

RESEARCH ARTICLE

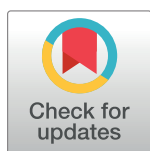
# Leveraging genetic interactions for adverse drug-drug interaction prediction

Sheng Qian <sup>1,2‡</sup>, Siqi Liang <sup>1,2‡</sup>, Haiyuan Yu <sup>1,2\*</sup>

**1** Department of Computational Biology, Cornell University, Ithaca, New York, United States of America, **2** Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, New York, United States of America

‡ These authors share first authorship on this work.

\* [haiyuan.yu@cornell.edu](mailto:haiyuan.yu@cornell.edu)



## Abstract

In light of increased co-prescription of multiple drugs, the ability to discern and predict drug-drug interactions (DDI) has become crucial to guarantee the safety of patients undergoing treatment with multiple drugs. However, information on DDI profiles is incomplete and the experimental determination of DDIs is labor-intensive and time-consuming. Although previous studies have explored various feature spaces for *in silico* screening of interacting drug pairs, their use of conventional cross-validation prevents them from achieving generalizable performance on drug pairs where neither drug is seen during training. Here we demonstrate for the first time targets of adversely interacting drug pairs are significantly more likely to have synergistic genetic interactions than non-interacting drug pairs. Leveraging genetic interaction features and a novel training scheme, we construct a gradient boosting-based classifier that achieves robust DDI prediction even for drugs whose interaction profiles are completely unseen during training. We demonstrate that in addition to classification power—including the prediction of 432 novel DDIs—our genetic interaction approach offers interpretability by providing plausible mechanistic insights into the mode of action of DDIs.

## OPEN ACCESS

**Citation:** Qian S, Liang S, Yu H (2019) Leveraging genetic interactions for adverse drug-drug interaction prediction. *PLoS Comput Biol* 15(5): e1007068. <https://doi.org/10.1371/journal.pcbi.1007068>

**Editor:** Avner Schlessinger, Icahn School of Medicine at Mount Sinai, UNITED STATES

**Received:** January 4, 2019

**Accepted:** May 3, 2019

**Published:** May 24, 2019

**Copyright:** © 2019 Qian et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This work was supported by National Institutes of Health (<https://www.nih.gov/>) grants (R01 GM124559, R01 GM125639, R01 DK115398) and National Science Foundation (<https://www.nsf.gov/>) grant (DBI-1661380) to HY. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author summary

Adverse drug-drug interactions are adverse side effects caused by taking two or more drugs together. As co-prescription of multiple drugs becomes an increasingly prevalent practice, affecting 42.2% of Americans over 65 years old, adverse drug-drug interactions have become a serious safety concern, accounting for over 74,000 emergency room visits and 195,000 hospitalizations each year in the United States alone. Since experimental determination of adverse drug-drug interactions is labor-intensive and time-consuming, various machine learning-based computational approaches have been developed for predicting drug-drug interactions. Considering the fact that drugs effect through binding and modulating the function of their targets, we have explored whether drug-drug interactions can be predicted from the genetic interaction between the gene targets of two drugs, which characterizes the unexpected fitness effect when two genes are simultaneously knocked out. Furthermore, we have built a fast and robust classifier that achieves accurate

**Competing interests:** The authors have declared that no competing interests exist.

prediction of adverse drug-drug interactions by incorporating genetic interaction and several other types of widely used features. Our analyses suggest that genetic interaction is an important feature for our prediction model, and that it provides mechanistic insight into the mode of action of drugs leading to drug-drug interactions.

## Introduction

Drug-drug interactions (DDIs) refer to the unexpected pharmacologic or clinical responses due to the co-administration of two or more drugs [1]. With the simultaneous use of multiple drugs becoming increasingly prevalent, DDIs have emerged as a severe patient safety concern over recent years [2]. According to The Center for Disease Control and Prevention (CDC), the percentage of Americans taking three or more prescription drugs in the past 30 days increased from 11.8% in 1988–1994 to 21.5% in 2011–2014, and the occurrence of polypharmacy, defined as the concurrent use of five or more drugs, increased from 4.0% to 10.9% within the same time period [3,4]. Polypharmacy is especially common among elderly people, affecting 42.2% of Americans aged 65 years and older, exposing them to a higher risk of adverse DDIs. Indeed, DDIs were estimated to be responsible for 4.8% of hospitalization in the elderly, a 8.4-fold increase compared to the general population [5]. Overall, DDIs contribute to up to 30% of all adverse drug events (ADEs) [6] and account for about 74,000 emergency room visits and 195,000 hospitalizations each year in the United States alone [3]. Therefore, it has become a medical imperative to identify and predict interacting drug pairs that lead to adverse effects.

In order to facilitate identification of interacting drug pairs, a number of *in vitro* and *in vivo* methods have been developed. For example, drug pharmacokinetic parameters and drug metabolism information collected from *in vitro* pharmacology experiments and *in vivo* clinical trials can be used to predict interacting drug pairs [7,8]. However, these methods are labor-intensive and time-consuming, and are thus not scalable to all unannotated drug pairs [9]. In the past decade, machine learning-based *in silico* approaches have become a new direction for predicting DDIs by leveraging the large amount of biological and phenotypic data of drugs available. The advantage of machine learning-based approaches lies in their ability to perform large-scale DDI prediction in a short time frame. So far, various features have been explored for building DDI prediction models, including similarity-based features and network-based features, among others. Similarity-based features characterize the similarity of the two drugs at question in terms of chemical structure, side effect profile, indication, target sequence, target docking, ATC group, etc. [10–24]. Network-based features exploit the topological properties of the drug-drug interaction network or the protein-protein interaction network, which relates to DDIs through drug-target associations [16,25–27]. While these methods have yielded important information about DDIs, few methods to date have been able to provide insight into the molecular mechanisms of drug-drug interactions.

To this end, in this study, we employ the genetic interaction between genes that encode the targets of two drugs as a novel feature for predicting interacting drug pairs that cause adverse drug reactions. We show that targets of adversely interacting drugs tend to have more synergistic genetic interactions than targets of non-interacting drugs. Exploiting this finding, we apply a machine learning framework (S1 Fig) and build a gradient boosting-based classifier for adverse DDI prediction by integrating genetic interaction and three widely used features—indication similarity, side effect similarity and target similarity. We show that our model provides accurate DDI prediction even for pairs of drugs whose interaction profiles are completely unseen during training. Furthermore, we find that excluding the genetic interaction features

significantly decreases the performance of our model. Through genetic interactions, our method provides insight into the mode of action of drugs that lead to adverse combinatory effects.

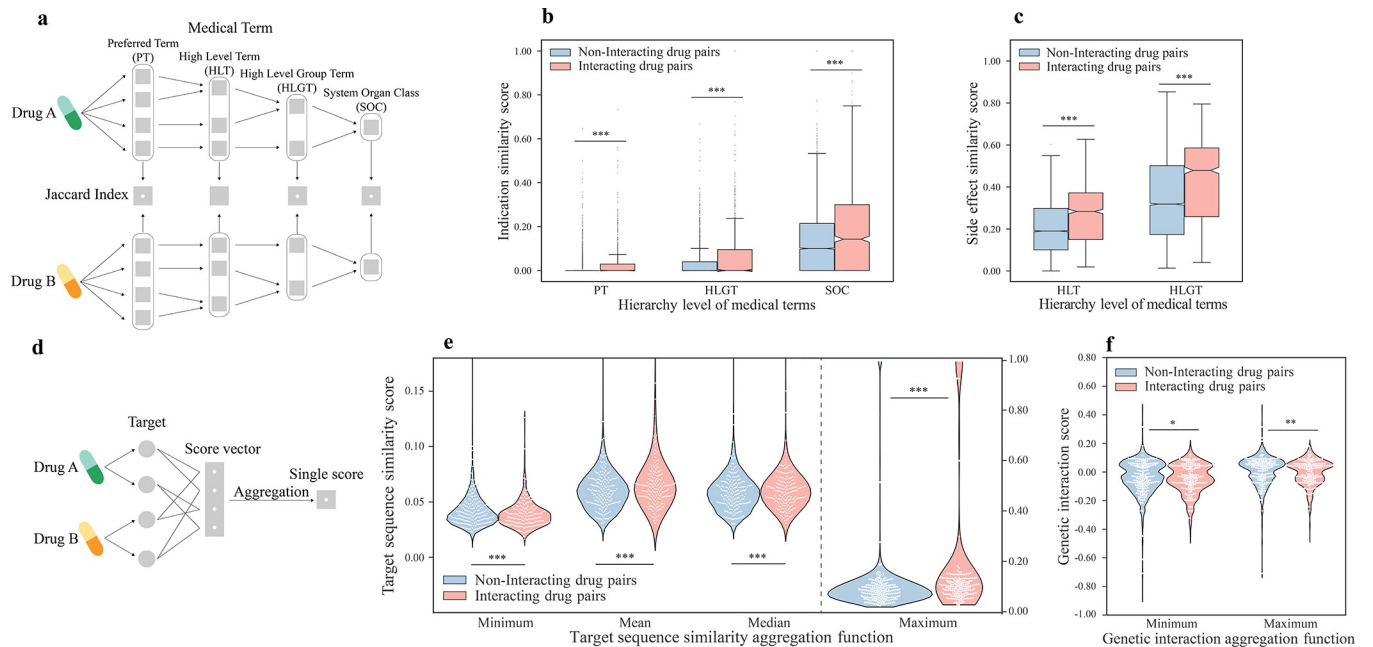
## Results

### Genetic interaction profiles provide complementary information for distinguishing interacting and non-interacting drugs

In order to explore the separating power of various features to distinguish adversely interacting drug pairs from non-interacting drug pairs, we constructed a high-confidence set of adversely interacting drug pairs from all DDIs labeled “the risk or severity of adverse effects can be increased” in DrugBank [28] (S1 Table). This resulted in a set of 117,045 adversely interacting drug pairs involving 2,261 drugs. 2,195,023 non-interacting drug pairs were generated by taking all other combinations of these drugs before removing any drug pair that has been reported in DrugBank, TWOSIDES [29] or a complete dataset of DDIs compiled from a variety of sources [30]. Furthermore, we required that all features, including indication similarity, side effect similarity, target sequence similarity and genetic interaction, should be available for each drug pair. After this filtering step, 1,113 adversely interacting drug pairs and 11,313 non-interacting drug pairs involving 262 drugs remained.

Interacting and non-interacting drug pairs exhibit different distributions in terms of the four groups of properties that we investigated. Indications and side effects of drugs were mapped to four levels of the MedDRA hierarchy [31] (Fig 1A). At every level, adversely interacting drugs are associated with significantly more similar side effects as well as indications than non-interacting drugs (Fig 1B and 1C, S2A and S2B Fig, S2 Table). On another front, target similarity was calculated by aligning the sequences of the protein targets with the Smith-Waterman algorithm [32]. Since a drug may have multiple protein targets, aggregation was performed by taking the minimum, mean, median or maximum alignment score for each drug pair (Fig 1D). As expected, the maximum, mean and median target similarity between targets of adversely interacting drug pairs are significantly higher than those of non-interacting drug pairs (Fig 1E, S2 Table). Interestingly, interacting drug pairs manifest a significantly lower minimum target similarity than non-interacting drug pairs (Fig 1E, S2 Table). This could be due to the fact that interacting drugs possess a higher number of protein targets combined, thereby having a higher chance of targeting vastly different targets (S3A Fig). These results establish indication similarity, side effect similarity and target similarity as informative predictors of adverse DDIs.

Genetic interaction refers to deviation from the expected phenotype when two genes are simultaneously mutated [33]. In short, the genetic interaction score quantifies the extent to which the fitness of a double mutant carrying mutations on two genes deviates from what is expected from the fitness defects of the corresponding single mutants. A negative score indicates synergistic genetic interaction, where the double mutant exhibits a fitness defect that is more extreme than expected from single mutants, while a positive score suggests buffering genetic interaction, where the double mutant exhibits a greater fitness than expected [34]. Since binding of drugs modulates the function of their targets, the genetic interaction between protein targets of two drugs might be associated with their joint effects. On this account, we investigated whether targets of adversely interacting drugs and targets of non-interacting drugs display divergent genetic interaction profiles. For each pair of drugs, we mapped their protein targets to the corresponding yeast homologs and obtained genetic interaction scores between the yeast genes from a global yeast genetic interaction network [35]. When the minimum, mean, median or maximum genetic interaction score was taken for targets of each drug



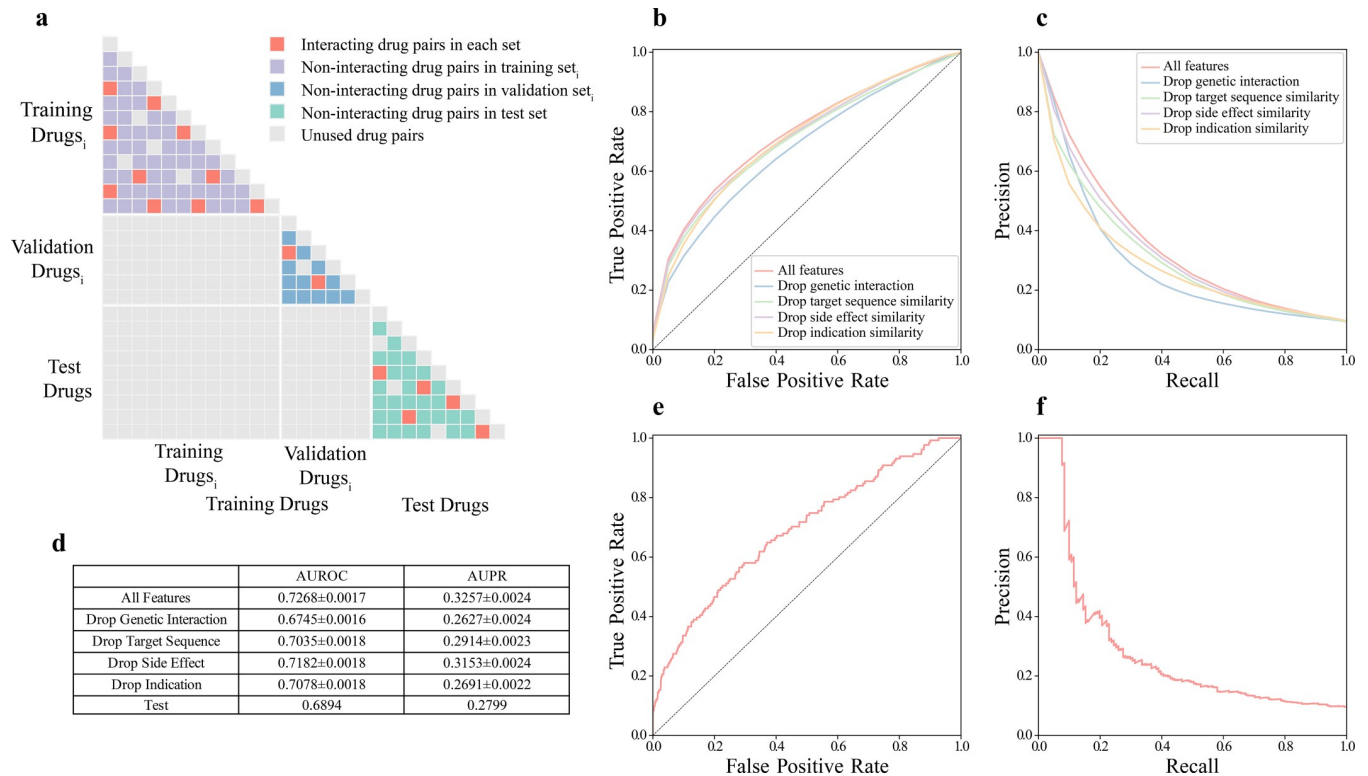
**Fig 1. Adversely interacting drug pairs and non-interacting drug pairs significantly differ with regard to the 11 features selected.** (a) Schematics of calculating indication similarity and side effect similarity features. (b) Indication similarity score of hierarchy level PT, HLT and SOC between two drugs. (c) Side effect similarity score of hierarchy level HLT and HLGT between two drugs. (d) Schematics of calculating target sequence similarity and genetic interaction features. Genetic interaction scores indicate the deviation from the expected phenotype when two genes are simultaneously knocked out, and were obtained from a global genetic interaction network in yeast by mapping targets of drugs to their yeast homologs. A negative score denotes synergistic interaction while a positive score indicates buffering interaction. (e) Minimum, mean, median and maximum target sequence similarity score between targets of two drugs. (f) Minimum and maximum genetic interaction score between targets of two drugs. Statistical significance was determined by the two-sided permutation test on the sample mean. PT, preferred term; HLT, high level term; HLGT, high level group term; SOC, system organ class. \*  $p < 0.001$ ; \*\*  $p < 0.0001$ ; \*\*\*  $p < 0.00001$ .

<https://doi.org/10.1371/journal.pcbi.1007068.g001>

pair, adversely interacting drugs showed significantly lower scores than non-interacting drugs irrespective of the aggregation function applied (Fig 1F, S2C Fig). This trend can be recapitulated using two recently published human genetic interaction datasets (S3B and S3C Fig). Furthermore, genetic interaction provides complementary information that is not captured by target similarity, indication similarity, or side effect similarity, as seen from their poor correlation (S4 Fig). Therefore, genetic interaction profiles of drug targets provide new information as a predictor of adverse DDIs.

### Building a machine learning model for predicting adverse DDIs

To divide drug pairs into a training set and a test set for building a machine learning model, most previous studies randomly split their data with a specified ratio [10,16,17,19,22,23,36,37], without considering the fact that drugs appearing in both sets may carry extra information about their interaction propensity. Considering the scenario of predicting interactions of drugs without prior information about their interaction profiles, this splitting scheme becomes inappropriate. To address this problem, we draw on a method that partitions drug pairs based on drugs [14,20,21]. All drugs in our constructed dataset were randomly split into “training drugs” and “test drugs” with a ratio of 2:1. The training set consists of all drug pairs where both drugs are “training drugs” and the test set comprises all drug pairs where both drugs are “test drugs” (Fig 2A). As a result, 475 interacting drug pairs and 4,802 non-interacting drug pairs involving 175 drugs went into the training set; 131 interacting drug pairs and 1,322 non-interacting drug pairs involving 87 drugs went into the test set.



**Fig 2. The train-test splitting scheme and model performance on the test set.** (a) The train-test splitting scheme. Drugs are randomly divided into “training drugs” and “test drugs” with ratio of 2:1. Training set only consists of drug pairs constituted by “training drugs” and test set only consists of drug pairs constituted by “test drugs”. Training drugs are further split into “training drugs<sub>i</sub>” and “validation drugs<sub>i</sub>” with the same splitting scheme to obtain training set<sub>i</sub> and validation set<sub>i</sub> in the training phase. For each iteration of hold-out validation, the classifier is fit with training set<sub>i</sub> and evaluated with validation set<sub>i</sub>. Purple squares represent non-interacting drug pairs in training set<sub>i</sub>. Blue squares represent non-interacting drug pairs in validation set<sub>i</sub>. Green squares represent non-interacting drug pairs in test set. Red squares represent interacting drug pairs in each set. Grey squares represent unused drug pairs. (b) Approximate receiver operating characteristic (ROC) curves on the training set. (c) Approximate precision-recall curves on the training set. (d) AUROCs and AUPRs on the training set and the test set. (e) Receiver operating characteristic (ROC) curve on the test set. (f) Precision-recall curve on the test set.

<https://doi.org/10.1371/journal.pcbi.1007068.g002>

To build a more interpretable model and speed up the training process, we applied a feature selection method known as group minimax concave penalty (MCP) [38] that has been previously employed on biological datasets [39]. This resulted in a final group of 11 features whose value distributions were all significantly different between adversely interacting drugs and non-interacting drugs (Fig 1B, 1E and 1F). An extreme gradient boosting (XGBoost) classifier [40] was then built because of its speed and outstanding performance in data science competitions. We optimized hyperparameters of the classifier using the tree-structured Parzen Estimator (TPE) approach [41], which has been shown to drastically improve the performance in a recent study predicting protein-protein interaction interfaces [42]. Notably, instead of doing cross-validation, we adopted the same drug-based splitting scheme on the training set for hold-out validation (Fig 2A). This enables the model to be best tuned for predicting interacting drug pairs without any prior information about the interaction profiles of the drugs involved. Indeed, a previous report by Liu et al. showed that classifier performance dropped significantly when evaluated on a test set consisted of pairs of drugs completely unseen in the training set if conventional cross-validation was performed [21], and this flaw in the generalizability of cross-validation performance has been shown to be true in general for pair-input data [43]. Our novel training strategy resulted in an average area under the receiver operating characteristic curve (AUROC) of 0.727 and an average area under the precision-recall curve

Table 1. Top 20 DDI predictions in the test set.

	ID1	ID2	Drug Name1	Drug Name2	Label	Probability
1	DB01076	DB01098	Atorvastatin	Rosuvastatin	1	0.9943
2	DB01076	DB01095	Atorvastatin	Fluvastatin	1	0.9938
3	DB00381	DB00421	Amlodipine	Spironolactone	1	0.9889
4	DB00880	DB00887	Chlorothiazide	Bumetanide	1	0.9704
5	DB00313	DB01356	Valproic Acid	Lithium cation	1	0.9614
6	DB00887	DB00999	Bumetanide	Hydrochlorothiazide	1	0.9463
7	DB00880	DB01119	Chlorothiazide	Diazoxide	1	0.8835
8	DB00421	DB01076	Spironolactone	Atorvastatin	1	0.8815
9	DB00421	DB00622	Spironolactone	Nicardipine	1	0.8753
10	DB00999	DB01119	Hydrochlorothiazide	Diazoxide	1	0.8723
11	DB00279	DB00755	Liothyronine	Tretinoin	0	0.8677
12	DB00162	DB00755	Vitamin A	Tretinoin	1	0.861
13	DB00477	DB01098	Chlorpromazine	Rosuvastatin	0	0.8601
14	DB00421	DB01119	Spironolactone	Diazoxide	0	0.8457
15	DB00313	DB00523	Valproic Acid	Alitretinoin	0	0.8419
16	DB05015	DB06176	Belinostat	Romidepsin	0	0.8378
17	DB01065	DB01069	Melatonin	Promethazine	1	0.8342
18	DB00622	DB01076	Nicardipine	Atorvastatin	1	0.7995
19	DB00755	DB00900	Tretinoin	Didanosine	0	0.7962
20	DB00162	DB01212	Vitamin A	Ceftriaxone	0	0.7858

<https://doi.org/10.1371/journal.pcbi.1007068.t001>

(AUPR) of 0.326 over 1,000 trials of hold-out validation on the training set (Fig 2B–2D). When evaluated on the test set, our classifier achieved an AUROC of 0.689 (Fig 2D and 2E) and an AUPR of 0.280 (Fig 2D and 2F), demonstrating the robustness of our model. As shown in Table 1, our classifier attained a precision of 100% on the top 10 predictions, and a precision of 65% on the top 20 predictions (Table 1). Since there is no gold-standard set of non-interacting drugs, it is plausible that our non-interacting drug pairs might actually contain adverse DDIs. Not surprisingly, some non-interacting drug pairs with the high predicted probabilities can be found with evidence supporting their possible adverse interactions. For example, the drug pair with a non-interacting label with the highest predicted interacting probability in the test set, liothyronine and tretinoin, has been indicated to potentially cause intracranial pressure increase and a higher risk of pseudotumor cerebri when taken together [44]. Furthermore, diazoxide and spironolactone, predicted with an interacting probability of 0.846, have been reported to induce asthma, cardiac hypertrophy and pulmonary edema according to FDA reports when co-administrated [45].

In order to showcase the competitiveness of the XGBoost algorithm, we implemented a number of alternative classification algorithms including support vector machine (SVM), random forest and the standard gradient boosting algorithm and performed the same prediction task using exactly the same dataset and features. We found that XGBoost achieved better or comparable performance than the other algorithms (S3 Table). Furthermore, XGBoost is substantially faster than its closest contenders in terms of performance, gradient boosting and random forest. These results highlight the advantage of XGBoost over other algorithms in both predictive performance and speed. To further demonstrate the efficacy of our method, we compared it against a previously published similarity-based method for DDI prediction [18] using our training and test sets. Our method exhibited a substantial advantage both in training and on the test set (S3 Table).

Table 2. Top 20 new adverse DDI predictions.

	ID1	ID2	Drug Name1	Drug Name2	Probability	In Twosides
1	DB00347	DB01189	Trimethadione	Desflurane	0.975	No
2	DB00136	DB00630	Calcitriol	Alendronic acid	0.9739	Yes
3	DB00417	DB01050	Phenoxymethylpenicillin	Ibuprofen	0.9699	Yes
4	DB00228	DB00347	Enflurane	Trimethadione	0.968	No
5	DB00347	DB00753	Trimethadione	Isoflurane	0.9668	No
6	DB00347	DB01236	Trimethadione	Sevoflurane	0.9656	No
7	DB00887	DB01586	Bumetanide	Ursodeoxycholic acid	0.9567	Yes
8	DB00532	DB01189	Mephenytoin	Desflurane	0.9531	No
9	DB00532	DB01236	Mephenytoin	Sevoflurane	0.9521	No
10	DB00228	DB00532	Enflurane	Mephenytoin	0.9466	No
11	DB01067	DB01083	Glipizide	Orlistat	0.9456	Yes
12	DB00731	DB01016	Nateglinide	Glyburide	0.943	Yes
13	DB01050	DB01053	Ibuprofen	Benzylpenicillin	0.9422	No
14	DB00532	DB00753	Mephenytoin	Isoflurane	0.9368	No
15	DB00421	DB01216	Spironolactone	Finasteride	0.9333	Yes
16	DB00162	DB00165	Vitamin A	Pyridoxine	0.926	No
17	DB01236	DB04930	Sevoflurane	Permethrin	0.925	No
18	DB00284	DB00731	Acarbose	Nateglinide	0.9199	Yes
19	DB00136	DB00273	Calcitriol	Topiramate	0.9132	Yes
20	DB00877	DB01586	Sirolimus	Ursodeoxycholic acid	0.9122	Yes

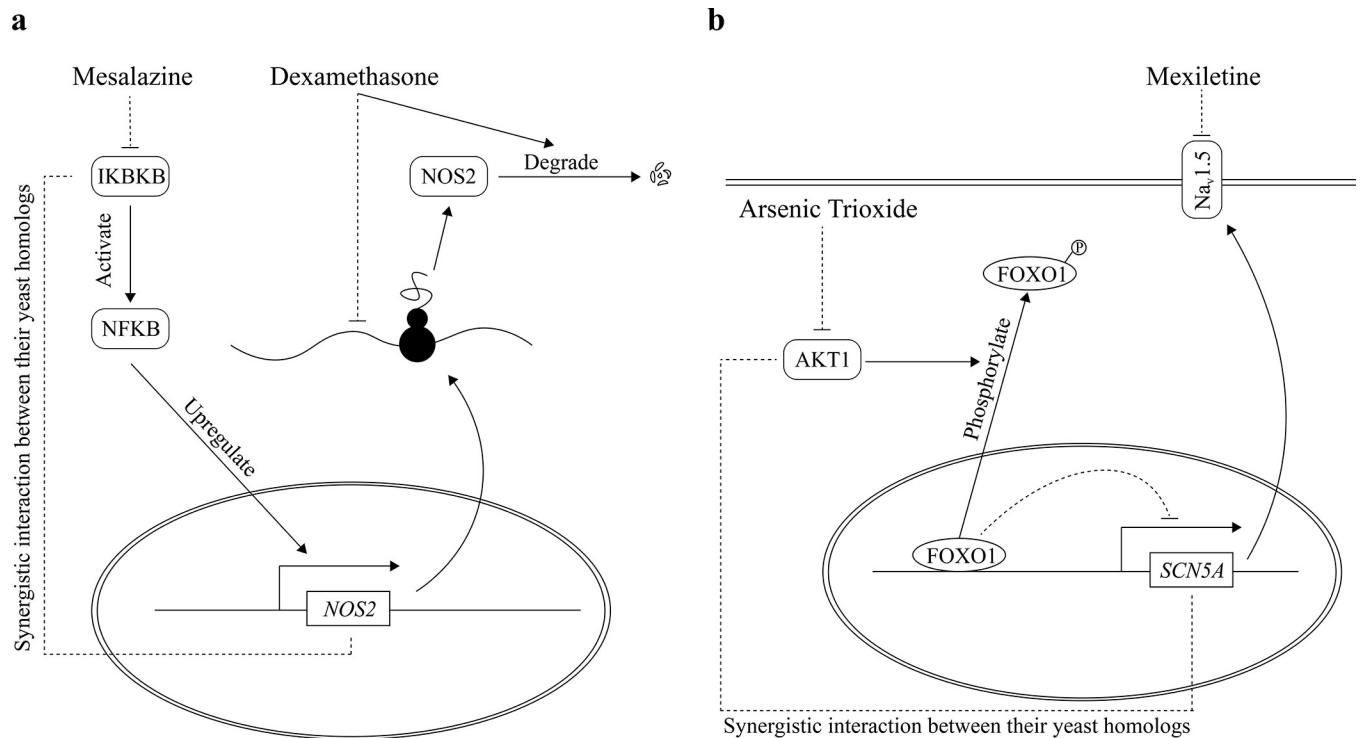
<https://doi.org/10.1371/journal.pcbi.1007068.t002>

To demonstrate the utility of our method, we obtained 5,039 drug pairs involving 295 drugs that had not been used for training and testing (S6 Fig). After refitting our model on all 12,426 drug pairs that were used to develop our method, we predicted 432 novel DDIs (S4 Table). Remarkably, out of the top 20 newly predicted adversely interacting drug pairs, 9 can be verified in the TWOSIDES database (Table 2), manifesting the reliability of our method.

### Genetic interaction provides mechanistic insight into drug-drug interactions

We investigated the contribution of genetic interaction features to classifier performance by building and tuning a new model without them. Excluding genetic interaction features significantly decreases classifier performance when either AUROC or AUPR is examined ( $P < 10^{-20}$  for both AUROC and AUPR, two-sided Welch's t-test). More interestingly, the performance drop is not as profound when other groups of features are excluded (Fig 2B–2D). Furthermore, prediction with genetic interaction features alone rendered significantly better performance than prediction with target similarity features alone ( $P < 10^{-20}$  for both AUROC and AUPR, two-sided Welch's t-test, S3 Table). These results establish genetic interaction as an important feature in our model for predicting DDIs, providing complementary information that other features cannot capture.

More importantly, genetic interaction can help us generate plausible mechanistic explanations for drug-drug interactions. For example, mesalazine and dexamethasone, both of which are anti-inflammatory drugs, are a pair of drugs in the test set that have been labeled as adversely interacting. Mesalazine can target the IKBKB protein, whereas dexamethasone can target NOS2, which plays important roles in nitric oxide signaling. In yeast, double knockout of *ATG1* and *TAH18*, the respective yeast homologs of *IKBKB* and *NOS2*, exhibits a more negative impact on cell viability than expected from single knockout phenotypes [35]. In human,



**Fig 3. Genetic interaction provides possible mechanistic insights into DDIs.** (a) Mesalazine inhibits IKBKB, a positive regulator of NF-κB activity, and NF-κB is a transcription factor which induces NOS2 transcription. Dexamethasone can inhibit the transcription of NOS2 and facilitate degradation of NOS2. The combined use of dexamethasone and mesalazine could potentially reduce the amount of NOS2 in cells to a large extent, which may affect neurotransmission, antimicrobial and antitumoral activities. (b) Mexiletine targets Nav1.5, a sodium channel encoded by SCN5A, while arsenic trioxide targets AKT1. The transcription of SCN5A is repressed by the transcriptional repressor FOXO1. AKT1 can activate the transcription of SCN5A by phosphorylating FOXO1. The combined use of mexiletine and arsenic trioxide could inactivate the transcription of SCN5A and at the same time block the existing sodium channel, which may largely reduce sodium influx in cardiac cells.

<https://doi.org/10.1371/journal.pcbi.1007068.g003>

IKBKB can phosphorylate the NF-κB inhibitor and activate NF-κB [46], which is a family of transcription factors involved in inflammation and immunity. Notably, the transcription of NOS2 is induced by NF-κB activity [47]. Mesalazine has been shown to inhibit IKBKB, thereby inhibiting the activation of NF-κB, while dexamethasone is a negative modulator of NOS2. A previous study has reported that dexamethasone can decrease NOS2 translation and facilitate NOS2 degradation in rat [48] (Fig 3A). The combined use of mesalazine and dexamethasone may largely reduce the amount of NOS2, potentially affecting neurotransmission, antimicrobial and antitumoral activities.

As another example, arsenic trioxide and mexiletine are a pair of drugs not labelled as adversely interacting in DrugBank, but predicted by our model to interact with high probability. As a chemotherapy drug for acute promyelocytic leukemia (APL), arsenic trioxide has been reported to decrease the activity of a serine/threonine-protein kinase AKT1 [49]. On the other side, mexiletine is a sodium channel blocker that has also been used as part of a prophylactic therapy to treat APL patients to reduce cardiac complications [50]. PKC1, the yeast homolog of AKT1, exhibits strong synergistic interaction with CCH1 [35,51], which is the homolog of SCN5A, the gene encoding the sodium channel Nav1.5 targeted by mexiletine. In human, the transcription of SCN5A is repressed by FOXO1, whose transcriptional repression activity is in turn inactivated by AKT1-dependent phosphorylation [52] (Fig 3B). Therefore, the simultaneous inhibition of AKT1 and the sodium channel by the two drugs may reduce sodium influx in cardiac cells to a greater extent, potentially causing undesired adverse effects.



Indeed, this pair of drugs is reported by TWOSIDES as interacting, providing additional supporting evidence to their adverse interaction.

## Discussion

In the past decade, many methods have been developed for predicting DDIs based on various types of features. In this study, we have incorporated a novel feature, namely genetic interaction, to build a gradient boosting-based model for fast and accurate adverse DDI prediction. We have shown that our classifier can robustly predict drug-drug interactions even for drugs whose interaction profiles are completely unseen during training. Furthermore, we have predicted 432 novel DDIs, with additional evidence supporting our top predictions, demonstrating the usefulness of our approach.

Most previous efforts of predicting DDIs suffer from an inability to make predictions for newly developed drugs due to train-test split based on drug pairs rather than drugs [10,16,17,19,22,23,36,37]. Three studies attempted to address this problem by dividing the entire dataset based on drugs [14,20,21]. However, they failed to do so during the training phase, resulting in an inflated performance on the training set. We have followed the drug-based train-test splitting scheme and have adopted a hold-out validation approach to avoid using overlapping drug sets for fitting the model and evaluating its performance. By doing so, we have achieved robust performance on the training set and the test set, which establishes the ability of our method to predict new DDIs for drugs whose interaction profiles are completely unknown.

By examining genetic interactions, our method provides mechanistic insights into how two drugs may interact in a detrimental fashion. The combined modulatory effect resulted from binding of two drugs to their respective targets might underlie adverse DDIs, and genetic interaction gives valuable information about the nature of such combined effect. Indeed, we have observed that genetic interaction features are indispensable to our classifier performance. Notably, target sequence similarity features and genetic similarity features capture conceptually different mechanisms by which DDIs can occur. While the former can capture dosage effects where two drugs target same or similar genes, as exemplified by prolonged QT interval caused by concomitant administration of terfenadine and ketoconazole, both of which are strong CYP3A4 inhibitors [53], the latter captures DDIs resulting from drug pairs targeting genes with an epistatic relationship. For example, asthma patients receiving leukotriene-modifying drugs often show attenuated response to  $\beta_2$ -agonists, including albuterol. This drug-drug interaction has been implicated to be associated with the epistasis between *ALOX5AP* and *LTA4H* [54].

Nevertheless, our work is limited by the lack of a global human genetic interaction network. As a surrogate for human genetic interactions, genetic interactions of yeast homologs were used in this study. Fortunately, large-scale human genetic interaction studies are coming into sight. Using a recently published dataset of human genetic interactions in K562 cells encompassing 222,784 gene pairs [55], we have found that the distribution of human genetic interaction scores vary significantly between adversely interacting drugs and non-interacting drugs (S3B Fig). Notably, the same trends could be recapitulated with a smaller dataset of genetic interactions [56] in the HEK293T cell line, demonstrating the generalizability of genetic interactions across different cell contexts (S3C Fig), although certain genetic interactions can exist in a cell type-dependent manner. For example, interactions between cancer driver genes are frequently specific to the cancer type [57]. In addition to DDI prediction, a similar machine learning method leveraging genetic interaction features can potentially be developed for predicting beneficial drug combinations. Indeed, current combination therapy for cancers have

typically been developed to induce synthetic lethal genetic interactions in cancer cells [58,59]. While there have been some efforts aimed at predicting synergistic drug effects [60,61] or directly predicting drug combinations for disease therapy, especially cancer treatment [62–64], incorporating cell type-specific genetic interaction data from the matching cell type can be crucial for developing combination therapies that specifically target certain cell types. With the continuous advancement of technologies for probing human genetic interactions including CRISPR interference, we anticipate that more comprehensive maps of human genetic interactions for multiple cell lineages will become available in the near future, which could illuminate predictions of adverse DDIs and beneficial drug combinations to a larger extent.

## Methods

### Data collection

We obtained DDI data from DrugBank (version 5.0.10) [28]. Among the 5 major interaction categories in DrugBank (**S1 Table**), we only considered the first category as they were clearly defined as adverse DDIs. Non-interacting drug pairs were constructed by taking all other combinations using the same set of drugs, removing drug pairs also appearing in other categories in DrugBank, TWOSIDES [29], or a complete dataset of DDIs [30] compiled from a number of sources. This minimizes the chance of having actual adverse DDIs in the non-interacting set given the absence of a gold standard set of non-interacting drug pairs. From DrugBank, we also collected human protein targets of drugs and their sequences.

Side effects were obtained from SIDER 4.1 [65] and OFFSIDES [29]. Both databases use UMLS concept IDs as their side effect identifiers. However, as reported by Zhang et al. [20], some side effect terms are similar, and synonyms could cause biases when calculating side effect similarity. To solve this problem, we obtained mapping from UMLS concept IDs to MedDRA concept IDs from the 2017AB release of UMLS [66]. Furthermore, we obtained the full MedDRA hierarchy from MedDRA (version 21.0) [31]. This allowed us to map UMLS concept IDs to different levels (PT, HLT, HLGT and SOC) of the MedDRA hierarchy. Similar to side effect data, indications of drugs were acquired from SIDER 4.1 [65] and mapped to the same 4 levels of the MedDRA hierarchy.

For genetic interactions, we obtained yeast genetic interactions from Costanzo et al. [35]. We first filtered all genetic interactions by a p-value cutoff of 0.05 and aggregated the scores of all combinations of alleles of each yeast gene pair by applying the arithmetic mean. Drug targets in the form of UniProt IDs were mapped to gene names by UniProt [67] and these human genes were mapped to their yeast homologs via SGD YeastMine [68]. For human gene pairs mapped to multiple yeast gene pairs, we obtained a single score for each human gene pair by applying the arithmetic mean.

### Feature extraction and the train-test split

For a drug pair (A,B), four groups of features were calculated (**Fig 1A and 1D**): indication similarity scores between A and B, side effect similarity scores between A and B, target sequence similarity scores between targets of drug A and targets of drug B, and genetic interaction scores between targets of drug A and targets of drug B. Indications and side effects of drugs were mapped to 4 different levels of the MedDRA hierarchy as described above. At each level, indication similarity was calculated by taking the Jaccard index between the respective indication vectors of drug A and drug B (**Fig 1A**). Similarly, side effect similarity was calculated by applying the same measure on the side effect vectors at the 4 different MedDRA hierarchy levels (**Fig 1A**). For genetic interactions, since each drug can have multiple targets, we obtained a single score for each drug pair by aggregating the genetic interaction scores of all their

corresponding target pairs using 4 different functions, namely taking the minimum, mean, median or maximum (Fig 1D). Similarly, the same 4 functions were used for constructing target similarity features, which were calculated from the target sequences with the Smith-Waterman algorithm using the scikit-bio Python library. The raw scores were normalized as described in Bleakley et al. [69]. Overall, 16 features belonging to 4 feature groups were constructed. Only drug pairs with all features available were considered when building the machine learning model. All drugs were randomly split into “training drugs” and “test drugs” with a 2:1 ratio. The training set consisted of all drug pairs where both drugs were “training drugs” and the test set consisted of all drug pairs where both drugs were “test drugs” (Fig 2A, S5 Table). We constrained the fraction of adversely interacting drug pairs in the training set and that in the test set to be fairly balanced. To obtain the optimal feature combination, we calculated all features for the training set and applied group minimax concave penalty (MCP) [38] with the ‘grpreg’ R package with default parameters. All subsequent training was done using this optimal set of features.

### Hyperparameter optimization and classifier training

The gradient boosting-based algorithm XGBoost [40] was used in this study. To find the best combination of hyperparameters for the XGBoost classifier, the tree-structured Parzen estimator (TPE) approach [41] was adopted. Because of the drug-based approach by which we split our dataset into training and test sets, we applied the same splitting scheme on the training set multiple times to obtain training set<sub>i</sub> and validation set<sub>i</sub> instead of simply using cross-validation. Each split on the training set can be seen as a hold-out validation, as we used training set<sub>i</sub> to fit the model and validated model performance on validation set<sub>i</sub>. We selected one minus the average AUPR of 50 trials of hold-out validation as the loss function to minimize for TPE, and we ran TPE for 2,000 iterations to obtain set of hyperparameters that minimized the loss function for our XGBoost classifier (S5 Fig). After finding the optimal set of hyperparameters, we fit the model on the complete training data.

### Model evaluation

Model performance on training set was evaluated by 1,000 runs of hold-out validation on the training set. For each hold-out validation, we fitted the model on training set<sub>i</sub> and obtained AUROC and AUPR. We averaged AUROC and AUPR over 1,000 runs of hold-out validation as measurements of the performance of the model. Approximate ROC curve and precision-recall curve (Fig 2B and 2C) were plotted by averaging the 1,000 ROC curves and 1,000 precision-recall curves respectively at every thousandth of a point on the x-axis. In order to evaluate the ability of the classifier to identify drug-drug interactions between drugs whose interaction profiles were completely unknown during training, the model was evaluated on the test set which had no overlap with the training set in terms of the drugs involved. Predictions were ranked according to their raw prediction scores to produce the ROC curve and the precision-recall curve.

### Making new predictions

To make novel adverse DDI predictions, we examined all combinations of drugs that appeared in DrugBank, excluding drug pairs where both drugs were involved in the first category of DDIs (S6 Fig), which we used for building the machine learning model. We then predicted 6,690 drug pairs involving 336 drugs for which all features could be calculated using the classifier retrained on the whole dataset. The probability cutoff that produced the maximum

averaged F1 score over 1,000 runs of hold-out validation on the training set was chosen for determining new DDI predictions.

## Supporting information

**S1 Fig. Schematics of our DDI prediction framework.** Four groups of features were calculated for each drug pair. Drug pairs were then divided into a training set and a test set. A gradient boosting-based model was built on the training set after feature selection. Model performance was evaluated on the training set using hold-out validation and also on the test set. We demonstrate the importance of our novel feature with a case study and provide novel DDI predictions at the end.

(TIF)

**S2 Fig. The distribution of adversely interacting drug pairs and non-interacting drug pairs in terms of the 5 unused features.** (a) Indication similarity score of hierarchy level HLT between two drugs. (b) Side effect similarity score of hierarchy levels PT and SOC between two drugs. (c) Mean and median genetic interaction score between targets of two drugs. Statistical significance was determined by the two-sided permutation test on the sample mean. \*

$p < 0.001$ ; \*\*  $p < 0.0001$ ; \*\*\*  $p < 0.00001$ .

(TIF)

**S3 Fig.** (a) The total number of protein targets between two drugs. (b) Minimum, mean, median and maximum human K562 cell line genetic interaction score between targets of two drugs. (Statistical significance determined by two-sided Mann-Whitney U test) (c) Minimum, mean, median and maximum human HEK293T cell line genetic interaction score between targets of two drugs. (Statistical significance determined by two-sided Mann-Whitney U test).

(TIF)

**S4 Fig. The correlation between genetic interaction features and other features.**

(TIF)

**S5 Fig. Values of hyperparameters of the XGBoost model over 2000 TPE iterations.**

(TIF)

**S6 Fig. Construction of a set of drug pairs used for new predictions.** (a) All combinations between drugs that appear in the first category in DrugBank and other drugs, as well as all pairwise combinations of drugs not in the first category, are taken for new predictions. Green squares represent drug pairs used for building the classifier. Grey squares represent unused drug pairs. Blue squares represent drug pairs used for new predictions. (b) Maximum target similarity feature distribution of drug pairs used for model building (green triangular section in (a)), drug pairs where one drug appears in the dataset used for model building (blue rectangular section in (a)), and drug pairs where neither drug appears in the dataset used or model building (blue triangular section in (a)).

(TIF)

**S1 Table. Five main DDI categories in DrugBank.**

(DOCX)

**S2 Table. Summary statistics including mean, standard error of the mean and p-value of each feature.** Statistical significance was determined by the two-sided permutation test on the sample mean.

(XLSX)

**S3 Table. Tab 1: Performance comparison of XGBoost with several other algorithms with and without genetic interaction features.** Tab 2: comparison of our method with Zhao and Cheng, 2014. Tab 3: model performance using only genetic interaction features of target sequence similarity features.

(XLSX)

**S4 Table. A list of 432 new adverse DDI predictions.**

(XLSX)

**S5 Table. A list of all drug pairs in the training set and a list of all drug pairs in the test set.**

(XLSX)

**S6 Table. Side effects, indications, human gene targets and their yeast homolog of all drugs that appear in the training set or the test set.**

(XLSX)

## Acknowledgments

The authors would like to thank G. Hooker, S. Chen and S. Wierbowski for helpful discussions.

## Author Contributions

**Conceptualization:** Haiyuan Yu.

**Data curation:** Sheng Qian, Siqi Liang.

**Formal analysis:** Sheng Qian, Siqi Liang.

**Funding acquisition:** Haiyuan Yu.

**Methodology:** Sheng Qian, Siqi Liang, Haiyuan Yu.

**Project administration:** Haiyuan Yu.

**Supervision:** Haiyuan Yu.

**Validation:** Sheng Qian, Siqi Liang.

**Visualization:** Sheng Qian.

**Writing – original draft:** Sheng Qian, Siqi Liang.

**Writing – review & editing:** Sheng Qian, Siqi Liang, Haiyuan Yu.

## References

1. Crowther NR, Holbrook AM, Kenwright R, Kenwright M. Drug interactions among commonly used medications: Chart simplifies data from critical literature review. *Can Fam Physician*. 1997; 43: 1972–1981. PMID: [9386884](https://pubmed.ncbi.nlm.nih.gov/9386884/)
2. Lu Y, Shen D, Pietsch M, Nagar C, Fadli Z, Huang H, et al. A novel algorithm for analyzing drug-drug interactions from MEDLINE literature. *Sci Rep*. Nature Publishing Group; 2015; 5: 17357. <https://doi.org/10.1038/srep17357> PMID: [26612138](https://pubmed.ncbi.nlm.nih.gov/26612138/)
3. Percha B, Altman RB. Informatics confronts drug-drug interactions. *Trends Pharmacol Sci*. Elsevier Ltd; 2013; 34: 178–184. <https://doi.org/10.1016/j.tips.2013.01.006> PMID: [23414686](https://pubmed.ncbi.nlm.nih.gov/23414686/)
4. National Center for Health Statistics. Health, United States, 2016: With Chartbook on Long-Term Trends in Health (US Department of Health and Human Services, Hyattsville, MD, 2017). Hyattsville;
5. Gu Q, Dillon CF, Burt VL. Prescription drug use continues to increase: U.S. prescription drug data for 2007–2008. *NCHS Data Brief*. 2010; 42: 1–8. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20854747>

6. Becker ML, Kallewaard M, Caspers PW, Visser LE, Leufkens HG, Stricker BHC. Hospitalisations and emergency department visits due to drug–drug interactions: a literature review. *Pharmacoepidemiol Drug Saf.* 2007; 16: 641–651. <https://doi.org/10.1002/pds.1351> PMID: 17154346
7. Brown HS, Ito K, Houston AG, Brian J. Prediction of in vivo drug-drug interactions from in vitro data: Impact of incorporating parallel pathways of drug elimination and inhibitor absorption rate constant. *Br J Clin Pharmacol.* 2005; 60: 508–518. <https://doi.org/10.1111/j.1365-2125.2005.02483.x> PMID: 16236041
8. Ohno Y, Hisaka A, Ueno M, Suzuki H. General framework for the prediction of oral drug interactions caused by CYP3A4 induction from in vivo information. *Clin Pharmacokinet.* 2008; 47: 669–680. <https://doi.org/10.2165/00003088-200847100-00004> PMID: 18783297
9. Duke JD, Han X, Wang Z, Subhadarshini A, Karnik SD, Li X, et al. Literature Based Drug Interaction Prediction with Clinical Assessment Using Electronic Medical Records: Novel Myopathy Associated Drug Interactions. *PLoS Comput Biol.* 2012; 8: e1002614. <https://doi.org/10.1371/journal.pcbi.1002614> PMID: 22912565
10. Vilar S, Harpaz R, Uriarte E, Santana L, Rabadan R, Friedman C. Drug-drug interaction through molecular structure similarity analysis. *J Am Med Inform Assoc.* 2012; 19: 1066–1074. <https://doi.org/10.1136/amiajnl-2012-000935> PMID: 22647690
11. Vilar S, Uriarte E, Santana L, Friedman C, Tatonetti NP. State of the art and development of a drug-drug interaction large scale predictor based on 3D pharmacophoric similarity. *Curr Drug Metab.* 2014; 15: 490–501. PMID: 25431152
12. Ferdousi R, Safdari R, Omidi Y. Computational prediction of drug-drug interactions based on drugs functional similarities. *J Biomed Inform.* 2017; 70: 54–64. <https://doi.org/10.1016/j.jbi.2017.04.021> PMID: 28465082
13. Zhang W, Chen Y, Liu F, Luo F, Tian G, Li X. Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. *BMC Bioinformatics.* 2017; 18: 18. <https://doi.org/10.1186/s12859-016-1415-9> PMID: 28056782
14. Abdelaziz I, Fokoue A, Hassanzadeh O, Zhang P, Sadoghi M. Large-scale structural and textual similarity-based mining of knowledge graph to predict drug-drug interactions. *Web Semant Sci Serv Agents World Wide Web.* 2017; 44: 104–117.
15. Ryu JY, Kim HU, Lee SY. Deep learning improves prediction of drug–drug and drug–food interactions. *Proc Natl Acad Sci.* 2018; 115: E4304–4311. <https://doi.org/10.1073/pnas.1803294115> PMID: 29666228
16. Kastrin A, Ferk P, Leskos B. Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PLoS One.* 2018; 13: e0196865. <https://doi.org/10.1371/journal.pone.0196865> PMID: 29738537
17. Gottlieb A, Stein GY, Oron Y, Ruppin E, Sharan R. INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol.* Nature Publishing Group; 2012; 8: 592. <https://doi.org/10.1038/msb.2012.26> PMID: 22806140
18. Cheng F, Zhao Z. Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J Am Med Inform Assoc.* 2014; 21: e278–e286. <https://doi.org/10.1136/amiajnl-2013-002512> PMID: 24644270
19. Luo H, Zhang P, Huang H, Huang J, Kao E, Shi L, et al. DDI-CPI, a server that predicts drug-drug interactions through implementing the chemical-protein interactome. *Nucleic Acids Res.* 2014; 42: W46–W52. <https://doi.org/10.1093/nar/gku433> PMID: 24875476
20. Zhang P, Wang F, Hu J, Sorrentino R. Label Propagation Prediction of Drug-Drug Interactions Based on Clinical Side Effects. *Sci Rep.* Nature Publishing Group; 2015; 5: 12339. <https://doi.org/10.1038/srep12339> PMID: 26196247
21. Liu L, Chen L, Zhang YH, Wei L, Cheng S, Kong X, et al. Analysis and prediction of drug–drug interaction by minimum redundancy maximum relevance and incremental feature selection. *J Biomol Struct Dyn.* Taylor & Francis; 2017; 35: 312–329. <https://doi.org/10.1080/07391102.2016.1138142> PMID: 26750516
22. Sridhar D, Fakhraei S, Getoor L. A probabilistic approach for collective similarity-based drug-drug interaction prediction. *Bioinformatics.* 2016; 32: 3175–3182. <https://doi.org/10.1093/bioinformatics/btw342> PMID: 27354693
23. Hameed PN, Verspoor K, Kusljic S, Halgamuge S. Positive-Unlabeled Learning for inferring drug interactions based on heterogeneous attributes. *BMC Bioinformatics.* 2017; 18: 140. <https://doi.org/10.1186/s12859-017-1546-7> PMID: 28249566
24. Takeda T, Hao M, Cheng T, Bryant SH, Wang Y. Predicting drug-drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics

- knowledge. *J Cheminform*. Springer International Publishing; 2017; 9: 16. <https://doi.org/10.1186/s13321-017-0200-8> PMID: 28316654
25. Huang J, Niu C, Green CD, Yang L, Mei H, Han JDJ. Systematic Prediction of Pharmacodynamic Drug-Drug Interactions through Protein-Protein-Interaction Network. *PLoS Comput Biol*. 2013; 9: e1002998. <https://doi.org/10.1371/journal.pcbi.1002998> PMID: 23555229
  26. Cami A, Manzi S, Arnold A, Reis BY. Pharmacointeraction Network Models Predict Unknown Drug-Drug Interactions. *PLoS One*. 2013; 8: e61468. <https://doi.org/10.1371/journal.pone.0061468> PMID: 23620757
  27. Park K, Kim D, Ha S, Lee D. Predicting pharmacodynamic drug-drug interactions through signaling propagation interference on protein-protein interaction networks. *PLoS One*. 2015; 10: e0140816. <https://doi.org/10.1371/journal.pone.0140816> PMID: 26469276
  28. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res*. Oxford University Press; 2018; 46: D1074–D1082. <https://doi.org/10.1093/nar/gkx1037> PMID: 29126136
  29. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-Driven Prediction of Drug Effects and Interactions. *Sci Transl Med*. 2012; 4: 125ra31. <https://doi.org/10.1126/scitranslmed.3003377> PMID: 22422992
  30. Ayvaz S, Horn J, Hassanzadeh O, Zhu Q, Stan J, Tatonetti NP, et al. Toward a complete dataset of drug-drug interaction information from publicly available sources. *J Biomed Inform*. Elsevier Inc.; 2015; 55: 206–217. <https://doi.org/10.1016/j.jbi.2015.04.006> PMID: 25917055
  31. Brown E, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf*. 1999; 20: 109–117. <https://doi.org/10.2165/00002018-199920020-00002> PMID: 10082069
  32. Zhao M, Lee WP, Garrison EP, Marth GT. SSW library: An SIMD Smith-Waterman C/C++ library for use in genomic applications. *PLoS One*. 2013; 8: e82138. <https://doi.org/10.1371/journal.pone.0082138> PMID: 24324759
  33. Mani R, Onge RP St., Hartman JL, Giaever G, Roth FP. Defining genetic interaction. *Proc Natl Acad Sci*. 2008; 105: 3461–3466. <https://doi.org/10.1073/pnas.0712255105> PMID: 18305163
  34. Boucher B, Jenna S. Genetic interaction networks: Better understand to better predict. *Front Genet*. 2013; 4: 290. <https://doi.org/10.3389/fgene.2013.00290> PMID: 24381582
  35. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science*. 2016; 353: aaf1420. <https://doi.org/10.1126/science.aaf1420> PMID: 27708008
  36. Vilar S, Uriarte E, Santana L, Tatonetti NP, Friedman C. Detection of Drug-Drug Interactions by Modeling Interaction Profile Fingerprints. *PLoS One*. 2013; 8: e58321. <https://doi.org/10.1371/journal.pone.0058321> PMID: 23520498
  37. Liu R, AbdulHameed MDM, Kumar K, Yu X, Wallqvist A, Reifman J. Data-driven prediction of adverse drug reactions induced by drug-drug interactions. *BMC Pharmacol Toxicol*. BMC Pharmacology and Toxicology; 2017; 18: 44. <https://doi.org/10.1186/s40360-017-0153-6> PMID: 28595649
  38. Breheny P, Huang J. Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. *Stat Comput*. 2015; 25: 173–187. <https://doi.org/10.1007/s11222-013-9424-2> PMID: 25750488
  39. Breheny P, Huang J. Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *Ann Appl Stat*. 2011; 5: 232–253. <https://doi.org/10.1214/10-AOAS388> PMID: 22081779
  40. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In *Proc 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016; 785–794.
  41. Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for Hyper-Parameter Optimization. In *Advances in Neural Information Processing Systems* (eds Shawe-Taylor, T et al). 2011; 2546–2554.
  42. Meyer MJ, Beltrán JF, Liang S, Fragoza R, Rumack A, Liang J, et al. Interactome INSIDER: a structural interactome browser for genomic studies. *Nat Methods*. 2018; 15: 107–114. <https://doi.org/10.1038/nmeth.4540> PMID: 29355848
  43. Park Y, Marcotte EM. A flaw in the typical evaluation scheme for pair-input computational predictions. *Nat Methods*. 2012; 9: 1134–1136. <https://doi.org/10.1038/nmeth.2259> PMID: 23223166
  44. LIOTHYRONINE SODIUM SR CAPSULES [Internet]. Available: [https://www.empowerpharmacy.com/drugs/liothyronine-sodium-sr-capsules.html#footnote22\\_c6u8axd](https://www.empowerpharmacy.com/drugs/liothyronine-sodium-sr-capsules.html#footnote22_c6u8axd)
  45. Craigle V. MedWatch: The FDA Safety Information and Adverse Event Reporting Program. *J Med Libr Assoc*. Rockville: Md: U.S. Food and Drug Administration; 2007; 95: 224–225.

46. Salmerón A, Janzen J, Soneji Y, Bump N, Kamens J, Allen H, et al. Direct phosphorylation of NF- $\kappa$ B p105 by the I $\kappa$ B kinase complex on serine 927 is essential for signal-induced p105 proteolysis. *J Biol Chem*. 2001; 276: 22215–22222. <https://doi.org/10.1074/jbc.M101754200> PMID: 11297557
47. Liu T, Zhang L, Joo D, Sun S-C. NF- $\kappa$ B signaling in inflammation. *Signal Transduct Target Ther*. 2017; 2: 17023. <https://doi.org/10.1038/sigtrans.2017.23> PMID: 29158945
48. Kunz D, Walker G, Eberhardt W, Pfeilschifter J. Molecular mechanisms of dexamethasone inhibition of nitric oxide synthase expression in interleukin 1 beta-stimulated mesangial cells: evidence for the involvement of transcriptional and posttranscriptional regulation. *Proc Natl Acad Sci*. 1996; 93: 255–259. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40217&tool=pmcentrez&rendertype=abstract> <https://doi.org/10.1073/pnas.93.1.255> PMID: 8552616
49. Guilbert C, Annis MG, Dong Z, Siegel PM, Jr WHM, Koren K. Arsenic Trioxide Overcomes Rapamycin-Induced Feedback Activation of AKT and ERK Signaling to Enhance the Anti-Tumor Effects in Breast Cancer. 2013; 8: e85995. <https://doi.org/10.1371/journal.pone.0085995> PMID: 24392034
50. Kazunori O, H Y, K S, S N, S F, K N, et al. Prolongation of the QT Interval and Ventricular Tachycardia in Patients Treated with Arsenic Trioxide for Acute Promyelocytic Leukemia. *Ann Intern Med*. 2001; 133: 881–885.
51. Srivas R, Shen JP, Yang CC, Sun SM, Li J, Gross AM, et al. A network of conserved synthetic lethal interactions for exploration of precision cancer therapy. *Mol Cell*. 2016; 63: 514–525. <https://doi.org/10.1016/j.molcel.2016.06.022> PMID: 27453043
52. Ballou LM, Lin RZ, Cohen IS. Control of Cardiac Repolarization by Phosphoinositide 3-kinase Signaling to Ion Channels. *Circ Res*. 2015; 16: 127–137. <https://doi.org/10.1161/CIRCRESAHA.116.303975> Control
53. Wiśniewska B, Tylutki Z, Wyszogrodzka G, Polak S. Drug-drug interactions and QT prolongation as a commonly assessed cardiac effect—comprehensive overview of clinical trials. *BMC Pharmacol Toxicol*. 2016; 17: 12. <https://doi.org/10.1186/s40360-016-0053-1> PMID: 26960809
54. Via M, Tcheurekdjian H, Burchard EG. Role of interactions in pharmacogenetic studies: leukotrienes in asthma. *Pharmacogenomics*. 2013; 14: 923–929. <https://doi.org/10.2217/pgs.13.70> PMID: 23746186
55. Horlbeck MA, Xu A, Wang M, Bennett NK, Park CY, Bogdanoff D, et al. Mapping the Genetic Landscape of Human Cells. *Cell*. Elsevier Inc.; 2018; 174: 953–967.e22. <https://doi.org/10.1016/j.cell.2018.06.010> PMID: 30033366
56. Shen JP, Zhao D, Sasik R, Luebeck J, Birmingham A, Bojorquez-gomez A, et al. Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. 2017; 14: 573–576. <https://doi.org/10.1038/nmeth.4225> PMID: 28319113
57. Park S, Lehner B. Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. *Mol Syst Biol*. 2015; 11: 824. <https://doi.org/10.15252/msb.20156102> PMID: 26227665
58. Deshpande R, Asiedu MK, Klebig M, Sutor S, Kuzmin E, Nelson J, et al. A Comparative Genomic Approach for Identifying Synthetic Lethal Interactions in Human Cancer. *Cancer Res*. 2013; 73: 6128–6137. <https://doi.org/10.1158/0008-5472.CAN-12-3956> PMID: 23980094
59. Vizeacoumar FJ, Arnold R, Vizeacoumar FS, Chandrashekar M, Buzina A, Young JTF, et al. A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. *Mol Syst Biol*. 2013;9. <https://doi.org/10.1038/msb.2013.54> PMID: 24104479
60. Wildenhain J, Spitzer M, Dolma S, Jarvik N, White R. Data Descriptor: Systematic chemical interaction datasets for prediction of compound synergism. *Sci Data*. 2016; 3: 160095. <https://doi.org/10.1038/sdata.2016.95> PMID: 27874849
61. Chen X, Ren B, Chen M, Wang Q, Zhang L, Yan G. NLLSS: Predicting Synergistic Drug Combinations Based on Semi-supervised Learning. *PLoS Comput Biol*. 2016; 12: e1004975. <https://doi.org/10.1371/journal.pcbi.1004975> PMID: 27415801
62. Iwata H, Sawada R, Mizutani S, Kotera M, Yamanishi Y. Large-Scale Prediction of Beneficial Drug Combinations Using Drug Efficacy and Target Profiles. *J Chem Inf Model*. 2015; 55: 2705–2716. <https://doi.org/10.1021/acs.jcim.5b00444> PMID: 26624799
63. Jeon M, Kim S, Park S, Lee H, Kang J. In silico drug combination discovery for personalized cancer therapy. *BMC Syst Biol*. 2018; 12: 16. <https://doi.org/10.1186/s12918-018-0546-1> PMID: 29560824
64. Li J, Ugalde-morales E, Wen WX, Decker B, Eriksson M, Torstensson A, et al. Differential Burden of Rare and Common Variants on Tumor Characteristics, Survival, and Mode of Detection in Breast Cancer. *Cancer Res*. 2018; 78: 6329–6338. <https://doi.org/10.1158/0008-5472.CAN-18-1018> PMID: 30385609
65. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. *Nucleic Acids Res*. 2016; 44: D1075–D1079. <https://doi.org/10.1093/nar/gkv1075> PMID: 26481350



66. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res.* 2004; 32: D267–D270. <https://doi.org/10.1093/nar/gkh061> PMID: 14681409
67. The UniProt Consortium. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* Oxford University Press; 2017; 45: D158–D169. <https://doi.org/10.1093/nar/gkw1099> PMID: 27899622
68. Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G, Hong EL, et al. YeastMine—an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit. *Database (Oxford)*. 2012; 2012: bar062. <https://doi.org/10.1093/database/bar062> PMID: 22434830
69. Bleakley K, Yamanishi Y. Supervised prediction of drug-target interactions using bipartite local models. *Bioinformatics.* 2009; 25: 2397–2403. <https://doi.org/10.1093/bioinformatics/btp433> PMID: 19605421