Bioinformatics, 31(12), 2015, 2007–2016
doi: 10.1093/bioinformatics/btv080
Advance Access Publication Date: 8 February 2015
Original Paper

Systems biology

Large-scale exploration and analysis of drug combinations

Peng Li[1,†], Chao Huang[1,†], Yingxue Fu[1,†], Jinan Wang[1,†], Ziyin Wu[1], Jinlong Ru[1], Chunli Zheng[1], Zihu Guo[1], Xuetong Chen[1], Wei Zhou[1], Wenjuan Zhang[1], Yan Li[2], Jianxin Chen[3], Aiping Lu[4] and Yonghua Wang[1]*

[1]Lab of Systems Pharmacology, Center of Bioinformatics, College of Life Science, Northwest A&F University, Yangling, Shaanxi, China, [2]School of Chemical engineering, Dalian University of Technology, Dalian, Liaoning, China, [3]Beijing University of Chinese Medicine, ChaoYang District, Beijing, China and [4]School of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong

*To whom correspondence should be addressed.
[†]The authors wish it to be known that, in their opinion, the first 4 authors should be regarded as Joint First Authors
Associate Editor: Igor Jurisica

Received on July 7, 2014; revised on January 17, 2015; accepted on February 3, 2015

Abstract

Motivation: Drug combinations are a promising strategy for combating complex diseases by improving the efficacy and reducing corresponding side effects. Currently, a widely studied problem in pharmacology is to predict effective drug combinations, either through empirically screening in clinic or pure experimental trials. However, the large-scale prediction of drug combination by a systems method is rarely considered.

Results: We report a systems pharmacology framework to predict drug combinations (PreDCs) on a computational model, termed probability ensemble approach (PEA), for analysis of both the efficacy and adverse effects of drug combinations. First, a Bayesian network integrating with a similarity algorithm is developed to model the combinations from drug molecular and pharmacological phenotypes, and the predictions are then assessed with both clinical efficacy and adverse effects. It is illustrated that PEA can predict the combination efficacy of drugs spanning different therapeutic classes with high specificity and sensitivity (AUC = 0.90), which was further validated by independent data or new experimental assays. PEA also evaluates the adverse effects (AUC = 0.95) quantitatively and detects the therapeutic indications for drug combinations. Finally, the PreDC database includes 1571 known and 3269 predicted optimal combinations as well as their potential side effects and therapeutic indications.

Availability and implementation: The PreDC database is available at http://sm.nwsuaf.edu.cn/lsp/predc.php.

Contact: yh_wang@nwsuaf.edu.cn

Supplementary Information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Drug combination therapies have been used for the treatment of complex diseases such as cancer and infection for over 30 years due to the advantage of higher efficacy, fewer side effects and less toxicity compared with single-drug treatment (Al-Lazikani *et al.*, 2012; Roemer and Boone, 2013; Zimmermann *et al.*, 2007). The main reason is that complex diseases normally involve physiological processes controlled in a combinatorial/systems fashion featured as

redundancy and multifunctionality, which limits the therapeutic opportunity of one gene-one drug applications (Fitzgerald *et al*., 2006; Shaheen *et al*., 2001). Despite the increasing number of drug combinations in use, many of them were found in clinic by experience, or experimentally derived by dose–response curves for each pair of drugs against a protein target. The mechanistic understanding of synergistic drug pairs remains largely elusive, which makes it difficult to propose new drug combinations.

Systematic surveys of combination drugs *in vitro* have been proposed to investigate the synergistic drug pairs such as the high-throughput screening method (Borisy *et al*., 2003; Lehár *et al*., 2009) and the 'multiplex screening for interacting compounds' (Tan *et al*., 2012). However, the large-scale experiments currently used to evaluate the drug combinations are very time consuming simply because they are severely dependent on the searches of a vast space of possible target combinations (Cokol *et al*., 2011; Winter *et al*., 2012). Alternatively, some computational approaches have been proposed which aimed at using network analysis and chemical biology data to identify novel combinatorial drugs (Chou, 2010; Tang *et al*., 2013; Zhao *et al*., 2011). But most of them are often limited in their capability to dissect the underlying molecular mechanisms, or to extract the information from a larger pharmacological space, or to associate the targets with multiple diseases for combinatorial drugs.

Generally, simultaneous administration of two or more medications may result in significant drug–drug interactions (DDIs), leading to a high risk of adverse effects for patients (Manzi and Shannon, 2005). DDIs may be pharmacokinetic or pharmacodynamic (Jonker *et al*., 2005; Zhang *et al*., 2009). Previous work concerns the prediction of DDIs, mainly relying on the pharmacokinetic properties of the compound such as its solubility (Boobis *et al*., 2002) or depending on pharmacodynamic constants (Huang *et al*., 2013; Li *et al*., 2007) or handling both pharmacokinetic and pharmacodynamic DDIs (Gottlieb *et al*., 2012). However, an up-to-date combined analysis integrating both the efficacy and adverse effects for known or novel drug pairs, which may provide the basis for future clinical trials, is still lacking.

Recently, we have developed a set of systems pharmacology strategies for systematic pursuit of optimal drug combinations. These works lay foundations for a more comprehensive understanding of pharmacological synergy in herbal medicine (Wang *et al*., 2012) and the combination rule of Traditional Chinese Medicine (Yao *et al*., 2013). Furthermore, a large-scale systematic analysis combining pharmacokinetics, chemogenomics, pharmacology and systems biology data was performed through computational methods and experimental validation, which results in a superior output of information for systematic drug design strategies for complex diseases (Li *et al*., 2014).

In this work, we propose a new systems pharmacology framework consisting of a new algorithm termed probability ensemble approach (PEA), through integrating the molecular chemical space, the pharmacological space, the gene annotations, in particular, the connectivity of biological networks, to predict effective drug combinations. In contrast to those previous studies mentioned above, the novelty of PEA is threefold: (i) to the best of our knowledge, PEA presents up to now the largest-scale unbiased prediction of effective drug combinations based on a complete set of drug-combination, drug-disease and drug–side effect relations; (ii) PEA also provides a quantitative assessment of the therapeutic indications and side effects for each combination in clinical usage; (iii) to show the predictive value of our approach, the predictions were benchmarked against independent datasets and further tested using cancer cell and antibacterial models.

2 Methods

2.1 Computational

We designed a novel algorithm termed PEA to integrate the molecular and pharmacological characteristics of drugs (Fig. 1). Given a pair of query drugs, we firstly calculated its six similarity features to a known drug pair including three drug-based and three target-based similarity measures and combine them using a Bayesian network into a likelihood ratio (LR) that represents its probabilistic similarity to the known interaction. We then defined a raw score, as an estimate of the overall similarity of the query drug pair, by summing its LRs to all the known drug pairs in each set [either effective drug combinations (EDCs) or undesirable drug–drug interactions (UDDIs)] above a threshold LR. The raw score can be further converted to a P value (ranging from 0 to 1) from a random raw score distribution. This resulting P value represents the probability of a given raw score that better to be observed from random data. Because we trained our models separately for EDCs and UDDIs, the likelihood that a query drug pair interacts is finally expressed as two probabilities P_1, P_2 which reflect their reliabilities to be an EDC or UDDI, respectively.

2.1.1 Drug–drug similarity measures

We defined and computed six drug–drug similarity measures including the chemical similarity, the similarities based on side effects, the anatomical therapeutic and chemical (ATC) classification system, and the similarities between drug targets, represented by sequence similarity, the distance on a protein–protein interaction (PPI) network and gene ontology (GO) semantic similarity. All similarity measures were normalized to be in the range [0, 1].

(1) *Chemical based*. The 2D chemical structures (Mol file format) of the drugs were downloaded from DrugBank (Knox *et al*., 2011). The hashed binary chemical fingerprints were computed using the Chemical Development Kit with default 2D parameters (Steinbeck *et al*., 2006). The fingerprints were used to compute the similarity score between two drugs using Tanimoto coefficient, that is, the size of the intersection divided by the size of the union of the two fingerprints.

(2) *ATC based*. The ATC classification system established by the WHO is used for the classification of drugs (Skrbo *et al*., 2004). This pharmaceutical coding system categorizes drugs according to the organ or system on which they act and their therapeutic, pharmacological and chemical characteristics. We extracted ATC codes of the drugs from DrugBank. Considering the hierarchical structure of ATC codes, we calculated the similarity scores between drugs using the semantic similarity algorithm (Wang *et al*., 2003):

$$S(c_i, c_j) = w(c_i, c_j)\exp\left(-\gamma \cdot d(c_i, c_j)\right) \quad (1)$$

where d (c_i, c_j) represents the shortest distance between ATC codes c_i and c_j in the hierarchical structure of the ATC classification system. $c(t_i)$ and $c(t_j)$ represent the weights of the corresponding ATC-codes, and are defined as the inverse of ATC-code frequencies, which means that more emphasis was put on specific codes rather than the general ones. γ is a predefined parameter (set to be 0.25 in this study).

(3) *Side-effect based*. Drug side effects were obtained from SIDER (Kuhn *et al*., 2010), a public resource containing drug side-effect information. We assigned a side-effect profile to each drug in our datasets, whose elements encode for the presence or absence of each of the side-effect keywords by 1 or 0,

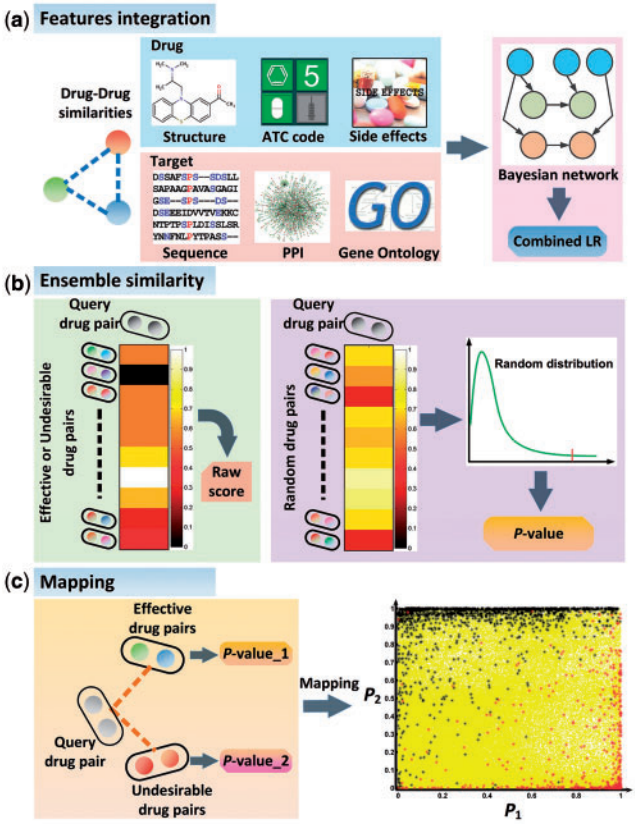

Fig. 1. A schematic overview of PEA model. (**a**) Computing the six drug–drug similarity measures and quantifying LR of query drug pairs according to the similarity to known drug pairs. (**b**) Calculating *P* value by the statistical model inferred from the reference database. (**c**) Assessing potential of query drug pairs using two probabilities P_1 and P_2

respectively. As described above, we defined and computed the similarity scores between drugs according to the Tanimoto coefficient between their side-effect profiles.

(4) *Sequence based*. Sequence-based similarity between two drug target proteins was calculated based on the drug target-centered systems (Sharan *et al.*, 2007). We defined a system similarity score (*S*-score) to describe the target sequence similarity for two drugs (d_i, d_j) as the following:

$$S(d_i, d_j) = \frac{< P_{i,j}, C_{i,j} >}{\sqrt{< C_{i,j}, C_{i,j} >}} \quad (2)$$

where $<\cdot,\cdot>$ is the inner product, $P_{i,j} = (P_{(1,1)}, P_{(1,2)}, \cdots, P_{(1,n)}, P_{(2,1)}, \cdots, P_{(2,n)}, \cdots, P_{(m,1)}, \cdots, P_{(m,n)})$ is a similarity vector, in which m and n are the target number of the drug i and j respectively, and $P_{(t,k)}$ is the sequence similarity between two target and calculated based on the Smith–Waterman sequence alignment score. $C_{i,j}$ is the indicator vector with the same length to $P_{i,j}$,

$$c_{t,k}(\lambda) = \begin{cases} 1 \text{ if } p_{t,k} \geq \lambda \\ 0 \text{ if } p_{t,k} < \lambda \end{cases} \quad (3)$$

(λ is a threshold where $P_{(t,k)}$ can be obtained by a probability of less than 0.05 at random. We constructed the random distribution of $P_{(t,k)}$ by calculating the 10 000 $P_{(t,k)}$ of two random proteins. The λ is 0.034 in this study. The normalized similarity score between two drug target-centered systems is given by dividing the *S*-score by the geometric mean of the scores obtained from the *S*-score of each drug target-centered system against itself.

(5) *PPI network based*. The similarity between each pair of drug target proteins in the human PPI network was calculated using the drug target-centered systems connection (Sharan *et al.*, 2007). We firstly defined a target-centered system for each drug, which includes drug targets and their first-step neighboring proteins in the PPI network. Finally, we defined a system connection score (*S*-score) to describe the connection between target-centered systems for two drugs (d_i, d_j) in the PPI network as the following:

$$S(d_i, d_j) = \frac{< P_{i,j}, C_{i,j} >}{\sqrt{< C_{i,j}, C_{i,j} >}} \quad (4)$$

where $<\cdot,\cdot>$ is the inner product, $P_{i,j} = (P_{(1,1)}, P_{(1,2)}, \cdots, P_{(1,n)}, P_{(2,1)}, \cdots, P_{(2,n)}, \cdots, P_{(m,1)}, \cdots, P_{(m,n)})$ is a similarity vector, in which m and n are the target number of the drug i and j respectively, and $P_{(t,k)}$ is the similarity between two target based on the PPI distance:

$$p_{(t,k)} = e^{\text{dis}(t,k)} \quad (5)$$

$C_{i,j}$ is the indicator vector with the same length to $P_{i,j}$, and

$$c_{t,k}(\lambda) = \begin{cases} 1 \text{ if } p_{t,k} \geq \lambda \\ 0 \text{ if } p_{t,k} < \lambda \end{cases} \quad (6)$$

λ is a threshold where $P_{(t,k)}$ can be obtained by a probability of less than 0.05 at random. We constructed the random distribution of $P_{(t,k)}$ by calculating the 10 000 $P_{(t,k)}$ of two random nodes in the PPI network. The λ is 2 in this study. The normalized similarity score between two drug target-centered systems

is given by dividing the system connection score by the geometric mean of the scores obtained from the *S*-score of each drug target-centered system against itself.

(6) *GO based*: Semantic similarity scores between drug targets were calculated according to Resnik (2011), using the csbl.go R package (Ovaska *et al*., 2008) selecting the option to use all three ontologies.

2.1.2 Features for drug pairs

We defined six features (F_1, F_2, F_3, F_4, F_5 and F_6) for drug pairs based on the above drug-drug similarities to quantitatively describe the similarities between drug pairs. To acquire the similarity of features between a query drug pair (d_1, d_2) and a known drug pair (d_1', d_2'), we first compute the drug–drug similarities S (d_1, d_1') and S (d_2, d_2') (and symmetrically S (d_1, d_2') and S (d_2, d_1')). Then, the two similarities are combined to a single feature similarity score by calculating their geometric mean (Perlman *et al*., 2011). The overall score is

$$F_i(dp, dp') = \max_{d_1,d_2 \neq d_1',d_2'} \sqrt{S(d_1, d_1') \cdot S(d_2, d_2')} \quad (7)$$

2.1.3 Bayesian network for integrating features

To integrate the six features for a drug pair, i.e. F_1, F_2, F_3, F_4, F_5 and F_6, we used a Bayesian networks approach, which was proved to be competent in predicting PPIs by integrating evidence from different sources (Jansen *et al*., 2003). Bayesian networks are a representation of the joint probability distribution among multiple variables. Formally, the feature F can be expressed as a LR, i.e. L (F), which is defined as the fraction of gold-standard positives having feature F divided by the fraction of negatives having F. For the six features F_1, F_2, F_3, F_4, F_5 and F_6, the LR of the combined evidence is the product based on the naive Bayesian network:

$$L(F_1 \ldots F_6) = \prod_{i=1}^{6} L(F_i) \quad (8)$$

This equation is produced as the six features are independent from the analysis of Pearson correlation coefficients for each pair of features (Supplementary Tables S1 and S2). For each feature F_i, its LR relates prior and posterior odds according to the Bayes rule:

$$L(F_i) = \frac{O_{\text{post}}}{O_{\text{prior}}} \quad (9)$$

where the 'prior' and 'posterior' odds are:

$$O_{\text{prior}} = \frac{P(\text{positive})}{P(\text{negative})} = \frac{P(\text{positive})}{1 - P(\text{positive})} \quad (10)$$

$$O_{post} = \frac{P(\text{positive}|F_i)}{P(\text{negative}|F_i)} \quad (11)$$

Among them, the terms 'prior' and 'posterior' refer to the situation before and after knowing the feature F_i. P (*positive*) and P (*negative*) are the odds that a drug pair is in the positive set and in the negative set, respectively. P (positive|F_i) and P (negative|F_i) are the probabilities that a drug pair is in the positive set and in the negative set after knowing that this drug pair has feature F_i. This leads to:

$$L(F_i) = \frac{P(F_i|\text{positive})}{P(F_i|\text{negative})} \quad (12)$$

Here P (F_i|positive) and P (F_i|negative) refer to the probabilities that a drug pair in the positives and the negatives has the feature F_i, respectively, and can be estimated by kernel density estimation (KDE) method.

2.1.4 Kernel density estimation

KDE is a non-parametric way of estimating the probability density function population. The probability $P_i(F_i|C=c)$ was estimated using Equation:

$$P_i(F_i|C=c) = \frac{1}{N_c h} \sum_{j=1}^{N_c} K(F_i, F_{j|i|c}) \quad (13)$$

$$K(a,b) = \frac{1}{\sqrt{2\pi}} e^{\frac{(a-b)^2}{2h^2}} \quad (14)$$

where K is a Gaussian function kernel with mean zero and variance 1, N_c represents the number of the input data F belonging to class c, $F_{j|i|c}$ is the feature value in the ith position of the j-th input $F = (F_1\ F_2 \ldots F_i \ldots F_n)$ in class c, and h is a bandwidth, or a smoothing parameter. To optimally estimate the conditional probabilities, h was optimized on the training dataset.

2.1.5 Raw score and P value

To obtain a good estimate of the overall similarity with the positive set for a drug pair, we first defined a raw score for this drug pair by summing its combined LRs relative to all N drug pairs in positive set with $L_j \geq L_{cut}$.

$$\text{Raw } score = \sum_{j=1}^{N} L_j(F_1 \ldots F_6), (L_j \geq L_{cut}) \quad (15)$$

We determined the threshold L_{cut} by leave-one-out analysis. By sampling across the range of L_{cut} choices, we chose the threshold which led to the highest F1 score in cross-validation. Scores below the *LR* threshold are discarded and do not contribute to the overall set similarity.

Then, a model for the random LRs of the raw scores was developed and fit. A random raw score was calculated by comparing a randomly selected drug pair to a random positive set (with the same size) which was randomly populated from the drugs in the real positive set. We here generated overall $n = 1 \times 10^6$ random raw scores, and the probability of obtaining the same or better raw score by random chance alone can be estimated by (Pearson, 1998):

$$\hat{P}(R \leq r) = 1 - \frac{k}{n} \quad (16)$$

where k is the amount of raw scores that are greater than r occurring in the random distribution.

2.2 Experimental

We used a 10-fold cross validation scheme to evaluate the accuracy of our prediction model. And we compared the performance of the PEA algorithm with a K-Nearest Neighbors algorithm (or KNN for short) which is a non-parametric method used for classification. We benchmarked the prediction results against a number of independent researches about drug combination. To further illustrate the power and potential of our large-scale approach, a set of *in vitro* experimental assays using anti-bacterial and cancer cell models were applied, in which the DDIs characterizing the additivity and synergism are quantitated based on the Chou–Talalay method (Chou, 2006). More details can be found in the Supplementary text.

3 MATERIALS

We collected drug combinations from three localities: (i) 300 combinations in the Drug Combination Database (DCDB) as of 2010

(Liu *et al.*, 2010); (ii) 97 combinations from the Therapeutic Target Database Version 4.3.02(Zhu *et al.*, 2010); and (iii) 1613 combinations from Pubmed literatures (downloaded on September 27, 2012) by data mining through keywords 'drug combination', 'drug interaction', 'multi-drug', 'additive', 'antagonism', 'antagonistic', 'infra-additive', 'potentiated', 'potentiative', 'potentiation', 'reductive', 'supra_additive', 'synergism', 'synergistic' and 'synergy', followed by the actual reading of all full texts (Jia *et al.*, 2009). We then discarded drug combinations that are redundant (the same combination extracted from different sources) and vague (the combinations with none or unclear efficacy). This led to 1571 drug combinations spanning 951 drugs. All these data are available at our website (http://sm.nwsuaf.edu.cn/lsp/predc.php). The DDIs were extracted from DrugBank version 3 (Knox *et al.*, 2011). The full list of the DDIs contains 11 085 DDIs involving 1110 drugs. We manually checked each DDIs and distinguished 1536 UDDIs from all these DDIs. The UDDIs are defined as those in which two drugs can cause an adverse effect, such as the increase of the toxicity or the decrease of the effect.

The 2D chemical structures of the drugs were downloaded in Mol file format from DrugBank (Knox *et al.*, 2011). Drug targets were obtained from the DCDB (Liu *et al.*, 2010) and DrugBank databases. Drug side effects were downloaded from the SIDER database (Kuhn *et al.*, 2010). Human PPI data were assembled from multiple sources (see Supplementary Text for details). Protein sequences and GO annotations (Ashburner *et al.*, 2000) were parsed from Uniprot (Jain *et al.*, 2009). Anatomical Therapeutic Chemical (ATC) codes of drugs were extracted from DrugBank (Knox *et al.*, 2011). Overall, we collected 656 EDCs spanning 375 drugs and 1536 UDDIs involving 313 drugs for which all six drug–drug similarity measures could be computed.

4 RESULTS

4.1 Performance evaluation

To quantitatively assess the performances of the PEA model with all six features or each single feature in predicting effective drug combinations, we used the 656 EDCs of our gold standard and performed a 10-fold cross validation accompanied with the receiver operating characteristic (ROC) curve analysis. As a result, the model with whole features (AUC = 0.90) exhibits better performance than those with single feature (AUC = 0.72–0.85) (Fig. 2a). Among the six features, ATC has the most predicting performance (AUC = 0.85). In order to check if ATC is a dominant contributor to the model, PEA was further trained with the remaining five features without considering the ATC one. Surprisingly, the resultant model shows a very similar performance in predictions (AUC = 0.89) compared to the whole-feature one, indicating that PEA is not biased by the ATC parameter. Moreover, considering the high negative/positive ratio ($238.6 \gg 1$) in samples, the precision recall (PR) curve (Fig. 2b) was also applied to evaluate these models for highlighting the differences of performance that might be lost in the ROC curve analysis (Davis and Goadrich, 2006). It turns out that the whole-feature PEA achieves AUC = 0.15 and a high precision of 91%, whereas the single-feature model never exceeds 50% precision, including ATC.

A combinational use of drugs could cause complex either pharmacodynamic or pharmacokinetic interactions, or both, which makes it difficult to characterize its effectiveness and side effects (Sun *et al.*, 2013) and may, meanwhile, bring up additional health problems (Zhao *et al.*, 2013). Therefore, we further evaluated whether an

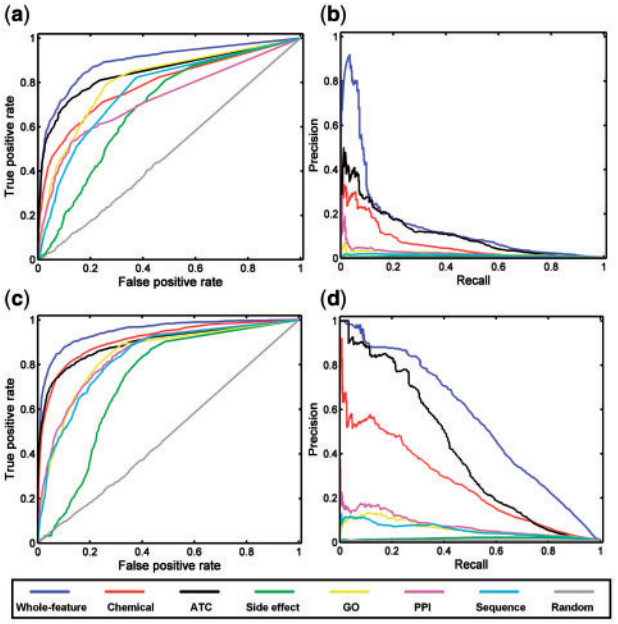

Fig. 2. Performance of PEA model. (**a**) ROC and (**b**) PR (precision-recall) curves for the whole-feature PEA (blue) and the six single features for predicting effective drug combinations (EDCs). (**c**) ROC and (**d**) PR curves for the whole-feature PEA (blue) and the six single features for predicting UDDIs

effective combination has possible adverse effects. Here, we applied the PEA model with 1536 UDDIs to quantitatively characterize a drug pair with respect to adverse effects. Resultantly, similar to the previous case, the whole-feature model exhibits the optimal predictions with AUC = 0.95 for ROC analysis and 0.36 (AUC) for PR analysis respectively, superior to any other models with single feature (Fig. 2c and 2d). The high prediction accuracy for the whole-feature PEA model indicates that the model is reasonable for predicting the adverse effects of drug combinations. In addition and not surprisingly, ATC is again not predictive for detecting the adverse effects of drug combinations based on the PR curves (Fig. 2d). In view of that the sparsely space of DDIs and the non-uniformly covered by experimental data, we have also calculated model accuracy for the drug combinations by dividing drugs into 'new drugs' and 'known drugs'. Drugs in the training set are called 'known' whereas those not in the training set are called 'new'. Overall, the result showed that our model have a better performance for the drug combinations consisted by known drugs (ROC AUC = 0.88 and 0.91 for EDCs and UDDIs, respectively) than those consisted by new drugs (ROC AUC = 0.76 and 0.71 for EDCs and UDDIs, respectively). Still, the model has a good predictive ability for the drug combinations consisted by 'new' drugs (Supplementary Table S3).

Further, in order to verify the newly developed PEA algorithm, the performance of this method to transform the predictive features to objective activity (combination or adverse effect) is compared with the K-nearest neighbor (KNN) method where k is determined by cross-validation and then pick maximum performance k. After 5 folds of cross-validations, k = 6 was determined since it has the best performance. Firstly, we calculated the overall similarity between a query drug pair and all known EDCs (UDDIs) by simply multiplying the six feature similarity scores (rather than using a Bayesian transformation as in PEA), and then assigned the highest similarity score to that query drug pair. The resulting ROC and PR AUC scores of the KNN model trained on the EDC set are 0.79 and 0.04, respectively, which have been considerably outperformed by the new PEA

algorithm. In addition, the model trained on UDDIs also obtains similar inferior results with the ROC and PR AUC scores of 0.85 and 0.13, respectively. All this demonstrates the feasibility of our method in dealing with drug combinations, which is thus further employed for predicting new effective drug pairs with undesirable adverse effects in consideration.

4.2 Integrative analysis of side-beneficial effects for drug combinations

To maximize the therapeutic effects and minimize the adverse effects of drug treatment, the recognition and quantification of the situation, in multiple dimensions, is a critical prerequisite. The integrative analysis of side-beneficial effects would enable us to analyze the relationships between them for drug combinations and also to evaluate the functional consequences of the same to different disease-associated treatment. Here, for each query drug pair, PEA model outputs two values: P_1 and P_2 to represent its effectiveness and perniciousness, respectively. To visualize these two quantitative standards, we first constructed a binary diagram of known EDCs (green dots) and UDDIs (red dots) (Fig. 3a). We observed that PEA has high recall rates with bigger P_1 and lower P_2, confirming P value as an appropriate measure of the odds of a real interaction. In addition, the disease distributions with respect to P_1 and P_2 indicate that current drug pairs (known EDCs) have higher P_1 (0.8–1.0) and lower P_2 (0.0–0.2), and mainly focus on infection, cancer, nervous system disease and cardiovascular disease. Conservatively, we here selected P value as 0.9 throughout our analysis. The EDCs and UDDIs are intensively apportioned in two distinct areas: the (I, II) quadrant ($P_1 > 0.9$) for EDCs and (II, III) quadrant ($P_2 > 0.9$) for UDDIs, respectively. The recall rates are 67 and 82% for the two quadrants, respectively, further proving the reliability and validity of PEA.

Next, we applied PEA to predict novel drug combinations. To visualize the prediction results, we drew a heat map of all the 156 520 unknown drug pairs with the two values: P_1 and P_2 (Fig. 3b). Figure 3b shows that drug pairs in the quadrant 1 are regarded as effective drug combinations. The ideal drug combinations mainly cluster in the quadrant 1' involving 3269 (2.1%) pairs with $P_1 > 0.9$ (a high probability of synergism) and $P_2 < 0.1$ (low probability of causing adverse effects). Here, the critical standard for initialing the P values for the quadrant n' ($n = 1, 3, 4$) is to assure the prediction accuracy with high confidence in each quadrant. In order to validate the quadrant 1' efficiency, we benchmarked the predictions against an independent dataset (see later for details) and also experimentally examined 44 predicted drug pairs from this region. As a result, 35 drug pairs (79.5%) are demonstrated effective (as seen in Supplementary Table S4). The drug combinations in quadrant 2 should be warranted in clinic as they are more likely to cause unwanted adverse effects although they might be effective. For example, voriconazole and itraconazole ($P_1 = 0.96$, $P_2 = 0.97$) are synergistic in treating infections caused by Fusarium (Spader *et al.*, 2013), but the combination of these two drugs may cause a dangerous abnormal heart rhythm (RxList, 2008). Drug pairs in quadrant 3 have a high probability to cause undesirable DDIs. These pairs strongly intend to produce adverse effects ($P_2 > 0.9$), but with low possibility to become an effective combination ($P_1 < 0.1$). Quadrant 4' components have both small P_1 and P_2 values, which suggests that there is a small chance of interactions between the two drugs in one drug pair. In addition, looking specifically at the 3269 optimal drug combinations, their associations with specific diseases is in good agreement with the interest of current combination therapies for complex and chronic diseases (Fig. 3c).

4.3 External literature validation

To validate the reliability of our method, we further check whether the predicted drug pairs were validated in external literatures which were not used to build the training dataset for the PEA model. The experiment types, effectiveness, diseases and adverse effects of the predicted drug combinations were manually collected from an unbiased survey of the literature (See Supplementary Table S5) Overall, we obtained detailed information about 642 novel drug

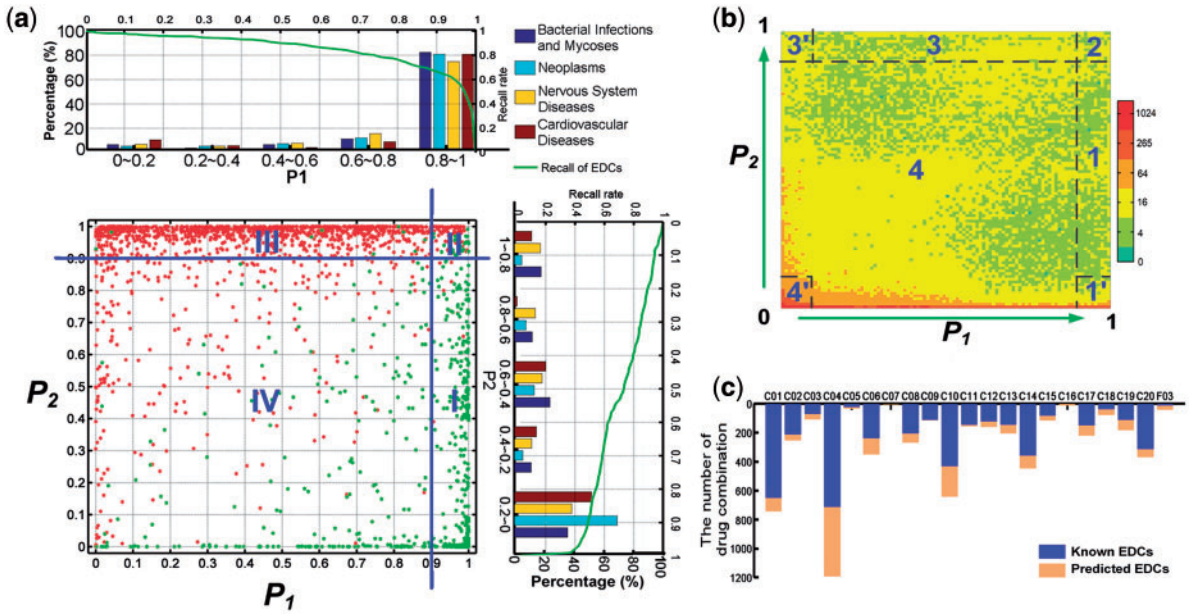

Fig. 3. Integrative analysis of side–beneficial effects for drug combinations. (**a**) The binary diagram of known EDCs (green dots) and UDDIs (red dots), the EDC recall rate ($> P_1$ or $< P_2$) and the disease distribution with respect to P_1 and P_2. (**b**) The heat map of all the 156 520 unknown drug pairs with respect to the two values: P_1 and P_2. The colors of the heat map represent the number of drug pairs with the specific P_1, P_2 score. (**c**) The number of known and predicted drug combinations in 21 MeSH (Medical Subject Headings) disease categories

combinations. Surprisingly, 84% of which are consistent with our predictions. Among the 537 effective predictions directly supported by the literature, 69% have already entered clinical trials and 45 drug combinations were proved to be effective in at least two experiment types (Fig. 4a). Moreover, 85% of the drug combinations validated by cell assays have a synergistic effect (Fig. 4b).

To put our finding into context, we divided the drugs into different drug classes based on their third level of ATC code and computed the proportion of verified drug combinations in predicted ones within and between drug classes. We then calculated the significance (by *P*-value using Fisher's exact test) of overrepresentation against the background incidence of the respective drug class. The *P*-values were then adjusted to control for multiple hypothesis testing, yielding *q*-values. More formally, *q*-values represent the minimum false discovery rate for which the connection will be regarded as significant. To find out which drug class connections have a high verified rate, we utilized the significant (q-value < 0.05) drug class connections to create a drug class combination network. We filtered connections that are involved with at least 10 predicted drug combinations and with a verified proportion greater than 0.3.

In total, we found 65 statistically significant drug-class connections between 41 drug classes, the majority of which are inter-class connections, indicating that most of the predicted drug combinations consist of drugs belonging to different drug classes (Fig. 4c). Analysis of the known EDCs in drug classes also shows similar results (Supplementary Fig. S1). The drug class J05A (direct acting antiviral drugs), which has 45 predicted drug combinations in it, has the highest proportion of the verified drug combinations (64%, q-value $= 7\times10^{-9}$). Following on are combinations between D07X (corticosteroids, other combinations) and R03A (adrenergics, inhalants) (60%, q-value $= 0.024$). Interestingly, there is actually no effective drug combinations (EDCs) consist of drugs belong to D07X and R03A in our gold-standard dataset (Supplementary Table S6). These results further indicated that the PEA model was not biased by the ATC code similarity between drugs, but provided a weighted similarity measurement which took all drug molecular and pharmacological features into account. There are 20 drug-class connections have a verified proportion greater than or equal to 50%, which indicates that a half or more than a half of our predictions consist of drugs between these drug classes have gained direct supports of certain literature. The full list of the verified proportion and *q*-values between or within drug classes is provided in Supplementary Table S6. The connection between drug classes L01X (other antineoplastic agents) and L01B (antimetabolites), which is the most connected one between drug classes (36%, q-value $= 2.3\times10^{-5}$), has 102 drug combinations in it. This result indicates that drugs in these two drug classes have a larger chance to combine with each other to create synergy. The L01X class mainly includes platinum compounds and protein kinase inhibitors, which have been used widely in cancer combination therapy (Dancey and Sausville, 2003; Lee *et al.*, 2008).

4.4 Experimental validation

In total, we examined 102 novel predicted drug pairs, resulting in the confirmation of 77 effective combinations (~75% of all tested drug pairs) (Supplementary Table S11). In the cancer model, we examined 55 predicted drug pairs against the human non-small cell lung cancer A549 cells. Resultantly, 39 of these pairs are found effective (~71% accuracy), among which 34 cases are synergistic and 5 are additive (Supplementary Table S7). For anti-bacterial model, 47 drug pairs were tested by *Staphylococcus aureus* and *Escherichia coli*. We validate that 38 pairs (~81% of all 47 antibacterial pairs) are effective (Supplementary Tables S8 and S9). Among them, 9 drug pairs exhibit activities (synergy or additivity) against both bacterial species. More details can be found in the Supplementary text.

5 DISCUSSION

A major challenge in predicting the drug combination in a large scale is to achieve the desired prediction accuracy while dealing with

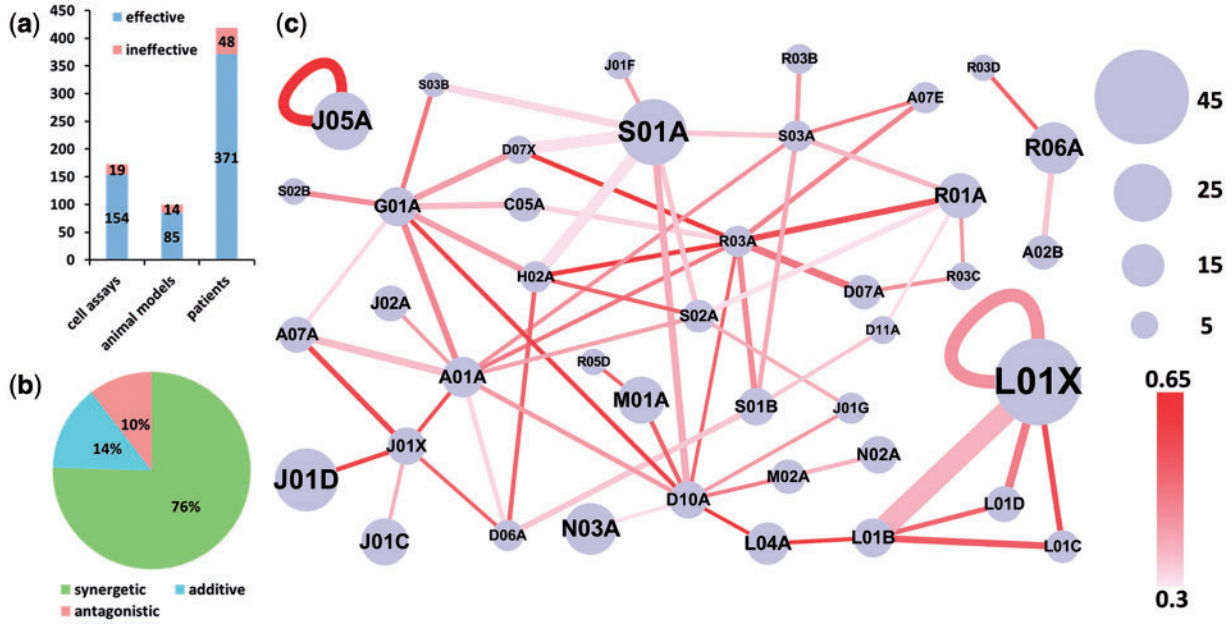

Fig. 4. Benchmarking against independent datasets. (**a**) The overview of the independent datasets of drug combinations derived from the literature. (**b**) The distribution of synergistic, additive and antagonistic effects in cell assays. (**c**) The drug class networks. Edge color depth corresponds to the proportion of verified drug combinations in predicted ones. Edge width corresponds to the number of drug combinations between two drug classes. Node size corresponds to the number of drugs in the drug class. The first letter of each ATC category denotes the top level, anatomical class

the complex conditions of missing data, different data types, unclear mechanisms, etc. One means to achieve this is to exploit a large database with sufficient information for drug combinations and a method that distinguishes the therapeutic efficacy from adverse effects that may make it feasible to more precisely select the optimal combinations. The incompleteness of target network information and the scarceness of the side effect features hinder the application of computational approaches to drug combinations (Brouwers *et al*., 2011). Compared with the present available models (Zhao *et al*., 2011), the PEA algorithm shows several major advantages, such as high training efficiency and extensive applicability. More particularly, PEA provides two quantitative indexes to describe the property of a drug combination, which is convenient and easy to understand.

PEA is specifically designed to accommodate those missing data and differing types of variables and handle unbalanced multiclass datasets. By integrating the weakly predictive features, such as Target sequence, Chemical structure, etc., PEA exhibits similar performances as the whole-feature model. The major reason is due to that Bayes algorithm extracts the underlying pattern of features by converting features to a common probabilistic framework so as to improve the performance of the ensemble-based scoring approach. In detail, Bayesian network algorithm transforms the features to a LR, which provides probabilistic scoring for drug combinations in an ensemble. The ensemble score is further converted to a *P* value based on a random raw score distribution to calculate the reliability for a predicted drug pair. In contrast to conventional method, where an optimal combination is often determined by comparing the similarity with a single known reference (Gottlieb *et al*., 2012), the ensemble approach compares a probability distribution of a query drug pair with all related combinations, resulting in the improvement of generalization ability of the PEA tool. However, the simple feature-enrich method (Zhao *et al*., 2011), as well as the one-nearest neighbor, may not be capable of dealing with this complex system due to the unadaptability and information loss. In addition, it should be noted that we just calculated the protein global similarity instead of the drug binding regions (Konc and Janežič, 2014). Our considerations are: (i) sequence similar proteins might have similar functions, since they might be homologs. However, similar drug binding regions may not reveal whether the two proteins are functionally similar; (ii) the analyses of drug binding regions are normally dependent of the crystal structures, which severely limits the algorithms for more general applications. And more importantly, even we have obtained the structures; we still need to define which amino acids are involved in drug binding, since different drugs might bind to different amino acids in most cases even in the same pocket for the same protein.

The present computational solution considers drug actions and their clinical effects in the context of molecular network systems, which provides information for further understanding of the molecular mechanism and pharmacological effect underlying drug combinations. Generally, due to the unclearness of the molecular mechanisms underlying combination therapies, most drug combinations are inferred based on sets of clinical rules derived from clinical experience or randomized clinical trials. For example, combination therapy for hypertension has been assigned some preferred combinations from various classes of antihypertensive medications, such as renin-angiotensin-aldosterone system (RAAS) inhibitor + diuretic, RAAS inhibitor + calcium channel blocker (CCB), and CCB + diuretics (Salahuddin *et al*., 2013). Although these conventional rules facilitate searching for valid combinations, they limit this search into drug entities with similar functions, that is, drugs in combinations often have the same first level of the ATC code. This is confirmed by the fact that majority (~70%) of known 1571 drug combinations in our dataset belong to the same therapeutic category (from the first to fifth level of the ATC code) (Supplementary Fig. S1).

To analyze whether PEA could overcome this limitation, we have counted the predicted EDCs in the database that are composed by drugs with different ATC classes (the first level). Extremely interesting, we found 43% of our high-confidence predictions (with $P_1 \geq 0.9$ and $P_2 \leq 0.1$) are such kind of combinations, indicating that our model is not restrained by the functional similarity between two drugs for a combination. The benchmark dataset also shows that 121 such combinations have been proved to be effective. Moreover, we have experimentally validated 10 novel EDCs that are combined by antibacterial and anticancer drugs (Supplementary Tables S6–S8). The results show that 80% pairs are synergistic to cancer models. For example, the antibacterial agent, tetracycline, as a bacterial 30 S ribosomal subunit inhibitor, displays synergistic cytotoxicity against A549 cells when being combined with some anticancer drugs, such as the DNA synthesis inhibitor fludarabine (CI 0.49), cross-linking reagent cisplatin (CI 0.69) and protein kinase inhibitor imatinib (CI 0.80). It is worth noting that the 'wet' experiment could only found ~16% (38/200) synergistic pairs in antifungal test (Cokol *et al*., 2011), which partially proves that the PEA algorithm carries out high efficiency prediction in practice.

Various approaches have been suggested to benefit-risk assessment during the development of new medicines (Eichler *et al*., 2008; Garrison *et al*., 2007), but little has been reported of being applied on the combinational drugs, though they have attracted more and more interests from researchers and industry in recent years. Drug combinations may overcome the side effects by countering network robustness and bypass compensation, and thereby increasing the clinical efficacy while minimizing the overlapping toxicity and allowing reduced dosage of each compound (Jia *et al*., 2009; Ramaswamy, 2007). The efficacy and side effect are normally coupled together, and the best end results actually depend on the side effects contributing to the overall therapeutic benefit. Thus, the decision to take a drug combination depends upon the benefit-risk weighting of all the potential risks and benefits, which up to date are still extremely difficult to predict on a global scale. The complex DDIs in multiple dimensions pose modeling challenges to classical linear approaches, as well as hinder the targeted experimental approaches due to a combinatorial explosion both in the pharmacological and molecular spaces.

In an effective combinatorial setting, the PEA algorithm has incorporated the clinical efficacy and adverse effect evaluation into the current model, resulting in two standards, i.e. P_1 and P_2. The combined evaluation for a predicted combination opens new avenues of drug combination and even guides the drug dosage. For example, the fludarabine + tetracycline combination with high P_1 (0.87) and low P_2 (0.36) shows a significant synergistic effect (CI 0.46) on A549 cells. By lowering the dosage of fludarabine (from 0.08611 μM to 0.03538 μM at the 50% cytotoxicity level), tetracycline also reduces the risk of opportunistic infections induced by fludarabine. More importantly, those combinations with low P_1 and high P_2 predicted by PEA should be very carefully applied in practice due to their uncertain influence on the efficacy and safety of drug co-administration. For example, the monoamine oxidase (MAO) inhibitor, tranylcypromine, may increase the vasopressor effect of the alpha1-agonist, midodrine, whose P_1 and P_2 are 0.055 and 0.993, respectively. Thus, concomitant use of these two drugs should be avoided.

One limitation for this work is that the dosage is not taken into account in the model. Till now, two methods are in common use for calculating the expected dose-response relationship for drug combination as compared to mono-therapy: Loewe additivity and Bliss independence (Fitzgerald *et al*., 2006). They both need experiments to characterize drug combinations. Therefore, following *in silico* studies should focus not only strategies for predicting drug combinations but also improving the efficiency of trials.

Funding

This work is supported by grants from Northwest A&F University, National Natural Science Foundation of China (11201049 and 31170796), Program for New Century Excellent Talents in University of Ministry of Education of China, Interdisciplinary Research Matching Scheme (IRMS) of Hong Kong Baptist University (RC-IRMS/12-13/02) and Hong Kong Baptist University Strategic Development Fund (SDF13-1209-P01).

Conflict of Interest: none declared.

REFERENCES

Al-Lazikani,B. *et al.* (2012) Combinatorial drug therapy for cancer in the post-genomic era. *Nat. Biotechnol.*, **30**, 679–692.

Ashburner,M. *et al.* (2000) Gene Ontology: tool for the unification of biology. *Nat. Genet.*, **25**, 25–29.

Bleakley,K. and Yamanishi,Y. (2009) Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*, **25**, 2397–2403.

Boobis,A. *et al.* (2002) In silico prediction of ADME and pharmacokinetics: report of an expert meeting organised by COST B15. *Eur. J. Pharm. Sci.*, **17**, 183–193.

Borisy,A.A. *et al.* (2003) Systematic discovery of multicomponent therapeutics. *Proc. Natl Acad. Sci.*, **100**, 7977–7982.

Brouwers,L. *et al.* (2011) Network neighbors of drug targets contribute to drug side-effect similarity. *PLoS One*, **6**, e22187.

Chou,T.-C. (2006) Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.*, **58**, 621–681.

Chou,T.-C. (2010) Drug combination studies and their synergy quantification using the Chou-Talalay method. *Cancer Res.*, **70**, 440–446.

Cokol,M. *et al.* (2011) Systematic exploration of synergistic drug pairs. Mol. Syst. Biol., **7**, 544.

Dancey,J. and Sausville,E.A. (2003) Issues and progress with protein kinase inhibitors for cancer treatment. *Nat. Rev. Drug Discov.*, **2**, 296–313.

Davis,J. and Goadrich,M. (2006) The relationship between Precision-Recall and ROC curves. In: *Proceedings of the 23rd international conference on Machine learning*. ACM, pp. 233–240.

Eichler,H.G. *et al.* (2008) Balancing early market access to new drugs with the need for benefit/risk data: a mounting dilemma. *Nat. Rev. Drug Discov.*, **7**, 818–826.

Fitzgerald,J.B. *et al.* (2006) Systems biology and combination therapy in the quest for clinical efficacy. *Nat. Chem. Biol.*, **2**, 458–466.

Garrison,L.P.,Jr. *et al.* (2007) Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. *Health Aff. (Millwood).* **26**, 684–695.

Gottlieb,A. *et al.* (2012) INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol. Syst. Biol.*, **8**, 592.

Huang,J. *et al.* (2013) Systematic prediction of pharmacodynamic drug–drug interactions through protein–protein-interaction network. *PLoS Comp. Biol.*, **9**, e1002998.

Jain,E. *et al.* (2009) Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinformatics*, **10**, 136.

Jansen,R. *et al.* (2003) A Bayesian networks approach for predicting protein–-protein interactions from genomic data. *Science*, **302**, 449–453.

Jia,J. *et al.* (2009) Mechanisms of drug combinations: interaction and network perspectives. *Nat. Rev. Drug Discov.*, **8**, 111–128.

Jonker,D.M. *et al.* (2005) Towards a mechanism-based analysis of pharmacodynamic drug–drug interactions in vivo. *Pharmacol. Ther.*, **106**, 1–18.

Knox,C. *et al.* (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.*, **39**, D1035–D1041.

Konc,J. and Janežič,D. (2014). Binding site comparison for function prediction and pharmaceutical discovery. *Curr. Opin. Struct. Biol.*, **25**, 34–39.

Kuhn,M. *et al.* (2010) A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.*, **6**, 343.

Lee,J.L. *et al.* (2008) A phase II study of docetaxel as salvage chemotherapy in advanced gastric cancer after failure of fluoropyrimidine and platinum combination chemotherapy. *Cancer Chemother. Pharmacol.*, **61**, 631–637.

Lehár,J. *et al.* (2009) Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.*, **27**, 659–666.

Li,L. *et al.* (2007) Drug–drug interaction prediction: a Bayesian meta-analysis approach. *Stat. Med.*, **26**, 3700–3721.

Li,P. *et al.* (2014) Systems pharmacology strategies for drug discovery and combination with applications to cardiovascular diseases. *J. Ethnopharmacol.*, **151**, 93–107.

Lin,D. (1998) An information-theoretic definition of similarity. *ICML*. Morgan Kaufman Publishers, San Francisco, pp. 296–304.

Liu,Y. *et al.* (2010) DCDB: drug combination database. *Bioinformatics*, **26**, 587–588.

Manzi,S.F. and Shannon,M. (2005) Drug interactions—a review. *Clin. Pediatric Emer. Med.*, **6**, 93–102.

Ovaska,K. *et al.* (2008) Fast Gene Ontology based clustering for microarray experiments. *BioData Mining*, **1**, 11.

Pauwels,E. *et al.* (2011) Predicting drug side-effect profiles: a chemical fragment-based approach. *BMC Bioinformatics*, **12**, 169.

Pearson,W.R. (1998) Empirical statistical estimates for sequence similarity searches. *J. Mol. Biol.*, **276**, 71–84.

Perlman,L. *et al.* (2011) Combining drug and gene similarity measures for drug-target elucidation. *J. Comput. Biol.*, **18**, 133–145.

Ramaswamy,S. (2007) Rational design of cancer–drug combinations. *N. Engl. J. Med.*, **357**, 299–300.

Resnik,P. (2011) Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language. *arXiv preprint arXiv:1105.5444*.

Roemer,T. and Boone,C. (2013) Systems-level antimicrobial drug and drug synergy discovery. *Nat. Chem. Biol.*, **9**, 222–231.

RxList,R. (2008) The Internet Drug Index, *Aldactone®,(spironolactone) Tablets, USP*, 1–8.

Salahuddin,A. *et al.* (2013) Combination therapy for hypertension 2013: an update. *J. Am. Soc. Hypertension JASH*, **7**, 401–407.

Shaheen,R.M. *et al.* (2001) Tyrosine kinase inhibition of multiple angiogenic growth factor receptors improves survival in mice bearing colon cancer liver metastases by inhibition of endothelial cell survival mechanisms. *Cancer Res.*, **61**, 1464–1468.

Sharan,R. *et al.* (2007) Network-based prediction of protein function. *Mol. Syst. Biol.*, 3.

Skrbo,A. *et al.* (2004) Classification of drugs using the ATC system (Anatomic, Therapeutic, Chemical Classification) and the latest changes]. *Med. Arh.*, **58**, 138.

Smith,T.F. *et al.* (1985) The statistical distribution of nucleic acid similarities. *Nucleic Acids Res.*, **13**, 645–656.

Spader,T.B. *et al.* (2013) Synergysm of voriconazole or itraconazole with other antifungal agents against species of Fusarium. *Rev. Iberoam Micol.*, **30**, 200–204.

Steinbeck,C. *et al.* (2006) Recent developments of the chemistry development kit (CDK)-an open-source java library for chemo-and bioinformatics. *Curr. Pharm. Des.*, **12**, 2111–2120.

Sun,X. *et al.* (2013) High-throughput methods for combinatorial drug discovery. *Sci. Transl. Med.*, **5**, 205rv201–205rv201.

Tan,X. *et al.* (2012) Systematic identification of synergistic drug pairs targeting HIV. *Nat. Biotechnol.*, **30**, 1125–1130.

Tang,J. *et al.* (2013) Target inhibition networks: predicting selective combinations of druggable targets to block cancer survival pathways. *PLoS Comp. Biol.*, **9**, e1003226.

Wang,X. *et al.* (2012) A systems biology approach to uncovering pharmacological synergy in herbal medicines with applications to cardiovascular disease. *Evid. Based Complement Alternat. Med.*, **2012**, 519031.

Wang,Y.C. *et al.* (2013) Network predicting drug's anatomical therapeutic chemical code. *Bioinformatics*, **29**, 1317–1324.

Winter,G.E. *et al.* (2012) Systems-pharmacology dissection of a drug synergy in imatinib-resistant CML. *Nat. Chem. Biol.*, **8**, 905–912.

Yao,Y. *et al.* (2013) Deciphering the combination principles of Traditional Chinese Medicine from a systems pharmacology perspective based on Mahuang Decoction. *J. Ethnopharmacol.*, **150**, 619–638.

Zhang,L. *et al.* (2009) Predicting drug–drug interactions: an FDA perspective. *AAPS J.*, **11**, 300–306.

Zhao,S. *et al.* (2013) Systems pharmacology of adverse event mitigation by drug combinations. *Sci. Transl. Med.*, **5**, 206ra140.

Zhao,X.-M. *et al.* (2011) Prediction of drug combinations by integrating molecular and pharmacological data. *PLoS Comp. Biol.*, **7**, e1002323.

Zhu,F. *et al.* (2010) Update of TTD: therapeutic target database. *Nucleic Acids Res.*, **38**, D787–D791.

Zimmermann,G.R. *et al.* (2007) Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov. Today*, **12**, 34–42.