Genetic Interaction-Based Biomarkers Identification for Drug Resistance and Sensitivity in Cancer Cells

Yue Han,[1,5] Chengyu Wang,[1,5] Qi Dong,[1,5] Tingting Chen,[1] Fan Yang,[1] Yaoyao Liu,[1] Bo Chen,[1] Zhangxiang Zhao,[1] Lishuang Qi,[1] Wenyuan Zhao,[1] Haihai Liang,[4] Zheng Guo,[1,2,3] and Yunyan Gu[1]

[1]Department of Systems Biology, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150086, China; [2]Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Department of Bioinformatics, Fujian Medical University, Fuzhou, China; [3]Fujian Key Laboratory of Tumor Microbiology, Fujian Medical University, Fuzhou, China; [4]Department of Pharmacology, College of Pharmacy, Harbin Medical University, Harbin, China

Cancer cells generally harbor hundreds of alterations in the cancer genomes and act as crucial factors in the development and progression of cancer. Gene alterations in the cancer genome form genetic interactions, which affect the response of patients to drugs. We developed an algorithm that mines copy number alteration and whole-exome mutation profiles from The Cancer Genome Atlas (TCGA), as well as functional screen data generated to identify potential genetic interactions for specific cancer types. As a result, 4,529 synthetic viability (SV) interactions and 10,637 synthetic lethality (SL) interactions were detected. The pharmacogenomic datasets revealed that SV interactions induced drug resistance in cancer cells and that SL interactions mediated drug sensitivity in cancer cells. Deletions of HDAC1 and DVL1, both of which participate in the Notch signaling pathway, had an SV effect in cancer cells, and deletion of DVL1 induced resistance to HDAC1 inhibitors in cancer cells. In addition, patients with low expression of both HDAC1 and DVL1 had poor prognosis. Finally, by integrating current reported genetic interactions from other studies, the Cancer Genetic Interaction database (CGIdb) (http://www. medsysbio.org/CGIdb) was constructed, providing a convenient retrieval for genetic interactions in cancer.

INTRODUCTION

Traditional cancer therapy predominantly involves chemotherapy, which provides limited improvements in patient survival. Patients frequently acquire resistance to chemotherapy, significantly decreasing the usefulness of anticancer drugs in the clinic.[1] As a new approach of precision medicine, target therapy has brought a new opportunity for cancer treatment. Target therapy refers to formulating the best treatment regimens for patients according to personalized genome characteristics, maximizing the treatment effect and minimizing patient injury.[2] There are hundreds of genomic alterations in tumor cells, such as copy number alterations (CNAs) and somatic mutations. Cancer patients, even with the same cancer type, have high heterogeneous genomic alterations,[3] which induce complex drug responses. It is still a great challenge to identify effective biomarkers to predict drug resistance and drug sensitivity for target

therapy. Inevitable drug resistance in target therapy is also a major obstacle. Tremendous efforts and progress have been made over the past few years to reveal the specific mechanism of drug resistance and sensitivity, including genetic interactions among gene alterations in the cancer genome.[4,5]

The generation of an unexpected phenotypic outcome when combining two alterations is referred to as a genetic interaction, including synthetic viability and synthetic lethality.[6] Synthetic viability (SV) describes the scenario in which the synthesis or combination of two gene effects remedy the effect of single-gene defects, such as cell death or a significant impairment of fitness, and enhance cell survival, which may be a potential mechanism for drug resistance.[5] Synthetic lethality (SL) describes the scenario in which single-gene defects are compatible with cell viability, but the combination of alterations in both genes induces cell death, which is generally used to identify drug-sensitive biomarkers or drug targets.[7] A well-known example of an SL effect is the selective impediment of PARP inhibitors on the growth of BRCA1/2-mutated cancer cells, and this SL effect has been successfully applied to clinical therapy.[8] Olaparib was approved by the US Food and Drug Administration (FDA) for treatment of BRCA1/2-mutated advanced ovarian cancer patients in 2014. However, many ovarian cancer patients carrying mutations of BRCA1/2 have resistance to olaparib in clinical applications.[9] Hu et al.[10] found that GPBP1 loss causes resistance to PARP inhibitors by regulating the expression of factors involved in homologous recombination (HR) according to a quantitative chemotherapy genetic interaction map.

Gene alterations in the cancer genome present specific patterns, such as co-occurrence or mutual exclusivity.[11] Mutual exclusivity indicates that both gene alterations rarely occur in the same patients, and mutual exclusivity has been used to identify synthetic lethal interactions in cancer cells. For example, Unni et al.[12] reported that an SL effect underlies the mutual exclusivity of oncogenic *KRAS* and *EGFR* mutations in lung adenocarcinoma. Co-occurring alterations, in which both gene alterations significantly and simultaneously occur in cancer patients, may have an SV effect on cancer cells, and these alterations have been validated in our previous work.[5] Recent advances in functional screening technology provide an avenue to explore genetic interactions in cancer cells.[13] However, most studies have focused on the identification of SL interactions and further detection of drug targets while ignoring SV interactions involved in the cancer genome. Thus, besides SL interactions, the present work identified and analyzed SV interactions in cancer cells using The Cancer Genome Atlas (TCGA) cancer genome data and validating by short hairpin RNA (shRNA), CRISPR, and yeast datasets. Moreover, the present study further detected drug-resistance biomarkers using SV interactions based on the hypothesis that cancer cells with alterations in partner genes are resistant to specific drugs if drug targets and partner genes have SV effects.

In the present study, CNAs and the mutation and expression profiles of 8,580 samples were integrated across 32 cancer types from TCGA to identify co-occurrence and mutual exclusivity gene pairs (Figures 1A and 1B). Reliable co-occurrence and mutual exclusivity interactions were then selected from functional screening data, including shRNA and CRISPR datasets (Figure 1C). The strategy was based on the notion that knockdown of one gene causes a selective enhancement (SV) or reduction (SL) in cell viability, with simultaneous alterations in another gene. In addition, the positive and negative interactions in yeast genetic interactions were also used to infer SV and SL relationships in human cancer cells, respectively. Ultimately, the co-occurrence and mutual exclusivity gene pairs verified in at least one type of dataset (shRNA, CRISPR, or yeast) were selected as candidate genetic interactions (Figure 1C). SV and SL interactions were validated by (1) showing the expected drug response detected by four pharmacogenomic datasets (Figure 1D) as well as (2) observing worse survival of patients with alterations in SV interactions and better survival of patients with alterations in SL interactions in the prognosis analysis (Figure 1E). Moreover, network analysis and pathway enrichment analysis were performed to investigate the functional relationship between genes with genetic interactions. The biomarkers identified by the present work will contribute to predict the mechanism of drug resistance or sensitivity in clinical application and will guide precise targeting of existing therapies.

RESULTS

Identification of Candidate SV Interactions and SL Interactions in Cancers

To identify significant co-occurring and mutual exclusive relationships among genomic alterations, mutation and CNA data from 8,580 samples of 32 different cancer types studied by the TCGA consortium were utilized (Figure 1A). Considering the heterogeneous distribution of alterations across genes and samples, a randomization procedure was applied to detect gene pairs with significant co-occurring or mutual exclusive alterations (see Materials and Methods; false discovery rate [FDR] < 0.05; Figure 1B). The statistics of co-occurring and mutual exclusive gene pairs within each cancer type are shown in Figure 2A.

The co-occurring and mutual exclusive gene pairs, which had a combined effect verified by at least one of the shRNA, CRISPR, and yeast datasets, were selected as candidate SV and SL interactions. According to the criteria to select candidate SV and SL interactions (see Materials and Methods), 15,928, 103,879, and 191 candidate SV interactions were separately identified in the shRNA, CRISPR, and yeast datasets (Figure 2B). In total, 11,001, 12,421 and 196 candidate SL interactions were separately identified in the shRNA, CRISPR, and yeast datasets (Figure 2C). The SL and SV interactions from 16 studies were integrated for follow-up research accompanied by our predictions. The information of 16 studies is provided in the Supplemental Materials and Methods. In total, 490 SV interactions and 29,171 SL interactions were selected from current studies (Table S1). An integrative confidence score combining scores from all evidence sources and our prediction were calculated to estimate the reliability of SL or SV genetic interaction (see Supplemental Materials and Methods). The quantitative score according to the experimental methods was annotated in evidence sources (Table S2).

SV Interactions and SL Interactions Associated with Drug Response

Four drug datasets (from CCLE, GDSC, Cancer Therapeutics Response Portal [CTRP], and NCI60) were used to identify drug biomarkers by investigating the SV and SL interactions in cancer cells. Our previous studies have demonstrated that drug response heterogeneity exists among tissue-specific cell lines.[14] In the present study, there were significant differences in drug sensitivity among different tissues according to ANOVA ($p < 0.05$; Figure S1). Therefore, the drug responses of SV and SL interactions were analyzed in tissue-specific cell lines, which limited the number of cell lines to three. For SV interactions, if one gene was targeted by a particular drug, the drug response measures (IC_{50} or AUC) in cell lines with alterations of the partner gene were significantly higher than those in the cell lines with the wild-type (WT) partner gene. For SL interactions, if one gene targeted was by a particular drug, the drug response measures (IC_{50} or AUC) in cell lines with alterations of the partner gene were significantly lower than those in the cell lines with the WT partner gene. The SV and SL interactions conformed to the expected drug response were different for each tissue (Figures S2A and S2C). The significant overlapping of drug-response-related SV (or SL) interactions identified by AUC and LN_IC_{50} (natural log of the fitted half maximal inhibitory concentration) values from GDSC indicated the reliability of the present study ($p < 0.001$; Figures S2B and S2D). A well-known example of SL is that ovarian cancer patients with *BRCA1/2* mutations are treated with *PARP* inhibitors. Figure S3 demonstrates that the *BRCA2* mutation was associated with sensitivity to olaparib, a *PARP* inhibitor, in cancer cells.
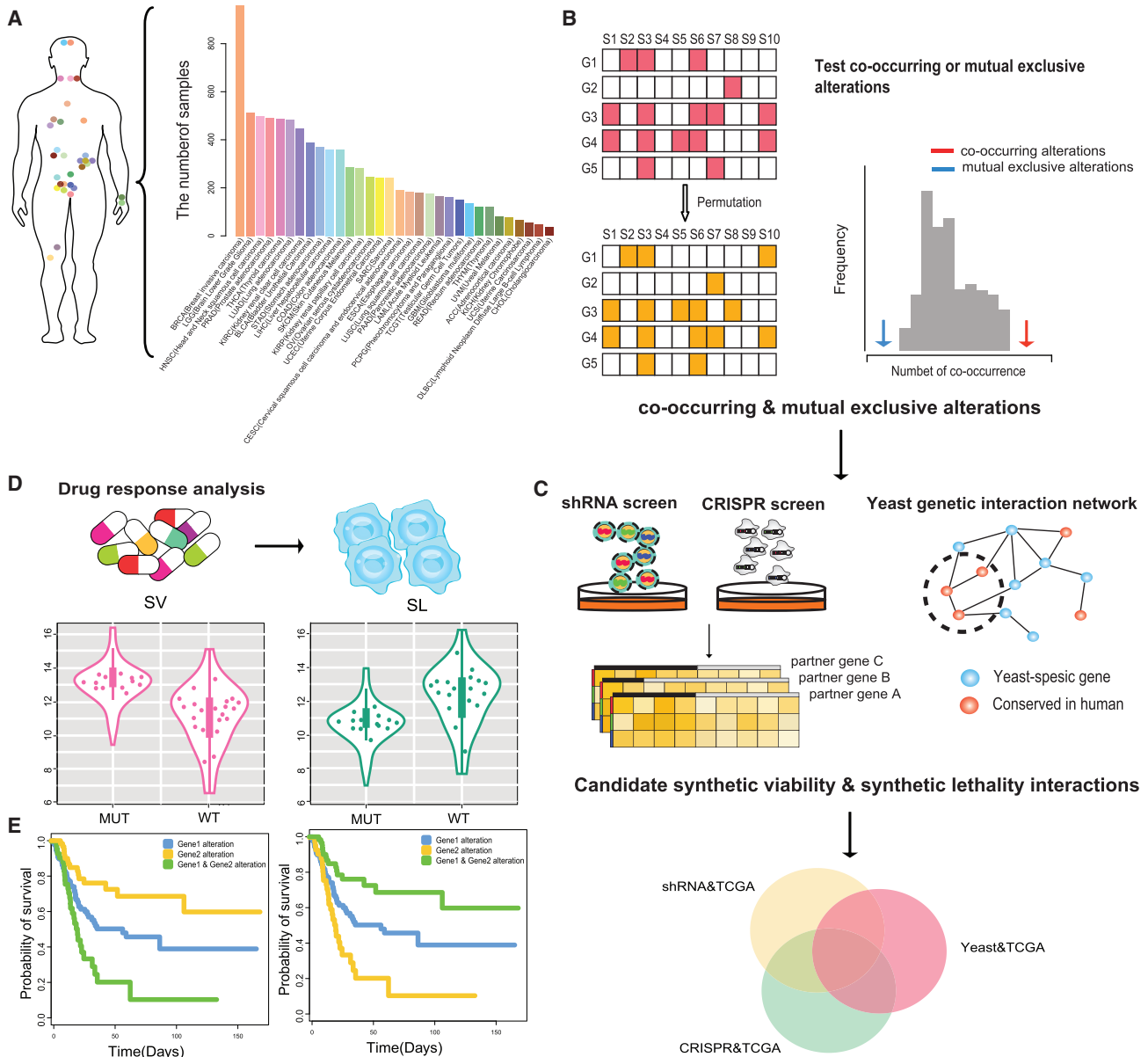
Figure 1. Workflow of the Present Study

(A) Statistics of 32 cancers in TCGA. (B) Identification of co-occurring and mutual exclusive interactions across 32 cancer types. (C) Identification of candidate genetic interactions using shRNA, CRISPR screening data, and yeast genetic interaction data. (D) Prediction of biomarkers for cancer cell drug sensitivity and resistance. (E) Prognostic analysis.

Here, we calculated the percentage of the SV (SL) interactions that conformed to our hypothesis relative to all significant SV (SL) interactions and defined it as consistency ratio. The consistency ratios of four drugs (including erlotinib, lapatinib, nilotinib, and paclitaxel) were simultaneously screened in four drug datasets, which are shown in Figures 3A and 3B. In the NCI60 dataset, no SV and SL interactions exhibited the expected drug response due to the small number of cell lines. As shown in Figure 3A, the SV interactions associated with

paclitaxel resistance had a higher consistency ratio in breast cancer, stomach cancer, and lung cancer. Overall, the response rate to paclitaxel is 50% as first-line chemotherapy, decreasing to 20%–30% when used as second- or third-line chemotherapy; nearly half of the patients with breast cancer do not respond to paclitaxel.[15] According to the present results, the breast cancer cell lines with alterations of *PSEN1* were related to paclitaxel resistance (Figure S4), which has been reported in esophageal cancer.[16]
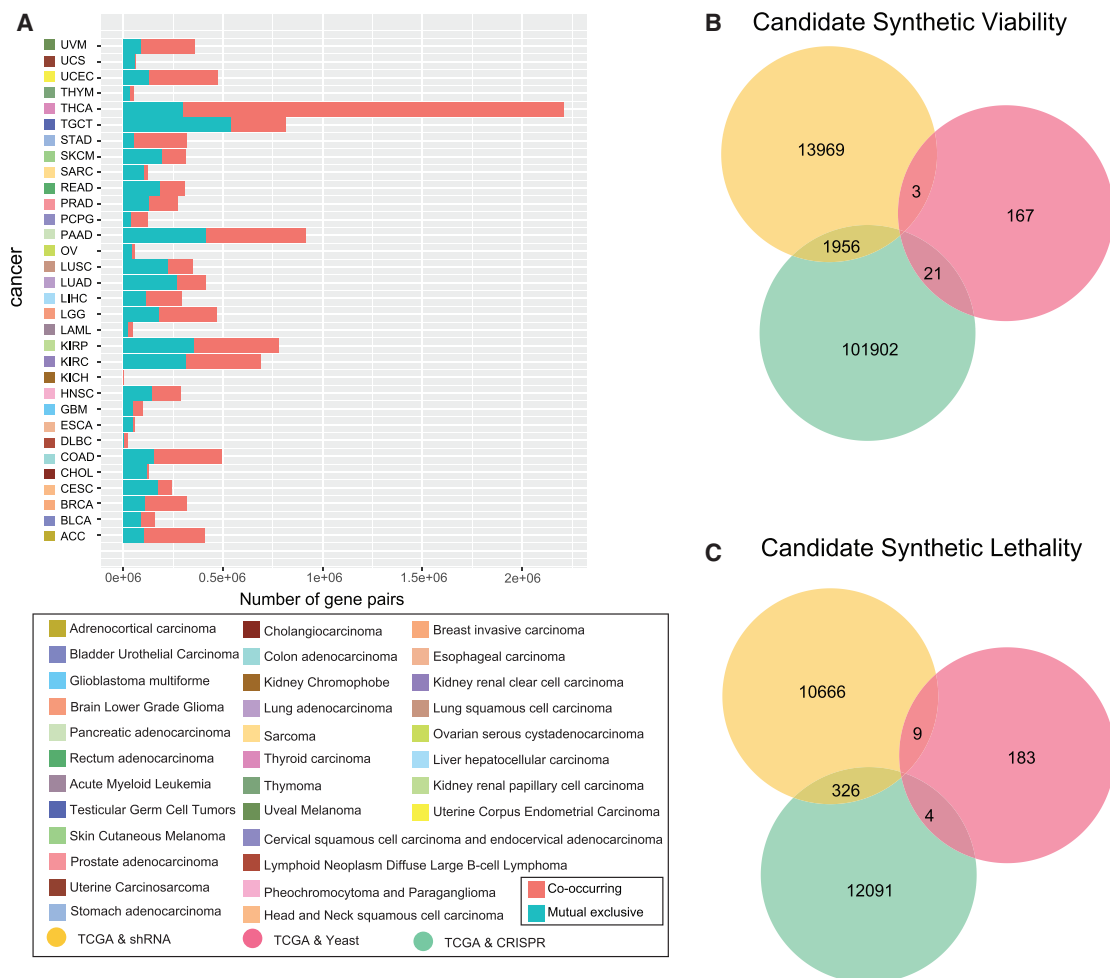
Figure 2. Identification of Candidate Synthetic Viable and Synthetic Lethal Interactions in Cancer

(A) Statistics of co-occurring and mutually exclusive gene pairs in different types of cancer. (B) Overlapping of the candidate synthetic viable interactions verified in the shRNA, CRISPR, and yeast datasets. (C) Overlapping of candidate synthetic lethal interactions verified in the shRNA, CRISPR, and yeast datasets.

Erlotinib is an oral *HER1/EGFR* tyrosine kinase inhibitor approved for patients with non-small-cell lung cancer. In the present results, erlotinib showed a high consistency ratio in lung cancer tissue (Figure 3B), and the consistency ratio was 100% in the CCLE, GDSC (AUC), and GDSC (LN_IC$_{50}$) datasets, p < 0.01 (Figure 3C). Cell lines with alterations of *EGFR* were more sensitive to erlotinib in the CCLE (Figure 3D), CTRP (Figure 3E), and GDSC (Figures 3F and 3G) datasets, which demonstrated the reliability of the present results. Together, these findings suggested that SV interactions are associated with drug resistance and that SL interactions are related to drug sensitivity.

Drug-Response-Related Genetic Interaction Network Shows Biological Characteristics

By integrating drug-resistance-related SV interactions and drug-sensitivity-related SL interactions, SV and SL networks, respectively, were constructed. The SV interaction network consisted of 2,516 genes and 4,529 target-partner gene relationships (Figure 4A), and the SL interac-

tion network consisted of 3,184 genes and 10,637 target-partner gene relationships (Figure 4B); within two networks, 80% of SV interactions and 90% of SL interactions had one-step (red edge) or two-step (pink edge) neighbors in the protein-protein interaction (PPI) network, which was not expected by random chance (p < 0.001). Random SV and SL networks were constructed by randomly selecting the same number of gene pairs from the cancer alteration profiles and counting the number of one- or two-step interactions that overlapped with the PPI network, which was performed 1,000 times. Also, we calculated the empirical p value. Both SL and SV networks showed scale-free characteristics (Figures 4C and 4D). The aforementioned results indicated that the SV and SL networks had biological functions.[17] The top 10 genes with the highest degrees in the SV and SL networks are shown in Figures 4E and 4F. Several drug targets, such as *AKT1*, *PARP1*, and *HDAC1*, had high degrees in the SV or SL network. Moreover, partner genes of these drug targets also had high degrees. For example, the *RHOA* partner gene in the SV network is a member of the Rho family of small GTPases,
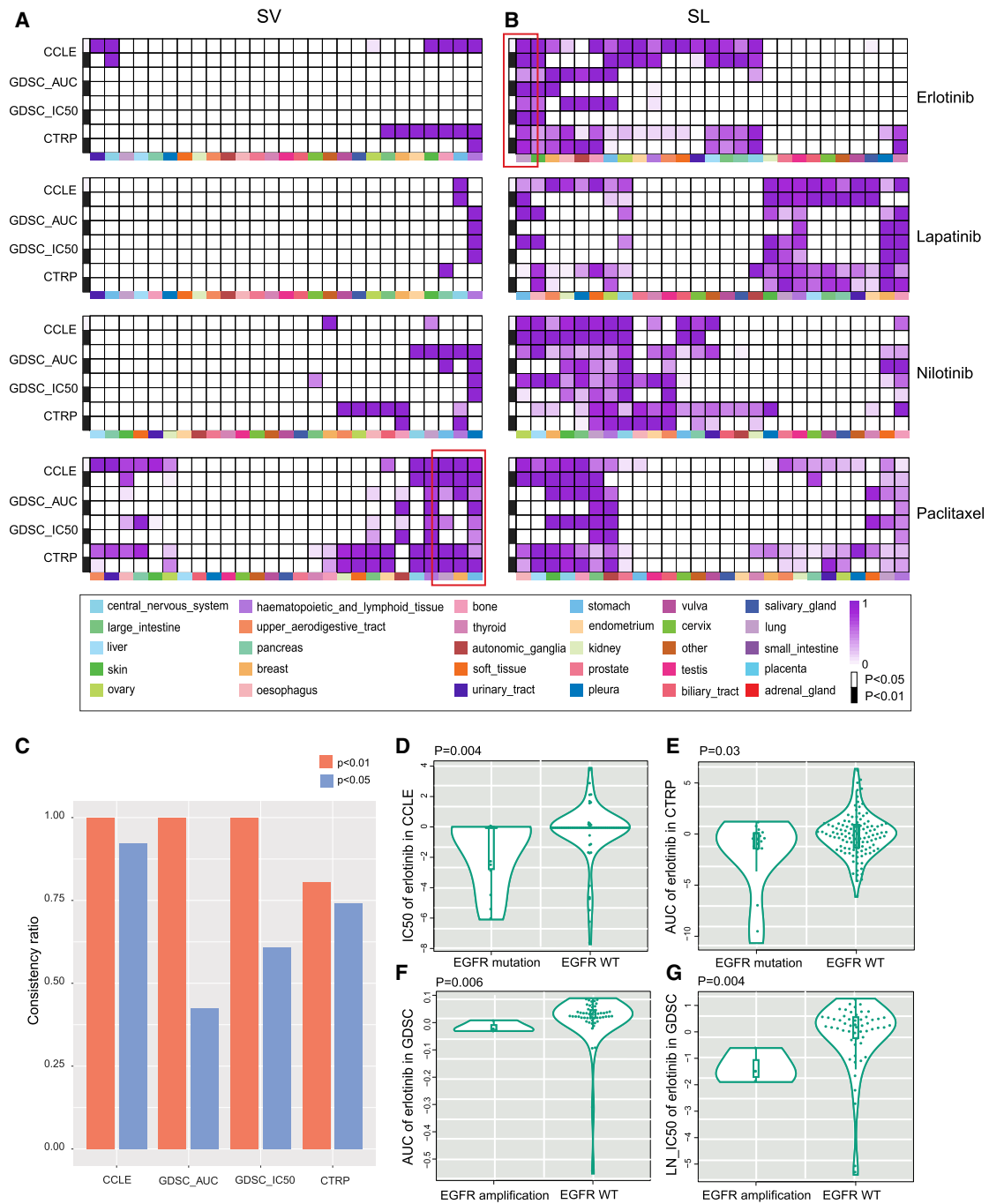
Figure 3. Genetic Interactions Related to Drug Response in Tissue-Specific Cell Lines

(A and B) Consistency ratio of (A) SV interactions related to drug resistance and (B) SL interactions related to drug sensitivity of four drugs ($p < 0.05$, Wilcoxon rank-sum test, white; and $p < 0.01$, Wilcoxon rank-sum test, black) in CCLE, CTRP, and GDSC datasets. Dark purple indicates a higher consistency ratio, and grids lacking purple indicate no detection. Different colors represent different tissues. The blank grid indicates that the consistency ratio cannot be tested because of a limited number of cell lines. (C) Consistency ratio of erlotinib in lung cancer tissue. (D–G) Cell lines with EGFR amplification or mutation are sensitive to the ERBB2-related drug, erlotinib, in CCLE, CTRP, and GDSC (AUC, LN_IC_{50}) datasets (Wilcoxon rank-sum test). (D) IC_{50} of erlotinib in CCLE. (E) AUC of erlotinib in CTRP. (F) AUC of erlotinib in GDSC. (G) LN_IC_{50} of erlotinib in GDSC.
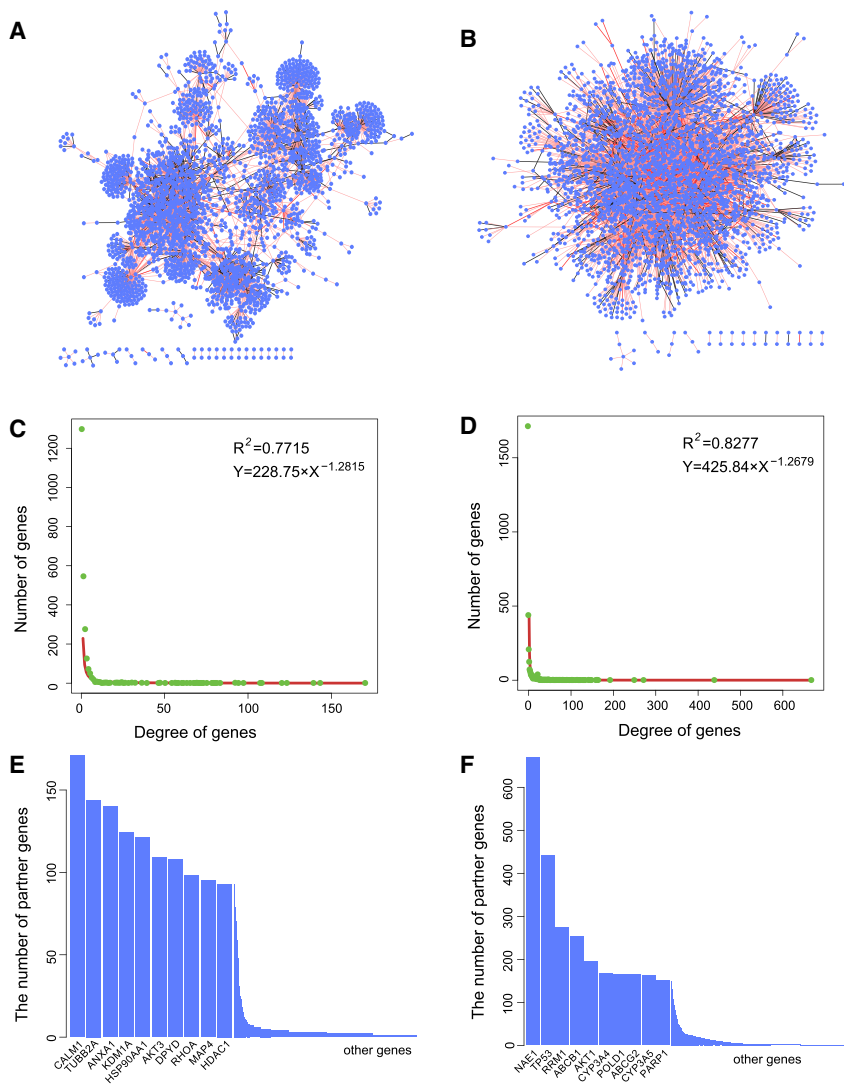
Figure 4. Analysis of Genetic Interaction Networks

(A) SV interaction network. (B) SL interaction network. Nodes depict genes, and edges represent SV interactions between genes. The light red line indicates that two genes with SV interactions also have direct protein-protein interactions, and the pink line indicates that two genes with SV or SL interactions have common neighbors in the protein-protein interaction network. (C) Distribution of the degree of genes in the SV network. (D) Distribution of the degree of genes in the SL network. (E) Histogram shows the degree of genes in the SV interaction network. (F) Histogram shows the degree of genes in the SL interaction network.

many genes interacted with *HDAC1*, whose alterations may induce resistance to a *HDAC1* inhibitor in cancer cells. The partner genes of *HDAC1* were significantly enriched in 35 pathways (hypergeometric test, p < 0.05; Figure S5), including Notch, Wnt, and mTOR signaling pathways (Figure 5A). Both *HDAC1* and the *DVL1* partner gene were important regulators in the Notch signaling pathway (Figure 5B), which governs the growth and proliferation of cancer cells. In the pharmacogenomic dataset of CTRP, the deletions of *DVL1* were related to *HDAC1* inhibitor (pandacostat) resistance in the cell lines of CNS tissues (p = 0.01, Wilcoxon rank-sum test; Figure 5C). Knockout of *HDAC1* by CRISPR demonstrated that cell lines with *DVL1* deletion survived better than those with WT *DVL1* (p = 0.003, Wilcoxon rank-sum test; Figure 5D). *HDAC1* is a member of the histone deacetylase/AcuC/AphA family whose main function is to remove acetyl from histone. *HDAC1* is an inhibitor of the CSL protein in the Notch pathway, which can form a transcriptionally active complex with the Notch1 intracellular domain (NICD) to activate downstream targets of the transcriptional suppressor family, such as HES1/5 and HERP (Figure 5B). If *HDAC1* is suppressed by pandacostat, transcriptional inhibition can be promoted. However, deletion of *DVL1* may compensate for transcriptional inhibition. *DVL1* inhibits the Notch1 receptor, and deletion of *DVL1* may upregulate Notch1, which produces anti-apoptotic signals by regulating the phosphatidylinositol 3-kinase (PI3K)-PKB/Akt pathway.[20] To confirm this conjecture, differentially expressed genes were examined in liver hepatocellular carcinoma (LIHC) samples containing both deletions of *HDAC1* and *DVL1* versus LIHC samples with WT *HDAC1* and *DVL1*. Notch1 was significantly upregulated in samples containing deletions of both *HDAC1* and *DVL1* (p = 0.001, t test; Figure S6). Thus, when *HDAC1* inhibitors promote transcriptional inhibition, the deletion of *DVL1* may lead to an increase in the activity of the Notch receptor, which promotes cell proliferation by regulating the mitogen-activated protein kinase (MAPK) signaling pathway. To test whether this SV interaction affects patient survival, a log-rank

which functions as molecular switches in signal transduction cascades. The present study predicted that alterations of *RHOA* may induce resistance to drugs, such as fluorouracil, in cancer cells. Several studies have reported that alterations of *RHOA* mediate tumor invasion and drug resistance in cancer. Misek et al.[18] found that *RHOA* contributes to *BRAF* inhibitor resistance in Sox9High/Sox10Low melanoma cells by regulating activation of the RhoA family signaling pathway. In addition, partner genes with high degrees in the SL network, such as *NAE1* and *TP53*, also played an important role in the development and progression of cancer. The present work predicted that alterations of *TP53* were related to drug sensitivity, which has been validated in other studies.[19] Thus, the partner genes of drug targets in the SL and SV networks may be regarded as candidate biomarkers for cancer therapies.

SV Interactions Induce Drug Resistance

Among the genes with the top 10 highest degrees in the SV network, *HDAC1* is a key drug target approved by the FDA. In the SV network,
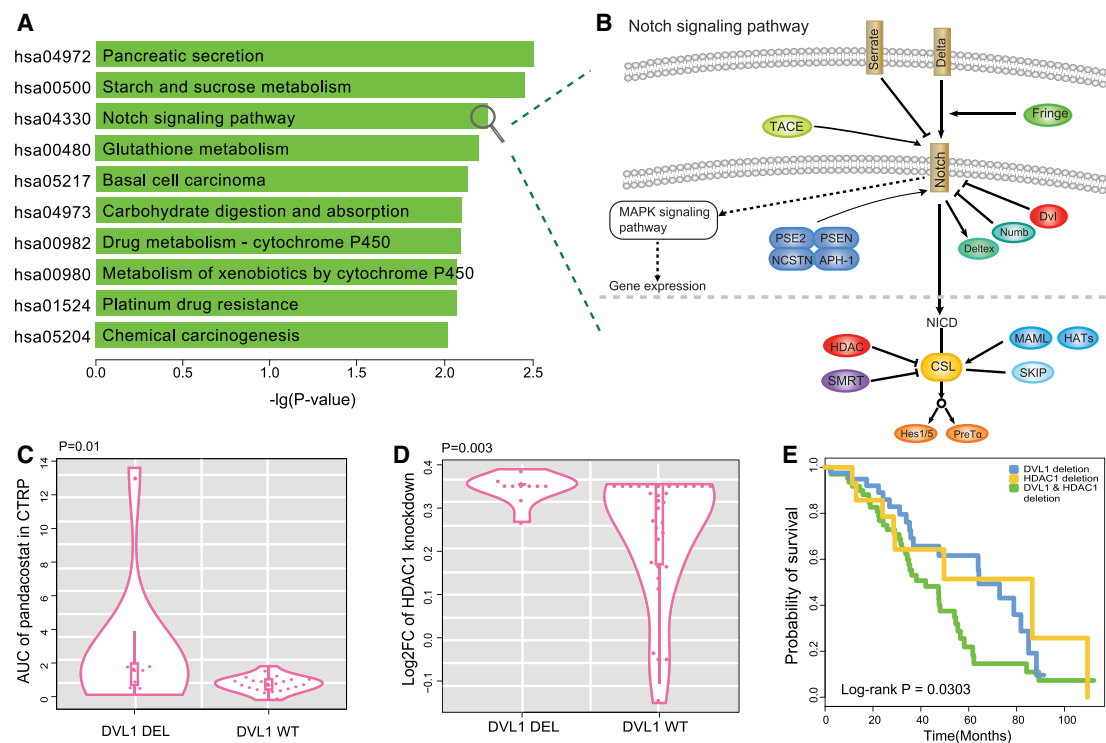
Figure 5. Functional Analysis of Drug-Resistance-Related SV Interactions

(A) KEGG pathways enriched with the partner genes of *HDAC1* in the SV interaction network. The y axis represents significantly enriched pathways, and the x axis is the negative $\log_{10}$-transformed hypergeometric test p value. (B) Notch signaling pathway. The pathway is composed of receptors, ligands, and the CSL DNA-binding protein. *DVL1* is located in functional extracellular domains, acting as an inhibitor for Notch receptors, and *HDAC1* plays a key role in regulating transcriptional activity by inhibiting CSL. (C) Deletions of *DVL1* were related to pandacostat resistance in cell lines compared to wild-type *DVL1* (p = 0.01, Wilcoxon rank-sum test). (D) Deletions of *DVL1* showed higher viability in *HDAC1* knockout cell lines compared to wild-type *DVL1* (p = 0.003, Wilcoxon rank-sum test). (E) The Kaplan-Meier overall survival analysis of patients in three groups as follows: *HDAC1* deletion, *DVL1* deletion, and *HDAC1* and *DVL1* deletion (p = 0.0303, log-rank test).

test was applied to analyze the survival data of patients from TCGA. In ovarian cancer, patients with deletions of both *HDAC1* and *DVL1* showed worse survival than patients with deletions of either *HDAC1* or *DVL1* (p = 0.0303, log-rank test).

SL Interactions Mediate Drug Sensitivity

AKT1, as a key gene encoding serine and threonine protein kinase in the PI3K/AKT pathway, interacted with gefitinib. Among genes that interacted with drugs in the SL network, *AKT1* had the second largest number of partner genes (Figure S7). Pathway enrichment analysis showed that the partner genes of *AKT1* were mainly enriched in the Neurotrophin signaling pathway, p53 signaling pathway, apoptosis pathway, and cell cycle (hypergeometric distribution, p < 0.05; Figures 6A and 6B). These results indicated that the alterations of the partner genes, which had an SL effect with deletion of *AKT1*, were related to gefitinib sensitivity in cancer cells. For example, *ARHGDIA*-deficient cell lines were related to gefitinib sensitivity in hematopoietic and lymphoid tissues (p = 0.05, Wilcoxon rank-sum test; Figure 6C). Knockout of *AKT1* by CRISPR in cells with *ARHGDIA* deletion showed worse survival than cells with WT *ARHGDIA* (p = 0.02, Wilcoxon rank-sum test; Fig-

ure 6D). Both *AKT1* and *ARHGDIA* participated in the Neurotrophin signaling pathway (Figure 6B), which regulates the growth, development, survival, and repair of the nervous system. *AKT1*, as the direct downstream target of PI3K, plays an important role in maintaining the survival of cerebellar cells.[21] *ARHGDIA*, a downstream gene of *p75NTR*, has been proven to regulate the survival of neurons.[22] Monje et al.[23] confirmed that neuron-mediated developmental mechanisms are recapitulated in cancers. Cancer patient prognostic data were used to test whether co-deletions of *AKT1* and *ARHGDIA* are beneficial to patient survival. As shown in Figure 6E, the prognosis of lung squamous cell carcinoma patients with *AKT1* and *ARHGDIA* deletions was better than that for patients with single-gene deletions (p = 0.0582, log-rank test). Thus, alterations of *ARHGDIA* in cancer cells are beneficial for gefitinib treatment, and patients with co-alterations of *AKT1* and *ARHGDIA* have better prognosis.

SL Interactions Suggest New Therapeutic Strategy

SL interactions provide a conceptual framework for the development of cancer-specific drugs.[24] In the present results, alterations of partner genes involved in the SL network were related to drug
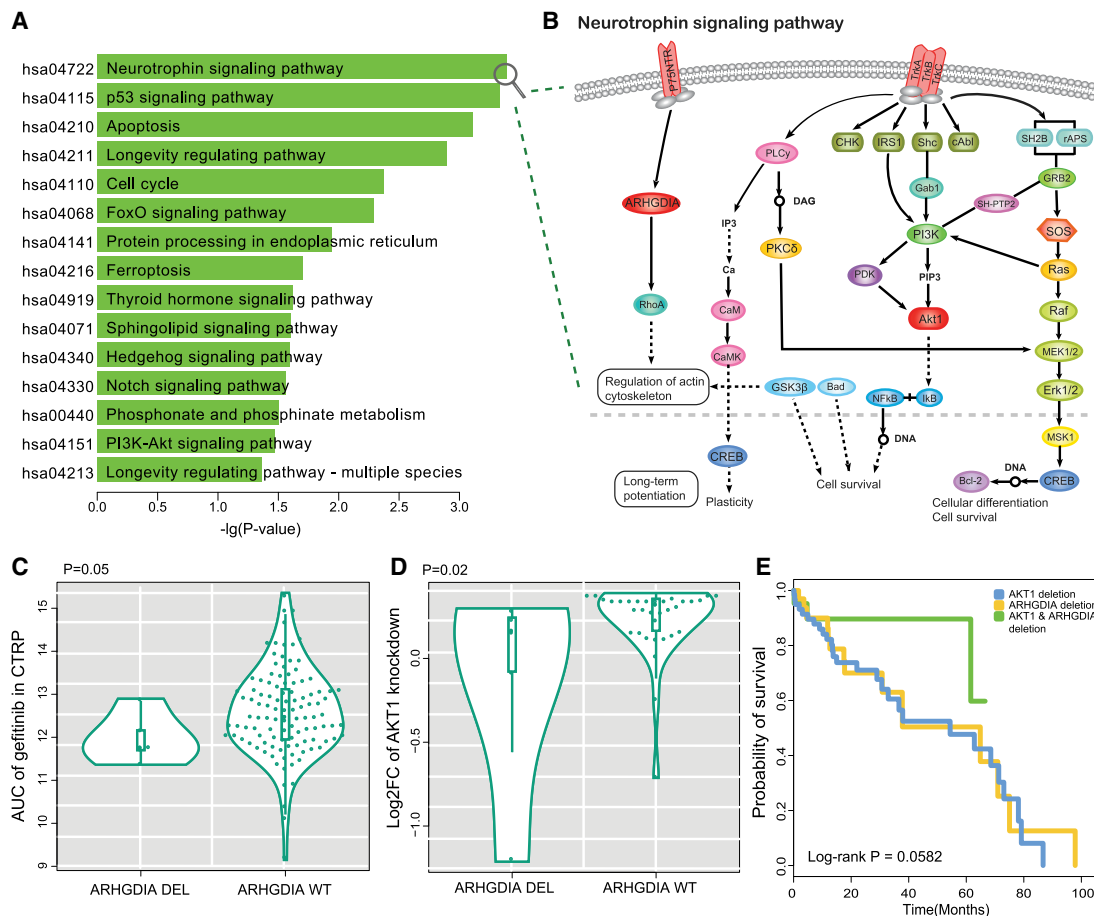
Figure 6. Functional Analysis of Drug-Sensitivity-Related SL Interaction Genes

(A) KEGG pathways enriched with the partner genes of *AKT1* in the SL interaction network. The y axis represents significantly enriched pathways, and the x axis represents the negative log$_{10}$-transformed hypergeometric test p value. (B) Neurotrophin signaling pathway is mainly regulated by two types of receptors, the Trk tyrosine kinase receptors and the p75 neurotrophin receptor (p75NTR). *AKT1* is the downstream effector of Trk, and *ARHGDIA* is the downstream gene of p75NTR. Both of these genes are important regulators for neuronal growth. (C) Deletions of *ARHGDIA* were related to gefitinib sensitivity in cell lines compared to wild-type *ARHGDIA* (p = 0.05, Wilcoxon rank-sum test). (D) Deletions of *ARHGDIA* showed lower viability in *AKT1* knockout cell lines compared to wild-type *ARHGDIA* (p = 0.02, Wilcoxon rank-sum test). (E) The Kaplan-Meier overall survival (OS) analysis of patients in three groups as follows: *AKT1* deletion, *ARHGDIA* deletion, and *AKT1* and *ARHGDIA* deletion (p = 0.0582, log-rank test).

sensitivity in tissue-specific cell lines, which indicated that drugs in the SL network may be new targeted therapeutics for specific cancer types (Figure 7A). To investigate whether drugs could be applied to specific cancers, the GI50 values for 90 drugs screened in breast cancer lines were evaluated.[25] As expected, breast cancer cell lines with *LAMP1* amplifications were related to paclitaxel sensitivity compared to WT cell lines (p = 0.02, Wilcoxon rank-sum test; Figure 7B). Knockout of *TUBB3* by CRISPR in cells with *LAMP1* amplifications showed worse survival than cells with WT *LAMP1* (p = 0.006, Wilcoxon rank-sum test; Figure 7C). These results indicated that breast cancer patients with amplifications of *LAMP1* may benefit from paclitaxel chemotherapy.

Database of Genetic Interaction in Cancer

In the present work, we report that CancerGeneticInteractiondb (CGIdb) is a comprehensive platform for genetic interaction in cancer. In addition to providing candidate SV and SL interactions predicted by the present methods, CancerGeneticInteractiondb also collected a large number of SV and SL interactions predicted by other studies (Table S1). With a user-friendly interface, the website is composed of four sections: Search, Browse, Data, and About. On the Search page, the user can enter genes of interest and search for SV or SL interactions containing those genes. Furthermore, the drug response of SV (SL) interactions can be viewed by clicking the Detail button. In the Browse page, users can search SV and SL interactions by tissue types, and this page also provides drug information. Users can download and upload data on the Data page. For user convenience, all SV and SL interactions are grouped by tissues and data sources. Additional details for CancerGeneticInteractiondb are provided on the About page (also see Supplemental Materials and Methods). The website for accessing the database is located at http://www.medsysbio.org/CGIdb.
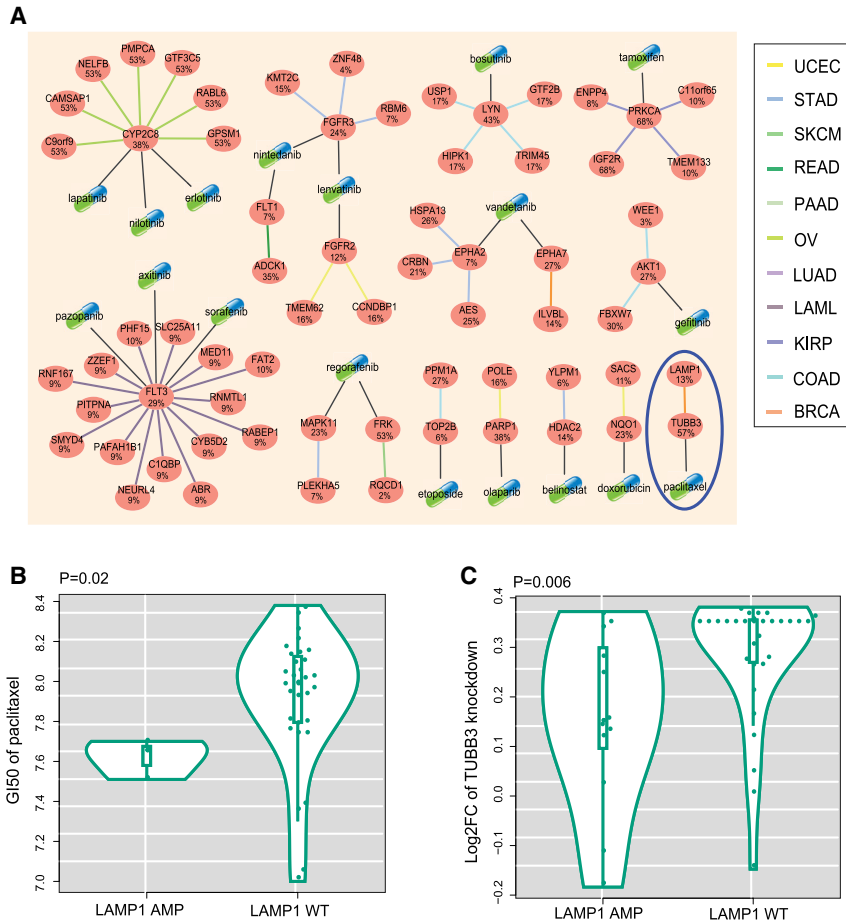
Figure 7. SL Interactions Suggest New Therapeutic Strategy

(A) Network of potential anti-cancer therapies. Ellipse nodes represent genes, and the gene alteration rates in the specific cancer types are shown below the gene symbols. Capsules represent drugs. The black lines between capsules and ellipse node depict drug-target relationship. The edge colors represent the cancer types within which the synthetic lethal interactions were detected. (B) Breast cancer cell lines with *LAMP1* amplifications had significantly lower GI50 of paclitaxel than breast cancer cell lines with wild-type *LAMP1* (p = 0.02, Wilcoxon rank-sum test). (C) Knockout of *TUBB3* by CRISPR in cancer cells with *LAMP1* amplification showed better survival than cells with wild-type *LAMP1* (p = 0.006, Wilcoxon rank-sum test).

patients and to make an improvement in overall survival, which was verified in the clinical prognostic data. Finally, the present study provided a comprehensive CGIdb database for a convenient retrieval of genetic interactions in cancer.

It is well known that amplifications and deletions correspond to the upregulation and downregulation of gene expression, respectively. However, the functional effect of mutation on drug response is unclear. Somatic mutations can be classified as frameshift mutations, missense mutations, and so on. It is common for researchers to assume that frameshift insertions and deletions (indels) are loss-of-function variants.[27] Missense mutations result in a single-nucleotide base change, which is difficult to use in predicting the effect on protein stability or PPI.[28,29] Investigating the functions of somatic mutation from structure, protein expression, and so on warrants our future work.

The different hypotheses of methods and original input data may result in low concordance of genetic interactions (Figures 2B and 2C). Ye et al.[30] collected and compared the SL interactions identified by five studies. The results show that very little SL interactions are overlapped among different studies, which provided an indispensable reason that the genetic interaction relationships found in each work were only a small part because of the different hypotheses. Thus, beyond the predictions from our work, we collected results of 16 studies for follow-up research (Table S1).

Numerous studies have focused on identifying genetic interactions by shRNA or CRISPR technology, which provide large-scale functional screens in cell lines for genetic interaction detection.[13,31–33] CRISPR, a new technology of genome editing, has been widely applied to loss-of-function screens for essential genes, which have uncovered both core and cell-line-specific fitness genes. Large-scale CRISPR screens have been performed, and they provide new data sources

DISCUSSION

The use of genetic data, pharmacological data, and functional screening datasets to gain valuable insights into the mechanism of drug sensitivity and drug resistance has been considered as the cornerstone of precision cancer medicine.[26] Cancer cells have frequent defects in specific genes that drive growth and metastasis, and alterations in different genes may induce a genetic interaction effect. Identification of genetic interaction with cancer-promoting genes represents a compelling approach for the development of cancer therapies. The present study developed a new strategy for identifying genetic interactions in cancer cells based on CNAs, mutations, and expression profiles across 8,580 samples from 32 cancer types from the TCGA dataset, which were further filtered applying the shRNA, CRISPR, and yeast datasets. The present results demonstrated that SV interactions induce drug resistance and that SL interactions mediate drug sensitivity in cancer cells. The results highlighted that *DVL1* deletion induces resistance to *HDAC1* inhibitors through disturbing the Notch signaling pathway, which was confirmed by gene expression profiles. Within SL interactions, alterations of partner genes were related to drug sensitivity in cancer cell lines, suggesting that these genes may be new biomarkers for cancer therapies, thereby maximizing the clinical effectiveness of drugs. An expected application of genetic interaction screens is to optimize the therapeutic regimen of

for identifying genetic interactions.[34–36] In the present study, the CRISPR screening dataset was added to further filter candidate SV and SL interactions. However, the shRNA or CRISPR studies were limited as follows: (1) lack of mutation information of cell lines—for example, the CCLE collection of 1,000 cell lines contained no acute myeloid leukemia (AML) cell line with IDH1 mutation, even though the rate of mutation in the AML patients is up to 10%;[37] and (2) the inevitable difference of tumor microenvironment between *in vitro* screening conditions and *in vivo* tumors. The present method used human primary tumor data to identify potential cancer-specific SV and SL interactions to complement the limitations of existing cell-line screening methods.

The distribution of drug response values in cell lines derived from different tissues was significantly different. Therefore drug sensitivity in cancer cell lines is tissue specific (Figure S1). By the drug sensitivity analysis on the tissue-specific cell-line datasets of the large-scale pharmacogenomic data, most identified SL interactions with significant drug sensitivity conform with the SL hypothesis and most identified SV interactions with significant drug resistance conform with the SV hypothesis (Figure 3). However, the fewer cell-line samples for specific cancer types limited the statistical power. A new approach is urgently needed to estimate and remove the viability of drug sensitivity of different cells.

The purpose of the present work was to elucidate the mechanisms of drug resistance and drug sensitivity to exploit cancer-specific vulnerabilities and expand the scope of precision oncology. The partner genes of SL or SV identified in the present work had a close functional relationship with drug-interacted targets shown in Figures 4A and 4B. *HDAC1*, with the highest degree in the SV network (Figure 4A), was selected as a case study in the present work. *HDAC1* is a transcription inhibitor and plays a key role in the regulation of eukaryotic gene expression accompanied by decreasing activity of transcription.[38] Many studies have shown that *HDAC1* may form a co-inhibitory complex with other chromatin regulatory proteins to regulate important biological processes, such as cell cycle and cell differentiation.[39] At present, several *HDAC* inhibitors have been approved by the FDA for treatment of cancer. In the present results, several partner genes of *HDAC1* induced drug resistance to *HDAC1* inhibitors and participated in the same pathway with *HDAC1*. For example, *DVL1* and *HDAC1* participated in the Notch signaling pathway, which relies on the proteolytic cascade to release the transcriptional activity of the intracellular domain.[40] When *HDAC1* is inhibited by drugs to promote transcriptional inhibition in the Notch signaling pathway, *DVL1* deletion activates the MAPK signaling pathway by the Notch receptor and further promotes cell proliferation. Thus, interactions between *HDAC1* inhibition and *DVL1* deletion may enhance cancer cell survival. Therefore, *DVL1* deletion may induce resistance to *HDAC1* inhibitors in cancer cells by disturbing the Notch signaling pathway, which was confirmed by the gene expression profile. *NOTCH1* was significantly upregulated in samples with deletions of both *HDAC1* and *DVL1* compared to samples with single-gene deletions. Importantly, ovarian cancer patients with deletions of both

HDAC1 and *DVL1* had worse survival than patients with deletions of either *HDAC1* or *DVL1*. Specific mechanisms were conjectured by bioinformatics statistics, but they should be verified by subsequent cell and animal experiments.

PARP1, as a well-known drug target, had high degrees in the SL network. *PARP* inhibitors are the first clinically approved drugs designed to exploit SL.[8] Therefore, it would be meaningful to expand the application of *PARP* inhibitors in the clinic. The present analysis identified the well-known SL interaction between *PARP* inhibition and *BRCA1* mutation. In the present results, alterations of the partner genes of PARP in the SL network were significantly related to the sensitivity of *PARP* inhibitors in cancer cell lines. For example, the mutation of *POLE* mediated sensitivity to the *PARP* inhibitor veliparib in lung cancer cell lines, which was previously characterized by Safiri et al.,[41] In addition, there were many newly discovered alterations of genes that enhanced cell-line sensitivity to *PARP* inhibitors, such as *DOCK3*, *CLASP2*, *CMTM6*, and *MAP4*. The new discoveries warrant additional detailed research, and experimental verification (cell and animal models) is required to unravel the mechanism of drug sensitivity, which will be the major focus in our future studies.

MATERIALS AND METHODS
Identification of Candidate Genetic Interactions
The present study used mutation, CNA, and expression profiles derived across 32 cancer types from the TCGA consortium, including 8,580 patient samples. High-frequency somatic mutations were identified via the MutSig algorithm from whole-exome sequencing data.[42] CNAs were determined by Genomic Identification of Significant Targets in Cancer (GISTIC).[43] Similar to Shi et al.,[44] the $\log_2$ ratio cutoff values >0.25 were defined as amplification. In contrast, the $\log_2$ ratio cutoff values < −0.25 were defined as deletion. To further optimize CNAs, Pearson's correlation test was used to select genes whose copy number and expression value had a positive correlation coefficient. The p value was corrected using the Benjamini-Hochberg (BH) correction for multiple tests, and the results with an FDR < 0.05 were reported in the present study.

Integration of mutation and CNA data allowed generation of alteration profiles for 32 cancer types. The alteration profile columns represent samples, and the rows represent altered genes (Figure 1B). To discover gene pairs with significant co-occurrence or mutual exclusivity, a permutation strategy that controls for alteration heterogeneity within and across tumor samples was applied.[3] Permutated genomic alteration matrices were implemented by the permatswap function in the R package vegan (http://vegan.r-forge.r-project.org/), which maintains the total number of alterations for each gene across samples as well as the total number of alterations per sample. In addition, due to the bias in alteration frequencies among cancer types, perturbations for different cancer samples were separately performed. The proportion of 10,000 permutations in which the observed co-occurrence was higher (P_{co}) or lower (P_{me}) than in the real data was taken as an empirical p value (Figure 1B). To avoid bias of limited number of permutation experiments,

the hypergeometric test was used to further filter co-occurrence and mutually exclusive gene pairs. Bonferroni was used to correct the p value, and FDR < 0.001 was reported as significant. In addition, we have removed the co-occurring gene pairs from the same region of CNAs.

shRNA-Based Screening

shRNA data were downloaded from The Project Achilles database, which provides shRNA depletion scores from pooled genomic library tests across 216 cancer cell lines. A gene-level composite score (shRNA score) was obtained by Analytic Technique for Assessment of RNAi by Similarity (ATARiS) (https://www.broadinstitute.org/achilles/datasets/all). Higher shRNA scores indicated enhancement of cell viability. The genetic background of these cell lines was obtained from the Cancer Cell Line Encyclopedia (CCLE; https://www.broadinstitute.org/ccle/home), including mutation and copy number variation. Genomic alteration profiles of cell lines were built. In the genomic alteration profile, the row represented the gene, and the column represented the cell line. For each pair of genes (G_1 and G_2), shRNA scores of G_1 knockdown were compared between the cell lines with and without G_2 alterations using one-sided Wilcoxon rank-sum test (Figure 1C).

CRISPR-Based Screening

High-throughput CRISPR-Cas9 screening data, including 63 cancer cell lines, were downloaded from the GenomeCRISPR database. To compare fitness phenotypes of CRISPR-Cas9 screens, all screens were reanalyzed through the Bayesian Analysis of Gene Essentiality (BAGEL).[45] The Bayesian factor serves as a quantitative measure of the essentiality of the gene in the screen, in which higher Bayesian factors indicate more essential genes. Negative Bayesian factors were used as fitness scores. A lower fitness score represents higher confidence that a given gene knockout causes a decrease in fitness.[17] CNA and mutation data of cell lines were obtained from CCLE. In cell lines with knockout of G_1, a one-sided Wilcoxon rank-sum test was used to test whether the fitness scores were significantly higher or lower in cell lines with and without alterations of G_2 (Figure 1C).

Yeast-Based Comparative Genomic Strategy

A genome-scale genetic interaction map in yeast was obtained from Costanzo et al.,[46] who generated quantitative genetic interaction profiles for approximately 75% of all genes in budding yeast. Deshpande et al.[47] developed a comparative genomic algorithm to identify cancer-related genetic interactions. The present study used a cutoff of $\varepsilon > 0.08$ with $p < 0.05$ (yeast p value) to detect yeast SV interactions and $\varepsilon < -0.08$ with $p < 0.05$ to detect yeast SL interactions (ε is a measure of genetic interaction strength) (Figure 1C). A positive ε represents an increase in cell fitness, and a negative ε represents a decrease in cell fitness. InParanoid7 was applied to map yeast genes to human genes. Only 1:1 orthologs were used in the present study.

Lastly, the co-occurrence and mutual exclusivity gene pairs verified in at least one type of the aforementioned three datasets were selected as candidate genetic interactions (Figure 1C).

Identifying SV and SL Interactions Related to Drug Response

The original pharmacological screening data were downloaded from CCLE,[48] Genomics of Drug Sensitivity in Cancer (GDSC),[49] Broad CTRP (https://www.broadinstitute.org/ctrp/), and NCI60.[50,51] The CCLE dataset contained IC_{50} values of 24 anti-cancer compounds across 501 cancer cell lines derived from 23 cancer types. The GDSC data contained IC_{50} values and the area under the drug inhibition curve (AUC) of 251 drugs detected in 1,012 cell lines from 30 cancer types. The CTRP data contained the AUC of 481 small molecules used to treat 823 cancer cell lines from 23 cancer types. The NCI60 data contained the concentration required for cell growth inhibition by 50% (GI50) of 4,329 anti-cancer compounds across 60 cancer cell lines. The information of drug and targeted genes or affected genes was collected from Drug Bank,[52] ChEMBL,[53] CCLE, and Catalogue of Somatic Mutations in Cancer (COSMIC).[54]

For a G_1 and G_2 gene pair, the IC_{50} (or AUC or GI50) of cell lines carrying G_2 alterations and cell lines containing WT G_2 were compared using a one-sided Wilcoxon rank-sum test if G_1 in the SV pair was targeted by a drug. Only cancer types with more than 3 cell lines were included for statistical analyses.

Generation of Genetic Interaction Network and Functional Analysis

The network of SV related to drug resistance and the network of SL related to drug sensitivity were constructed. Cytoscape software was used to visualize the networks (https://cytoscape.org/). PPI data was downloaded from the Pathway Commons database.[55] A hypergeometric distribution model was used to test whether the partner genes in the SV (SL) network were significantly enriched in biological pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.[56]

Prognosis Analysis

The mutation and CNA profiles from TCGA were analyzed to examine the prognostic value embedded in the SV and SL networks. The patients were divided into three groups according to the status of genes in the SV (SL) interactions as follows: G_1 alteration only, G_2 alteration only, and G_1 and G_2 alteration groups. The overall survival time of the three patient groups for specific cancer types was tested by using log-rank test, and the results were represented by Kaplan-Meier plots.

CGIdb Construction

CGIdb is freely available at http://www.medsysbio.org/CGIdb with no registration or login. CGIdb was implemented by HTML and Django, which is a high-level Python web framework. The dynamic backend of CGIdb was realized through the Linux, Apache, MySQL, and Python (LAMP) architecture. The interface component was designed using the Cascading Styling Sheet (CSS) and JavaScript, which have been tested in Google Chrome and Firefox browsers.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2019.07.003.

AUTHOR CONTRIBUTIONS

Y.G. and Z.G. guided the research study. Y.H., C.W., and Q.D. completed data processing, analysis, and draft of the paper. C.W. and Z.Z. constructed the database. T.C., F.Y., Y.L., and B.C. extracted data and visualized the results. L.Q., W.Z., and H.L. provided valuable suggestions for the manuscript.

ACKNOWLEDGMENTS

REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365, 1687–1717.

2. Santhosh, S., Kumar, P., Ramprasad, V., and Chaudhuri, A. (2015). Evolution of targeted therapies in cancer: opportunities and challenges in the clinic. Future Oncol. 11, 279–293.

3. Park, S., and Lehner, B. (2015). Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. Mol. Syst. Biol. 11, 824.

4. Sinha, S., Thomas, D., Chan, S., Gao, Y., Brunen, D., Torabi, D., Reinisch, A., Hernandez, D., Chan, A., Rankin, E.B., et al. (2017). Systematic discovery of mutation-specific synthetic lethals by mining pan-cancer human primary tumor data. Nat. Commun. 8, 15580.

5. Gu, Y., Wang, R., Han, Y., Zhou, W., Zhao, Z., Chen, T., Zhang, Y., Peng, F., Liang, H., Qi, L., et al. (2018). A landscape of synthetic viable interactions in cancer. Brief. Bioinform. 19, 644–655.

6. Ashworth, A., Lord, C.J., and Reis-Filho, J.S. (2011). Genetic interactions in cancer progression and treatment. Cell 145, 30–38.

7. Li, X.J., Mishra, S.K., Wu, M., Zhang, F., and Zheng, J. (2014). Syn-lethality: an integrative knowledge base of synthetic lethality towards discovery of selective anticancer therapies. Biomed Res. Int. 2014, 196034.

8. Lord, C.J., and Ashworth, A. (2017). PARP inhibitors: synthetic lethality in the clinic. Science 355, 1152–1158.

9. Lin, K.K., Harrell, M.I., Oza, A.M., Oaknin, A., Ray-Coquard, I., Tinker, A.V., Helman, E., Radke, M.R., Say, C., Vo, L.T., et al. (2019). BRCA reversion mutations in circulating tumor DNA predict primary and acquired resistance to the PARP inhibitor rucaparib in high-grade ovarian carcinoma. Cancer Discov. 9, 210–219.

10. Hu, H.M., Zhao, X., Kaushik, S., Robillard, L., Barthelet, A., Lin, K.K., Shah, K.N., Simmons, A.D., Raponi, M., Harding, T.C., and Bandyopadhyay, S. (2018). A quantitative chemotherapy genetic interaction map reveals factors associated with PARP inhibitor resistance. Cell Rep. 23, 918–929.

11. Ochoa, S., Martinez-Perez, E., Zea, D.J., Molina-Vila, M.A., and Marino-Buslje, C. (2017). Co-mutation and exclusion analysis in human tumors, a means for cancer biology studies and treatment design. bioRxiv, https://doi.org/10.1101/182501.

12. Unni, A.M., Lockwood, W.W., Zejnullahu, K., Lee-Lin, S.Q., and Varmus, H. (2015). Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. eLife 4, e06907.

13. Thompson, J.M., Nguyen, Q.H., Singh, M., and Razorenova, O.V. (2015). Approaches to identifying synthetic lethal interactions in cancer. Yale J. Biol. Med. 88, 145–155.

14. Wang, R., Han, Y., Zhao, Z., Yang, F., Chen, T., Zhou, W., Wang, X., Qi, L., Zhao, W., Guo, Z., and Gu, Y. (2017). Link synthetic lethality to drug sensitivity of cancer cells. Brief. Bioinform. Published online December 28, 2017. https://doi.org/10.1093/bib/bbx172.

15. Jiang, Y.Z., Yu, K.D., Peng, W.T., Di, G.H., Wu, J., Liu, G.Y., and Shao, Z.M. (2014). Enriched variations in TEKT4 and breast cancer resistance to paclitaxel. Nat. Commun. 5, 3802.

16. Meng, F., Qian, L., Lv, L., Ding, B., Zhou, G., Cheng, X., Niu, S., and Liang, Y. (2016). miR-193a-3p regulation of chemoradiation resistance in oesophageal cancer cells via the PSEN1 gene. Gene 579, 139–145.

17. Rauscher, B., Heigwer, F., Henkel, L., Hielscher, T., Voloshanenko, O., and Boutros, M. (2018). Toward an integrated map of genetic interactions in cancer cells. Mol. Syst. Biol. 14, e7656.

18. Misek, S.A., Appleton, K.M., Dexheimer, T.S., Lisabeth, E.M., Lo, R., Larsen, S.D., Gallo, K., and Neubig, R.R. (2018). Rho-mediated MRTF and YAP1 activation contributes to BRAF inhibitor resistance in Sox9High/Sox10Low melanoma cells. bioRxiv, https://doi.org/10.1101/381806.

19. Liu, Y., Zhang, X., Han, C., Wan, G., Huang, X., Ivan, C., Jiang, D., Rodriguez-Aguayo, C., Lopez-Berestein, G., Rao, P.H., et al. (2015). TP53 loss creates therapeutic vulnerability in colorectal cancer. Nature 520, 697–701.

20. Nair, P., Somasundaram, K., and Krishna, S. (2003). Activated Notch1 inhibits p53-induced apoptosis and sustains transformation by human papillomavirus type 16 E6 and E7 oncogenes through a PI3K-PKB/Akt-dependent pathway. J. Virol. 77, 7106–7112.

21. Chao, M.V. (2003). Neurotrophins and their receptors: a convergence point for many signalling pathways. Nat. Rev. Neurosci. 4, 299–309.

22. Ciavardelli, D., Silvestri, E., Del Viscovo, A., Bomba, M., De Gregorio, D., Moreno, M., Di Ilio, C., Goglia, F., Canzoniero, L.M., and Sensi, S.L. (2010). Alterations of brain and cerebellar proteomes linked to Aβ and tau pathology in a female triple-transgenic murine model of Alzheimer's disease. Cell Death Dis. 1, e90.

23. Venkatesh, H., and Monje, M. (2017). Neuronal activity in ontogeny and oncology. Trends Cancer 3, 89–112.

24. Reinhardt, H.C., Jiang, H., Hemann, M.T., and Yaffe, M.B. (2009). Exploiting synthetic lethal interactions for targeted cancer therapy. Cell Cycle 8, 3112–3119.

25. Daemen, A., Griffith, O.L., Heiser, L.M., Wang, N.J., Enache, O.M., Sanborn, Z., Pepin, F., Durinck, S., Korkola, J.E., Griffith, M., et al. (2013). Modeling precision treatment of breast cancer. Genome Biol. 14, R110.

26. Berlow, N., Haider, S., Wan, Q., Geltzeiler, M., Davis, L.E., Keller, C., and Pal, R. (2014). An integrated approach to anti-cancer drug sensitivity prediction. IEEE/ACM Trans. Comput. Biol. Bioinform. 11, 995–1008.

27. Hu, J., and Ng, P.C. (2012). Predicting the effects of frameshifting indels. Genome Biol. 13, R9.

28. Pandurangan, A.P., Ochoa-Montaño, B., Ascher, D.B., and Blundell, T.L. (2017). SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res. 45 (W1), W229–W235.

29. Brender, J.R., and Zhang, Y. (2015). Predicting the effect of mutations on protein-protein binding interactions through structure-based interface profiles. PLoS Comput. Biol. 11, e1004494.

30. Ye, H., Zhang, X., Chen, Y., Liu, Q., and Wei, J. (2016). Ranking novel cancer driving synthetic lethal gene pairs using TCGA data. Oncotarget 7, 55352–55367.

31. Vizeacoumar, F.J., Arnold, R., Vizeacoumar, F.S., Chandrashekhar, M., Buzina, A., Young, J.T., Kwan, J.H. Sayad, A., Mero, P., Lawo, S., et al. (2013). A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. Mol. Syst. Biol. 9, 696.

32. Wang, T., Yu, H., Hughes, N.W., Liu, B., Kendirli, A., Klein, K., Chen, W.W., Lander, E.S., and Sabatini, D.M. (2017). Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. Cell 168, 890–903.e15.

33. Han, K., Jeng, E.E., Hess, G.T., Morgens, D.W., Li, A., and Bassik, M.C. (2017). Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat. Biotechnol. 35, 463–474.

34. Shen, J.P., Zhao, D., Sasik, R., Luebeck, J., Birmingham, A., Bojorquez-Gomez, A., Licon, K., Klepper, K., Pekin, D., Beckett, A.N., et al. (2017). Combinatorial

CRISPR-Cas9 screens for de novo mapping of genetic interactions. Nat. Methods *14*, 573–576.

35. Bertomeu, T., Coulombe-Huntington, J., Chatr-Aryamontri, A., Bourdages, K.G., Coyaud, E., Raught, B., Xia, Y., and Tyers, M. (2017). A high-resolution genome-wide CRISPR/Cas9 viability screen reveals structural features and contextual diversity of the human cell-essential proteome. Mol. Cell. Biol. *38*, 00302-17.

36. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

37. Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al.; Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

38. Datta, M., Staszewski, O., Raschi, E., Frosch, M., Hagemeyer, N., Tay, T.L., Blank, T., Kreutzfeldt, M., Merkler, D., Ziegler-Waldkirch, S., et al. (2018). Histone deacetylases 1 and 2 regulate microglia function during development, homeostasis, and neurodegeneration in a context-dependent manner. Immunity *48*, 514–529.e6.

39. Micheli, L., D'Andrea, G., Leonardi, L., and Tirone, F. (2017). HDAC1, HDAC4, and HDAC9 bind to PC3/Tis21/Btg2 and are required for its inhibition of cell cycle progression and cyclin D1 expression. J. Cell. Physiol. *232*, 1696–1707.

40. Kovall, R.A., Gebelein, B., Sprinzak, D., and Kopan, R. (2017). The canonical Notch signaling pathway: structural and biochemical insights into shape, sugar, and force. Dev. Cell *41*, 228–241.

41. Safiri, S., and Ayubi, E. (2017). Smoking history predicts sensitivity to PARP inhibitor veliparib in patients with advanced non-small cell lung cancer: methodological issues. J. Thorac. Oncol *12*, e130–e131.

42. Forrest, W.F., and Cavet, G. (2007). Comment on "The consensus coding sequences of human breast and colorectal cancers." Science *317*, 1500.

43. Beroukhim, R., Getz, G., Nghiemphu, L., Barretina, J., Hsueh, T., Linhart, D., Vivanco, I., Lee, J.C., Huang, J.H., Alexander, S., et al. (2007). Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc. Natl. Acad. Sci. USA *104*, 20007–20012.

44. Shi, Z.-Z., Jiang, Y.-Y., Hao, J.-J., Zhang, Y., Zhang, T.-T., Shang, L., Liu, S.-G., Shi, F., and Wang, M.-R. (2014). Identification of putative target genes for amplification within 11q13.2 and 3q27.1 in esophageal squamous cell carcinoma. Clin. Transl. Oncol. *16*, 606–615.

45. Hart, T., and Moffat, J. (2016). BAGEL: a computational framework for identifying essential genes from pooled library screens. BMC Bioinformatics *17*, 164.

46. Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. Science *327*, 425–431.

47. Deshpande, R., Asiedu, M.K., Klebig, M., Sutor, S., Kuzmin, E., Nelson, J., Piotrowski, J., Shin, S.H., Yoshida, M., Costanzo, M., et al. (2013). A comparative genomic approach for identifying synthetic lethal interactions in human cancer. Cancer Res. *73*, 6128–6136.

48. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607.

49. Yang, W., Soares, J., Greninger, P., Edelman, E.J., Lightfoot, H., Forbes, S., Bindal, N., Beare, D., Smith, J.A., Thompson, I.R., et al. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. *41*, D955–D961.

50. Rees, M.G., Seashore-Ludlow, B., Cheah, J.H., Adams, D.J., Price, E.V., Gill, S., Javaid, S., Coletti, M.E., Jones, V.L., Bodycombe, N.E., et al. (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat. Chem. Biol. *12*, 109–116.

51. Shankavaram, U.T., Varma, S., Kane, D., Sunshine, M., Chary, K.K., Reinhold, W.C., Pommier, Y., and Weinstein, J.N. (2009). CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. BMC Genomics *10*, 277.

52. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., et al. (2011). DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res. *39*, D1035–D1041.

53. Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., and Overington, J.P. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. *40*, D1100–D1107.

54. Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., et al. (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. *43*, D805–D811.

55. Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. *39*, D685–D690.

56. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45* (D1), D353–D361.

Supplemental Information

Genetic Interaction-Based Biomarkers

Identification for Drug Resistance

and Sensitivity in Cancer Cells

Yue Han, Chengyu Wang, Qi Dong, Tingting Chen, Fan Yang, Yaoyao Liu, Bo Chen, Zhangxiang Zhao, Lishuang Qi, Wenyuan Zhao, Haihai Liang, Zheng Guo, and Yunyan Gu

Table S1. Statistics of the SV and SL interactions from other studies

Source (PMID)	SL	SV
27438146	107	
26427375	843	
27453043	5065	
24025726	98	
23728082	100	
26516187	19952	
24104479	200	
26227665	23	40
28319113	168	10
26451775	1309	
Boettcher et al.	57	80
28481362	95	178
26781748	846	
27557495	464	
23563794	211	199
28319085	30	

Table S2. Quantitative scores assigned to SV and SL according to the experimental methods annotated by evidence sources

Method	Score
Mutant & Mutant	0.9
CRISPR	0.9
Low-throughput	0.8
RNA interference & Mutant	0.75
Bi-specifie RNA interference	0.5
RNA interference & Drug inhibition	0.5
High-throughput	0.5
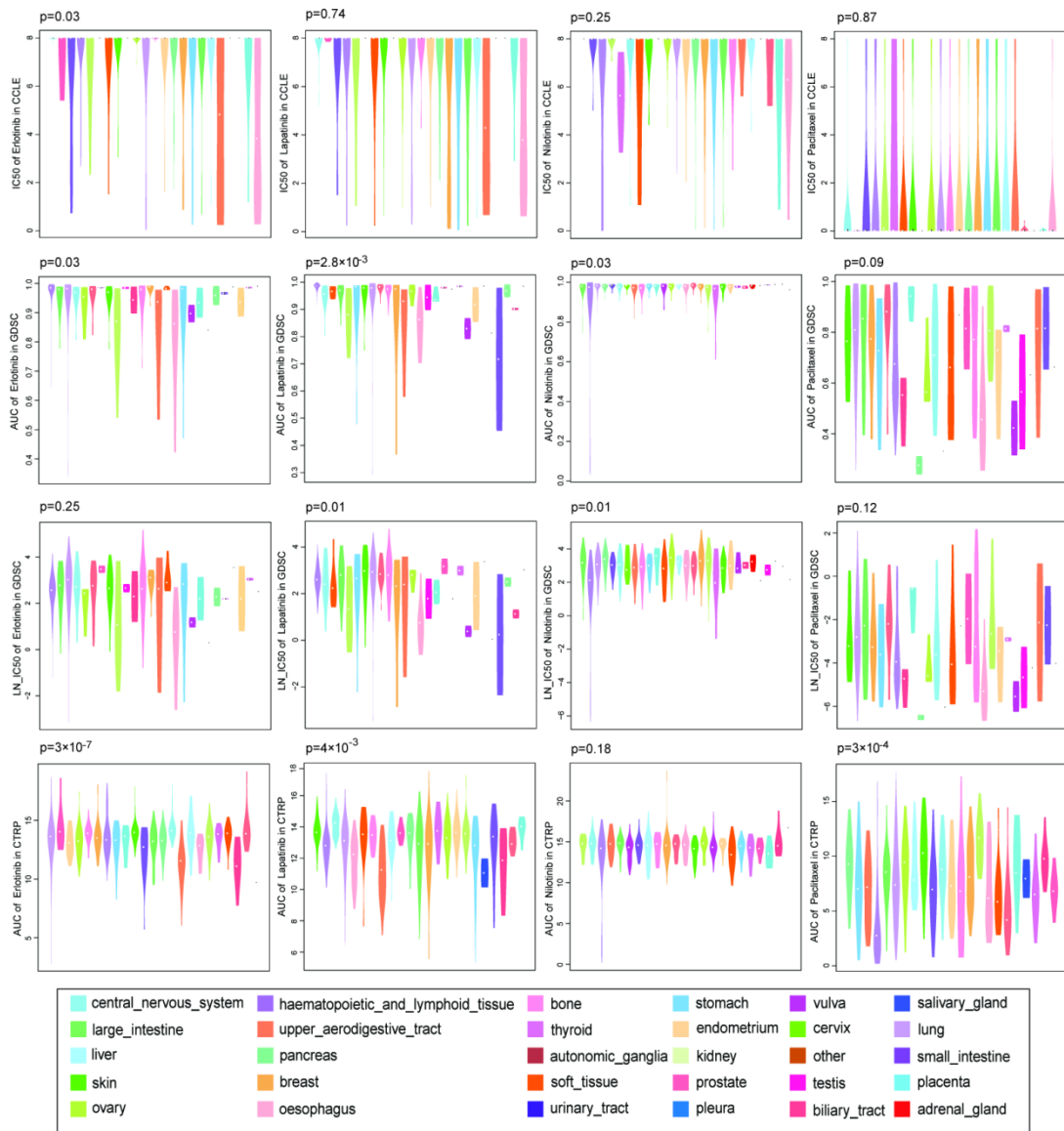
Figure S1. Distribution of drug sensitivity scores among tissues from CCLE, GDSC and CTRP.

Analysis of Variance(ANOVA) was used to test whether the the drug sensitivity score are different among kinds of tissue specific cells.
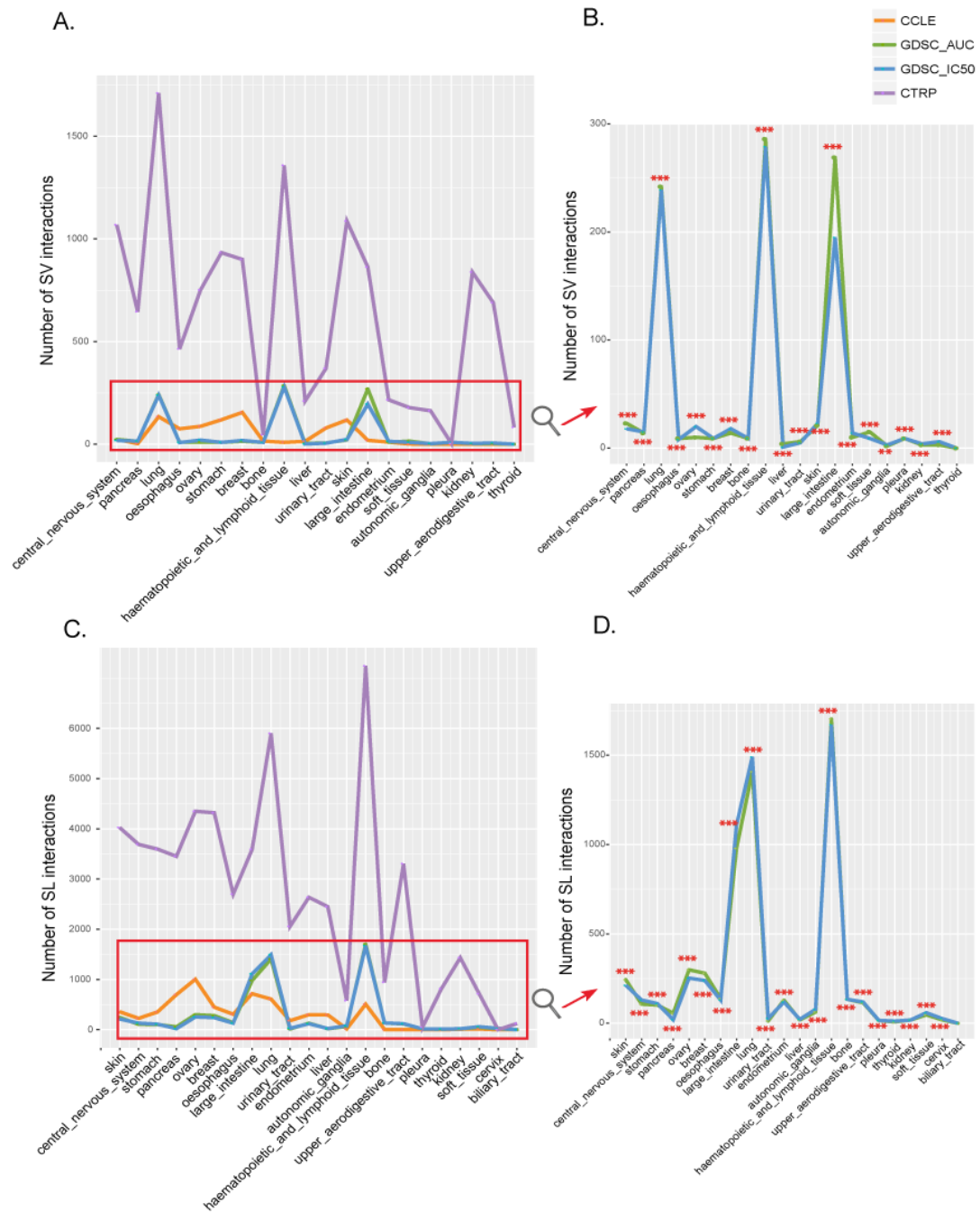
Figure S2. Statistics of the SV and SL interactions in tissue-specific cell lines. (A) Statistics of SV interactions related to drug resistance in 20 tissues. (B) Statistics of SV interactions identified by GDSC (AUC) and GDSC (LN_IC50). Red asterisks represent significant overlapping of SV interactions between GDSC (AUC) and GDSC (LN_IC50) (*P* <0.001, hypergeometric test). (C) Statistics of

SL interactions related to drug sensitivity in 22 tissues. (D) Statistics of SL interactions identified by GDSC (AUC) and GDSC (LN_IC50). Red asterisks represent significant overlapping of SL interactions between GDSC (AUC) and GDSC (LN_IC50) ($P <0.001$, hypergeometric test).
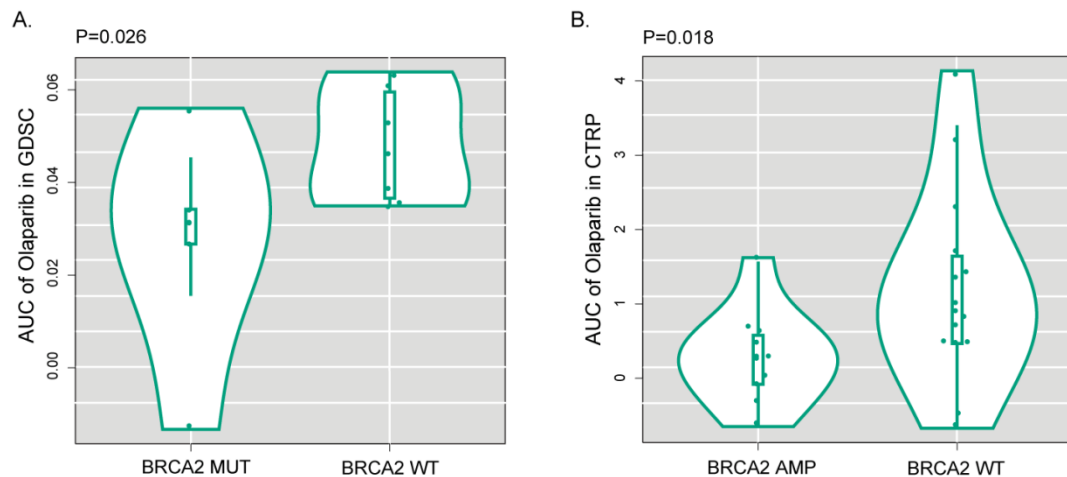
Figure S3. *BRCA2* alterations were related to olaparib sensitivity in cancer cell lines. (A) Cell lines with *BRCA2* mutations are sensitive to olaparib in GDSC. The difference of AUC between the cell lines with or without mutations in *BRCA2* was tested by one-sided Wilcoxon rank sum test (*P*=0.026). (B) Cell lines with *BRCA2* mutations are sensitive to olaparib in CTRP (*P*=0.018, Wilcoxon rank sum test).
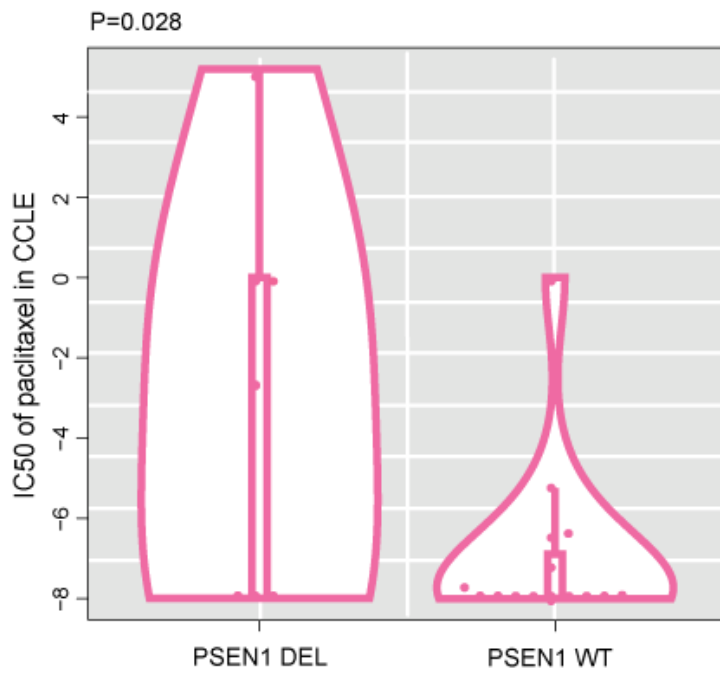
Figure S4. *PSEN1* deletions were related to paclitaxel resistance in CCLE (*P*=0.028, Wilcoxon rank sum test).
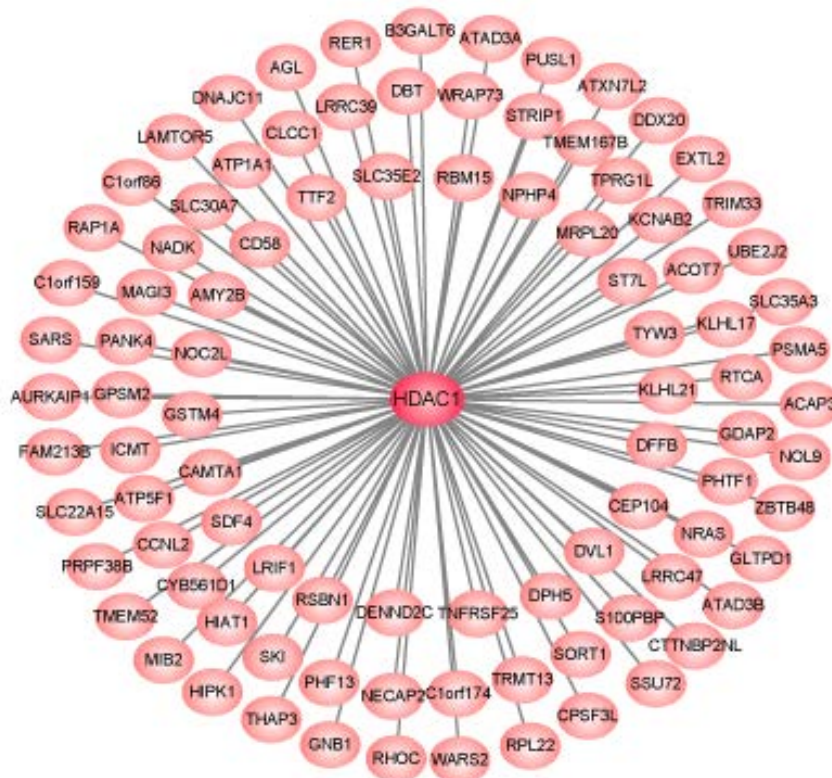
Figure S5. *HDAC1* subnetwork derived from the SV network.

Dark pink node represents HDAC1 and light pink nodes represent the partner genes of HDAC1 in SV network.
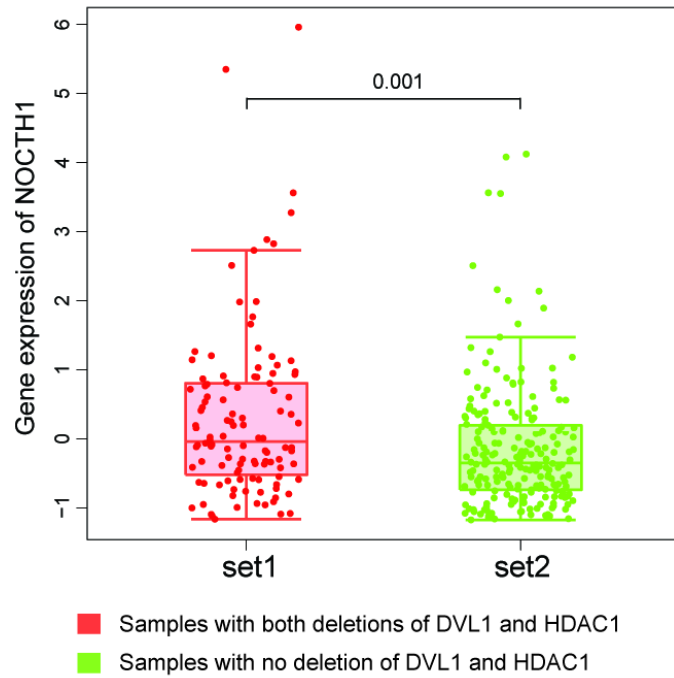
Figure S6. Distribution of *NOTCH1* expression values in liver hepatocellular carcinoma (LIHC). Red points represent the samples with both deletions of DVL1 and HDAC1. Green points represent the samples with no deletion of DVL1 and HDAC1. The difference of expression between the samples with or without deletion in DVL1 and HDAC1 was tested by T-test (*P*=0.001).

Figure S7. *AKT1* subnetwork derived from the SL network. Dark green node represents AKT1 and light pink nodes represent the partner genes of AKT1 in SL network.

Supplementary methods

Information for other studies

SV or SL interactions were integrated from 16 studies (Table S1). The first main source of data was manually curated SV and SL interactions from a CRISPR screen.[1-4] SV and SL interactions identified from other high-throughput screening experiments, such as shRNA, bispecific shRNA, and combinatorial RNAi/drug screens, were also utilized.[5, 6] Second, several SV and SL interactions were obtained using bioinformatics statistical methods, such as permutation and t-test statistics.[7, 8] Ye *et al.* ranked novel cancer-driving synthetic lethal gene pairs using the Chi-square test [9], and *Kranthi et al.* [10] applied network information centrality . Third, additional SV and SL interactions were identified by applying the DAISY algorithm. DAISY combines multiple approaches into a single screen to identify SL interactions.[11] Several studies have identified SL interactions based on mutual exclusivity of gene alterations accompanied by high-throughput screening experiments, such as shRNA screens.[12, 13] In addition, previous studies have generated strategies to identify SV and SL interactions from other species relying on genetic homology or structure.[14, 15] Finally, SL interactions from SLDB were included in the present study.[16]

Integrative confidence scores

The SV and SL interactions analyzed in the present study were obtained from different types of sources, including computational predictions, biochemical assays, and text mining results. In addition, biochemical assays were based on different experimental technologies and platforms, such as shRNA, CRISPR and drug inhibition. Because multiple types of evidence are conducive to the identification of SV (SL) interactions, an integrative confidence score combining scores from these evidence sources can provide an overall estimation of the reliability of a SV (SL)

interaction. In principle, we supposed that (i) the contribution of experimental evidence to the confidence score is more significant than the contribution of predictive algorithms or text mining and that (ii) the SV (SL) interactions supported by more evidence sources should be beneficial to the confidence score. The scoring procedures were divided into two steps, i.e., quantification and integration. A large number of SV (SL) interactions collected from other studies had only qualitative annotation evidence (such as "high-throughput" or "low-throughput"), or technological descriptions of wet-lab experiments (such as "CRISPR screening" or "shRNA screening"). Thus, it was necessary to assign quantitative scores to these SV (SL) interactions before the calculation of integrative scores. Similar to the scoring scheme from SLDB, the quantitative scores were assigned based on the experimental methods (Table S2). For instance, "Mutant & Mutant" indicated that the pair of SV (SL) genes were disturbed by transgenic or genetic deletions. Moreover, "RNA interference & Mutant" indicated that one gene was perturbed by RNAi and that the other was perturbed via mutation. In summary, the SV (SL) interactions obtained from low-throughput experiments were considered to be more reliable than the results from high-throughput experiments due to the lower false positive rate. Therefore, a higher confidence score was assigned to low-throughput evidence than high-throughput evidence. Compared to other RNA interference experiments (such as shRNA, siRNA and dsRNA), the CRISPR screen had lower off-target effects, which were assigned higher confidence scores similar to mutation and transfection experiments (Table S2).

The following formula was utilized to combine the individual scores:

$$s = 1 - \prod_{i-1}^{n}(1 - p_i)$$

where s represents the integrative score corresponding to the experimental evidence; p_i is the individual score; and n is the total number of experimental supporting evidence.
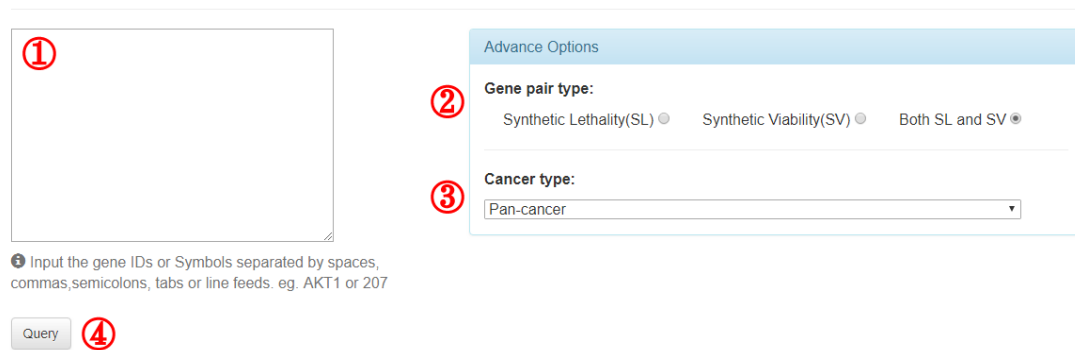
Database guide

The URL of the database is http://www.medsysbio.org/CGIdb.

1. Search

In the Search page, users can key in gene ID or symbol to perform a search of corresponding genetic interactions (1), which can be further screened by advance options (2-3).

2. Search Results

When user queries the genes to the database, the CGIdb will provide search results on this page. The search results are divided into two main parts: (i) information box, which includes the basic information of your selected gene, and the external link to NCBI for more detail (1); (ii) The SL/SV pairs list is provided on the bottom of page (2). And user can click the button (3) to view the details of the gene pairs, including drugs effect and protein-protein interaction network. The results can be exported as CSV format (4).
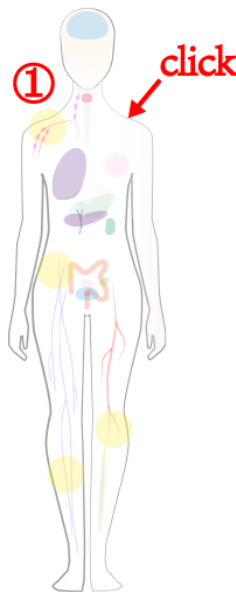
3. Search Details

The details of results are divided in three parts. (i) Distribution of gene alteration in TCGA data are provided. CGIdb shows the altered samples (mutation and copy number alteration) of genes and statistic of analysis results (1-2). (ii) In the middle of page, user can get results of the drug effect. In cell lines with drug targeting the searching gene, a one-sided Wilcoxon rank sum test is used to test whether the drug response measures, such as IC50, are significantly higher or lower in cells lines with and without alterations of the partner genes in the genetic interaction (3). Detailed information about the pharmacodynamic data are shown in the right table (4-5). (iii) On the bottom of the detail page, CGIdb provides the visualized network of protein-protein interactions derived from PathwayCommon. User can export network as jpg or png file (6-7).

Gene Alteration Distribution in TCGA data

①

	TCGA.EK.A2R7.01	TCGA.FU.A3EO.01	TCGA.C5.A3HD.01	TCGA.EK.A2R9.01	TCGA.Q1.A6DT.01	TCGA.Q1.A73P.01
AKT1						
EML5						

② ☐ mutation ☐ amplification ☐ deletion

ⓘ Statistic of Analysis Result:

Gene1 ID	Gene1 Symbol	Gene2 ID	Gene2 Symbol	P-value in Yeast'data	P-value in shRNA'data	P-value in CRISPR'data	P-value of Hypergeometric Test
207	AKT1	161436	EML5	NA	NA	0.008	0.00E+00

Drug Effect in CTRP

③ gefitinib

④

	Database	Drug	Dose Response Parameters	P-value
☐	CCLE	No significant drug effect	IC50	pvalue
☐	COSMIC	a-443654 ▾	AUC	pvalue
☐	COSMIC	akt inhibitor ▾	IC50	0.022
☑	CTRP	gefitinib ▾	AUC	0.017
☐	NCI60	No significant drug effect	GI50	pvalue

⑤

Network in PathwayCommon

⑥

🖫 Save Network ▾   🔍 100% ▾   ⚙ Layout ▾

⑦

EML5

NTRK1

AKT1

4. Browse

In the Browse page, users can easily filter SV and SL pairs classified by tissue types. The human body on the left includes optional tissue types (1). Click on the tissue on the model to filter the genetic interactions of the tissue. The detailed results are

showed in the right table (2).

5. Data

In the data page, users can download or upload data, We provide all SV and SL pairs, which are classified by different sources and tissue types (1-3).

References

1.   Boettcher. M, Tian. R, Blau. J, Markegard. E, Wu. D, Biton. A, Zaitlen. N,

McCormick. F, Kampmann. M, and McManus. MT. (2017). Decoding directional genetic dependencies through orthogonal CRISPR/Cas screens. bioRxiv.

2. Shen JP, Zhao D, Sasik R, Luebeck J, Birmingham A, Bojorquez-Gomez A, Licon K, Klepper K, Pekin D, Beckett AN, et al. (2017). Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. Nature methods. *14*, 573-6.

3. Du D, Roguev A, Gordon DE, Chen M, Chen SH, Shales M, Shen JP, Ideker T, Mali P, Qi LS, et al. (2017). Genetic interaction mapping in mammalian cells using CRISPR interference. Nature methods. *14*, 577-80.

4. Han K, Jeng EE, Hess GT, Morgens DW, Li A, and Bassik MC. (2017). Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nature biotechnology. *35*, 463-74.

5. Vizeacoumar FJ, Arnold R, Vizeacoumar FS, Chandrashekhar M, Buzina A, Young JT, Kwan JH, Sayad A, Mero P, Lawo S, et al. (2013). A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. Molecular systems biology. *9*, 696.

6. Laufer C, Fischer B, Billmann M, Huber W, and Boutros M. (2013). Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. Nature methods. *10*, 427-31.

7. Park S, and Lehner B. (2015). Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. Molecular systems biology. *11*, 824.

8. Wang X, and Simon R. (2013). Identification of potential synthetic lethal genes to p53 using a computational biology approach. BMC medical genomics. *6*, 30.

9. Ye H, Zhang X, Chen Y, Liu Q, and Wei J. (2016). Ranking novel cancer driving synthetic lethal gene pairs using TCGA data. Oncotarget. *7*, 55352-67.

10. Kranthi T, Rao SB, and Manimaran P. (2013). Identification of synthetic lethal pairs in biological systems through network information centrality. Molecular bioSystems. *9*, 2163-7.

11. Cunningham CE, Li S, Vizeacoumar FS, Bhanumathy KK, Lee JS, Parameswaran S, Furber L, Abuhussein O, Paul JM, McDonald M, et al.

(2016). Therapeutic relevance of the protein phosphatase 2A in cancer. Oncotarget. *7*, 61544-61.

12.   Srihari S, Singla J, Wong L, and Ragan MA. (2015). Inferring synthetic lethal interactions from mutual exclusivity of genetic events in cancer. Biology direct. *10*, 57.

13.   Wappett M, Dulak A, Yang ZR, Al-Watban A, Bradford JR, and Dry JR. (2016). Multi-omic measurement of mutually exclusive loss-of-function enriches for candidate synthetic lethal gene pairs. BMC genomics. *17*, 65.

14.   Jacunski A, Dixon SJ, and Tatonetti NP. (2015). Connectivity Homology Enables Inter-Species Network Models of Synthetic Lethality. PLoS computational biology. *11*, e1004506.

15.   Srivas R, Shen JP, Yang CC, Sun SM, Li J, Gross AM, Jensen J, Licon K, Bojorquez-Gomez A, Klepper K, et al. (2016). A Network of Conserved Synthetic Lethal Interactions for Exploration of Precision Cancer Therapy. Molecular cell. *63*, 514-25.

16.   Guo J, Liu H, and Zheng J. (2016). SynLethDB: synthetic lethality database toward discovery of selective and sensitive anticancer drug targets. Nucleic acids research. *44*, D1011-7.