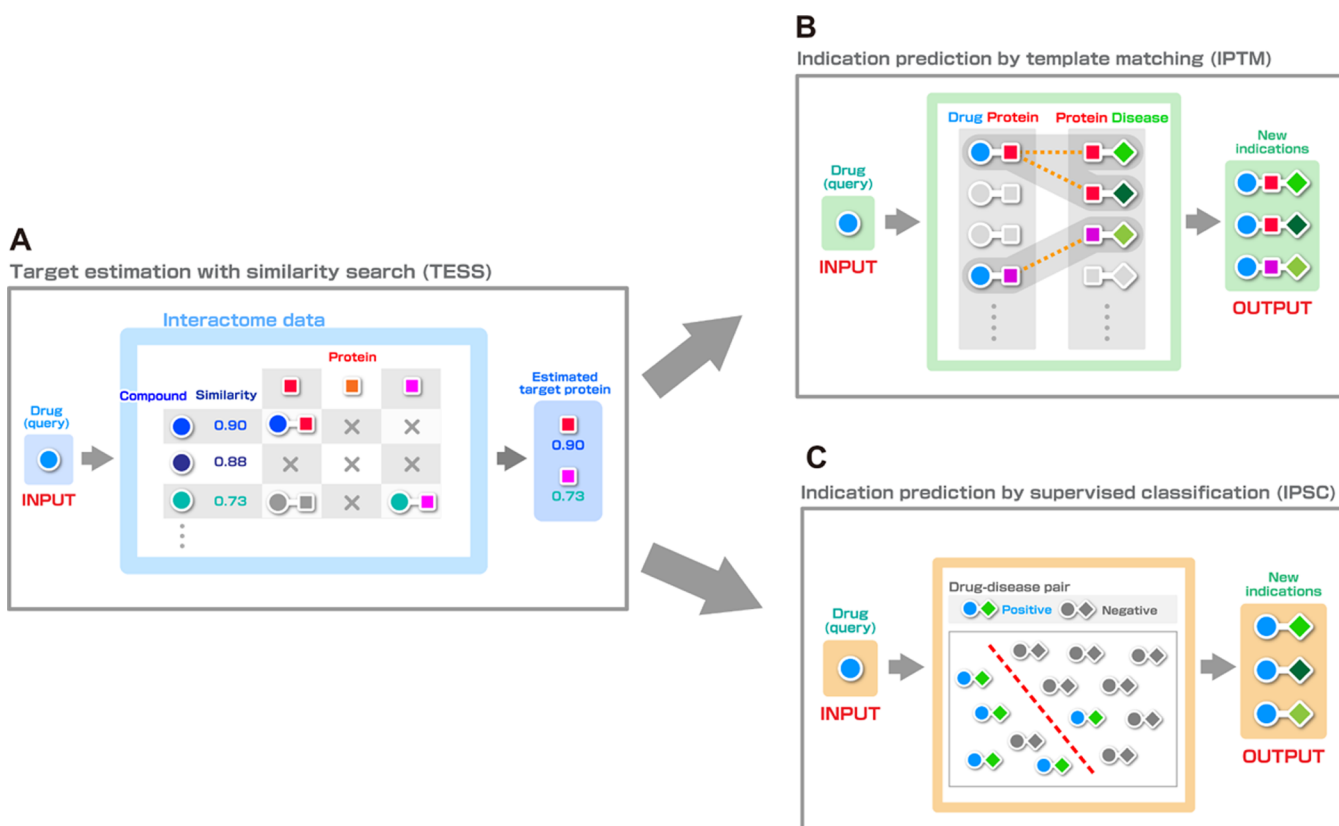
Figure 2. Workflow of our proposed methods for drug target estimation and drug indication prediction. (A) TESS performs drug target estimation by similarity search using large-scale chemical-protein interactome data. (B) IPTM performs drug indication prediction by template matching, based on the comparison between drug target profiles and disease-target association template. (C) IPSC performs drug indication prediction by supervised classification, based on preknowledge about partially known drug−disease associations.

scores are set to one if the matched target proteins are known targets of the query drug, while the prediction scores are set to the TESS score if the matched target proteins are newly predicted by TESS based on chemical similarity or phenotypic similarity. Finally, we predict drug−disease pairs with prediction scores above a certain threshold to be candidates for new drug indications. The threshold is set to the upper 5th percentile of the distribution of all prediction scores.

The originality of the IPTM method is that the prediction outputs can be represented by a drug−target−disease network.

Indication Prediction by Supervised Classification (IPSC). Here we propose IPSC, another new method to predict drug indications using supervised classification (see Figure 2C). We consider predicting drug indications based on target proteins of the query drug (including the primary target and off-targets) and various molecular features of diseases (e.g., disease-causing genes, diagnostic markers, disease-related pathways, and environmental factors) in a supervised framework with partially known drug−disease associations. Our proposed IPSC method is described in detail below.

Drug indication prediction is regarded as a problem of drug−disease network inference and can be formulated by supervised classification of drug−disease pairs. A pair made up of drug Y and disease Z is represented by (Y, Z). Each drug−disease pair (Y, Z) is represented by a feature vector as $\Phi(Y, Z)$. We use a function, $f(Y, Z) = \mathbf{w}^T \Phi(Y, Z)$, to predict whether the drug−disease pair (Y, Z) is an associating pair or not. The weight vector $\mathbf{w}$ is optimized on the basis of a learning set with label information. The learning set consists of drug−disease pairs $(Y_i,$

$Z_j)$ $(i = 1, 2, ..., n_Y; j = 1, 2, ..., n_Z)$, where the drug−disease pairs are known to be association pairs (positive pairs) or not (negative pairs). n_Y is the number of drugs, and n_Z is the number of diseases in the learning set.

The target profile of drug Y is represented by a p-dimensional binary vector as follows: $Y = (y_1, y_2, ..., y_p)^T$, where $y_k \in \{0, 1\}$, $k = 1, ..., p$. In the same manner, the molecular profile of disease Z is represented by a q-dimensional binary vector as follows: $Z = (z_1, z_2, ..., z_q)^T$, where $z_l \in \{0, 1\}$, $l = 1, ..., q$.

We represent each drug−disease pair as a high-dimensional feature vector using the tensor product between Y and Z as follows:

$$\Phi(Y, Z) = Y \otimes Z = (y_1 z_1, ..., y_1 z_q, ..., y_p z_1, ..., y_p z_q)^T$$

where $\Phi(Y, Z)$ is a $(p \times q)$-dimensional feature vector. This tensor product descriptor is similar to that in a previous work on a different biological problem.[40]

We consider a learning set of drug−disease pairs and association labels $(\Phi(Y_i, Z_j), a_{ij})$, $a_{ij} \in \{+1, -1\}$ $(i = 1, 2, ..., n_Y; j = 1, 2, ..., n_Z)$. The weight vector $\mathbf{w}$ of the linear logistic regression is learned with L_1 regularization as follows:

$$\min_{\mathbf{w}} \left\| \mathbf{w} \right\|_1 + C \sum_{i=1}^{n_Y} \sum_{j=1}^{n_Z} \log(1 + \exp(-a_{ij} \mathbf{w}^T \Phi(Y_i, Z_j)))$$

where $\|\cdot\|_1$ is L_1 norm (the sum of absolute values) and C is a regularization parameter to control the sparsity penalty. To handle the overfitting problem, we introduced L_1 regularization in the learning process. L_1 regularization has an effect that

makes the weights of uninformative features zeros without a loss of classification accuracy. Recently, efficient learning algorithms of linear models for high-dimensional feature vectors have been developed and commonly used in many practical applications.[41] The dependent variable is an association label: yes (+1) or no (−1). We used all possible drug—disease pairs excluding positive pairs as negative pairs to train the model. We did not consider the difference between target activation and target inhibition, because the number of therapeutic targets with such information is very limited.

Performance Evaluation Protocol for Drug Target Estimation. We evaluate the performance of our proposed TESS method for drug target prediction by performing the following 5-fold CV experiment.

First, we split all drugs in the gold standard drug—target interaction data set into five subsets. Second, we use each subset of drugs as test drugs. Third, we remove compounds that are identical to test drugs from the chemical—protein interactome data. Fourth, we perform TESS to predict targets of test drugs using the reconstructed chemical—protein interactome data. Finally, we evaluate the prediction accuracy for test drugs.

The performance is compared with a previous method known as the similarity ensemble approach (SEA).[42,43] Commercial software and databases [e.g., the MDL Drug Data Report (MDDR) database] were used in the SEA algorithm in the original paper; therefore, the SEA algorithm was implemented using public databases as an approximation in this study. More detailed explanations can be found in the Supporting Information.

Performance Evaluation Protocol for Drug Indication Prediction. Considering various practical situations for drug indication prediction, we evaluate the performance of our proposed IPTM and IPSC methods by performing the following two types of 5-fold CV experiments: pair-based CV and drug-based CV.

We perform the pair-based CV as follows. First, we randomly split all drug—disease pairs in the gold standard data set into five subsets of roughly equally sizes in an independent manner. Second, we use one subset of drug—disease pairs as a test set and use the remaining four subsets of drug—disease pairs as a training set. Third, we optimize a predictive model based on drug—disease pairs in the training set. Finally, we apply the predictive model to drug—disease pairs in the test set. Note that drugs and diseases in test pairs sometimes overlap with those in the training set.

We perform the drug-based CV as follows. First, we randomly split drugs in the gold standard data set into five subsets of drugs. Second, we use one subset of drugs as test drugs and use the remaining four subsets of drugs as training drugs. Third, we learn a predictive model based on drug—disease pairs consisting of training drugs and all diseases. Finally, we compute the prediction scores for drug—disease pairs consisting of test drugs versus all diseases. Note that drugs in the test pair set are completely different from those in the training pair set.

We perform the above CV procedures on two gold standard data sets: the Full data set and the Common data set. The performance is compared with previous methods: PREDICT,[21] PreDR,[23] and Analyses of Phenotypic and Molecular Data (APMD).[25] More detailed explanations can be found in the Supporting Information.

RESULTS

Performance Evaluation of the Proposed Method for Drug Target Prediction. We tested the proposed TESS method on its ability to reconstruct 21 646 known drug—target interactions involving 2787 drugs and 1497 target proteins. We compared the performance of TESS with that of the previous SEA method which is also based on large-scale compound—protein interactions.[42,43] We compared three different descriptors, KCF-S,[30] CDK,[29] and ECFPs,[27] in the framework of each method. The performance depends on the descriptors of chemical structures; thus, we tested different chemical descriptors. As phenotypic effect information is available only for approved drugs, we did not test the phenotypic effects in this subsection. We evaluated the performance of the methods by two types of CV: pair-based CV and drug-based CV (see Methods for more details).

We evaluated the prediction accuracy using receiver operating characteristic (ROC) and precision-recall (PR) curves. The ROC curve is a plot of true positive rates (correctly predicted pairs) against false positive rates (incorrectly predicted pairs), and the area under the ROC curve (AUC) returns 1 for perfect inference and 0.5 for random inference. The PR curve is a plot of precision (positive predictive value) against recall (sensitivity), and the area under the PR curve (AUPR) returns 1 for perfect inference and returns the ratio of positive samples to all samples for random inference. We calculated the AUC and AUPR scores in two different ways: using global evaluation and using local evaluation. In the global evaluation, we calculated the AUC and AUPR scores for all drug—target pairs at one time; whereas in the local evaluation, we calculated the AUC and AUPR scores of each drug for all target proteins and took their averages over all drugs. For drugs in the Full data set, missing features were imputed by zero-elements. In practice, we are often faced with the situation in which drugs of interest do not have all features. Note that the cross-validation experiment for the Full data set simulates such a practical situation. We calculated the AUC/AUPR scores by applying the predictive model to the imputed features.

Table 1 shows the results of the CV experiments. For all three descriptors, TESS provided higher AUC and AUPR

Table 1. AUC and AUPR Scores in the CV Experiments for Drug Target Estimation

method	TESS			SEA		
descriptor	KCF-S	CDK	ECFP	KCF-S	CDK	ECFP
Global Evaluation						
AUC	0.917	0.896	0.909	0.782	0.779	0.772
AUPR	0.203	0.189	0.203	0.115	0.130	0.115
Local Evaluation						
AUC	0.962	0.956	0.960	0.872	0.873	0.865
AUPR	0.482	0.453	0.489	0.342	0.314	0.343

scores than SEA. In a comparison among chemical descriptors, KCF-S and ECFP were comparable, followed by CDK. Although the AUC and AUPR scores were lower for global evaluation compared with local evaluation, the tendency of the difference among the three descriptors was same for the global and local evaluations. AUPR tends to be lower when the number of positive samples is much smaller than that of negative samples. In practical situations, it is difficult to

Table 2. AUC and AUPR Scores in the CV Experiments for Drug Indication Prediction[a]

	proposed method: IPTM				proposed method: IPSC				previous methods		
	known	known + new: chem sim	known + new: phen sim	known + new: chem sim + phen sim	known	known + new: chem sim	known + new: phen sim	known + new: chem sim + phen sim	PREDICT	PreDR	APMD
Global Evaluation											
Full data set											
Pair-based CV											
AUC	0.660	0.671	0.677	0.683	0.753	0.904	0.853	0.913	0.588	0.584	0.761
AUPR	0.042	0.051	0.049	0.049	0.208	0.414	0.407	0.472	0.038	0.012	0.409
Drug-based CV											
AUC	0.659	0.668	0.675	0.681	0.707	0.887	0.843	0.896	0.582	0.564	0.748
AUPR	0.041	0.050	0.048	0.048	0.153	0.353	0.309	0.415	0.027	0.011	0.355
Common data set											
Pair-based CV											
AUC	0.841	0.848	0.868	0.871	0.868	0.920	0.927	0.931	0.668	0.614	0.929
AUPR	0.096	0.123	0.117	0.118	0.326	0.480	0.621	0.632	0.078	0.027	0.618
Drug-based CV											
AUC	0.839	0.848	0.866	0.870	0.798	0.883	0.915	0.913	0.633	0.596	0.910
AUPR	0.095	0.120	0.115	0.115	0.265	0.426	0.598	0.602	0.049	0.023	0.618
Local Evaluation											
Full data set											
Pair-based CV											
AUC	0.701	0.732	0.720	0.736	0.725	0.846	0.772	0.853	0.554	0.556	0.626
AUPR	0.123	0.145	0.137	0.138	0.288	0.440	0.395	0.467	0.092	0.037	0.146
Drug-based CV											
AUC	0.699	0.730	0.719	0.734	0.715	0.832	0.767	0.847	0.599	0.562	0.622
AUPR	0.122	0.145	0.137	0.138	0.280	0.431	0.390	0.459	0.110	0.036	0.144
Common data set											
Pair-based CV											
AUC	0.831	0.842	0.847	0.850	0.875	0.893	0.898	0.901	0.602	0.588	0.769
AUPR	0.184	0.230	0.213	0.213	0.469	0.543	0.562	0.591	0.171	0.061	0.340
Drug-based CV											
AUC	0.831	0.842	0.848	0.850	0.856	0.883	0.893	0.892	0.653	0.589	0.768
AUPR	0.187	0.232	0.213	0.214	0.455	0.549	0.573	0.586	0.201	0.057	0.348

[a]Abbreviations chem sim and phen sim represent chemical similarity and phenotypic similarity, respectively; "chem sim + phen sim" represents the integration of drug targets estimated by chemical similarities and phenotypic similarities; "known" indicates that only known drug−target interactions are used; "known + new" means that known and newly predicted drug−target interactions are used.

correctly distinguish a few positive samples from many negative samples. Therefore, AUPR may represent the practical difficulty rather than AUC. Taken together, these results suggest that TESS with the KCF-S or ECFP descriptors is the best combination to predict drug−target interactions.

Performance Evaluation of the Proposed Methods for Drug Indication Prediction. We tested the proposed IPTM and IPSC methods on their abilities to reconstruct known drug−disease associations in two gold standard data sets: the Full data set and the Common data set. We compared the performance with the previous methods: PREDICT,[21] PreDR,[23] and APMD.[25] We evaluated the performance of the methods using two types of CV: pair-based CV and drug-based CV (see Methods for more details).

The left four columns in Table 2 show the results of the CV experiments for IPTM. In the "known" column, only known drug targets are used as predictors; whereas in the "known + new" columns, known and newly estimated drug targets are used as predictors. The AUC and AUPR scores with estimated drug target information were higher than those with known drug target information only. This finding implies that the performance for drug indication prediction was improved by additionally using estimated drug target information obtained

by TESS with chemical similarity, phenotypic similarity, or both. This tendency was retained for both pair-based and drug-based CV and for both Full and Common data sets. These results suggest that it is particularly important to take into account all potential drug target proteins for drug indication prediction.

The middle four columns in Table 2 show the results of the CV experiments for IPSC. The AUC and AUPR scores of IPSC were higher than those of IPTM. One explanation for this observation is that IPTM depends on the disease−target association template, but information on therapeutic target proteins is not available for all diseases in the gold standard data set. In contrast, IPSC can predict drug indications for diseases of unknown therapeutic targets by using other molecular features of diseases (e.g., disease-causing genes, diagnostic markers, disease-related pathways, and environmental factors), so the coverage of IPSC is higher than that of IPTM. Therefore, the resulting prediction accuracy of IPTM is lower than that of IPSC. However, it should be noted that IPTM could provide more interpretable prediction results, which will be illustrated in the next subsection.

In the comparison among data sources used for drug target estimation, phenotypic similarity worked better than chemical
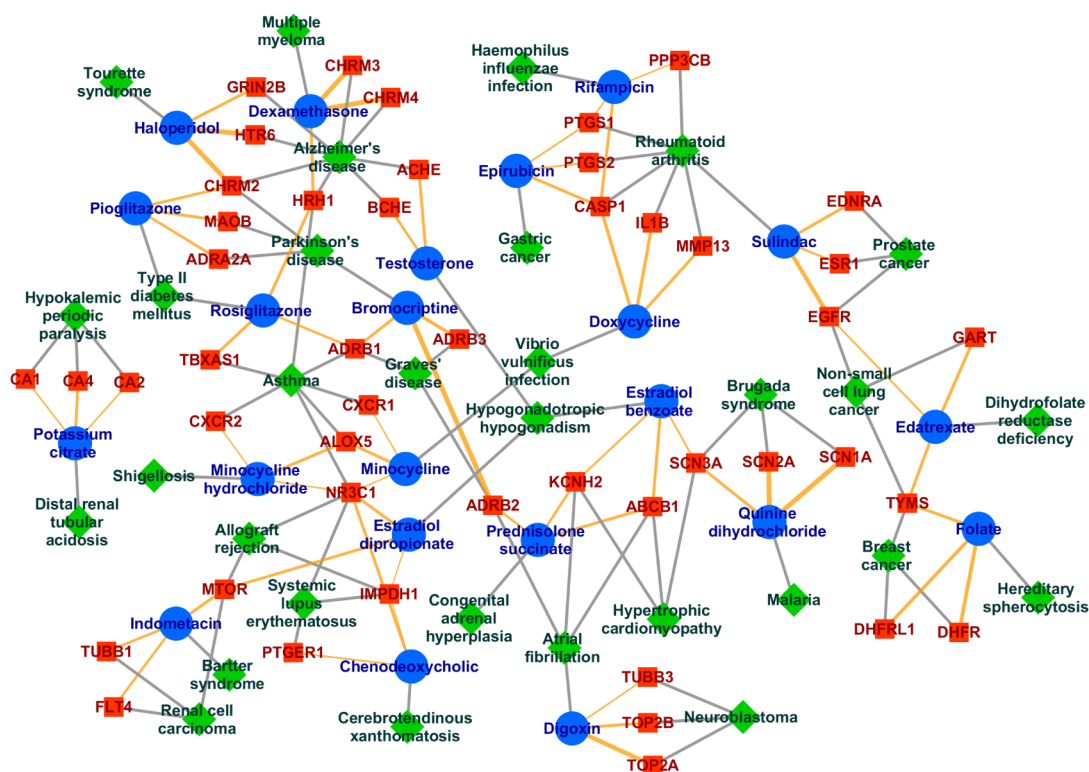
Figure 3. Small part of the drug−target−disease network predicted by IPTM. Blue circles, red rectangles, and green rhombuses indicate drugs, proteins, and diseases, respectively. Orange lines and gray lines indicate newly predicted interactions and known associations, respectively. The width of each orange line represents the prediction score.

Table 3. Examples of High Scoring Predictions by IPTM Based on Chemical and Phenotypic Similarities

drug	original indication	new indication	protein	score	confirmed ref
haloperidol lactate	Tourette syndrome	Alzheimer's disease	GRIN2B	1.000	62
bromocriptine mesilate	Parkinson's disease	Graves' disease	ADRB1	1.000	63
prednisolone succinate	congenital adrenal hyperplasia	atrial fibrillation	ABCB1	1.000	64
estradiol benzoate	hypogonadotropic hypogonadism	hypertrophic cardiomyopathy	ABCB1	1.000	65
indometacin sodium	Bartter syndrome	renal cell carcinoma	MTOR	1.000	66
quinine dihydrochloride	malaria	Brugada syndrome	SCN3A	1.000	67
estradiol dipropionate	hypogonadotropic hypogonadism	allograft rejection	NR3C1	1.000	68
folate sodium	hereditary spherocytosis	breast cancer	TYMS	1.000	69
rosiglitazone maleate	type II diabetes mellitus	asthma	TBXAS1	1.000	70
minocycline	*Vibrio vulnificus* infection	asthma	ALOX5	1.000	71
chenodeoxycholic acid	cerebrotendinous xanthomatosis	systemic lupus erythematosus	NR3C1	1.000	72
dexamethasone phosphate	multiple myeloma	alzheimer's disease	HRH1	1.000	73
epirubicin hydrochloride	gastric cancer	rheumatoid arthritis	CASP1	0.996	74
minocycline hydrochloride	shigellosis	asthma	ALOX5	0.996	71
doxycycline	*Vibrio vulnificus* infection	rheumatoid arthritis	MMP13	0.996	75
rifampicin	*Haemophilus influenzae* infection	rheumatoid arthritis	CASP1	0.995	76
potassium citrate	distal renal tubular acidosis	hypokalemic periodic paralysis	CA4	0.985	77
edatrexate	dihydrofolate reductasedeficiency	nonsmall cell lung cancer	TYMS	0.941	78
sulindac	rheumatoid arthritis	prostate cancer	EDNRA	0.936	79
digoxin	atrial fibrillation	neuroblastoma	TOP2B	0.925	80

similarity for the Common data set, while chemical similarity worked better or almost the same as phenotypic similarity for the Full data set. One possible explanation is that the number of drugs with phenotypic effect information is smaller than that of drugs with chemical structure information. For both Full and Common data sets, the best performance was achieved by integrating known and estimated drug targets based on both chemical similarities and phenotypic similarities in most cases. This tendency was observed in both IPTM and IPSC. The

integrative model works the best in AUC, but the known + chemical similarity model works the best in AUPR.

The right three columns in Table 2 show the results of the CV experiments for the previous methods. The proposed methods IPTM and IPSC outperformed the previous methods PREDICT, PreDR, and APMD in most cases. This result suggests that our proposed methods are more useful than the previous methods and that the use of large-scale chemical−
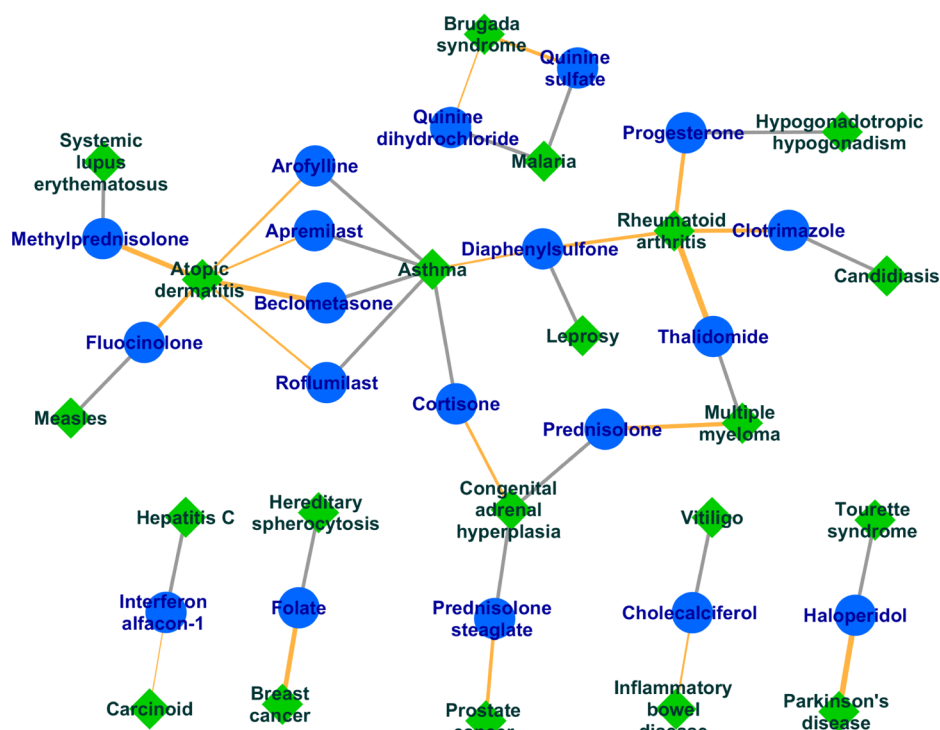
Figure 4. Small part of the drug−disease network predicted by IPSC. Blue circles and green rhombuses indicate drugs and diseases, respectively. Orange lines and gray lines indicate newly predicted associations and known associations, respectively. The width of each orange line represents the prediction score.

Table 4. Examples of High Scoring Predictions by IPSC Based on Chemical and Phenotypic Similarities

drug	original indication	new indication	score	confirmed ref
thalidomide	multiple myeloma	rheumatoid arthritis	4.714	81
haloperidol decanoate	Tourette syndrome	Parkinson's disease	4.642	82
beclometasone dipropionate	asthma	atopic dermatitis	4.267	83
methylprednisolone aceponate	systemic lupus erythematosus	atopic dermatitis	4.181	84
folate sodium	hereditary spherocytosis	breast cancer	4.054	69
prednisolone	congenital adrenal hyperplasia	multiple myeloma	3.718	85
clotrimazole	candidiasis	rheumatoid arthritis	3.635	86
progesterone	hypogonadotropic hypogonadism	rheumatoid arthritis	3.527	87
fluocinolone acetonide	measles	atopic dermatitis	3.375	88
quinine sulfate hydrate	malaria	Brugada syndrome	3.310	67
diaphenylsulfone	leprosy	rheumatoid arthritis	3.220	89
prednisolone steaglate	congenital adrenal hyperplasia	prostate cancer	3.217	90
cortisone acetate	asthma	congenital adrenal hyperplasia	2.858	91
arofylline	asthma	atopic dermatitis	2.596	92
apremilast	asthma	atopic dermatitis	2.380	93
roflumilast	asthma	atopic dermatitis	2.380	94
diaphenylsulfone	leprosy	asthma	2.345	95
cholecalciferol	vitiligo	inflammatory bowel disease	2.227	96
quinine dihydrochloride	malaria	Brugada syndrome	1.955	67
interferon alfacon-1	hepatitis C	carcinoid	1.677	97

protein interactome data is meaningful for drug indication prediction in practice.

Large-Scale Prediction of Novel Drug Indications. Finally, we conducted a comprehensive prediction of novel drug indications for 8270 drugs registered in Japan, USA, and Europe by applying IPTM and IPSC with drug target profiles consisting of known drug targets and drug targets newly estimated by TESS. We used the full disease−target association template for IPTM and learned a supervised classifier for IPSC based on all gold standard data as training data. For prediction results of novel drug indications, IPTM predicted 196 048 drug−disease associations involving 6531 drugs and 250 diseases and IPSC predicted 434 065 drug−disease associations involving 6301 drugs and 762 diseases.

Figure 3 shows a small part of the drug−target−disease association network predicted by IPTM with merged target protein information estimated from chemical and phenotypic similarities. Other prediction results are shown in the Supporting Information. An important feature of IPTM is that it can provide mechanistic interpretations for the predicted
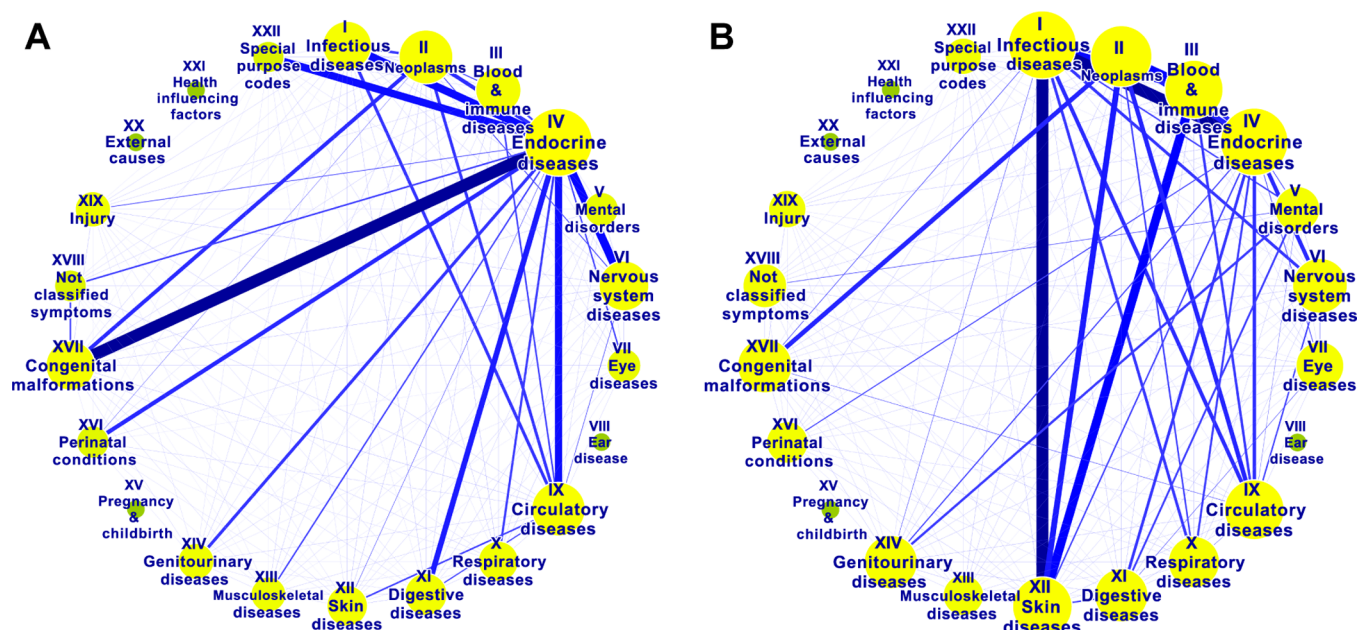
Figure 5. Distribution of drugs repositioned from the original disease class to other disease classes by (A) IPTM with chemical similarity and (B) IPTM with phenotypic similarity. Nodes represent the ICD disease chapters with chapter number and short chapter name. Edge width indicates the ratio of the number of repositioned drugs between two disease chapters against the number of drugs in each disease chapter. Node size indicates the sum of the edge widths of each node. Full names of ICD disease chapters are as follows. Chapter I: certain infectious and parasitic diseases (A00−B99); chapter II: neoplasms (C00−D48); chapter III: diseases of the blood, blood-forming organs, and certain disorders involving the immune mechanism (D50−D89); chapter IV: endocrine, nutritional, and metabolic diseases (E00−E90); chapter V: mental and behavioral disorders (F00−F99); chapter VI: diseases of the nervous system (G00−G99); chapter VII: diseases of the eye and adnexa (H00−H59); chapter VIII: diseases of the ear and mastoid process (H60−H95); chapter IX: diseases of the circulatory system (I00−I99); chapter X: diseases of the respiratory system (J00−J99); chapter XI: diseases of the digestive system (K00−K93); chapter XII: diseases of the skin and subcutaneous tissue (L00−L99); chapter XIII: diseases of the musculoskeletal system and connective tissue (M00−M99); chapter XIV: diseases of the genitourinary system (N00−N99); chapter XV: pregnancy, childbirth, and the puerperium (O00−O99); chapter XVI: certain conditions originating in the perinatal period (P00−P96); chapter XVII: congenital malformations, deformations, and chromosomal abnormalities (Q00−Q99); chapter XVIII: symptoms, signs, and abnormal clinical and laboratory findings not elsewhere classified (R00−R99); chapter XIX: injury, poisoning, and certain other consequences of external causes (S00−T98); chapter XX: external causes of morbidity and mortality (V01−Y98); chapter XXI: factors influencing health status and contact with health services (Z00−Z99); chapter XXII: codes for special purposes (U00−U99).

drug indications, i.e., the target proteins used for the prediction of new drug indications. In the network shown in the figure, those proteins are represented as red rectangles, which are connected to both to drugs (blue circles) and diseases (green rhombuses). Table 3 shows examples of high scoring predictions that were confirmed by the literature.

Figure 4 shows a small part of the drug−disease association network predicted by IPSC with merged target protein information estimated from chemical and phenotypic similarities. Other prediction results are shown in the Supporting Information. Only drug−disease interactions are indicated in the figure. Unlike IPTM, it is difficult to see the prediction process of IPSC. However, the advantage of IPSC is its higher prediction accuracy with a larger number of predictable drugs and higher coverage of a wide range of diseases. Table 4 shows examples of high scoring predictions that were confirmed by the literature.

Distribution of Repositioned Drugs Across Different ICD-10 Disease Classes for Prediction with Chemical Similarity and Phenotypic Similarity. Figure 5A and B shows distributions of drugs repositioned from the original disease class to other disease classes by IPTM with chemical similarity and phenotypic similarity, respectively. Original drug indications and newly predicted indications were classified according to ICD-10[37] disease classes, referred to as "ICD chapter." In the networks shown in the figures, nodes represent

ICD disease chapters with their chapter numbers and short chapter names and width of the lines indicates the ratio of the number of repositioned drugs between two disease chapters against the number of drugs in each disease chapter. The detailed numbers of repositioned drugs by IPTM with chemical similarity and phenotypic similarity are shown in Table S1 and Table S2, respectively.

Figure 5A shows the distribution of the repositioned drugs between different ICD disease chapters in the prediction result based on chemical similarity. The largest number of repositioned drugs was observed between Chapter IV and Chapter XVII (congenital malformations, deformations, and chromosomal abnormalities), followed by that between Chapters IV and VI (diseases of the nervous system) and between Chapters I and IV. Figure 5B shows the distribution of the repositioned drugs between different ICD disease chapters in the prediction result based on phenotypic similarity. The largest number of repositioned drugs was observed between Chapter I (certain infectious and parasitic diseases) and Chapter XII (diseases of the skin and subcutaneous tissue), followed by between Chapters I and IV (endocrine, nutritional, and metabolic diseases) and between Chapters II (neoplasms) and XII (diseases of the skin and subcutaneous tissue). These results suggest that chemical similarity-based predictions and phenotypic similarity-based predictions are complementary to each other.
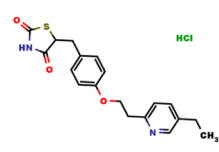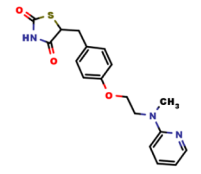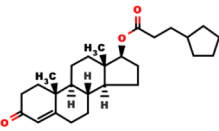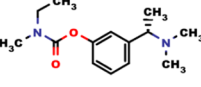
Figure 6. Examples of prediction based on chemical similarity and phenotypic similarity.

Examples of Drug Indications Predicted with Chemical Similarity and Phenotypic Similarity. Figure 6A shows an example of the predictions by IPTM with target protein information estimated by chemical similarity. Pioglitazone hydrochloride is a drug used to treat type II diabetes and its primary target is a peroxisome proliferator-activated receptor (PPARγ). Pioglitazone hydrochloride was predicted to work for Parkinson's disease because monoamine oxidase B (MAOB) was estimated as an off-target protein of pioglitazone hydrochloride by TESS based on chemical similarity. This prediction is reasonable from the viewpoint of disease mechanism. Levels of dopamine are decreased in the brain of patients with Parkinson's disease.[44] MAOB functions as a dopamine degradation enzyme. Therefore, inhibition of MAOB increases dopamine levels in the brain. Some drugs for Parkinson's disease are MAOB inhibitors. We investigated the validity of the prediction using independent resources and confirmed that the application of pioglitazone hydrochloride for the treatment of Parkinson's disease based on its inhibition of MAOB was reported in the literature.[45] Among compounds that interact with MAOB, the most chemically similar one to pioglitazone hydrochloride was a compound denoted as YASAKCUCGLMORW-UHFFFAOYSA-N (InChIKey), shown in the right most column in Figure 6A. This estimation was performed using chemical structure similarity; therefore, the chemical structures of pioglitazone hydrochloride and YASAKCUCGLMORW-UHFFFAOYSA-N are very similar. These results suggest that drug indication prediction with chemical similarity is brought from compounds that have similar chemical structure to query drugs.

Figure 6B shows another example of the predictions by IPTM with target protein information estimated by phenotypic similarity. Testosterone is a steroid hormone drug used to treat hormone deficiency diseases such as hypogonadotropic hypogonadism. Testosterone was predicted to work for Alzheimer's disease because acetylcholinesterase (AChE) was estimated to interact with testosterone by TESS based on phenotypic similarity. This prediction is reasonable from the viewpoint of the disease mechanism. Acetylcholine is decreased in the brain of patients with Alzheimer's disease.[46] Some drugs for Alzheimer's disease are AChE inhibitors. We investigated the validity of the prediction using independent resources and confirmed that the use of testosterone to treat Alzheimer's

disease has also been reported in the literature.[47] Among compounds which interact with AChE, the most phenotypically similar one to testosterone was rivastigmine, which is a known drug for Alzheimer's disease, shown in right side column in Figure 6B. This estimation was performed using phenotypic similarity; therefore, the chemical structures of testosterone and rivastigmine are quite different. These results suggest that drug indication prediction with phenotypic similarity is not dependent on the chemical structures of drugs.

■ DISCUSSION

In this study, we proposed new computational methods for predicting new drug indications on the basis of a global drug−target interaction network revealed by large-scale chemical−protein interactome data. The originality of the proposed methods lies in interpretability of the drug indication prediction process, in applicability to a wide range of diseases, and in high prediction accuracy with state-of-the-art machine learning methods. The cross-validation experiments demonstrated that the proposed IPTM and IPSC methods outperformed the previous methods for drug indication prediction. One explanation regarding the better performance is that the prediction is performed on the basis of the target protein interaction profiles consisting of both known and newly predicted target proteins by TESS because most phenotypic effects of drugs are caused by interactions between drug molecules and all possible target proteins (including the primary targets and off-targets). The proposed methods are expected to be useful for systematic drug repositioning.

One of the properties of IPTM is that it can provide the information on the mechanism of drug actions corresponding to the prediction results in terms of target proteins, as shown in drug−target−disease association networks. This property is important for the practical use of the prediction methods, although it lowers the prediction accuracy to some extent compared with IPSC. A variety of computational methods have been proposed for systematic drug repositioning,[12−19,21−25] but most of the previous methods cannot provide any biological interpretation about the prediction results. Some previous works were conducted from the viewpoints of drug targets. For example, known target proteins of drugs were linked to disease susceptibility genes via protein−protein interactions.[18] However, the coverage of known target proteins and the range of

diseases are limited, and the susceptibility genes of diseases are not always the therapeutic targets of diseases. The docking score profiles of drug−protein pairs were linked to drug indications or side effects,[20] but information on the protein binding pocket is required. Thus, the practical applicability may be limited because the 3D structures of most pharmaceutically useful membrane proteins (e.g., GPCRs) remain unknown.

The identification of all possible target proteins of drugs is crucial for a better understanding of polypharmacology. A variety of computational methods in the ligand-based framework have been proposed for predicting drug−target interactions on a genomewide scale.[42,48−55] However, the predictive models of most previous methods are learned based on known drug−target interactions involving only known drugs; therefore, the size of learning interaction set is relatively small. The learning algorithms of the previous methods are computationally expensive; therefore, the previous methods except for SEA were not computationally feasible for learning millions of compound−protein interactions used in this study. An important feature of TESS is the capacity for large-scale drug target prediction with large amounts of chemical−protein interactome data in the ligand-based framework. SEA is useful, particularly for proteins for which many ligands are available, but it does not work for proteins with a few ligands because the prediction is based on the average of ligand similarities. In contrast, TESS is applicable for proteins for which at least one ligand is available and it is computationally efficient. Another approach for drug target prediction is the docking-based framework based on docking simulations with protein 3D structures.[56−60] The docking-based methods can be used for predicting the ligands of orphan proteins, but the protein 3D structures are required. The ligand-based framework and the docking-based framework are complementary to each other; thus, the integration of the two frameworks would be an interesting research direction.

In the similarity search for drug target prediction, the computation of drug−compound similarity is crucial. In this study, we used two similarity measures: chemical similarity and phenotypic similarity. Our results show that chemical similarity and phenotypic similarity are complementary to each other; therefore, the integration of the predictions results based on different similarities is beneficial. Recently, large-scale drug-induced gene expression profiles have been published,[12,61] which is an important resource for evaluating drug similarity. The inclusion of these gene expression data will be one of our important future works.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00330.

Detailed explanations concerning the implementation of previous methods are presented in the "Implementation of the previous methods" section. Detailed prediction results and data sets can be found in the "Detailed results and datasets" section. Figure S1 shows the results for IPTM with target information estimated only by chemical similarity. Figure S2 shows the results for IPTM with target information estimated only by phenotypic similarity. Figure S3 shows the results for IPSC with target information estimated only by chemical similarity. Figure S4 shows the results for IPSC with

target information estimated only by phenotypic similarity. Figure S5 shows the distributions of highest similarity from both the chemical structures and phenotypic effects of drugs. Table S1 and Table S2 show the numbers of drugs repositioned from the original disease class to other disease classes by IPTM with chemical similarity and phenotypic similarity, respectively (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: yamanishi@bioreg.kyushu-u.ac.jp. Phone: +81-92-642-6699. Fax: +81-92-642-6692.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) DiMasi, J. A.; Hansen, R. W.; Grabowski, H. G. The Price of Innovation: New Estimates of Drug Development Costs. *J. Health Econ.* **2003**, *22*, 151−185.

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

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

(4) Novac, N. Challenges and Opportunities of Drug Repositioning. *Trends Pharmacol. Sci.* **2013**, *34*, 267−272.

(5) Whitebread, S.; Hamon, J.; Bojanic, D.; Urban, L. Keynote Review: *in Vitro* Safety Pharmacology Profiling: an Essential Tool for Successful Drug Development. *Drug Discovery Today* **2005**, *10*, 1421−1433.

(6) Blagg, J. Structure-activity Relationships for in Vitro and in Vivo Toxicity. *Annu. Rep. Med. Chem.* **2006**, *41*, 353−368.

(7) Lin, S. F.; Xiao, K. T.; Huang, Y. T.; Chiu, C. C.; Soo, V. W. Analysis of Adverse Drug Reactions Using Drug and Drug Target Interactions and Graph-based Methods. *Artif. Intell. Med.* **2010**, *48*, 161−166.

(8) Hopkins, A. L.; Mason, J. S.; Overington, J. P. Can We Rationally Design Promiscuous Drugs? *Curr. Opin. Struct. Biol.* **2006**, *16*, 127−136.

(9) Nacher, J. C.; Schwartz, J. M. Modularity in Protein Complex and Drug Interactions Reveals New Polypharmacological Properties. *PLoS One* **2012**, *7*, e30028.

(10) Csermely, P.; Korcsmáros, T.; Kiss, H. J.; London, G.; Nussinov, R. Structure and Dynamics of Molecular Networks: a Novel Paradigm of Drug Discovery: a Comprehensive Review. *Pharmacol. Ther.* **2013**, *138*, 333−408.

(11) Terstappen, G. C.; Reggiani, A. *in Silico* Research in Drug Discovery. *Trends Pharmacol. Sci.* **2001**, *22*, 23−26.

(12) Lamb, J.; Crawford, E. D.; Peck, D.; Modell, J. W.; Blat, I. C.; Wrobel, M. J.; Lerner, J.; Brunet, J.-P.; Subramanian, A.; Ross, K. N.; Reich, M.; Hieronymus, H.; Wei, G.; Armstrong, S. A.; Haggarty, S. J.; Clemons, P. A.; Wei, R.; Carr, S. A.; Lander, E. S.; Golub, T. R. The Connectivity Map: Using Gene-expression Signatures to Connect Small Molecules, Genes, and Disease. *Science* **2006**, *313*, 1929−1935.

(13) Hu, G.; Agarwal, P. Human Disease-drug Network Based on Genomic Expression Profiles. *PLoS One* **2009**, *4*, e6536.

(14) Iorio, F.; Bosotti, R.; Scacheri, E.; Belcastro, V.; Mithbaokar, P.; Ferriero, R.; Murino, L.; Tagliaferri, R.; Brunetti-Pierri, N.; Isacchi, A.;

di Bernardo, D. Discovery of Drug Mode of Action and Drug Repositioning from Transcriptional Responses. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14621–14626.

(15) Sirota, M.; Dudley, J. T.; Kim, J.; Chiang, A. P.; Morgan, A. A.; Sweet-Cordero, A.; Sage, J.; Butte, A. J. Discovery and Preclinical Validation of Drug Indications Using Compendia of Public Gene Expression Data. *Sci. Transl. Med.* **2011**, *3*, 96ra77–96ra77.

(16) Chiang, A. P.; Butte, A. J. Systematic Evaluation of Drug-disease Relationships to Identify Leads for Novel Drug Uses. *Clin. Pharmacol. Ther.* **2009**, *86*, 507–510.

(17) Ye, H.; Tang, K.; Yang, L.; Cao, Z.; Li, Y. Study of Drug Function Based on Similarity of Pathway Fingerprint. *Protein Cell* **2012**, *3*, 132–139.

(18) Zhao, S.; Li, S. A Co-module Approach for Elucidating Drug-disease Associations and Revealing Their Molecular Basis. *Bioinformatics* **2012**, *28*, 955–961.

(19) Ye, H.; Liu, Q.; Wei, J. Construction of Drug Network Based on Side Effects and Its Application for Drug Repositioning. *PLoS One* **2014**, *9*, e87864.

(20) Luo, H.; Chen, J.; Shi, L.; Mikailov, M.; Zhu, H.; Wang, K.; He, L.; Yang, L. DRAR-CPI: a Server for Identifying Drug Repositioning Potential and Adverse Drug Reactions via the Chemical-protein Interactome. *Nucleic Acids Res.* **2011**, *39*, W492–W498.

(21) Gottlieb, A.; Stein, G. Y.; Ruppin, E.; Sharan, R. PREDICT: a Method for Inferring Novel Drug Indications with Application to Personalized Medicine. *Mol. Syst. Biol.* **2011**, *7*, 496–496.

(22) Yang, L.; Agarwal, P. Systematic Drug Repositioning Based on Clinical Side-effects. *PLoS One* **2011**, *6*, e28025.

(23) Wang, Y.; Chen, S.; Deng, N.; Wang, Y. Drug Repositioning by Kernel-Based Integration of Molecular Structure, Molecular Activity, and Phenotype Data. *PLoS One* **2013**, *8*, e78518.

(24) Napolitano, F.; Zhao, Y.; Moreira, V. M.; Tagliaferri, R.; Kere, J.; D'Amato, M.; Greco, D. Drug Repositioning: a Machine-learning Approach Through Data Integration. *J. Cheminf.* **2013**, *5*, 30.

(25) Iwata, H.; Sawada, R.; Mizutani, S.; Yamanishi, Y. Systematic Drug Repositioning for a Wide Range of Diseases with Integrative Analyses of Phenotypic and Molecular Data. *J. Chem. Inf. Model.* **2015**, *55*, 446–459.

(26) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a Large-scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100–D1107.

(27) Rogers, D.; Hahn, M. Extended-connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.

(28) Steinbeck, C.; Han, Y.; Kuhn, S.; Horlacher, O.; Luttmann, E.; Willighagen, E. The Chemistry Development Kit (CDK): an Open-source Java Library for Chemo- and Bioinformatics. *J. Chem. Inf. Model.* **2003**, *43*, 493–500.

(29) Steinbeck, C.; Hoppe, C.; Kuhn, S.; Floris, M.; Guha, R.; Willighagen, E. L. Recent Developments of the Chemistry Development Kit (CDK) - an Open-source Java Library for Chemo- and Bioinformatics. *Curr. Pharm. Des.* **2006**, *12*, 2111–2120.

(30) Kotera, M.; Tabei, Y.; Yamanishi, Y.; Moriya, Y.; Tokimatsu, T.; Kanehisa, M.; Goto, S. KCF-S: KEGG Chemical Function and Substructure for Improved Interpretability and Prediction in Chemical Bioinformatics. *BMC Syst. Biol.* **2013**, *7* (Suppl 6), S2.

(31) Günther, S.; Kuhn, M.; Dunkel, M.; Campillos, M.; Senger, C.; Petsalaki, E.; Ahmed, J.; Urdiales, E. G.; Gewiess, A.; Jensen, L. J.; Schneider, R.; Skoblo, R.; Russell, R. B.; Bourne, P. E.; Bork, P.; Preissner, R. SuperTarget and Matador: Resources for Exploring Drug-target Relationships. *Nucleic Acids Res.* **2007**, *36*, D919–D922.

(32) 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.

(33) Roth, B. L.; Lopez, E.; Patel, S.; Kroeze, W. K. The Multiplicity of Serotonin Receptors: Uselessly Diverse Molecules or an Embarrassment of Riches? *Neuroscientist* **2000**, *6*, 252–262.

(34) Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for Representation and Analysis of Molecular Networks Involving Diseases and Drugs. *Nucleic Acids Res.* **2010**, *38*, D355–D360.

(35) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a Web-accessible Database of Experimentally Determined Protein-ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.

(36) Qin, C.; Zhang, C.; Zhu, F.; Xu, F.; Chen, S. Y.; Zhang, P.; Li, Y. H.; Yang, S. Y.; Wei, Y. Q.; Tao, L.; Chen, Y. Z. Therapeutic Target Database Update 2014: a Resource for Targeted Therapeutics. *Nucleic Acids Res.* **2014**, *42*, D1118–D1123.

(37) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*; World Health Organization: Geneva, 1992.

(38) Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; Yamanishi, Y. KEGG for Linking Genomes to Life and the Environment. *Nucleic Acids Res.* **2007**, *36*, D480–D484.

(39) Papadakis, M. A.; McPhee, S. J.; Rabow, M. W. *CURRENT Medical Diagnosis and Treatment 2014*; McGraw-Hill Medical: New York, 2013.

(40) Tabei, Y.; Pauwels, E.; Stoven, V.; Takemoto, K.; Yamanishi, Y. Identification of Chemogenomic Features from Drug-target Interaction Networks Using Interpretable Classifiers. *Bioinformatics* **2012**, *28*, i487–i494.

(41) Fan, R.-E.; Chang, K.-W.; Hsieh, C.-J.; Wang, X.-R.; Lin, C.-J. LIBLINEAR: a Library for Large Linear Classification. *J. Mach. Learn. Res.* **2008**, *9*, 1871–1874.

(42) 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.

(43) 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.

(44) Lotharius, J.; Brundin, P. Pathogenesis of Parkinson's Disease: Dopamine, Vesicles and Alpha-synuclein. *Nat. Rev. Neurosci.* **2002**, *3*, 932–942.

(45) Quinn, L. P.; Crook, B.; Hows, M. E.; Vidgeon-Hart, M.; Chapman, H.; Upton, N.; Medhurst, A. D.; Virley, D. J. The PPARgamma Agonist Pioglitazone Is Effective in the MPTP Mouse Model of Parkinson's Disease Through Inhibition of Monoamine Oxidase B. *Br. J. Pharmacol.* **2008**, *154*, 226–233.

(46) Anand, P.; Singh, B.; Singh, N. A Review on Coumarins as Acetylcholinesterase Inhibitors for Alzheimer's Disease. *Bioorg. Med. Chem.* **2012**, *20*, 1175–1180.

(47) Cherrier, M. M.; Matsumoto, A. M.; Amory, J. K.; Asthana, S.; Bremner, W.; Peskind, E. R.; Raskind, M. A.; Craft, S. Testosterone Improves Spatial Memory in Men with Alzheimer Disease and Mild Cognitive Impairment. *Neurology* **2005**, *64*, 2063–2068.

(48) Nagamine, N.; Sakakibara, Y. Statistical Prediction of Protein Chemical Interactions Based on Chemical Structure and Mass Spectrometry Data. *Bioinformatics* **2007**, *23*, 2004–2012.

(49) 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–i240.

(50) Faulon, J.-L.; Misra, M.; Martin, S.; Sale, K.; Sapra, R. Genome Scale Enzyme-metabolite and Drug-target Interaction Predictions Using the Signature Molecular Descriptor. *Bioinformatics* **2008**, *24*, 225–233.

(51) Jacob, L.; Vert, J.-P. Protein-ligand Interaction Prediction: an Improved Chemogenomics Approach. *Bioinformatics* **2008**, *24*, 2149–2156.

(52) Bleakley, K.; Yamanishi, Y. Supervised Prediction of Drug-target Interactions Using Bipartite Local Models. *Bioinformatics* **2009**, *25*, 2397–2403.

(53) 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.

(54) Yamanishi, Y.; Pauwels, E.; Saigo, H.; Stoven, V. Extracting Sets of Chemical Substructures and Protein Domains Governing Drug-target Interactions. *J. Chem. Inf. Model.* **2011**, *51*, 1183−1194.

(55) Tabei, Y.; Yamanishi, Y. Scalable Prediction of Compound-protein Interactions Using Minwise Hashing. *BMC Syst. Biol.* **2013**, *7* (Suppl 6), S3.

(56) Yang, L.; Wang, K.-J.; Wang, L.-S.; Jegga, A. G.; Qin, S.-Y.; He, G.; Chen, J.; Xiao, Y.; He, L. Chemical-protein Interactome and Its Application in Off-target Identification. *Interdiscip. Sci.: Comput. Life Sci.* **2011**, *3*, 22−30.

(57) Yang, L.; Wang, K.-J.; 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.

(58) Yang, L.; Chen, J.; Shi, L.; Hudock, M. P.; Wang, K.; He, L. Identifying Unexpected Therapeutic Targets via Chemical-protein Interactome. *PLoS One* **2010**, *5*, e9568.

(59) Yang, L.; Chen, J.; He, L. Harvesting Candidate Genes Responsible for Serious Adverse Drug Reactions from a Chemical-protein Interactome. *PLoS Comput. Biol.* **2009**, *5*, e1000441.

(60) Li, Y. Y.; An, J.; Jones, S. J. M. A Computational Approach to Finding Novel Targets for Existing Drugs. *PLoS Comput. Biol.* **2011**, *7*, e1002139.

(61) Vempati, U. D.; Chung, C.; Mader, C.; Koleti, A.; Datar, N.; Vidović, D.; Wrobel, D.; Erickson, S.; Muhlich, J. L.; Berriz, G.; Benes, C. H.; Subramanian, A.; Pillai, A.; Shamu, C. E.; Schürer, S. C. Metadata Standard and Data Exchange Specifications to Describe, Model, and Integrate Complex and Diverse High-Throughput Screening Data from the Library of Integrated Network-based Cellular Signatures (LINCS). *J. Biomol. Screening* **2014**, *19*, 803−816.

(62) Devanand, D. P.; Pelton, G. H.; Cunqueiro, K.; Sackeim, H. A.; Marder, K. A 6-month, Randomized, Double-blind, Placebo-controlled Pilot Discontinuation Trial Following Response to Haloperidol Treatment of Psychosis and Agitation in Alzheimer's Disease. *Int. J. Geriatr. Psychiatry* **2011**, *26*, 937−943.

(63) Demura, R.; Demura, H.; Odagiri, E.; Shizume, K. A Case of Acromegaly Associated with Graves' Disease. a Possible Role of Endogenous TRH and an Effect of Bromocriptine on GH and TSH Secretion. *Endocrinol. Jpn.* **1984**, *31*, 801−807.

(64) Koyama, T.; Tada, H.; Sekiguchi, Y.; Arimoto, T.; Yamasaki, H.; Kuroki, K.; Machino, T.; Tajiri, K.; Zhu, X. D.; Kanemoto-Igarashi, M.; Sugiyasu, A.; Kuga, K.; Nakata, Y.; Aonuma, K. Prevention of Atrial Fibrillation Recurrence with Corticosteroids after Radiofrequency Catheter Ablation: a Randomized Controlled Trial. *J. Am. Coll. Cardiol.* **2010**, *56*, 1463−1472.

(65) Chen, Y.; Zhang, Z.; Hu, F.; Yang, W.; Yuan, J.; Cui, J.; Hao, S.; Hu, J.; Zhou, Y.; Qiao, S. 17 ß-estradiol Prevents Cardiac Diastolic Dysfunction by Stimulating Mitochondrial Function: a Preclinical Study in a Mouse Model of a Human Hypertrophic Cardiomyopathy Mutation. *J. Steroid Biochem. Mol. Biol.* **2015**, *147*, 92−102.

(66) Ou, Y.-C.; Yang, C.-R.; Cheng, C.-L.; Raung, S.-L.; Hung, Y.-Y.; Chen, C.-J. Indomethacin Induces Apoptosis in 786-O Renal Cell Carcinoma Cells by Activating Mitogen-activated Protein Kinases and AKT. *Eur. J. Pharmacol.* **2007**, *563*, 49−60.

(67) Mehrotra, S.; Juneja, R.; Naik, N.; Pavri, B. B. Successful Use of Quinine in the Treatment of Electrical Storm in a Child with Brugada Syndrome. *J. Cardiovasc. Electrophysiol.* **2011**, *22*, 594−597.

(68) Zou, Y.; Steurer, W.; Klima, G.; Obrist, P.; Margreiter, R.; Brandacher, G. Estradiol Enhances Murine Cardiac Allograft Rejection under Cyclosporin and Can Be Antagonized by the Antiestrogen Tamoxifen. *Transplantation* **2002**, *74*, 354−357.

(69) Chen, P.; Li, C.; Li, X.; Li, J.; Chu, R.; Wang, H. Higher Dietary Folate Intake Reduces the Breast Cancer Risk: a Systematic Review and Meta-analysis. *Br. J. Cancer* **2014**, *110*, 2327−2338.

(70) Zhao, Y.; Huang, Y.; He, J.; Li, C.; Deng, W.; Ran, X.; Wang, D. Rosiglitazone, a Peroxisome Proliferator-activated Receptor-γ Agonist, Attenuates Airway Inflammation by Inhibiting the Proliferation of Effector T Cells in a Murine Model of Neutrophilic Asthma. *Immunol. Lett.* **2014**, *157*, 9−15.

(71) Naura, A. S.; Kim, H.; Ju, J.; Rodriguez, P. C.; Jordan, J.; Catling, A. D.; Rezk, B. M.; Elmageed, Z. Y. A.; Pyakurel, K.; Tarhuni, A. F.; Abughazleh, M. Q.; Errami, Y.; Zerfaoui, M.; Ochoa, A. C.; Boulares, A. H. Minocycline Blocks Asthma-associated Inflammation in Part by Interfering with the T Cell Receptor-nuclear Factor κB-GATA-3-IL-4 Axis Without a Prominent Effect on Poly(ADP-ribose) Polymerase. *J. Biol. Chem.* **2013**, *288*, 1458−1468.

(72) Suwannaroj, S.; Lagoo, A.; McMurray, R. W. Suppression of Renal Disease and Mortality in the Female NZB X NZW F1Mouse Model of Systemic Lupus Erythematosus (SLE) by Chenodeoxycholic Acid. *Lupus* **2001**, *10*, 562−567.

(73) Dziedzic, T.; Wybranska, I.; Dembinska-Kiec, A.; Klimkowicz, A.; Slowik, A.; Pankiewicz, J.; Zdzienicka, A.; Szczudlik, A. Dexamethasone Inhibits TNF-alpha Synthesis More Effectively in Alzheimer's Disease Patients than in Healthy Individuals. *Dementia Geriatr. Cognit. Disord.* **2003**, *16*, 283−286.

(74) Shi, J.; Diao, Z.; Zhou, J.; Zhu, J.; Yuan, H.; You, X.; Liu, Y.; Zheng, D. Epirubicin Potentiates Recombinant Adeno-associated Virus Type 2/5-mediated TRAIL Expression in Fibroblast-like Synoviocytes and Augments the Antiarthritic Effects of RAAV2/5-TRAIL. *Arthritis Rheum.* **2012**, *64*, 1345−1354.

(75) Sreekanth, V. R.; Handa, R.; Wali, J. P.; Aggarwal, P.; Dwivedi, S. N. Doxycycline in the Treatment of Rheumatoid Arthritis-a Pilot Study. *J. Assoc. Physicians India* **2000**, *48*, 804−807.

(76) Borg, A. A.; Davis, M. J.; Fowler, P. D.; Shadforth, M. F.; Dawes, P. T. Rifampicin in Early Rheumatoid Arthritis. *Scand. J. Rheumatol.* **1993**, *22*, 39−42.

(77) Fiore, D. M.; Strober, J. B. Treatment of Hypokalemic Periodic Paralysis with Topiramate. *Muscle Nerve* **2011**, *43*, 127−129.

(78) Gralla, R. J. Edatrexate Studies in Non-small Cell Lung Cancer. *Lung Cancer* **1995**, *12* (Suppl 1), S187−S191.

(79) Du, J.; Guo, Y.; Bao, Y.; Xing, M.; Mahmoud, A. M.; Che, Z.; Chen, Z.; Yang, W. Anticancer Activities of Sulindac in Prostate Cancer Cells Associated with C-Jun NH2-terminal Kinase 1/ß-catenin Signaling. *Oncol. Lett.* **2014**, *8*, 313−316.

(80) Svensson, A.; Azarbayjani, F.; Bäckman, U.; Matsumoto, T.; Christofferson, R. Digoxin Inhibits Neuroblastoma Tumor Growth in Mice. *Anticancer Res.* **2005**, *25*, 207−212.

(81) Huizinga, T. W.; Dijkmans, B. A.; van der Velde, E. A.; van de Pouw Kraan, T. C.; Verweij, C. L.; Breedveld, F. C. An Open Study of Pentoxyfylline and Thalidomide as Adjuvant Therapy in the Treatment of Rheumatoid Arthritis. *Ann. Rheum. Dis.* **1996**, *55*, 833−836.

(82) Hudson, C. J.; Seeman, P.; Seeman, M. V. Parkinson's Disease: Low-dose Haloperidol Increases Dopamine Receptor Sensitivity and Clinical Response. *Parkinson's Dis.* **2014**, *2014*, 684973.

(83) Aylett, S. E.; Atherton, D. J.; Preece, M. A. The Treatment of Difficult Atopic Dermatitis in Childhood with Oral Beclomethasone Dipropionate. *Acta Derm Venereol Suppl (Stockh)* **1992**, *176*, 123−125.

(84) Torrelo, A. Methylprednisolone Aceponate 0.1Lichenified Eczema in a 3-year-old Child with Chronic Atopic Dermatitis. *J. Eur. Acad. Dermatol. Venereol.* **2012**, *26* (Suppl 6), 18−19.

(85) Ogawa, Y.; Suzuki, K.; Sakai, A.; Iida, S.; Ogura, M.; Tobinai, K.; Matsumoto, M.; Matsue, K.; Terui, Y.; Ohashi, K.; Ishii, M.; Mukai, H. Y.; Ando, K.; Hotta, T. Phase I/II Study of Bortezomib-melphalan-prednisolone for Previously Untreated Japanese Patients with Multiple Myeloma. *Cancer Sci.* **2013**, *104*, 912−919.

(86) Dennison, W. B.; Loeser, R. F.; Turner, R. A.; Johnson, J. A.; Wells, H. B. A Double Blind Placebo Controlled Trial of Low Dose Clotrimazole in Rheumatoid Arthritis. *J. Rheumatol.* **1990**, *17*, 1003−1007.

(87) Cuchacovich, M.; Wurgaft, R.; Mena, M. A.; Valenzuela, C.; Gatica, H.; Tchernitchin, A. N. Intraarticular Progesterone Inhibits 3H-dexamethasone Binding to Synovial Cells from Patients with

Rheumatoid Arthritis. a Study by Dry Radioautographic Technique. *J. Rheumatol.* **1991**, *18*, 962−967.

(88) Dohil, M. A.; Alvarez-Connelly, E.; Eichenfield, L. F. Fluocinolone Acetonide 0.01 Atopic Dermatitis in Infants as Young as 3 Months of Age. *Pediatr. Dermatol.* **2009**, *26*, 262−268.

(89) Chang, D. J.; Lamothe, M.; Stevens, R. M.; Sigal, L. H. Dapsone in Rheumatoid Arthritis. *Semin. Arthritis Rheum.* **1996**, *25*, 390−403.

(90) Venkitaraman, R.; Lorente, D.; Murthy, V.; Thomas, K.; Parker, L.; Ahiabor, R.; Dearnaley, D.; Huddart, R.; De Bono, J.; Parker, C. A Randomised Phase 2 Trial of Dexamethasone Versus Prednisolone in Castration-resistant Prostate Cancer. *Eur. Urol.* **2015**, *67*, 673−679.

(91) Dötsch, J.; Hohenberger, I.; Riepe, F. G.; Sippell, W. G.; Dörr, H. G. Serum Cortisol and Cortisone Levels in Newborns with Congenital Adrenal Hyperplasia Before the Start of Therapy. *J. Endocrinol. Invest.* **2005**, *28*, 413−416.

(92) Ferrer, L.; Alberola, J.; Queralt, M.; Brazís, P.; Rabanal, R.; Llenas, J.; Puigdemont, A. Clinical Anti-inflammatory Efficacy of Arofylline, a New Selective Phosphodiesterase-4 Inhibitor, in Dogs with Atopic Dermatitis. *Vet. Rec.* **1999**, *145*, 191−194.

(93) Samrao, A.; Berry, T. M.; Goreshi, R.; Simpson, E. L. A Pilot Study of an Oral Phosphodiesterase Inhibitor (apremilast) for Atopic Dermatitis in Adults. *Arch. Dermatol.* **2012**, *148*, 890−897.

(94) Heo, J. Y.; Cho, Y. S.; Cheon, H. G. Topical Effects of Roflumilast on 1-chloro-2,4-dinitrobenzene-induced Atopic Dermatitis-like Skin Lesions in NC/Nga Mice. *Pharmazie* **2010**, *65*, 906−912.

(95) Berlow, B. A.; Liebhaber, M. I.; Dyer, Z.; Spiegel, T. M. The Effect of Dapsone in Steroid-dependent Asthma. *J. Allergy Clin. Immunol.* **1991**, *87*, 710−715.

(96) Kosmowska-Miskow, A. The Role of Vitamin D3 in Inflammatory Bowel Diseases. *Adv. Clin. Exp. Med.* **2014**, *23*, 497−504.

(97) Dahan, L.; Bonnetain, F.; Rougier, P.; Raoul, J.-L.; Gamelin, E.; Etienne, P.-L.; Cadiot, G.; Mitry, E.; Smith, D.; Cvitkovic, F.; Coudert, B.; Ricard, F.; Bedenne, L.; Seitz, J.-F. Phase III Trial of Chemotherapy Using 5-fluorouracil and Streptozotocin Compared with Interferon Alpha for Advanced Carcinoid Tumors: FNCLCC-FFCD 9710. *Endocr.-Relat. Cancer* **2009**, *16*, 1351−1361.