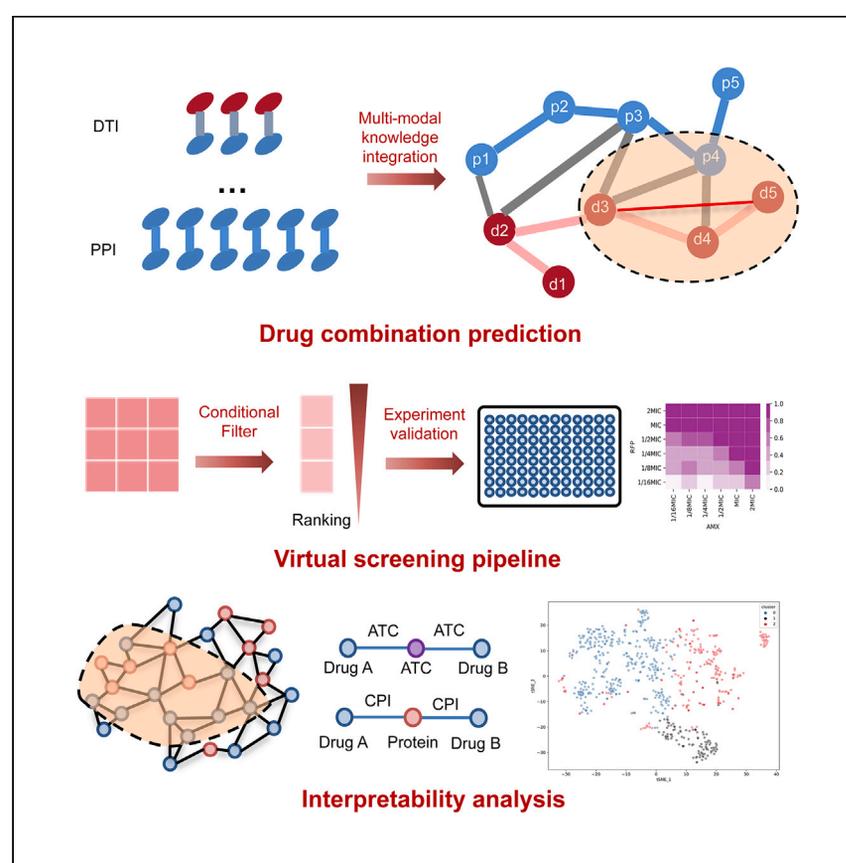


Article

Integrating multi-modal deep learning on knowledge graph for the discovery of synergistic drug combinations against infectious diseases



Ye et al. propose a framework that integrates knowledge graph representation learning and the technique of community discovery to predict drug combination therapies for infectious diseases. By incorporating multi-modal information, the framework enhances predictive performance and hit rates in virtual screening scenarios.

Qing Ye, Ruolan Xu, Dan Li, ..., Shibo He, Chang-Yu Hsieh, Tingjun Hou

kimhsieh@zju.edu.cn (C.-Y.H.)
tingjunhou@zju.edu.cn (T.H.)
cjm@zju.edu.cn (J.C.)

Highlights

A computational framework is proposed for accelerating drug combination discovery

Multi-modal information enhances the prediction of synergistic drug combinations

Community discovery techniques could improve hit rates in virtual screening

The proposed rigorous protocol enables analysis of underlying mechanisms



Article

Integrating multi-modal deep learning on knowledge graph for the discovery of synergistic drug combinations against infectious diseases

Qing Ye,^{1,3} Ruolan Xu,^{1,3} Dan Li,^{1,3} Yu Kang,¹ Yafeng Deng,² Feng Zhu,¹ Jiming Chen,^{1,*} Shibo He,¹ Chang-Yu Hsieh,^{1,*} and Tingjun Hou^{1,4,*}

SUMMARY

The threat to global health posed by unpredictable infections and increasing antimicrobial resistance necessitates the urgent development of drug combination therapies (DCBs) for infectious diseases. Substantial efforts have been devoted to perfecting predictions for DCBs, but data scarcity and poor model interpretability continue to present significant barriers to the development of novel DCBs. To address these issues, here we propose a framework for predicting DCBs by combining knowledge graph representation learning and the technique of community discovery for complex networks. Within this framework, we demonstrate that multi-modal information and multiple types of DCBs could significantly facilitate the predictive performance and improve hit rates in realistic virtual screening scenarios. The high hit rate of 85% for experimental validation strongly supports the proposal that our approach could effectively harness useful information hidden in highly complex biological networks and accelerate *in silico* discovery of pairwise DCBs for infectious diseases and beyond.

INTRODUCTION

A wide adoption of knowledge graph (KG)-based deep learning (DL) methods has yielded significant progress in pharmaceutical research, particularly in drug repurposing and drug-target discovery.^{1–3} By integrating huge amounts of biomedical data, KG enables DL methods to learn latent features from highly complex biomedical networks and thus improve the prediction performance for drug-discovery tasks. The topological structure of KG can also be utilized by network analysis approaches to identify hidden patterns and draw insights that are essential for drug-discovery tasks.^{4–6} In fact, these techniques not only accelerate the discovery of drugs and targets, but combining KG-based DL approaches with network analysis also enables broad generalization and improves model interpretability for various downstream tasks in drug discovery.

In this study, we propose a methodology combining the KG-based DL techniques with network analysis aiming to accelerate drug combination (DCB) discovery against infectious disease. Infections are a dominant contributor to the global health crisis. Since early 2020, the SARS-CoV-2 coronavirus disease 2019 (COVID-19) pandemic has posed a major threat to global health. Antimicrobial resistance, one of the leading public health threats of the 21st century, is another urgent issue that needs to be solved.^{7,8} With such issues in mind, development of new therapeutic treatments against infectious diseases are urgently needed. However, the

¹Zhejiang University, Hangzhou 310058, Zhejiang, China

²CarbonSilicon AI Technology Co., Ltd, Hangzhou 310018, Zhejiang, China

³These authors contributed equally

⁴Lead contact

*Correspondence: kimhsieh@zju.edu.cn (C.-Y.H.), tingjunhou@zju.edu.cn (T.H.), cjm@zju.edu.cn (J.C.)

<https://doi.org/10.1016/j.xcrp.2023.101520>



process of developing clinically trialed new drugs is still costly, complex, and time consuming.⁹ Hence, DCBs constitute an important and timely therapeutic strategy for repurposing existing drugs during pandemics and overcoming antimicrobial resistance. Synergistic DCBs of two or more drugs with distinct underlying mechanisms for improving clinical outcomes provide multiple advantages, such as improving efficacy, reducing toxicity, overcoming drug resistance, and repurposing existing drugs for monotherapy.^{10,11} Conventionally, trials of effective DCBs are prioritized by experts and then validated by biological experiments under the setting of high-throughput screening.^{12–14} However, owing to the exceedingly large combinatorial space of DCBs, exhaustive experimental trials are prohibitive, with limited resources and time. Therefore, it is necessary to develop cost-effective and mechanism-driven strategies to identify effective DCBs quickly.

In silico machine learning (ML) methods have become a viable alternative to experimental screening and have been increasingly applied to predict and investigate novel DCBs. Generally, computational prediction of DCBs can be classified into three categories: fingerprint-based (FP-based) approaches, molecular graph-based (MG-based) approaches, and knowledge graph-based (KG-based) approaches.¹⁵ FP-based approaches only adopt molecular fingerprints derived from the chemical structures of drugs as the model input,^{16–19} and MG-based approaches utilize the chemical structures as the model input directly and learn an embedding of DCBs via DL.^{20,21} However, the predictive ability of these two types of methods is often limited by insufficient data of DCBs and the lack of interpretability. With an explosive growth of multi-omics data, KG-based methods have become more popular as they make more holistic predictions by taking into account auxiliary facts conveyed by data from various sources. Traditionally, the multi-omics data and chemical structures of drugs are encoded by some handcrafted rules (e.g., similarity matrices, one-hot encoding) as the feature representation for classification modeling based on, for example, support vector machine (SVM) or random forest (RF).^{22–25} Recently, DL approaches based on multi-omics data have been applied for the prediction of synergistic combinations. For example, Jin et al. proposed a DL framework named ComboNet to utilize networks of drug-target interactions (DTIs) and target-disease associations, which enabled an effective *in silico* search for synergistic combinations against SARS-CoV-2.²¹ Despite much effort having been devoted to extracting effective information from KG, several challenges remain in these KG-based methods. First, the predictive performance of existing approaches is still not satisfactory. The traditional methods depend on handcrafted features to utilize omics data, which limits their ability to capture complex patterns from KG. As a result, most approaches only utilized disease-related DCBs in realistic screening scenarios, while our experiments demonstrate that different types of DCBs can further enhance the prediction by combining KG-based DL framework and network analysis. Second, DL approaches such as ComboNet face the problem of adaptability, since the modules in a complex DL-based framework designed for specific data such as DTIs are not easily transferable to other types of data such as biological pathways. Third, most existing DL algorithms lack transparent interpretability to elucidate the underlying mechanisms for effective DCBs, which is a crucial transition step from computational modeling to scientific insights, experimental validation, and acceleration of DCB discovery.²⁶

In response to these challenges, we introduced KG-CombPred (KG-based DCB prediction) as an end-to-end network embedding framework for drug pair scoring, integrating large-scale omics data (e.g., ATC codes of drugs, gene ontology) with KG to overcome the limitation of data for DCB and combining graph representation learning to improve predictability and interpretability. The performance assessment

of KG-CombPred on two general datasets and three disease-specific datasets clearly demonstrates that our method outperformed the reported baseline methods. In addition, we develop a feasible framework combining KG-CombPred with community discovery (CD), in which the disease is defined as a community on KG, to facilitate in the virtual screening (VS) of DCBs against specific diseases. We devised a VS scenario for COVID's DCB and revealed that by combining KG-CombPred and CD, the precision rates of VS were significantly increased by 38%. We then experimentally validated the framework on tuberculosis (TB) DCB screening, and the experimental results clearly show that KG-CombPred-VS achieved an 85% success rate (17 hits from 20 candidates). Furthermore, we implemented several interpretable analysis schemes based on KG-CombPred to promote the exploration of DCB mechanisms.

RESULTS

Drug combination prediction with KG-based deep learning

Leveraging multi-modal data, such as chemical and biological information, is essential for DCB prediction in order to improve prediction performance and enhance interpretability. In this study, we introduce KG-CombPred, a general framework for attributed multiplex heterogeneous network representation learning, aiming to explore the potential of multi-modal KG in DCB prediction. The KG-CombPred architecture consists of two components (Figure 1). The first component is a hierarchical random walk to generate a corpus, which is a collection of node sequences representing the complex connectivity patterns within the KG. The second component, called heterogeneous skip-gram, focuses on learning the representation for each node in the graph based on the generated corpus. Specifically, the embedding of each node is designed to have two parts: the base embedding and the edge embedding. The base embedding is shared across different edge types and represents the inherent properties of the node that are independent of the specific relationship type. The edge embedding, on the other hand, is specific to each edge type and is aggregated from the embeddings of neighboring nodes (defined for that particular edge type). It captures the contextual information of the node with respect to its neighboring nodes and the specific relationship type. By combining the base embedding and the aggregated edge embeddings, a comprehensive representation for each node on each edge type is obtained. Compared with traditional knowledge graph embedding (KGE) methods, KG-CombPred (see [experimental procedures](#)) models information from different perspectives by using both graph structure information and node attributes with an attention mechanism. To illustrate the wide applicability and performance of KG-CombPred, comparative evaluation was performed on two general datasets and three disease-specific datasets (see [experimental procedures](#); [Tables S1](#) and [S2](#)). The general datasets contain DCBs for all kinds of diseases, whereas the disease-specific datasets are evaluated on DCBs for a specific disease. For a comprehensive evaluation, we compared the performance of KG-CombPred with four standard feature-based ML algorithms (SVM, RF, deep neural network [DNN], and graph convolutional network [GCN]) and two mainstream KGE approaches (i.e., DistMult and ComplEx). These baseline approaches behave robustly and have been widely employed in the scenarios of *in silico* drug discovery including DCB prediction.^{16,17,21,27,28} [Table S3](#) provides a comprehensive summary of these methods. For general datasets, the results ([Figure 2](#)) indicate that the KG-based approaches (DistMult, ComplEx, and KG-CombPred) are obviously superior to the DCB method that employs solely chemical information, where the KG-CombPred yields the best performance. Compared to the highest-performing DCB-based approaches, KG-CombPred improved areas under the receiver-operating characteristic curve (ROC-AUC) and precision-recall curve

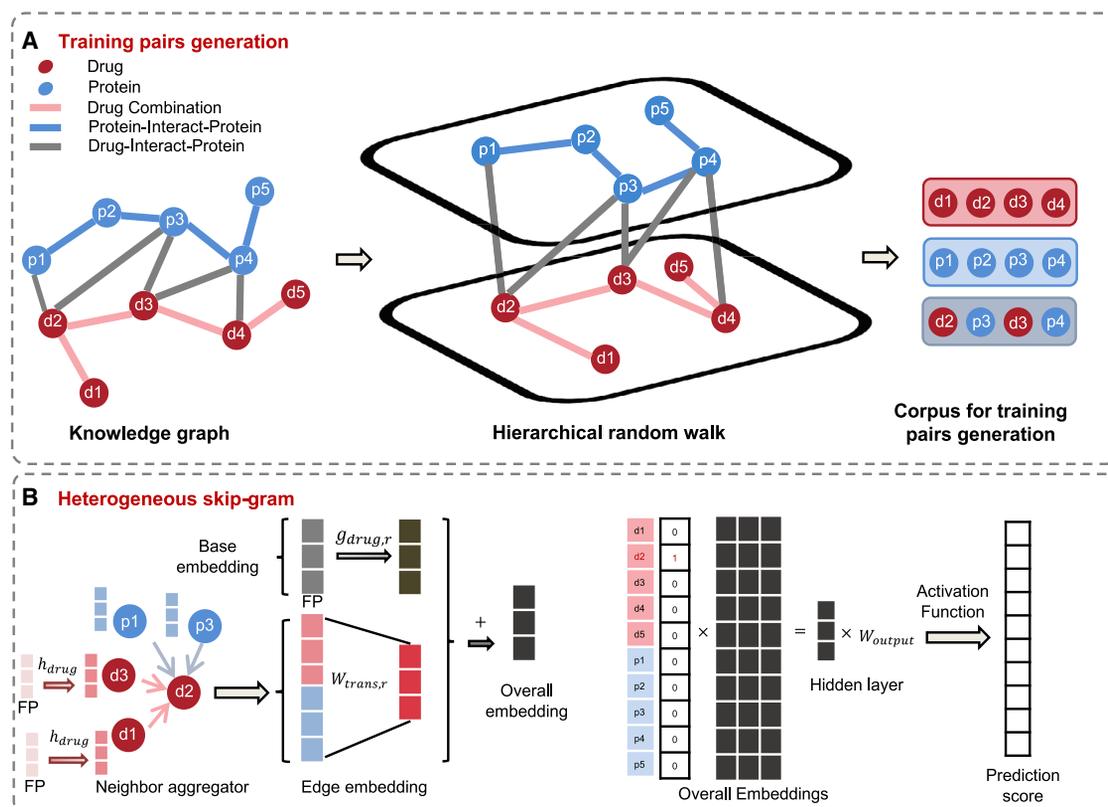


Figure 1. Drug combination scoring via KG-CombPred

(A) Training pairs generation. The KG is composed of drugs (red nodes) and proteins (green nodes), including drug combinations (red lines), protein-protein interactions (green lines), and DTIs (blue lines). Hierarchical random walk performs random walk operations on each type of edge separately, thus generating a corpus for node representation learning.

(B) Heterogeneous skip-gram. The overall embedding of a certain node v_i on each edge type r can be split into two parts: base embedding and edge embedding. The base embedding is shared between different edge types, and the edge embedding is aggregated from neighbors' edge embeddings. For example, the node of a drug owns two types of edges (drug combinations and DTIs) and the corresponding edge embeddings with the same dimension. These two types of edge embedding are concentrated and reshape the dimension by a trainable matrix $W_{trans,r}$. It should be noted that the initial base embedding and edge embedding for the nodes of drugs are defined as a parameterized function of molecular fingerprint (FP), while the other types of nodes are randomly initialized.

(PR-AUC) by 15% and 14% and 6% and 12% on dataset I and dataset II, respectively. KG-CombPred also outperformed popular KGE approaches, boosting ROC-AUC by 6%–10% and PR-AUC by 5%–7%. The superiority of KG-CombPred manifests the benefits of utilizing auxiliary information in KG for predicting potentially useful DCBs.

In realistic scenarios, DCB prediction for specific diseases is more prevalent. However, there is a shortage of DCBs for a specific disease. As a result, the predictive performance of DCB-based methods on such training data is restricted. As shown in Figure 2, the prediction performance of DCB-based models was poor. Yet we discovered that prediction accuracy can be substantially improved by augmenting prior data (i.e., KG). For the COVID and TB datasets, KG-CombPred exhibited the highest performance, with ROC-AUC and PR-AUC improved by 8% and 13% and 4% and 9%, respectively, when compared to the best-performing DCB-based approaches. For the HIV dataset, DistMult demonstrated the best performance, with ROC-AUC and PR-AUC enhanced by 5% compared to the top-performing Mol-GCN approach among DCB-based methods. To further facilitate DCB prediction, we collected various types of DCB data from the Continuous Drug Combination Database for disease-specific

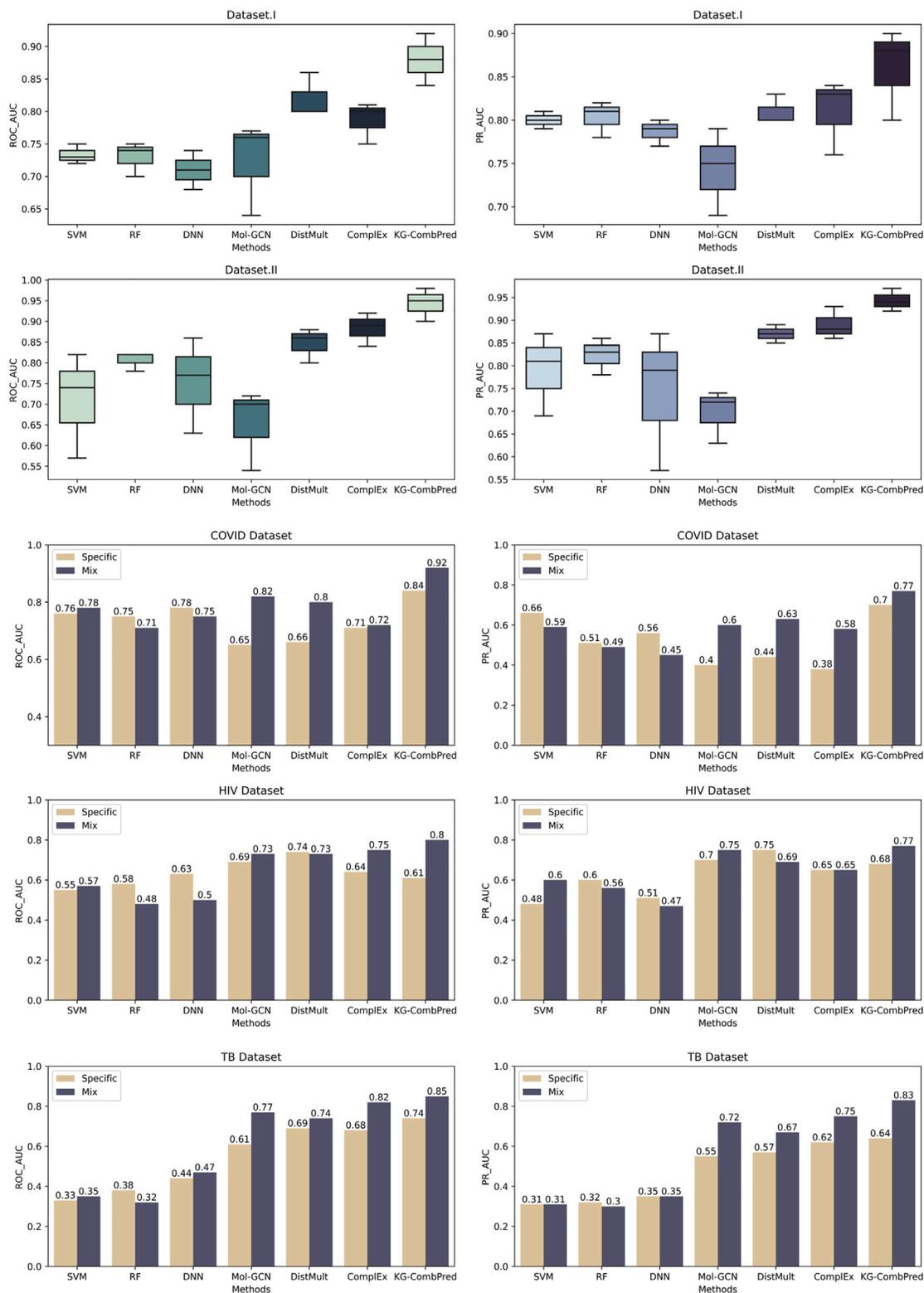


Figure 2. Prediction performance on the general datasets

n = 3 independent experiments. Box plots show the median as the center lines, upper and lower quartiles as the box limits, and whiskers as the maximum and minimum values, while dots represent outliers. "Specific" indicates that only disease-specific drug combinations (DCBs) are utilized for training, whereas "Mix" indicates that all DCBs are used for training.

datasets.²⁹ The results indicate that the prediction performance of KG-based methods is enhanced by the addition of other types of DCB data, particularly our proposed KG-CombPred, which improved prediction performance by more than 10% across all three datasets. When it comes to DCB-based methods, however, the added data almost never lead to better prediction results and can even lead to a slight decline in some datasets. This could be because DCB-based methods rely solely on chemical information for training, which, due to the complexity of DCB mechanism of action, cannot be fully represented by chemical information, and the addition of other types of DCBs instead confuses the model. However, for KG-based techniques, information such as DTI and protein-protein interaction (PPI) facilitate the model in scoring DCBs from the perspective of the mechanism of action, hence enhancing the prediction performance of disease-specific DCBs.

KG-CombPred combined with community discovery for virtual screening

In addition to determining whether there is a synergistic effect between two treatments, linking a DCB to a specific disease is an even higher barrier for the VS of DCBs. Owing to the polypharmacology of drugs and the complex effects of DCBs on biological systems, a DCB commonly has a synergistic effect on multiple diseases. However, the current DCB prediction methods do not take disease-related information into account, which can be a barrier for VS in real-world scenarios. In this aspect of characterizing diseases, the KG's wealth of multi-omics information is naturally advantageous. In the KG, diseases are represented as a community consisting of functional proteins, medicines, gene phenotypes, and so forth that are associated with the disease (Figure 3A). Thus, DCBs for a certain disease can be confined to the linkages between drug nodes contained inside the disease community. Yet it is extremely difficult to identify a comprehensive disease community based purely on expert knowledge, given that the functional proteins, medicines, and so forth gathered through literature survey and data mining are incomplete. In this study, we propose a framework to enable the discovery of disease communities in the KG by combining the community detection method OSSE and KG-CombPred, thus achieving a new state-of-the-art success rate for VS of DCBs for a specific disease.³⁰ By leveraging expert knowledge, we can identify some initial disease seed nodes. Considering that there may be overlaps between different diseases in terms of, e.g., drugs and proteins, the problem of disease community detection is defined as an overlapping community detection problem based on the initial seed nodes. OSSE expands these seeds using the personalized PageRank.³¹ The PageRank algorithm measures the importance of each node within the graph in consideration of their neighbors, connections, their importance, and the influence each one of those neighbors has. As a result, we obtain an extended disease community, which may not be entirely accurate but is adequate for narrowing the VS scope.

To assess the impact of disease community detection on VS, we performed VS in a real-world situation for COVID-19 DCBs. By combining 36 COVID-related compounds with 7,892 drugs in the KG, 28,344 pairs of candidate DCBs were generated, including 48 COVID DCBs already reported in the literature. All DCBs were ranked based on the scores provided by KG-CombPred. Next, the prediction scores for DCBs that were not categorized into the COVID group indicated by the disease discrimination model were lowered, resulting in a decrease in their rankings. For comparison, we also trained a DCB-based disease discrimination model, which used the RF algorithm based on molecular fingerprints to perform binary classification of COVID-related DCBs and other types of DCBs. We then calculated the ranking metrics of VS results for KG-CombPred, KG-CombPred with RF, and KG-CombPred with CD. As demonstrated in Figure 3B, when the RF disease

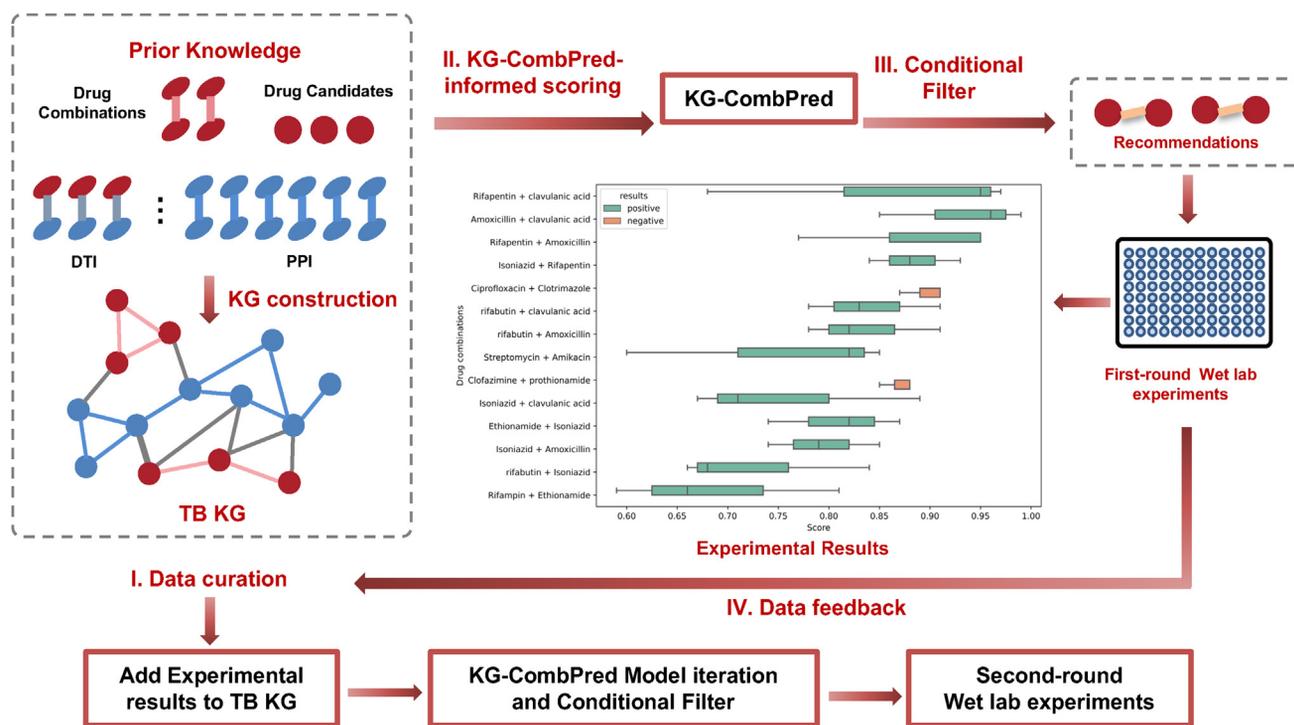


Figure 4. Discovery of drug combinations against TB via KG-CombPred-VS

The strategy comprises four steps. (I) Data curation. Publicly available data are collected and cleaned for disease-specific KG construction and virtual screening (VS), including reported drug combinations, prior knowledge (e.g., DTI, PPI), and promising drug candidates for potentially new pairwise drug combinations. (II) KG-CombPred-informed scoring for combination ranking. Curated TB KG is used to train the KG-CombPred model three times with different initializations. (III) Conditional filter for ranking score optimization and recommendations. Three modules including chemical similarity, community discovery, and model prediction uncertainty are applied to filter predicted results and thus generate reliable recommendations. (IV) Feedback from wet-lab experiments to refine KG for model iteration. With the updated TB KG by adding the results from the experiments, the steps II and III are repeated to generate new recommendations for the next round of wet-lab experiments.

was designed to generate accurate and reliable recommendations. The conditional filter could factor in multi-view information, providing an expert customizable module for recommendations. Inspired by the rationale for DCBs given by Jia et al.,³³ pharmacodynamically synergistic DCBs could be attributed to their complementary actions. Naturally, the chemical similarity between the two compounds of a TB combination toward two different targets could be hypothesized to be low. To validate our hypothesis, we calculated the chemical similarities for the collected 58 TB DCBs. The results indicate that the chemical similarities for all combinations are under 0.4, suggesting that the similarity filter with a threshold of 0.4 is reasonable (Figure S1). After the similarity filtering, 874 DCBs remained for further analyses. Uncertainty estimation of predictions is another crucial step toward providing reliable recommendations as well as an informative measure that can be exploited an algorithm to maintain the diversity of predictions. Deep ensembles, a widely applied method in drug discovery,^{21,34} was adopted to give uncertainty estimation as the standard deviation of the predictive distribution in this study. Acquiring the true label on a drug pair with a high uncertainty for its predictions is informative (referred to as exploration) and helps us to build an accurate model in a data-efficient manner. We also wish to experimentally evaluate the predictions of DCBs that are expected to give high DCB scores (referred to as exploitation), such that we can find effective DCBs to fulfill the main purpose of this project. To achieve a balance between exploration and exploitation, acquisition function (α) is defined as the function of upper confidence bound (UCB): $\alpha(x) = \hat{\mu}(x) + \kappa\hat{\sigma}(x)$, where $\hat{\mu}(x)$ represents the estimated

mean of the predictive distribution and $\hat{\sigma}(x)$ represents the estimated uncertainty. Here, κ is set to 1. For more detailed information we refer readers to the work by Bertin et al.,³⁴ which provides a comprehensive introduction and evaluation about uncertainty estimation. Thus, to maximize ranking scores, three modules are implemented: chemical similarity, CD, and model prediction uncertainty. Finally, 31 DCBs were recommended as the potentially effective synergy in the first round of VS (Round I) (Table S4 and Figure S5). Considering the conditions of the experiment and the availability of the compounds in our lab, 14 TB DCB candidates were sent to assays against *M. smegmatis in vitro*, in which 12 DCBs were validated successfully in the first round of the wet-lab experiments. The synergy of these combinations was assessed by the fractional inhibitory concentration³⁵ (see experimental procedures). Among these, seven drug pairs are novel (i.e., have not been previously reported), indicating that KG-CombPred-VS could enhance the existing DCB searching for data enrichment and enables the discovery of novel synergistic anti-TB DCBs. In particular, five DCBs have been previously reported to exhibit anti-TB synergistic effects,^{36–40} suggesting that KG-CombPred-VS could facilitate the process of data collection when few data are initially available. After this, we reproduced KG-CombPred-VS utilizing the KG that had been enriched with newly discovered DCBs. In this round, six DCB candidates were sent to the assays *in vitro*, in which five DCBs were validated successfully in the first round of the wet-lab experiments (Table S5 and Figure S5). In the second round of VS, we additionally conducted a cytotoxic assay for two highly effective DCBs to demonstrate that the observed synergy of these combinations is not solely attributable to their high cytotoxicity (Figure S6). Among these, two drug pairs with anti-TB synergistic effects^{41,42} have been previously reported and three drug pairs are novel. Since previous studies have shown that DCBs tend to cause antagonisms instead of forming synergies,^{43,44} the number of false positives could be high in the ranked list of recommendations. However, by integrating KG-CombPred, CD, and expertly calibrated conditional filters, our wet-lab experiments demonstrate that the DCB VS has a highly successful hit rate of 85%.

Interpretability of KG-CombPred

Having evaluated the predictive power and effectiveness of the proposed strategy, we turned to the mechanism interpretation. The network-based analysis and visualization of the drug embeddings for DCBs were combined to justify the effectiveness and illustrate the interpretability of KG-CombPred. First, we selected a few drug pairs (with known mechanisms) in the Food and Drug Administration (FDA) dataset to investigate whether we can infer the mechanisms from the KG-CombPred framework. To this end, we visually inspected the distribution of the DCBs in the network of DTIs. As shown in Figure 5A, the relations of DCBs tend to link nodes with a relative long separation distance in the DTI network, where the corresponding linked protein nodes serve different biological functions. For example, sunitinib, a receptor tyrosine kinase inhibitor used for the treatment of renal cell carcinoma, could be used in combination with hydrochlorothiazide for treating hypertension caused by sunitinib. This finding indicates that the two drugs in combination tend to have a complementary exposure relationship to DTIs. This is one specific example to corroborate that KG-CombPred along with properly compiled KG may capture the underlying mechanism of DCBs. For a more systematic and large-scale analysis, we then conducted a clustering-based investigation. Typically, synergistic DCBs could be divided into three groups: anti-counteractive actions, complementary actions, and facilitating actions.³³ Here, we represented the visualization of the embeddings of the DCBs for the FDA dataset via t-distributed stochastic neighbor embedding (t-SNE) and classified these DCBs into three groups by *k*-means clustering (Figure 5B and Note S1). Through literature

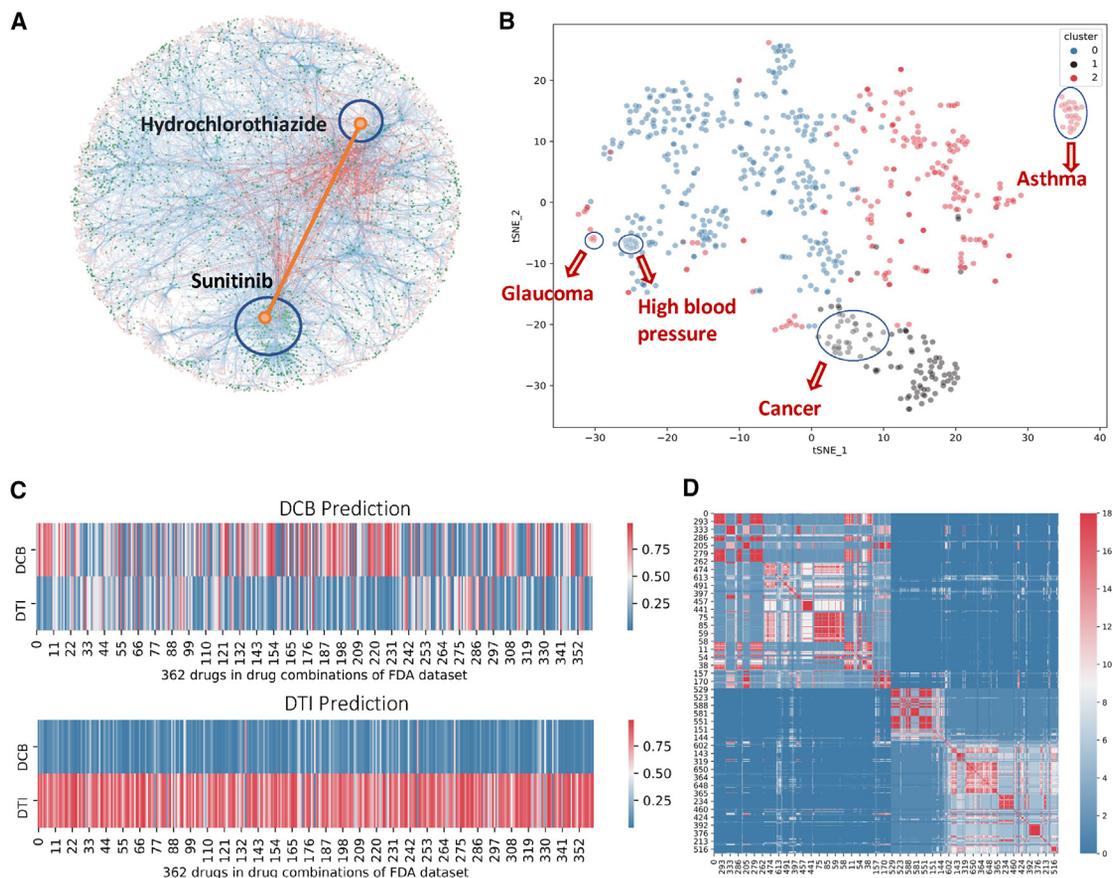


Figure 5. KG-CombPred learns mechanisms of drug combinations

(A) Visualization of the FDA dataset, where the pink nodes represent the drugs and the green nodes represent the proteins. The blue lines represent the DTIs and the red lines represent the drug combinations. (B) Visualization of the DCBs in the FDA dataset via t-SNE, where the color of the nodes represents the corresponding cluster by k-means clustering. (C) Heatmaps of the drugs' attention scores on edge type of DCB and DTI in the FDA dataset used for the DCB and DTI training pairs, respectively. (D) Heatmap of the mechanism similarities for the DCBs in the FDA dataset.

identification, we found that the clustered DCBs (circled on Figure 5B) were used to treat the same diseases, implying that KG-CombPred learned the pharmacological patterns of DCBs (Table S6). For example, DCBs for the treatment of asthma and glaucoma all belong to cluster 2, and their mechanisms could be attributed to facilitating actions.^{45–49} Many DCBs in cluster 1 are for cancer treatments, and most of these combinations are known to work via the anti-counteractive principle, implying that the two drugs in a pair act on different targets in related or cross-talking pathways.^{50,51} Many DCBs in cluster 0 are for high blood pressure and their combination mechanisms could be attributed to complementary actions, suggesting that the targets regulated by the two drugs in a combination are involved in multiple points of a pathway.^{52–54} Given the complicated modes of DCBs, we then further calculated the mode similarity (see experimental procedures) to describe the pharmacological patterns for more precise clustering (Figure 5D). Through this approach, the mechanisms of unexplored/predicted DCBs could be inferred by comparing the mode similarity with well-explored DCBs. Besides interpreting KG-CombPred from a perspective of pharmacology, we next made a model-based interpretation. Here, we visualized the attention score of each drug in the DCBs of the FDA dataset for DCB prediction and DTI prediction, respectively (Figure 5C). For the DTI prediction, the node of a drug mainly pays

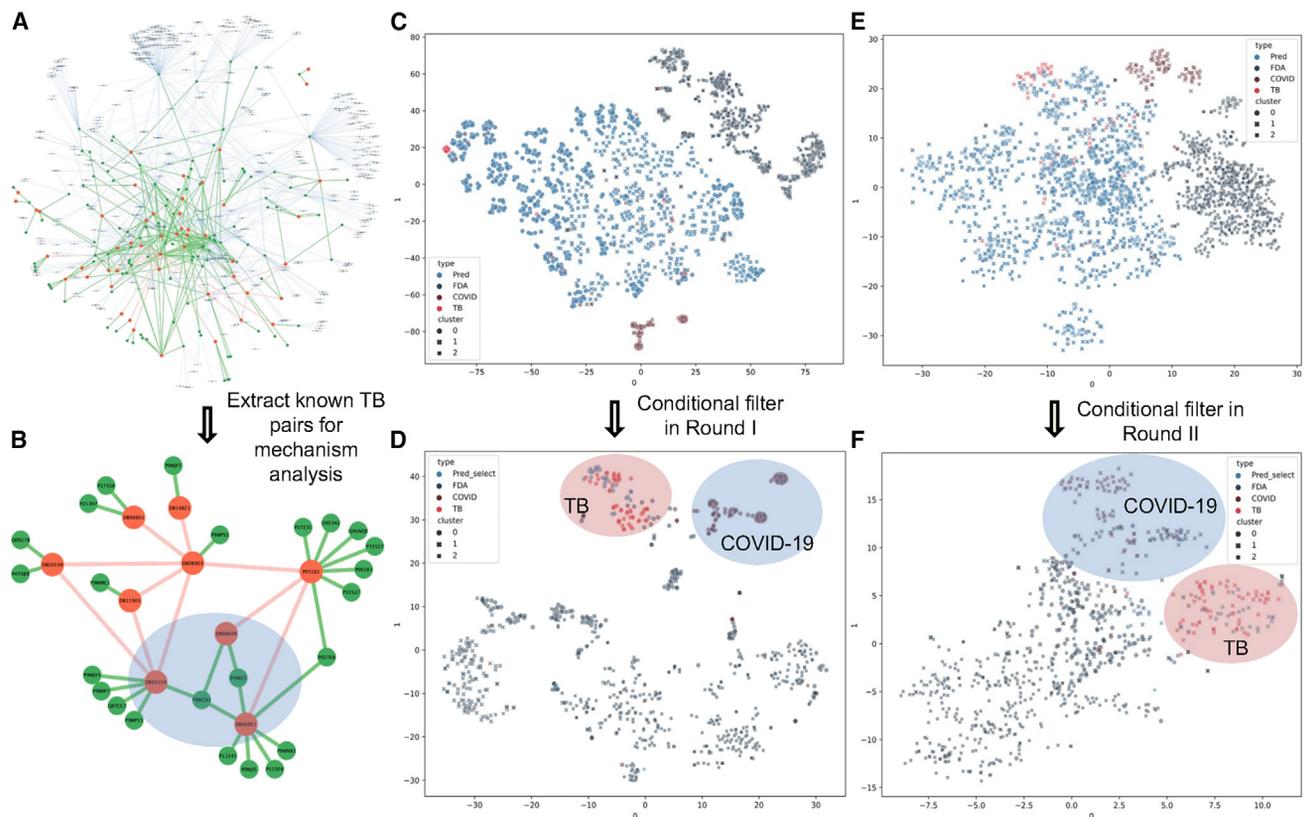


Figure 6. KG-CombPred-VS generates reliable recommendations for TB drug combinations

(A and B) Visualization of the TB DCBs and corresponding DTIs, where the red nodes represent the drugs and the green nodes represent the proteins. The green lines represent the DTIs and the red lines represent the DCBs.

(C) Visualization of the DCBs including the FDA, COVID-19, TB, and all TB candidates by the model trained in round I.

(D) Visualization of the DCBs including the FDA, COVID-19, TB, and selected TB candidates by KG-CombPred-VS in round I.

(E) Visualization of the DCBs of the FDA, COVID-19, updated TB, and all TB candidates by KG-CombPred-VS in round II.

(F) Visualization of the DCBs of the FDA, COVID-19, updated TB, and selected TB candidates by KG-CombPred-VS in round II.

The color of the nodes represents the type, and the marker of the node represents the corresponding cluster by *k*-means clustering.

attention to the message delivered from the related DTIs while for the DCB prediction, the node of a drug aggregates the information from both DTIs and DCBs with roughly the same contributions. In other words, the attention scores of DTIs and DCBs almost contribute equally to our model. This phenomenon could be attributed to the larger data size of DTIs (~15,000 pairs) compared with DCBs (768 pairs) in the compiled KG. Overall, these analyses indicate that the DCB prediction could benefit from KG and thus facilitate the mechanism exploration of DCBs via the attention scores of DTIs.

In the second part, we applied the analysis tools displayed in the first part on TB DCBs as an example to demonstrate the effectiveness of KG-CombPred-VS in realistic VS scenarios for DCB discovery. To explore the mechanism of TB DCBs, we visualized the relationships between drugs and drug targets and extracted several known TB DCBs for mechanism analysis. As convincingly shown in Figure 6B, most drugs in the combinations act on different targets. To uncover further details of the mechanisms for the newly discovered TB combinations, we applied mode similarity to investigate the combination of ethionamide and rifapentine, found in round II of the wet-lab experiment, as an example (Figure S2). The mode-similarity analysis suggests that the combination of ethionamide and rifapentine follows the same working principle for the combination of rifampicin and isoniazid, which can be

attributed to making effects by both targeting RNA polymerase and mycolic acid. Furthermore, to demonstrate the effectiveness of KG-CombPred-VS, we visualized the distribution of the DCBs of different types in the two rounds of VS. Based on conditional filtering, the selected DCBs are clustered in the area close to the known TB DCBs (Figures 6C and 6D). In addition, the visualizations (Figures 6E and 6F) also indicate that KG-CombPred could clearly distinguish three different types of DCBs.

DISCUSSION

For the treatment of emerging epidemic infections, DCB therapy could utilize “old drugs” to enable a more rapid, more effective, and less toxic therapeutic strategy. For instance, paxlovid, the first oral medication authorized for COVID-19 infection therapy, is a combination of nirmatrelvir and ritonavir that has previously been used in HIV combination therapy. Despite the fact that various computational approaches have been employed for DCB predictions, they are largely ineffective on unseen data, in part owing to the limited numbers of experimentally validated data available for model training. In this study, we present a KG-based framework that combines the deep graph learning model KG-CombPred and the technique of CD OSSE to predict pairwise DCBs. The predictive capability of KG-CombPred has been thoroughly validated on five benchmarks, and our approach outperformed six baselines by 6%–15% on ROC-AUC. With our approach, multidisciplinary information and multiple types of DCBs could be leveraged to strengthen the predictive performance of specific diseases, whereas this is not feasible with DCB-based methods (i.e., RF and Mol-GCN). More importantly, we utilize CD in an innovative way to represent diseases in graphs, thereby establishing connections between diseases and DCBs and increasing the hit rates of VS by 38%. The high successful hit rate of 85% in two-round VS assessed by the wet-lab experiments demonstrates that our KG-based framework provides reliable and novel recommendations of synergistic drug pairs *in vitro* and may potentially accelerate the discovery of effective DCBs in realistic drug-development projects. Furthermore, to make predictions more reliable and trustworthy, a rigorous protocol for analysis has been proposed to investigate the underlying mechanisms of DCBs against infectious disease. We illuminate that the predictions of different types benefit from different prior information and that DCB prediction is greatly facilitated by not only DCB data but also DTI data. Moreover, we emphasize that the mode similarity derived by KG-CombPred is a promising approach to decode the mechanism of DCBs. Despite providing various advantages for DCB prediction, KG-CombPred has several constraints. In comparison with DCB-based methods, KG-based approaches could integrate prior knowledge to facilitate prediction. However, prior knowledge must be transformed into a format consistent with KG entities, and this pre-processing step is highly dependent on expert knowledge. In addition, the cold-start problem is a hurdle for KG-based approaches, meaning that making inference on a completely new compound with no known correlations (in the KG) is inapplicable. Consequently, it should be noted that KG connectivity is essential for CD. Irrespective of these aforementioned concerns, KG-CombPred satisfies the urgent need of DCB prediction for infectious disease and represents a novel approach to integrating prior biological information with graph representation learning to accelerate the discovery of infectious DCBs that can further inform combination therapy.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests should be directed to the lead contact (T. Hou) at tingjunhou@zju.edu.cn.

Materials availability

This study did not generate new unique reagents.

Data and code availability

The source data and codes of KG-CombPred are available on the Zenodo repository at <https://zenodo.org/record/7692097>.

Dataset curation

In this study, we focus on anti-infectious pairwise DCBs. To evaluate the predictive performance of KG-CombPred, we constructed six DCB benchmarks with the corresponding knowledge graphs. Typically, the knowledge graph could consist of DTIs, drug-DCBs, and protein-protein interactions, in which nodes can be drugs or proteins and edges are indicators for the interactions or similarities between the connected nodes. In this study, all drugs are represented as DrugBank ID or CID of PubChem, and all proteins are represented as Uniprot ID to align entities in KG.

General dataset I, FDA

The FDA dataset includes 681 pairwise DCBs²⁴ collected from the clinical data in the Drug Combination Database (DCDB v2.0), Therapeutic Target Database (TTD v5.1.01),⁵⁵ and FDA Electronic Orange Book. For prior information enrichment, 217,160 human protein-protein interactions and 15,051 DTIs were collected from multiple data sources.

General dataset II, DCGCR

The DCGCR dataset²⁵ includes 173 pairwise DCBs from DCDB v2.0. To incorporate drug-target network patterns and pharmacological patterns, 449 drug-protein interactions and 139 ATC codes were collected from DrugBank.

Disease-specific dataset I, COVID-19

The COVID-19 dataset²¹ includes 179 SARS-CoV-2 from NCATS Open-Data Portal, Riva et al.,⁵⁶ Bobrowski et al.,¹⁹ and Tan et al.⁵⁷

Disease-specific dataset II, HIV

The HIV dataset includes 114 HIV DCBs collected from Gordon et al.⁵⁸ and ChEMBL.

Disease-specific dataset III, TB

The TB dataset contains 58 TB DCBs, 28,052 DTIs, and 44 TB single-agent data. Specifically, the TB DCBs were collected from seven publications.^{35,59–64}

The detailed descriptions of KG for disease-specific datasets are listed in [Table S2](#). For general datasets, negative samples are generated pairwise by drugs in the dataset and sampled with the same number as positives. To minimize the impact of data variability on the results, 3-fold cross-validation was used to compare the predictive performances of our method and other state-of-the-art methods. Because of the scarcity of DCB data, we randomly divided each disease-specific dataset into training and test sets in a ratio of 4:1.

KG-CombPred architecture

We formulate the problem of pairwise DCB prediction as a classification task. Given a pair of drugs *a* and *b* with chemical structures and prior information, we aim to predict their probability of being a DCB with synergistic effect. The KG-CombPred architecture ([Figure 1](#)) adapted from GATNE⁶⁵ can be decomposed into two modules: training pairs generation and heterogeneous skip-gram.

Training pairs generation is illustrated in Figure 1A. Given a heterogeneous network $G = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, each node $v_i \in \mathcal{V}$ is associated with some types of feature vectors and each edge is denoted as $e_{ij}^{(r)} \in \mathcal{E}$, where r corresponds to a certain edge type/relation. $\mathcal{A} = \{x_i | v_i \in \mathcal{V}\}$ is the set of node features for all nodes, where x_i is the associated node feature of node v_i . The pair generation proceeds as follows. First, hierarchical random walk performs random walk on each type of edge separately and thus generates a corpus from KG for training pair generation. Specifically, given a view r on the network, i.e., $G_r = (\mathcal{V}, \mathcal{E}_r, \mathcal{A})$ and a meta-path scheme $\mathcal{T} : \mathcal{V}_1 \rightarrow \mathcal{V}_2 \rightarrow \dots \rightarrow \mathcal{V}_t \rightarrow \dots \rightarrow \mathcal{V}_l$, where l is the length of the meta-path scheme, the transition probability at step t is defined as follows:

$$p(v_i | v_j, \mathcal{T}) = \begin{cases} \frac{1}{|N_{i,r} \cap \mathcal{V}_{t+1}|} & (v_i, v_j) \in \mathcal{E}_r, v_j \in \mathcal{V}_{t+1}, \\ 0 & (v_i, v_j) \in \mathcal{E}_r, v_j \notin \mathcal{V}_{t+1}, \text{ where } v_i \in \mathcal{V} \text{ and } N_{i,r} \text{ denotes the} \\ 0 & (v_i, v_j) \notin \mathcal{E}_r, \end{cases}$$

neighborhood of node v_i on edge type r . For example, with applying KG containing three types of data including DCBs, DTIs and PPIs shown in Figure 1A, the corpus could be $\{d1, d2, d3, d4\}$ for type of DCB, $\{p1, p2, p3, p4\}$ for type of PPI, and $\{d2, p3, d3, p4\}$ for type of DTI. The strategy of hierarchical random walk ensures that the semantic relationships between different types of nodes can be properly incorporated into skip-gram model.⁶⁶

The heterogeneous skip-gram is depicted in Figure 1B. Generally, the overall embedding of a certain node v_i on each edge type r could be split into two parts: base embedding and edge embedding. The base embedding of node v_i is randomly initialized and shared between different edge types. The edge embedding is aggregated from neighbors' edge embeddings. We denote the edge embedding as $u_{i,r}$, and concatenate all the edge embeddings for node v_i as U_i with size s -by- m , where s is the dimension of edge embeddings and m is the number of edge types:

$$U_i = (U_{i,1}, U_{i,2}, \dots, U_{i,m}). \quad (\text{Equation 1})$$

To incorporate the chemical structures of drugs into representation learning, the initial base embedding for the nodes of drugs is defined as a parameterized function of molecular fingerprint (FP) as

$$b_i = h_{\text{drug}}(x_i), \quad (\text{Equation 2})$$

where h_{drug} is a transformation function and x_i refers to the FP. Similarly, the initial edge embeddings $u_{i,r}^{(0)}$ for the nodes of drugs on edge type r is also parameterized as the function of FP as

$$u_{i,r}^{(0)} = g_{\text{drug},r}(x_i), \quad (\text{Equation 3})$$

where $g_{\text{drug},r}$ is also a transformation function that transforms the feature to an edge embedding of drug v_i on each edge type r and x_i refers to the FP, while the base embedding and edge embedding of other types of nodes are randomly initialized. Additionally, a self-attention mechanism is used to compute the coefficients $a_{i,r}$ of linear combination of vectors in U_i on edge type r as

$$a_{i,r} = \text{softmax}(w_r^T \tanh(W_r U_i))^T, \quad (\text{Equation 4})$$

where w_r and W_r are the trainable parameters for edge type r with size d_a and $d_a \times s$, respectively. Thus, the overall embedding of node v_i for edge type r is

$$v_{i,r} = b_i + \alpha_r M_r^T U_i a_{i,r}, \quad (\text{Equation 5})$$

where b_i is the base embedding for node v_i , α_r is a hyper-parameter denoting the importance of edge embeddings toward the overall embedding, and M_r is a

trainable transformation matrix. In this way, KG-CombPred could be applied to unobserved data and thus is suitable for realistic applications (e.g., new drugs/compounds for VS of DCBs).

To model the heterogeneous neighborhood of a node, the heterogeneous skip-gram model⁶⁶ is introduced by maximizing the probability of having the heterogeneous context for node v_i . Supposing that the random walk with length l on edge type r follows a path $P = (v_{p1}, \dots, v_{pl})$ such that $(v_{pt}, v_{pt}) \in \mathcal{E}_t (t = 2 \dots l)$, denote v_{pt} 's context as

$$C = \{v_{pk} | v_{pk} \in P, |k - t| \leq c, t \neq k\}, \quad (\text{Equation 6})$$

where c is the radius of the window size. Thus, given a node v_i with its context C of a path, the objective is to minimize the following negative log likelihood:

$$-\log P_\theta(\{v_j | v_j \in C | v_i\}) = \sum_{v_j \in C} -\log P_\theta(v_j | v_i), \quad (\text{Equation 7})$$

where θ denotes all parameters. Specifically, the probability of v_j given v_i is defined as

$$P_\theta(v_j | v_i) = \frac{\exp(c_j^T \cdot v_{i,r})}{\sum_{v_k \in \mathcal{V}_t} \exp(c_k^T \cdot v_{i,r})}, \quad (\text{Equation 8})$$

where $v_j \in \mathcal{V}_t$, c_k is the context embedding of node v_k and v_i is the overall embedding of node v_i for edge type r . Finally, heterogeneous negative sampling is used to approximate the objective function $-\log P_\theta(v_j | v_i)$ for each node pair (v_j, v_i) as

$$E = -\log \sigma(c_j^T \cdot v_{i,r}) - \sum_{l=1}^L \mathbb{E}_{v_k \sim P_t(v)} \left[\log \sigma(-c_l^T \cdot v_{i,r}) \right], \quad (\text{Equation 9})$$

where $\sigma(x) = 1/(1 + \exp(-x))$ is the sigmoid function, L is the ratio of the number of negative samples to the number of positive training samples, and v_k is randomly drawn from a noise distribution $P_t(v)$ defined on node v_i 's corresponding node set \mathcal{V}_t .

Model building

In this study, six methods are effective and easily deployable on diverse scenarios of DCB discovery for performance comparison in terms of ROC-AUC and PR-AUC: SVM, RF, DNN, Mol-GCN, DistMult,⁶⁷ and ComplEx.⁶⁸ Specifically, the SVM, RF, and DNN models were taken from Scikit-learn,⁶⁹ the Mol-GCN model was constructed with DeepPurpose,⁷⁰ and the DistMult and ComplEx model was constructed with AmpliGraph (<https://doi.org/10.5281/zenodo.4792436>). The four DCB-based algorithms were trained on DCB data only, while the two KGE model and KG-CombPred were trained on additional omics data. In the training process, (1) the DCB dataset was firstly split into the training set and the test set in each fold according to the scenarios (for details see "dataset curation"). For each type of data in the KG, a random walk is conducted using a hyper-parameter called "window size" (also called "length L " in random walk). Starting from each node in the respective data type, neighbor nodes within the specified "window size" are selected to form training pairs. (2) Both the base embedding and edge embedding are initialized with uniformly distributed values between -1.0 and 1.0 . Specifically, the embedding for node of drug is a linear function (in Equations 2 and 3) of Morgan fingerprint calculated by RDKit. For each algorithm, the best parameter set was identified using Bayesian optimization (Optuna python package, version 2.10.0).⁷¹ More details about these baselines are provided in supplemental information. The performance of each method was evaluated on the validation set by the ROC-AUC).

KG-CombPred with community discovery for VS

To evaluate the impact of disease community discoveries on VS, we conducted a realistic screening of DCBs for COVID. By combining 36 COVID-related compounds with 7,892 drugs in the KG, 28,344 pairs of candidate DCBs were generated, including 48 COVID DCBs already identified in the literature. We then inferred whether the DCB belonged to COVID through CD and RF, respectively. The community discovery algorithm OSSE was implemented with CDLIB.⁷² The initial seeds consisted of 36 identified COVID-related drugs. Precision @ n is the VS evaluation metric, which refers to the proportion of COVID-relevant DCBs among the top n candidates. The shortest path on a graph is calculated by NetworkX (<https://networkx.org/>).

TB assays *in vitro*

Strain and reagents

Considering the serious infectivity and pathogenicity of *Mycobacterium tuberculosis* (Mtb), *M. smegmatis* mc²155 was selected as model bacteria to replace Mtb in this study. Besides kanamycin sulfate purchased from Sangon Biotech, all compounds were purchased from MedChemExpress (purity \geq 98% by high-performance liquid chromatography).

Cultivation of *M. smegmatis*

M. smegmatis were grown in Luria-Bertani medium, and 0.05% Tween-80 (Sigma-Aldrich) was added to disperse the strain and avoid agglomeration. *M. smegmatis* were incubated in a 220 rpm shaker incubator at 37°C, and the growth was monitored by a BioTek Eon spectrophotometer (BioTek, Winooski, VT) until the logarithmic phase.

In vitro activity test using *M. smegmatis*

The compounds were dissolved in DMSO as a reserve solution, while sulfate compounds were dissolved in purified water as reserve solutions. All compounds were tested on *M. smegmatis* by the broth microdilution method. The 96-well plate was incubated at 37°C for 48 h, and the absorbance under OD_{600 nm} was measured to determine the minimum inhibitory concentration (MIC₅₀). MIC₅₀ is defined as the concentration at which 50% bacterial growth can be inhibited in contrast to the drug-free control plates. The data were analyzed using GraphPad Prism software version 6.0 (GraphPad, San Diego, CA, USA).

Drug activity combination assay

One drug was diluted horizontally along a 96-well plate with multiple concentrations of MIC₅₀ (see Table S7 and Figure S3), and the other was diluted along the longitudinal direction of the plate, with 5 μ L of drug A and 5 μ L of drug B being added to the respective well. In addition, the negative control without drugs and the positive control with positive drugs were set in the 96-well plate. The assay was performed with three biological replicates.

To evaluate the effectiveness of a DCB, we analyzed the synergy between the drugs in a combination using the fractional inhibitory concentration. Specifically, three dose responses are measured for each pairwise interaction: (1) the dose response for drug X; (2) the dose response for drug Y; and (3) the dose response for an equipotent mixture of drug X and drug Y. In this study, the dose response is assessed based on the MIC₅₀ (minimum inhibitory concentration at 50% bacterial growth). As seen in Figure S4, the simulated 2D checkerboard assay serves as a visual demonstration of both synergistic (left) and antagonistic (right) pairwise DCBs. In each

assay, one drug is incrementally increased along each axis, and the resulting growth inhibition (represented by a color bar ranging from white to purple) is recorded for each concentration combination. In the squares ranging from $1/16$ MIC to $1/2$ MIC, if any score is >0.5 the DCB is considered to be synergistic.

TB cytotoxic assay

Cell lines and growth conditions

Mouse embryo fibroblast NIH3T3 cells were cultivated in DMEM (Bio-Channel) supplemented with 10% newborn calf serum (Evergreen) and 1% antibiotic-antimycotic at 37°C in 5% CO₂ in a humidified incubator. Before the tests, the cell lines were passaged twice after throwing.

Drug cytotoxic activity combination assay

The cells were seeded at a density of 4,000 cells/well on 96-well plates and placed in an incubator with 5% CO₂ at 37°C. After 24 h, compounds were plated as a 6 × 6 dose-response combination matrix, and two rows of cells were separately treated with serial diluted compounds. After 3 days, 10 μL of 5 mg/mL MTT solution was added into each well and incubated for an additional 4 h, after which 100 μL of triplex 10% SDS-5% isobutyl alcohol-0.012 mol/L HCl (w/v/v) solution was added to dissolve the formazan crystals. The absorbance at 570 nm was measured with the reference wavelength at 650 nm using a spectrophotometer (BioTek Eon). The assay was performed with three biological replicates.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xcrp.2023.101520>.

ACKNOWLEDGMENTS

This work was financially supported by the National Key Research and Development Program of China (2022YFF1203003), the National Natural Science Foundation of China (22220102001), the Fundamental Research Funds for the Central Universities (226-2023-00111, 226-2022-00107), and the Natural Science Foundation of Zhejiang Province of China (LD22H300001).

AUTHOR CONTRIBUTIONS

Conceptualization, Q.Y., C.-Y.H., S.H., and T.H.; methodology, Q.Y., R.X., and D.L.; analysis, Q.Y., R.X., D.L., Y.D., Y.K., and F.Z.; writing, Q.Y., R.X., and C.-Y.H.; supervision, J.C., S.H., and T.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: April 18, 2023

Revised: June 23, 2023

Accepted: July 7, 2023

Published: July 31, 2023

REFERENCES

- Mohamed, S.K., Nováček, V., and Nounu, A. (2020). Discovering protein drug targets using knowledge graph embeddings. *Bioinformatics* 36, 603–610. <https://doi.org/10.1093/bioinformatics/btz600>.
- Sosa, D.N., and Altman, R.B. (2022). Contexts and contradictions: a roadmap for computational drug repurposing with knowledge inference. *Briefings Bioinf.* 23. <https://doi.org/10.1093/bib/bbac268>.
- Mohamed, S.K., Nounu, A., and Nováček, V. (2021). Biological applications of knowledge graph embedding models. *Briefings Bioinf.* 22, 1679–1693. <https://doi.org/10.1093/bib/bbaa012>.
- Zhou, Y., Liu, Y., Gupta, S., Paramo, M.I., Hou, Y., Mao, C., Luo, Y., Judd, J., Wierbowski, S., Bertolotti, M., et al. (2023). A comprehensive SARS-CoV-2-human protein-protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nat. Biotechnol.* 41, 128–139. <https://doi.org/10.1038/s41587-022-01474-0>.
- Gogleva, A., Polychronopoulos, D., Pfeifer, M., Poroshin, V., Ughetto, M., Martin, M.J., Thorpe, H., Bornot, A., Smith, P.D., Sidders, B., et al. (2022). Knowledge graph-based recommendation framework identifies drivers of resistance in EGFR mutant non-small cell lung cancer. *Nat. Commun.* 13, 1667. <https://doi.org/10.1038/s41467-022-29292-7>.
- Fang, J., Zhang, P., Zhou, Y., Chiang, C.-W., Tan, J., Hou, Y., Stauffer, S., Li, L., Pieper, A.A., Cummings, J., and Cheng, F. (2021). Endophenotype-based in silico network medicine discovery combined with insurance record data mining identifies sildenafil as a candidate drug for Alzheimer's disease. *Nat. Aging* 1, 1175–1188. <https://doi.org/10.1038/s43587-021-00138-z>.
- Antimicrobial Resistance Collaborators, Ikuta, K.S., Sharara, F., Swetschinski, L., Aguilar, G.R., Gray, A., Han, C., Bisignano, C., Rao, P., and Wool, E. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399, 629–655. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0).
- O'Neill, J. (2016). Tackling drug-resistant infections globally: final report and recommendations.
- Zheng, W., Sun, W., and Simeonov, A. (2018). Drug repurposing screens and synergistic drug-combinations for infectious diseases. *Br. J. Pharmacol.* 175, 181–191. <https://doi.org/10.1111/bph.13895>.
- Sun, W., Sanderson, P.E., and Zheng, W. (2016). Drug combination therapy increases successful drug repositioning. *Drug Discov. Today* 21, 1189–1195. <https://doi.org/10.1016/j.drudis.2016.05.015>.
- Kitano, H. (2007). A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* 6, 202–210. <https://doi.org/10.1038/nrd2195>.
- Mathews Griner, L.A., Guha, R., Shinn, P., Young, R.M., Keller, J.M., Liu, D., Goldlust, I.S., Yasgar, A., McKnight, C., Boxer, M.B., et al. (2014). High-throughput combinatorial screening identifies drugs that cooperate with ibrutinib to kill activated B-cell-like diffuse large B-cell lymphoma cells. *Proc. Natl. Acad. Sci. USA* 111, 2349–2354. <https://doi.org/10.1073/pnas.1311846111>.
- Crystal, A.S., Shaw, A.T., Sequist, L.V., Friboulet, L., Niederst, M.J., Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., et al. (2014). Patient-derived models of acquired resistance can identify effective drug combinations for cancer. *Science* 346, 1480–1486. <https://doi.org/10.1126/science.1254721>.
- He, L., Kuleskiy, E., Saarela, J., Turunen, L., Wennerberg, K., Aittokallio, T., and Tang, J. (2018). Methods for high-throughput drug combination screening and synergy scoring. In *Cancer systems biology* (Springer), pp. 351–398. https://doi.org/10.1007/978-1-4939-7493-1_17.
- Cantrell, J.M., Chung, C.H., and Chandrasekaran, S. (2022). Machine learning to design antimicrobial combination therapies: promises and pitfalls. *Drug Discov. Today* 27, 1639–1651. <https://doi.org/10.1016/j.drudis.2022.04.006>.
- Mason, D.J., Stott, I., Ashenden, S., Weinstein, Z.B., Karakoc, I., Meral, S., Kuru, N., Bender, A., and Cokol, M. (2017). Prediction of antibiotic interactions using descriptors derived from molecular structure. *J. Med. Chem.* 60, 3902–3912. <https://doi.org/10.1021/acs.jmedchem.7b00204>.
- Mason, D.J., Eastman, R.T., Lewis, R.P.I., Stott, I.P., Guha, R., and Bender, A. (2018). Using machine learning to predict synergistic antimalarial compound combinations with novel structures. *Front. Pharmacol.* 9, 1096. <https://doi.org/10.3389/fphar.2018.01096>.
- Zakharov, A.V., Varlamova, E.V., Lagunin, A.A., Dmitriev, A.V., Muratov, E.N., Fourches, D., Kuz'min, V.E., Poroikov, V.V., Tropsha, A., and Nicklaus, M.C. (2016). Qsar modeling and prediction of drug–drug interactions. *Mol. Pharm.* 13, 545–556. <https://doi.org/10.1021/acs.molpharmaceut.5b00762>.
- Bobrowski, T., Chen, L., Eastman, R.T., Itkin, Z., Shinn, P., Chen, C.Z., Guo, H., Zheng, W., Michael, S., Simeonov, A., et al. (2021). Synergistic and antagonistic drug combinations against SARS-CoV-2. *Mol. Ther.* 29, 873–885. <https://doi.org/10.1016/j.ymthe.2020.12.016>.
- Zagidullin, B., Wang, Z., Guan, Y., Pitkänen, E., and Tang, J. (2021). Comparative analysis of molecular fingerprints in prediction of drug combination effects. *Briefings Bioinf.* 22, bbab291. <https://doi.org/10.1093/bib/bbab291>.
- Jin, W., Stokes, J.M., Eastman, R.T., Itkin, Z., Zakharov, A.V., Collins, J.J., Jaakkola, T.S., and Barzilay, R. (2021). Deep learning identifies synergistic drug combinations for treating COVID-19. *Proc. Natl. Acad. Sci. USA* 118, e2105070118. <https://doi.org/10.1073/pnas.2105070118>.
- Bai, L.-Y., Dai, H., Xu, Q., Junaid, M., Peng, S.-L., Zhu, X., Xiong, Y., and Wei, D.-Q. (2018). Prediction of effective drug combinations by an improved naïve bayesian algorithm. *Int. J. Mol. Sci.* 19, 467. <https://doi.org/10.3390/ijms19020467>.
- Zhao, X.-M., Iskar, M., Zeller, G., Kuhn, M., Van Noort, V., and Bork, P. (2011). Prediction of drug combinations by integrating molecular and pharmacological data. *PLoS Comput. Biol.* 7, e1002323. <https://doi.org/10.1371/journal.pcbi.1002323>.
- Cheng, F., Kovács, I.A., and Barabási, A.L. (2019). Network-based prediction of drug combinations. *Nat. Commun.* 10, 1197. <https://doi.org/10.1038/s41467-019-09186-x>.
- Ding, P., Shen, C., Lai, Z., Liang, C., Li, G., and Luo, J. (2020). Incorporating Multisource Knowledge To Predict Drug Synergy Based on Graph Co-regularization. *J. Chem. Inf. Model.* 60, 37–46. <https://doi.org/10.1021/acs.jcim.9b00793>.
- Aittokallio, T. (2022). What are the current challenges for machine learning in drug discovery and repurposing? *Expert Opin. Drug Discov.* 17, 423–425. <https://doi.org/10.1080/17460441.2022.2050694>.
- Zou, J., Ji, P., Zhao, Y.-L., Li, L.-L., Wei, Y.-Q., Chen, Y.-Z., and Yang, S.-Y. (2012). Neighbor communities in drug combination networks characterize synergistic effect. *Mol. Biosyst.* 8, 3185–3196. <https://doi.org/10.1039/c2mb25267h>.
- Chandrasekaran, S., Cokol-Cakmak, M., Sahin, N., Yilancioglu, K., Kazan, H., Collins, J.J., and Cokol, M. (2016). Chemogenomics and orthology-based design of antibiotic combination therapies. *Mol. Syst. Biol.* 12, 872. <https://doi.org/10.15252/msb.20156777>.
- Shtar, G., Azulay, L., Nizri, O., Rokach, L., and Shapira, B. (2022). CDCDB: A large and continuously updated drug combination database. *Sci. Data* 9, 263. <https://doi.org/10.1038/s41597-022-01360-z>.
- Whang, J.J., Gleich, D.F., and Dhillon, I.S. (2013). Overlapping community detection using seed set expansion. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pp. 2099–2108. <https://doi.org/10.1145/2505515.2505535>.
- Andersen, R., Chung, F., and Lang, K. (2006). Local Graph Partitioning Using Pagerank Vectors. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pp. 475–486. <https://doi.org/10.1109/FOCS.2006.44>.
- Lane, T., Russo, D.P., Zorn, K.M., Clark, A.M., Korotcov, A., Tkachenko, V., Reynolds, R.C., Perryman, A.L., Freundlich, J.S., and Ekins, S. (2018). Comparing and Validating Machine Learning Models for Mycobacterium tuberculosis Drug Discovery. *Mol. Pharm.* 15, 4346–4360. <https://doi.org/10.1021/acs.molpharmaceut.8b00083>.
- Jia, J., Zhu, F., Ma, X., Cao, Z., Cao, Z.W., Li, Y., Li, Y.X., and Chen, Y.Z. (2009). Mechanisms of drug combinations: interaction and network perspectives. *Nat. Rev. Drug Discov.* 8, 111–128. <https://doi.org/10.1038/nrd2683>.

34. Bertin, P., Rector-Brooks, J., Sharma, D., Gaudelet, T., Anighoro, A., Gross, T., Martinez-Pena, F., Tang, E.L., Regep, C., and Hayter, J. (2022). Recover: sequential model optimization platform for combination drug repurposing identifies novel synergistic compounds in vitro. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2202.04202>.
35. Cokol, M., Kuru, N., Bicak, E., Larkins-Ford, J., and Aldridge, B.B. (2017). Efficient measurement and factorization of high-order drug interactions in Mycobacterium tuberculosis. *Sci. Adv.* 3, e1701881. <https://doi.org/10.1126/sciadv.1701881>.
36. Nadler, J.P., Berger, J., Nord, J.A., Cofsky, R., and Saxena, M. (1991). Amoxicillin-clavulanic acid for treating drug-resistant Mycobacterium tuberculosis. *Chest* 99, 1025–1026. <https://doi.org/10.1378/chest.99.4.1025>.
37. Sharma, S.K., Sharma, A., Kadiravan, T., and Tharyan, P. (2013). Rifamycins (rifampicin, rifabutin and rifapentine) compared to isoniazid for preventing tuberculosis in HIV-negative people at risk of active TB. *Cochrane Database Syst. Rev.* 2013. CD007545-294. <https://doi.org/10.1002/14651858.CD007545.pub2>.
38. Schless, J.M., Allison, R.F., Inglis, R.M., White, E.F., and Topperman, S. (1965). The Use of Ethionamide in Combined Drug Regimens in the Re-treatment of Isoniazid-Resistant Pulmonary Tuberculosis. *Am. Rev. Respir. Dis.* 91, 728–737. <https://doi.org/10.1164/arrd.1965.91.5.728>.
39. Jabès, D., Della Bruna, C., Rossi, R., and Olliaro, P. (1994). Effectiveness of rifabutin alone or in combination with isoniazid in preventive therapy of mouse tuberculosis. *Antimicrob. Agents Chemother.* 38, 2346–2350. <https://doi.org/10.1128/AAC.38.10.2346>.
40. Heifets, L.B. (1982). Synergistic effect of rifampin, streptomycin, ethionamide, and ethambutol on Mycobacterium intracellulare. *Am. Rev. Respir. Dis.* 125, 43–48. <https://doi.org/10.1164/arrd.1982.125.1.43>.
41. Gangadharam, P.R., Perumal, V.K., Jairam, B.T., Rao, P.N., Nguyen, A.K., Farhi, D.C., and Iseman, M.D. (1987). Activity of rifabutin alone or in combination with clofazimine or ethambutol or both against acute and chronic experimental Mycobacterium intracellulare infections. *Am. Rev. Respir. Dis.* 136, 329–333. <https://doi.org/10.1164/ajrccm/136.2.329>.
42. Saini, V., Ammerman, N.C., Chang, Y.S., Tasneen, R., Chaisson, R.E., Jain, S., Nueremberger, E., and Grosset, J.H. (2019). Treatment-shortening effect of a novel regimen combining clofazimine and high-dose rifapentine in pathologically distinct mouse models of tuberculosis. *Antimicrob. Agents Chemother.* 63, e00388-19–e00319. <https://doi.org/10.1128/AAC.00388-19>.
43. Brochado, A.R., Telzerow, A., Bobonis, J., Banzhaf, M., Mateos, A., Selkrig, J., Huth, E., Bassler, S., Zamarréno Beas, J., Zietek, M., et al. (2018). Species-specific activity of antibacterial drug combinations. *Nature* 559, 259–263. <https://doi.org/10.1038/s41586-018-0278-9>.
44. Cokol, M., Chua, H.N., Tasan, M., Mutlu, B., Weinstein, Z.B., Suzuki, Y., Nergiz, M.E., Costanzo, M., Baryshnikova, A., Giaever, G., et al. (2011). Systematic exploration of synergistic drug pairs. *Mol. Syst. Biol.* 7, 544. <https://doi.org/10.1038/msb.2011.71>.
45. Palmqvist, M., Ibsen, T., Mellén, A., and Lötvall, J. (1999). Comparison of the relative efficacy of formoterol and salmeterol in asthmatic patients. *Am. J. Respir. Crit. Care Med.* 160, 244–249. <https://doi.org/10.1164/ajrccm.160.1.9901063>.
46. Teach, S.J., Gill, M.A., Togias, A., Sorkness, C.A., Arbes, S.J., Jr., Calatroni, A., Wildfire, J.J., Gergen, P.J., Cohen, R.T., Pongracic, J.A., et al. (2015). Preseasonal treatment with eitheromalizumab or an inhaled corticosteroid boost to prevent fall asthma exacerbations. *J. Allergy Clin. Immunol.* 136, 1476–1485. <https://doi.org/10.1016/j.jaci.2015.09.008>.
47. Tashkin, D.P., Donohue, J.F., Mahler, D.A., Huang, H., Goodwin, E., Schaefer, K., Hanrahan, J.P., and Andrews, W.T. (2009). Effects of arformoterol twice daily, tiotropium once daily, and their combination in patients with COPD. *Respir. Med.* 103, 516–524. <https://doi.org/10.1016/j.rmed.2008.12.014>.
48. Holló, G., Bozkurt, B., and Irkeç, M. (2009). Brinzolamide/timolol fixed combination: a new ocular suspension for the treatment of open-angle glaucoma and ocular hypertension. *Exp. Opin. Pharmacother.* 10, 2015–2024. <https://doi.org/10.1517/14656560903124388>.
49. Konstas, A.G.P., Holló, G., Mikropoulos, D., Tsironi, S., Haidich, A.-B., Embeslidis, T., Georgiadou, I., Irkeç, M., and Melamed, S. (2010). Twenty-four-hour intraocular pressure control with bimatoprost and the bimatoprost/timolol fixed combination administered in the morning, or evening in exfoliative glaucoma. *Br. J. Ophthalmol.* 94, 209–213. <https://doi.org/10.1136/bjo.2008.155317>.
50. Araujo, J.C., Mathew, P., Armstrong, A.J., Braud, E.L., Posadas, E., Lonberg, M., Gallick, G.E., Trudel, G.C., Paliwal, P., Agrawal, S., and Logothetis, C.J. (2012). Dasatinib combined with docetaxel for castration-resistant prostate cancer: results from a phase 1-2 study. *Cancer* 118, 63–71. <https://doi.org/10.1002/cncr.26204>.
51. Pectasides, D., Pectasides, M., Farmakis, D., Kostopoulou, V., Nikolaou, M., Gaglia, A., Koumpou, M., Mylonakis, N., Xiros, N., Economopoulos, T., and Raptis, S.A. (2005). Comparison of docetaxel and docetaxel-irinotecan combination as second-line chemotherapy in advanced non-small-cell lung cancer: a randomized phase II trial. *Ann. Oncol.* 16, 294–299. <https://doi.org/10.1093/annonc/mdt053>.
52. Raji, L., Chiou, X.-C., Owens, R., and Wrigley, B. (1985). Therapeutic implications of hypertension-induced glomerular injury: comparison of enalapril and a combination of hydralazine, reserpine, and hydrochlorothiazide in an experimental model. *Am. J. Med.* 79, 37–41. [https://doi.org/10.1016/0002-9343\(85\)90078-6](https://doi.org/10.1016/0002-9343(85)90078-6).
53. Wilson, W.R., Okun, R., Tetreault, L., and Fallis, N. (1963). Methyl dopa and hydrochlorothiazide in primary hypertension: controlled clinical trial of drugs singly and in combination. *JAMA* 185, 819–825. <https://doi.org/10.1001/jama.1963.03060110023012>.
54. Kool, M.J., Lustermaans, F.A., Breed, J.G., Struyker Boudier, H.A., Hoeks, A.P., Reneman, R.S., and Van Bortel, L.M. (1995). The influence of perindopril and the diuretic combination amiloride+ hydrochlorothiazide on the vessel wall properties of large arteries in hypertensive patients. *J. Hypertens.* 13, 839–848. <https://doi.org/10.1097/00004872-199508000-00004>.
55. Yang, H., Qin, C., Li, Y.H., Tao, L., Zhou, J., Yu, C.Y., Xu, F., Chen, Z., Zhu, F., and Chen, Y.Z. (2016). Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res.* 44, D1069–D1074. <https://doi.org/10.1093/nar/gkv1230>.
56. Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Pache, L., Burgstaller-Muehlbacher, S., De Jesus, P.D., Teriete, P., Hull, M.V., et al. (2020). Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* 586, 113–119. <https://doi.org/10.1038/s41586-020-2577-1>.
57. Tan, X., Hu, L., Luquette, L.J., Gao, G., Liu, Y., Qu, H., Xi, R., Lu, Z.J., Park, P.J., and Elledge, S.J. (2012). Systematic identification of synergistic drug pairs targeting HIV. *Nat. Biotechnol.* 30, 1125–1130. <https://doi.org/10.1038/nbt.2391>.
58. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O’Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>.
59. Lamprecht, D.A., Finin, P.M., Rahman, M.A., Cumming, B.M., Russell, S.L., Jonnal, S.R., Adamson, J.H., and Steyn, A.J.C. (2016). Turning the respiratory flexibility of Mycobacterium tuberculosis against itself. *Nat. Commun.* 7, 12393–12414. <https://doi.org/10.1038/ncomms12393>.
60. Reddy, M.V., Srinivasan, S., and Gangadharam, P.R. (1994). In vitro and in vivo synergistic effect of isoniazid with streptomycin and clofazimine against Mycobacterium avium complex (MAC). *Tuber. Lung Dis.* 75, 208–212. [https://doi.org/10.1016/0962-8479\(94\)90010-8](https://doi.org/10.1016/0962-8479(94)90010-8).
61. Ramón-García, S., González del Río, R., Villarejo, A.S., Sweet, G.D., Cunningham, F., Barros, D., Ballell, L., Mendoza-Lozana, A., Ferrer-Bazaga, S., and Thompson, C.J. (2016). Repurposing clinically approved cephalosporins for tuberculosis therapy. *Sci. Rep.* 6, 1–14. <https://doi.org/10.1038/srep34293>.
62. Dshpande, D., Srivastava, S., Chapagain, M., Magombedze, G., Martin, K.R., Cirrincione, K.N., Lee, P.S., Koeuth, T., Dheda, K., and Gumbo, T. (2017). Ceftazidime-avibactam has potent sterilizing activity against drug-resistant tuberculosis. *Sci. Adv.* 3, e1701102. <https://doi.org/10.1126/sciadv.1701102>.
63. Makarov, V., Lechartier, B., Zhang, M., Neres, J., van der Sar, A.M., Raadsen, S.A., Hartkoorn, R.C., Ryabova, O.B., Vocat, A., Decosterd, L.A., et al. (2014). Towards a new combination therapy for tuberculosis with next generation benzothiazinones. *EMBO Mol. Med.* 6, 372–383. <https://doi.org/10.1002/emmm.201303575>.

64. Larkins-Ford, J., Greenstein, T., Van, N., Degefu, Y.N., Olson, M.C., Sokolov, A., and Aldridge, B.B. (2021). Systematic measurement of combination-drug landscapes to predict in vivo treatment outcomes for tuberculosis. *Cell Syst.* 12, 1046–1063.e7. <https://doi.org/10.1016/j.cels.2021.08.004>.
65. Cen, Y., Zou, X., Zhang, J., Yang, H., Zhou, J., and Tang, J. (2019). Representation learning for attributed multiplex heterogeneous network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1358–1368. <https://doi.org/10.1145/3292500.3330964>.
66. Dong, Y., Chawla, N.V., and Swami, A. (2017). metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 135–144. <https://doi.org/10.1145/3097983.3098036>.
67. Nickel, M., Tresp, V., and Kriegel, H.-P. (2011). A three-way model for collective learning on multi-relational data. *ICML*. <https://doi.org/10.5555/3104482.3104584>.
68. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., and Bouchard, G. (2016). Complex embeddings for simple link prediction. *International conference on machine learning*, 2071–2080. <https://doi.org/10.5555/3045390.3045609>.
69. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., and Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12, 2825–2830. <https://doi.org/10.5555/1953048.2078195>.
70. Huang, K., Fu, T., Glass, L.M., Zitnik, M., Xiao, C., and Sun, J. (2021). DeepPurpose: a deep learning library for drug–target interaction prediction. *Bioinformatics* 36, 5545–5547. <https://doi.org/10.1093/bioinformatics/btaa1005>.
71. Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A Next-Generation Hyperparameter Optimization Framework, pp. 2623–2631.
72. Rossetti, G., Milli, L., and Cazabet, R. (2019). CDLIB: a python library to extract, compare and evaluate communities from complex networks. *Appl. Netw. Sci.* 4, 52–26. <https://doi.org/10.1007/s41109-019-0165-9>.