Electronic supplementary information for

Neighbor communities in drug combination network characterize synergistic effect

Jun Zou*, Pan Ji, Ying-Lan Zhao, Lin-Li Li, Yu-Quan Wei, Yu-Zong Chen, Sheng-Yong Yang

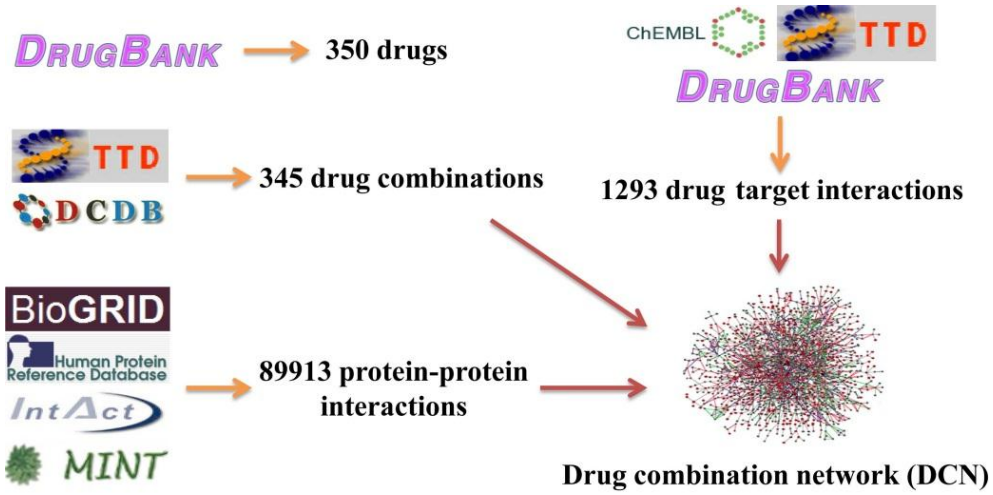

Fig. S1 Drug combination network construction. A list of 345 human drug combinations was obtained from public drug databases TTD and DCDB. This list includes 350 unique drugs, whose pharmacology information, including drug category, description, and indication, was retrieved from database DrugBank. The target information of the 350 unique drugs was retrieved from database ChEMBL, TTD, and DrugBank. A total of 1293 drug-target interactions with evidence of experimental activity were obtained. 89913 unique protein-protein interactions with experimental verification were extracted from database BioGrid, HPRD, IntAct and MINT.

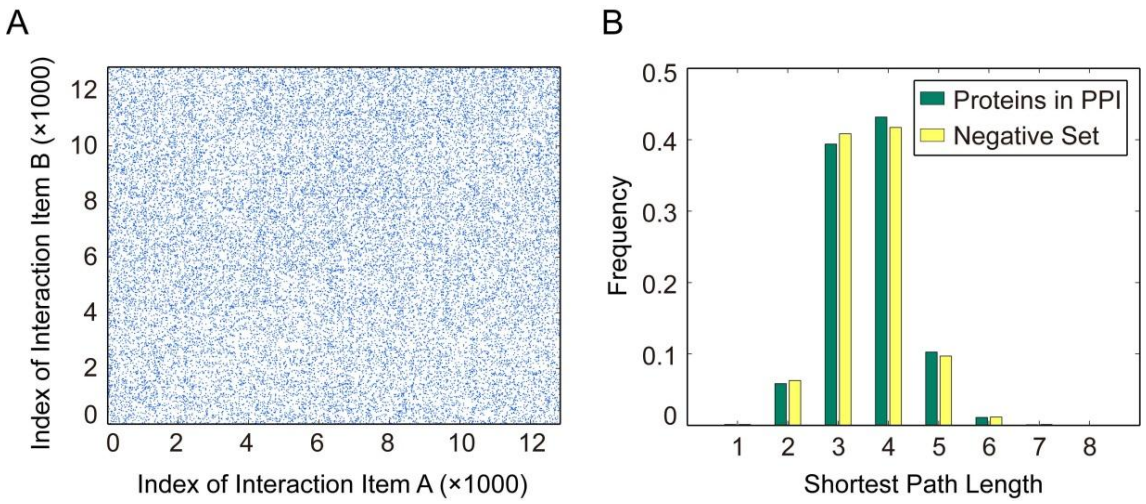

Fig. S2 Evaluation of the network properties of negative set. (A) Uniform distribution in the PPI network space of proteins of the randomly generated negative set can be observed. (B) The distribution of the shortest path length of proteins of the negative set is coincident with that of PPI network. Thus no bias was introduced in the generation of negative set.

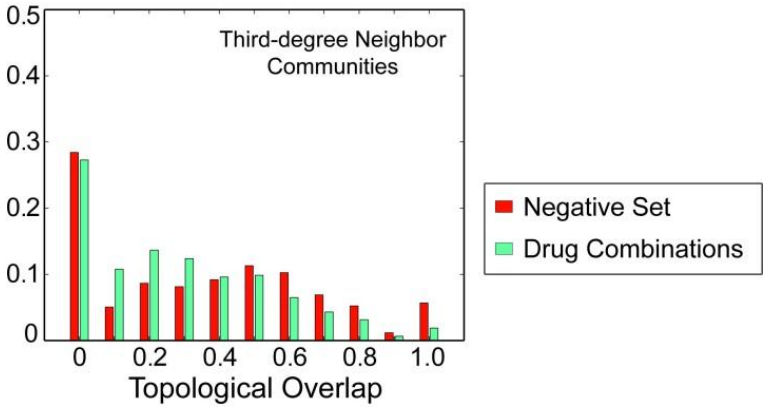

Fig. S3 Difference in topological overlap property of the third-degree neighbor communities between drug combinations and negative set.

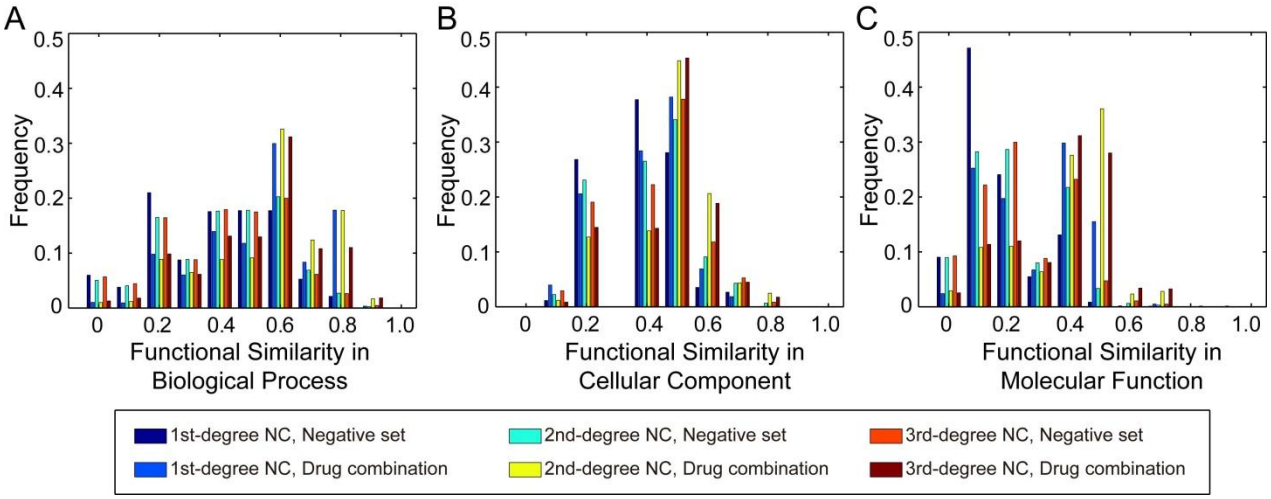

Fig. S4 Comparison of the frequency of the functional properties of neighbor communities (NC) between drug combinations and negative set, using different Gene Ontology namespace, i.e. biological process, cellular component, and molecular function. Here neighbor communities with functional property value equal to zero were not considered. It was found that all three degree neighbor communities of drug combinations have almost the same difference in the distribution of functional properties with that of the negative set. It has been observed that the second-degree neighbor communities of drug combinations have the highest difference in the distribution of functional properties with that of the negative set (Fig. 5A). The potential explanation for this observation might be that most drug targets (79%) share the second-degree neighbor communities (Fig. 2A). In such a circumstance, more drug combinations will have second-degree neighbor communities with functional property value not equal to zero, and have the highest difference in the distribution of functional properties.

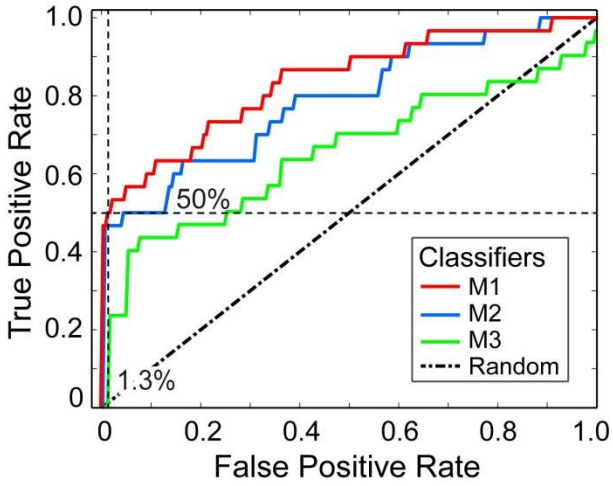

Fig. S5 Performance of using topological and functional properties of neighbor communities in distinguishing synergistic drug combinations from the other negative set with antagonistic mechanism. Receiver operating characteristic curves were used to represent the prediction performance of the SVM classifiers in leave-one-out cross-validation, where the diagonal dashed line denotes random prediction. The antagonistic drug combinations were used as the control negative set. The best SVM classifier M1 has an area under the ROC curve (AUC) of 0.85, and identifies 50% of the true positives with only 1.3% of the false positives. We believe that this prediction performance could be improved further with more available antagonistic drug combinations.

Figure 7A

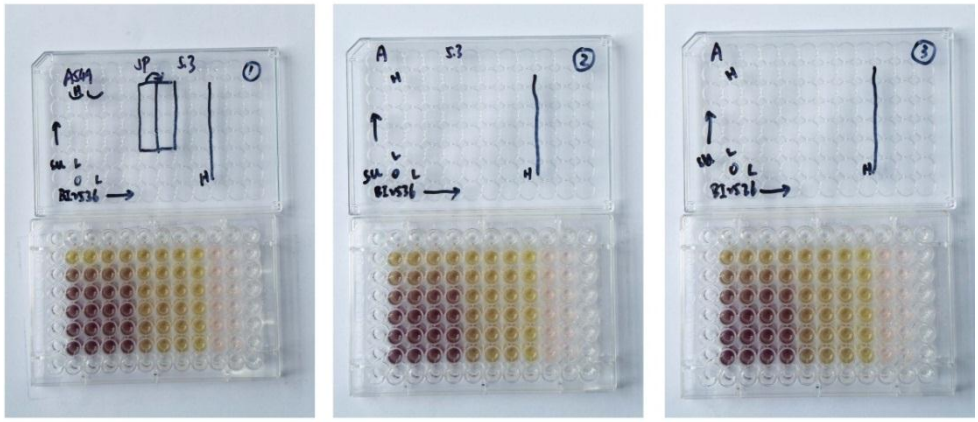

Figure 7B

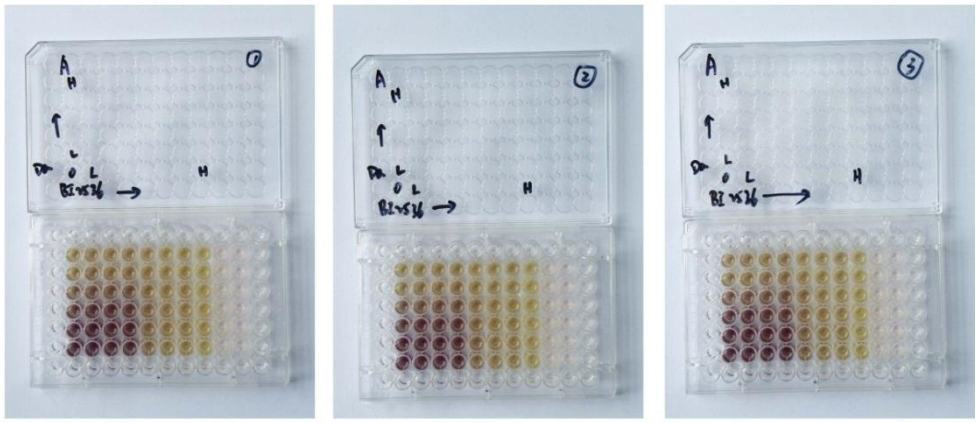

Figure 7C

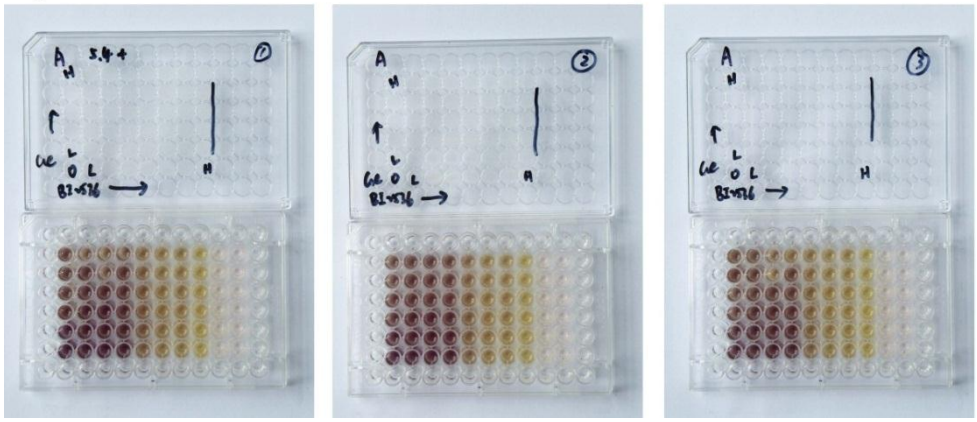

Fig. S6 Cell-based assessments of the effects of predicted drug combinations. The source data for Figure 7.

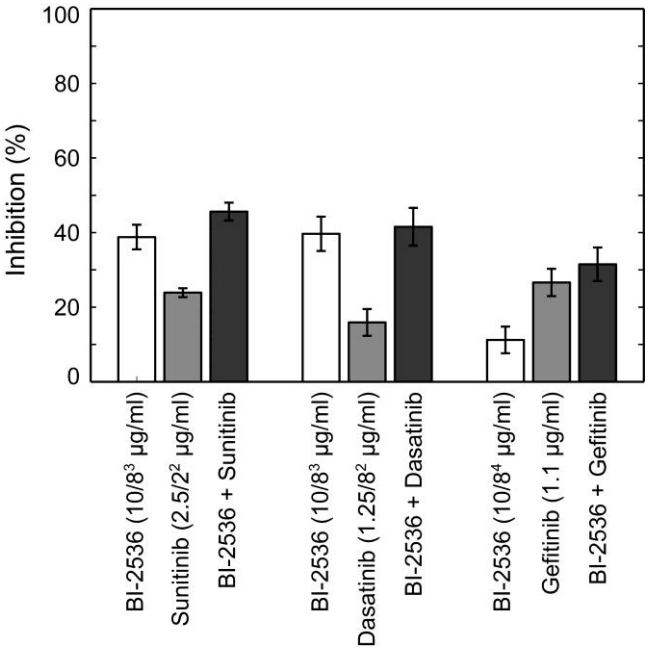

Fig. S7 Cell-based assessments of the toxicity of predicted drug combinations. Normal human lung fibroblast MRC-5 cell line was obtained from American Type Culture Collection, and maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum. BI-2536, dasatinib, gefitinib, and sunitinib were added to the culture medium with concentrations that could produce synergistic effect. The viability of MRC-5 cells following drug treatments was determined using MTT assay method (n=3). It was found that the drug combinations do not result in increased toxicity.

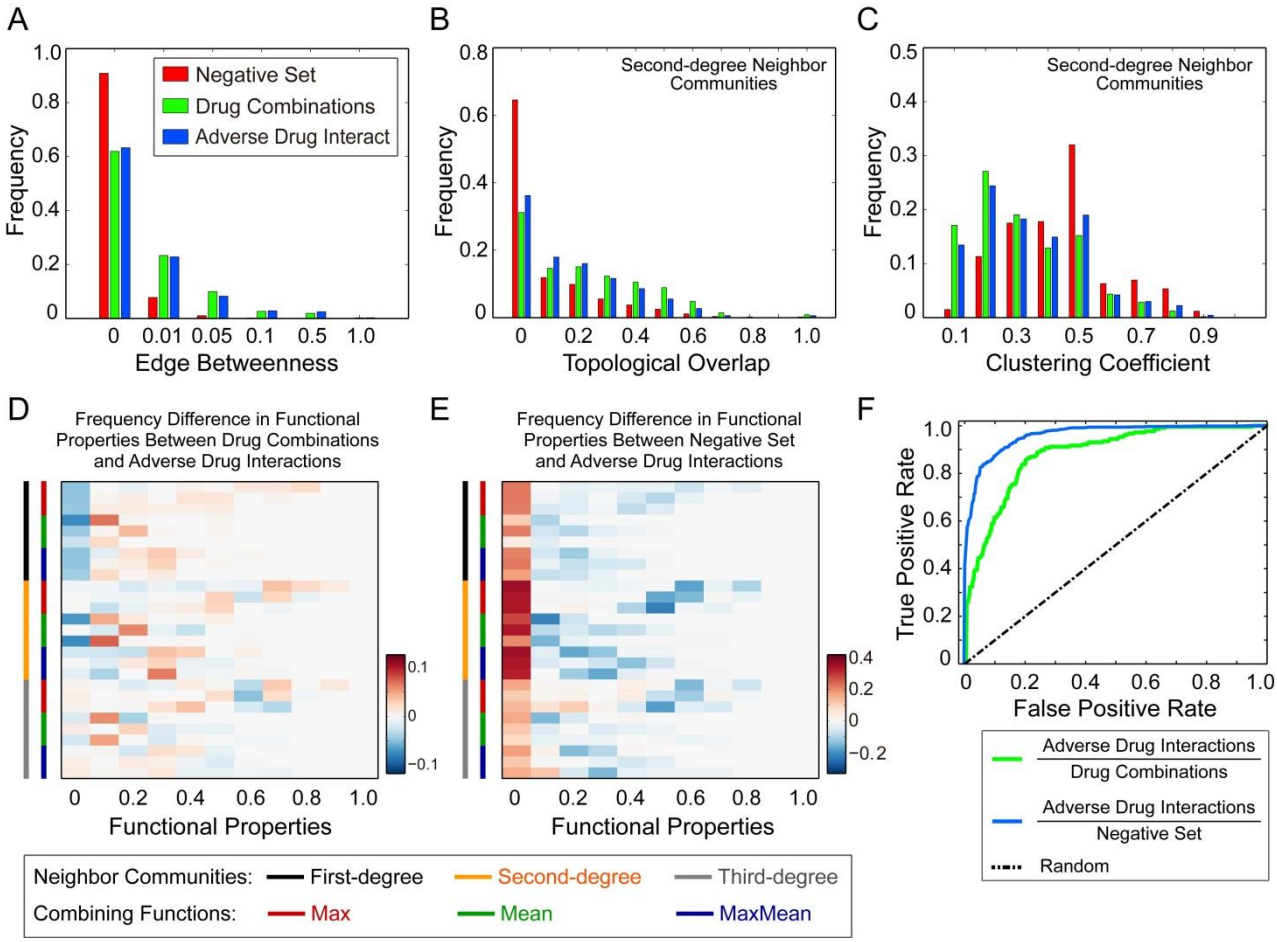

Fig. S8 (A-E) Difference in topological and functional properties of neighbor communities between adverse drug-drug interactions (Table S10), drug combinations and negative set. The topological properties include (A) edge betweenness centrality, (B) topological overlap, and (C) clustering coefficient. Systematical study of the frequency difference in the functional properties of neighbor communities (D) between adverse drug interactions and drug combinations, (E) between adverse drug interactions and negative set. (F) Performance of using topological and functional properties of neighbor communities in adverse drug interaction prediction. Receiver operating characteristic curves for the prediction performance of the SVM classifiers in leave-one-out cross-validation, where the diagonal dashed line denotes random prediction.

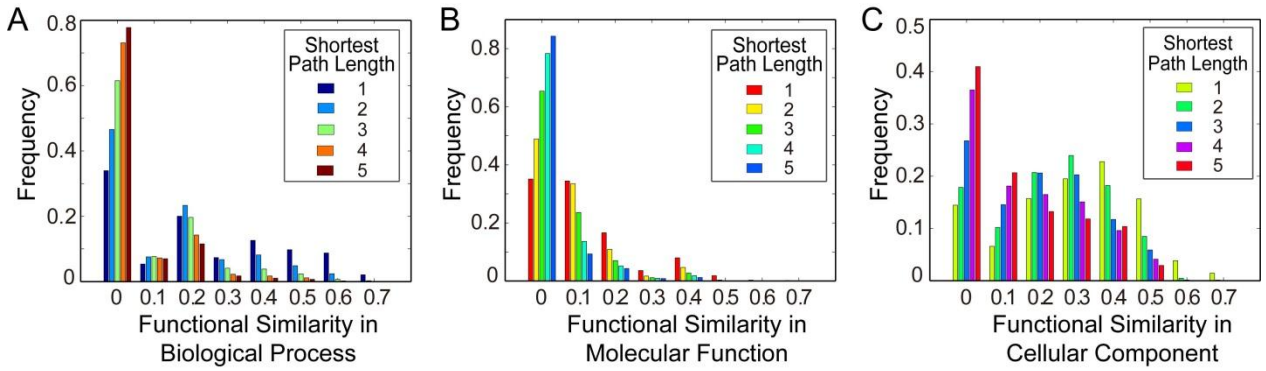

Fig. S9 The relationships between topological properties (i.e. shortest path length) and functional properties. It is assumed that two proteins will have more similar functions if they have shorter distance in network, but this does not guarantee that shortest path length can be used to predict their functional similarity, and vice versa. Thus, topological and functional properties should be integrated.

<File: Dataset_S1.zip>

Dataset S1 The associated Matlab source code and a test dataset for running this new computational method.

<File: Table_S1.xlsx>

Table S1 The list of 345 human drug combinations. The information of drug combinations was retrieved from public drug databases TTD and DCDB. Only these have pharmacodynamic synergistic or additive effect were compiled in this study.

<File: Table_S2.xlsx>

Table S2 Pharmacology information of 350 unique drugs in the drug combination list.

Their pharmacology information, including drug category, description and indication, was extracted from database DrugBank.

<File: Table_S3.xlsx>

Table S3 The list of 1293 drug-target interactions with evidence of experimental activity. The target information of the 350 unique drugs was obtained from databases DrugBank, ChEMBL, and TTD.

<File: Table_S4.xlsx>

Table S4 The list of 89913 unique protein-protein interactions (PPI). Experimentally verified PPI were extracted from database BioGrid, IntAct, HPRD, and MINT. Different types of protein identifiers were mapped using PICR, and repeated records were discarded.

<File: Table_S5.xlsx>

Table S5 The negative set of 1512 unique random drug pairs with target proteins generated as the control to evaluate the performance of drug combination prediction.

<File: Table_S6.xlsx>

Table S6 Systematical study of the contributions of integrating different topological and functional properties to the accuracy of drug combination prediction. Leave-one-out cross-validation was employed to prevent the over fitting problem. More negative (lower) score indicates higher possibility to be a drug combination. The detailed information for how different topological and functional properties were integrated and which properties used by SVM

classifiers was given at the bottom of this table.

<File: Table_S7.xlsx>

Table S7 The list of 30 drug combinations with antagonistic mechanism. This list was compiled using information from public drug databases TTD and DCDB. The target information of the 49 unique drugs was retrieved from database DrugBank, ChEMBL, and TTD. A total of 137 drug-target interactions with experimental evidence were obtained.

<File: Table_S8.xlsx>

Table S8 Confirming that drug targets are expressed in A549 cells. Literature search were carried out and it was found that the drug targets are expressed in A549 cell line. In addition, two publicly available database resources were also used to confirm the presence of these drug targets: (1) 'Gene Expression Atlas' (http://www.ebi.ac.uk/gxa/) that contains gene expression data, and (2) 'Human Protein Atlas' (http://www.proteinatlas.org/) that provides the expression information of proteins.

<File: Table_S9.xlsx>

Table S9 Pathway analysis of the most frequent neighbor communities of the targets of combination drugs. The most frequent neighbor communities of drug combinations can be mapped to signaling pathways retrieved from KEGG database. The therapeutic indications here include analgesics and anticonvulsants, antihypertensive agents, anti-inflammatory agents and anti-allergic agents.

<File: Table_S10.xlsx>

Table S10 The list of 445 adverse drug interactions. Based on the 350 unique drugs in our drug combination list, we retrieved 445 drug-drug interactions with adverse effects from DrugBank database. A total of 595 drug-target interactions with experimental evidence were further obtained.