Article

Computational Drug Repurposing Based on a Recommendation System and Drug–Drug Functional Pathway Similarity

Mengting Shao [1,2], **Leiming Jiang [1]**, **Zhigang Meng [2,*]** and **Jianzhen Xu [1,*]**

1 Computational Systems Biology Laboratory, Department of Bioinformatics, Shantou University Medical College (SUMC), Shantou 515041, China; 19mtshao@stu.edu.cn (M.S.); 18lmjiang@stu.edu.cn (L.J.)

2 Department of Computer Science, College of Computer Engineering and Applied Mathematics, Changsha University, Changsha 410005, China

* Correspondence: z20110956@ccsu.edu.cn (Z.M.); jzxu01@stu.edu.cn (J.X.); Tel.: +86-754-88900491 (J.X.)

Abstract: Drug repurposing identifies new clinical indications for existing drugs. It can be used to overcome common problems associated with cancers, such as heterogeneity and resistance to established therapies, by rapidly adapting known drugs for new treatment. In this study, we utilized a recommendation system learning model to prioritize candidate cancer drugs. We designed a drug–drug pathway functional similarity by integrating multiple genetic and epigenetic alterations such as gene expression, copy number variation (CNV), and DNA methylation. When compared with other similarities, such as SMILES chemical structures and drug targets based on the protein–protein interaction network, our approach provided better interpretable models capturing drug response mechanisms. Furthermore, our approach can achieve comparable accuracy when evaluated with other learning models based on large public datasets (CCLE and GDSC). A case study about the Erlotinib and OSI-906 (Linsitinib) indicated that they have a synergistic effect to reduce the growth rate of tumors, which is an alternative targeted therapy option for patients. Taken together, our computational method characterized drug response from the viewpoint of a multi-omics pathway and systematically predicted candidate cancer drugs with similar therapeutic effects.

Keywords: drug functional similarity; drug repurposing; pathway activities; recommender system

1. Introduction

Drug resistance is a major problem in cancer treatment where cancer cells escape from the effects of anticancer compounds and patients suffer from recurrence and more aggressive forms of cancer. Although the exact mechanism is still not completely understood, accumulated evidence indicates that resistance to cancer therapy is mediated by the complex interplay between several key factors, such as intratumoral heterogeneity, metabolic reprogramming, and cancer microenvironment [1]. Therefore, it is urgent to develop novel therapy options, such as the combination of known drugs, to combat cancer cells. However, the development and validation of new drugs are time-consuming, expensive, and often prone to failure. Studies have shown that each drug takes an average of 10–15 years and more than USD 2 billion to create; but, the success rate is less than 10% [2].

Drug repurposing is used to identify new uses of existing or research drugs that are beyond their original indications [3]. Classic examples include Minoxidil (initially used for hypertension; now for hair loss), Viagra (initially used for angina; now to treat erectile dysfunction and pulmonary hypertension), and Rituximab (initially used for chronic lymphocytic leukemia and rheumatoid arthritis; now for non-Hodgkin's lymphoma) [4]. Recently, bromocriptine has been approved for the treatment of type 2 diabetes in the U.S. due to the new finding that it is also a central dopamine agonist [5]. Compared with traditional drug discovery, the repositioning of known drugs has the advantages of lower cost and higher safety; thus, they may demonstrate a new way to accelerate the drug development process [6].

Many computational repositioning methods have been developed. Some researchers exploited the structural similarity of drug molecules or shared protein targets, whereas others leveraged prior knowledge of drug, target and disease interaction, and introduced systems biology approaches for drug repurposing. For example, we and others have developed network-based methods for finding molecules inhibiting tumor progression [7–9]. In this study, we adopted a recommendation system-based framework to prioritize drugs. The recommender system is widely applied in e-commerce to predict users' preferences for selling products. Recently, it has demonstrated great promise in various bioinformatics problems, such as predicting protein subcellular localization, RNA–protein interactions, and drug repurposing [10–13]. In this study, we combined various cancer multi-omics datasets, such as gene profiling, DNA methylation and CNV data, and enriched them into KEGG pathways. Then, based on these functional features, a recommender system is utilized to predict drug response. Our results indicated that this model can achieve both comparable accuracy and clear biological mechanisms when evaluated on two large cancer drug response datasets (CCLE and GDSC).

2. Results and Discussion

2.1. Workflow Overview

We developed a pipeline to prioritize candidate drugs based on a drug functional similarity. A flow diagram of the pipeline is shown in Figure 1. Our method consists of three main steps. (A) Download mRNA expression data, DNA methylation data, and copy number variation data from the GDSC and CCLE databases. Then, use single sample gene set enrichment analysis (ssGSEA) to project each type of omics data onto the KEGG pathway, thus obtaining multi-omics pathway activity profiles. (B) Download compounds' activity data, and then predict the drug response activity through a recommendation system based on multi-omics pathway features. (C) Construct a drug functional similarity between the drugs by integrating the mRNA, methylation, and CNV pathway activity profiles and the predicted drug response activity. This functional similarity is compared with compound structural similarity and target distance similarity.

2.2. Evaluation of Predictive Drug Response Results

For evaluation, we compared the above pipeline (multi-omics pathway) with the following two scenarios: (1) only mRNA expression profiles are used to infer pathway activity, then feed the same recommendation system (mRNA-pathway); (2) the mRNA expression profile, but not pathway activity, is used directly in the same recommendation system (mRNA expression). NDCG is a generally accepted index for evaluating ranking recommendations. It ranges from 0 to 1, where 1 means that the model predicted the drug's ranking accurately. As shown in Table 1, both the multi-omics pathway setting and mRNA-pathway setting are significantly better than the mRNA expression setting in both the GDSC and CCLE datasets. At the same time, these two scenarios also achieved a relatively smaller sum of squared error than when mRNA expression data were used, which indicated better robustness of the learning models.

This analysis suggests that pathway-based drug response prediction is better than gene-based prediction in the recommendation system. Furthermore, multi-omics pathway settings can achieve comparable accuracy with the single-omics (mRNA profile) pathway-based setting.
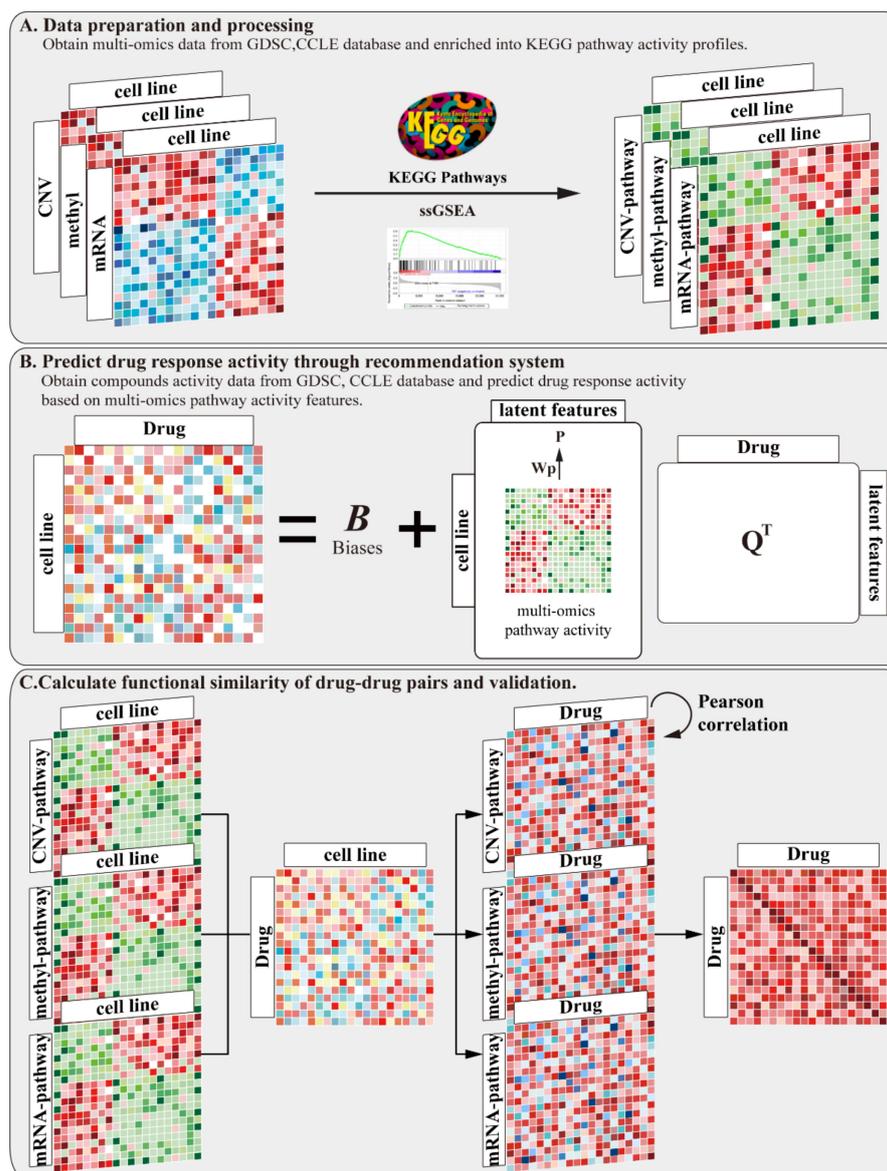
Figure 1. A pipeline to prioritize candidate compounds based on a drug functional similarity. Our method includes three main steps: (**A**) inferring multi-omics pathway activity profiles; (**B**) predicting drug response activity through recommendation system based on multi-omics pathway activity profiles; and (**C**) calculating drug–drug functional similarity and evaluating with other drug similarities to validate our result.

Table 1. Performance and robustness comparison on drug response.

	GDSC Multi-omics Pathway	GDSC mRNA-Pathway	GDSC mRNA Expression	CCLE Multi-omics Pathway	CCLE mRNA-Pathway	CCLE mRNA Expression
NDCG	0.815	0.816	0.381	0.976	0.978	0.798
Sum of squared error	1.176	1.173	3.540	0.728	0.662	2.633

2.3. Drugs' Effects on Biological Pathway Levels

Inspecting the multi-omics pathway activity profiles for each drug can reveal its functional effect in cells. In addition, comparing the multi-omics pathway activity profiles

of each drug pair can elucidate their mechanistic similarity, which provides a functional understanding for their mutual replacement.

The heatmap in Figure 2 shows the Pearson correlation of the top 10 drug repurposing pairs with the activity levels of multi-omics pathways. For example, (5Z)-7-Oxozeaenol and GSK2126458/omipalisib form a drug repurposing pair in GDSC datasets. Previous reports found that (5Z)-7-Oxozeaenol inhibits TAK1 [14] and GSK2126458 is the inhibitor of PI3K [15]. They have the highest correlation with "Toxoplasmosis pathway", "Tuberculosis pathway", and "Salmonells infection pathway" from the viewpoint of methylation, which may indicate that the two drugs have similar effects on the methylation of key genes in these pathways. In addition, for the L-685458 and ZD-6474/Vandetanib drug pair in the CCLE datasets, L-685458 is a kind of gamma-secretase inhibitor [16], and ZD–6474/Vandetanib efficiently suppresses RET kinase, including the vascular endothelial growth factor receptor and epidermal growth factor receptor signaling [17]. These two drugs both have the highest correlation with the "Bacterial invasion of epithelial cells pathway", "ErbB signaling pathway", and "Small cell lung cancer pathway" at the methylation level. Again, this result revealed that both L-685458 and ZD-6474 are involved in the methylation process of genes located in KEGG pathways.
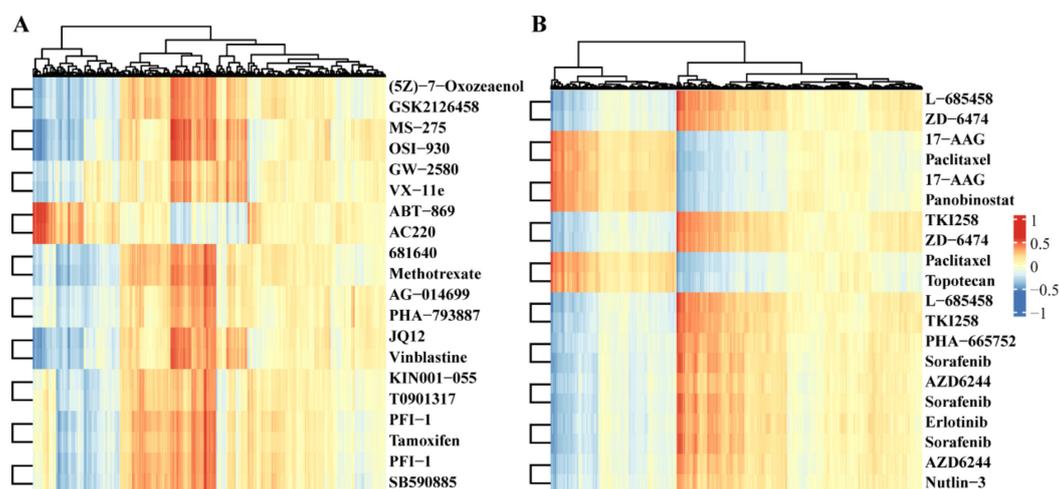

Figure 2. Heatmap of Pearson correlation coefficients of drugs in the activity of the multi-omics pathway. (**A**) Top 10 drug pairs based on functional similarity in GDSC database, in which each row represents a drug, and each column represents multi-omics pathways. The values are correlation coefficients. Drug names are listed at the right most columns. (**B**) Top 10 drug pairs based on functional similarity in the CCLE database.

Furthermore, these data also suggest that even though the drug targets are very different, their pathway activity in cells may have similar patterns, indicating that the drugs share common regulatory mechanisms. The activity pattern of the drug pairs across multi-omics pathways describes their functional similarity, which also lays the foundation for drug repurposing.

2.4. Comparison of Drug Pair Similarities

According to the activity patterns of the drug pairs across all the multi-omics pathways, we can assign each drug pair a functional similarity score. Table 2 lists the top 10 drug pairs' similarities in the GDSC and CCLE datasets. For comparison, we also investigated the similarities based on the chemical structures and protein targets of these drug pairs in GDSC and CCLE datasets. As shown in Table 2, the functional similarity is higher than other similarities. These results suggest that drug responses at biological pathway levels are more likely to reflect their biological influences and explain why the two drugs can induce similar effects on the cell; i.e., it is not based on their molecular structure or the closeness

of their protein targets in a PPI network, but rather due to their similar functional effects on the multiple-omics pathways. Furthermore, it confirmed that the functional similarity between drugs can prioritize their mutual replacement.

Table 2. Top 10 drug pairs' similarities in GDSC and CCLE datasets.

Database	Drug1	Drug2	Functional Sim	SMILES Sim	PPI Sim
GDSC	(5Z)-7-Oxozeaenol	GSK2126458	0.998365459	0.162790698	1.08E−70
GDSC	MS-275	OSI-930	0.998193161	0.362694301	3.30E−77
GDSC	GW-2580	VX-11e	0.997189963	0.242063492	3.71E−84
GDSC	ABT-869	AC220	0.996300765	0.257462687	0.9
GDSC	681640	Methotrexate	0.996231842	0.236734694	6.46E−66
GDSC	AG-014699	PHA-793887	0.996134942	0.223529412	5.37E−72
GDSC	JQ12	Vinblastine	0.995927571	-	1.54E−91
GDSC	KIN001-055	T0901317	0.99574713	0.161111111	4.90E−75
GDSC	PFI-1	Tamoxifen	0.995377083	0.208092486	2.38E−66
GDSC	PFI-1	SB590885	0.995238003	0.194029851	2.74E−83
CCLE	L-685458	ZD-6474	0.994678	0.11349	1.18E−67
CCLE	17-AAG	Paclitaxel	0.994097	0.272071	0.9
CCLE	17-AAG	Panobinostat	0.99304	0.075472	3.97E−71
CCLE	TKI258	ZD-6474	0.992772	0.227439	0.9
CCLE	Paclitaxel	Topotecan	0.992513	0.291483	0.9
CCLE	L-685458	TKI258	0.989722	0.129946	2.38E−66
CCLE	PHA-665752	Sorafenib	0.98863	0.157855	0.9
CCLE	AZD6244	Sorafenib	0.98829	0.169047	0.9
CCLE	Erlotinib	Sorafenib	0.987973	0.256565	0.9
CCLE	AZD6244	Nutlin-3	0.987127	0.137107	0.9

2.5. Case Study

The drug repurposing pair, Erlotinib and OSI-906, has a functional similarity of 0.978, indicating their similar therapeutic effects. Erlotinib potently blocks EGFR kinase activity and suppresses downstream signaling pathways, such as PI3K–AKT and MAPK [18]. In addition, these signaling pathways are promoted by other receptors, including IGF1R [19]. During embryonic development and postnatal growth, IGF1R is sensitized by ligands IGF-1 and -2 [20,21]. Researchers have indicated that the components of the IGF family are often abnormally expressed in cancers and activate tumorigenesis [20,22]. IGF1R overexpression is also associated with poor survival in some tumor types [23–26]. Furthermore, cancers have evolved a compensating mechanism for IGF1R inhibition [27]. A variant form of the insulin receptor (INSR-A) is induced by IGF-2 and insulin, which enhances proliferation and cell survival [28]. Thus, co-inhibition of IGF1R and INSR may provide enhanced antitumor activity [29,30].

Linsitinib (OSI-906) suppresses both IGF1R and INSR tyrosine kinase. Previous studies indicated it can inhibit proliferation in a variety of tumor cell lines and xenograft models [31,32]. Single-drug treatment of linsitinib in patients with solid tumors, such as melanoma and adrenocortical carcinoma, has proven its antitumor activity [33–35]. From drug mechanistic analysis, it can be seen that although Erlotinib and OSI-906 inhibit different protein targets, their influences in cells converge on the same downstream pathway. This explains why they can be mutually repurposed.

In addition, this repositioning drug pair also implies that the combined use of both drugs may be a good therapeutic option, especially for overcoming drug resistance. The literature review confirms this conclusion. The combined IGF1R/INSR and EGFR blockade has shown enhanced inhibition of common downstream signaling pathways, and suppressed resistance to single-receptor blockades [22,30,36]. Preclinical investigations among NSCLC, breast, pancreatic, and colorectal cancer patients have indicated that the combined administration of IGF1R/INSR and EGFR inhibitors lead to additive effects on

tumor growth [37–41]. Furthermore, IGF-2 is induced in erlotinib-resistant tumors and small-molecule IGF1R TKI sensitized the tumors to the EGFR inhibitor [42].

Finally, we set the functional similarity of multi-omics pathways that are >0.95 as a threshold to select candidate repurposing drug pairs. In total, we obtained 1015 out of 31125 drug pairs in the GDSC dataset, and 53 pairs out of 276 pairs in the CCLE dataset (Supplementary Table S1).

3. Materials and Methods

3.1. Data Sources and Data Processing

3.1.1. Chemical Compounds Activity Data

GDSC drug data: The drug activity data of 250 unique drugs in 904 cancer cell lines were collected from the GDSC database (http://www.cancerrxgene.org/gdsc1000/GDSC1 000_WebResources/Home.html). This page provides several Excel sheets referred to in the Iorio et al. paper [43] as Supplementary Materials. Drug responses were provided as the natural logarithm of half the maximal inhibitory concentration in a data matrix (file:TableS4A.xlsx, tab "TableS4A-IC50s"), which we subsequently converted to negative logarithms with a base 10, ensuring that the drug activity levels are expressed as the negative log of the half-maximal inhibitory concentration [$-\log 10(IC50)$]. A higher value represents more drug sensitivity in a cell line.

CCLE drug data: CCLE drug sensitivity data in the form of IC50 values were extracted from the file "CCLE_NP24.2009_Drug_data_2015.02.24.csv", which is available on the CCLE website [44] and includes IC50 data for 24 anti-cancer drugs. Afterwards, the negative number of logarithms with a base 10 of the IC50 values was calculated for the drug activity values. In total, there were 24 unique drugs in 402 cell lines that we used for our analyses.

3.1.2. Multi-Omics Expression Data

GDSC multi-omics data: Raw cell line expression array data were downloaded from ArrayExpress (file: E-MTAB-3610) [43], in which data is measured by the Affymetrix Human Genome U219 Array chip, and after RMA normalization of R packages 'affy' [45], the expression data of each gene in each cell line were obtained. The methylation data were extracted from the average pre-processed β-values for each of all CpG islands (file: F2_METH_CELL_Data.txt), which were also available on the download portal. For the CpG islands, we used the GPL13534–11288 reference to match genes to the CpG islands in the dataset. The copy number variation data were extracted from PICNIC [46], and the absolute copy numbers were derived from the Affymetrix SNP6.0 array data (file: cnv_abs_copy_number_picnic_20191101.csv). Thus, we derived the estimated value of the copy number of each gene. For each omics profile of the GDSC database, we converted the cell names to cosmic IDs based on the annotation file of the GDSC cell lines.

CCLE multi-omics data: Omics data (expression, methylation, and copy number variation) describing the cancer cell lines were acquired via bulk download from the Cancer Cell Line Encyclopedia (CCLE) (https://portals.broadinstitute.org/ccle/data, access date: 2020/08/05) [47]. In agreement with the original publications [44,47], expression data were obtained through Affymetrix U133+2 arrays and processed to obtain gene-centric RMA-normalized mRNA expressions (file: CCLE_Expression_Entrez_2012-09-29.gct) [48]. Raw Affymetrix CEL files were converted to a single value for each probe set using the Robust Multi-Array Average (RMA) and normalized using quantile normalization. Methylation data were derived by quantifying CpG islands using Reduced Representation Bisulfite Sequencing (file: CCLE_RRBS_tss_CpG_clusters_20181022.txt.gz). Copy number variation (CNV) data were acquired from the Affymetrix SNP6.0 arrays (file: CCLE_copynumber_byGene_2013-12-03.txt.gz). Copy numbers were normalized by comparingto the most similar HapMap normal samples [49]. Segmentation of the normalized log2 (CN/2) ratios was achieved using the circular binary segmentation (CBS) algorithm [47,50]. For mRNA expression profiles with duplicate gene IDs, we aggregated

their expression values by their means for our analysis, and for methylation profiles with duplicate gene IDs, we aggregated their values by sum.

3.2. Inferring Multi-Omics Pathway Activity Profiles

First, we obtained 250 selected gene sets (C2) of the KEGG pathway from the previous research [51] and then used single sample gene set enrichment analysis (ssGSEA) with R packages "GSVA" [52,53] against KEGG pathway data to convert mRNA expression data, methylation level data, and copy number variation data into a pathway activity profile, which was accomplished by a "single sample" extension of GSEA [54], defining an enrichment score based on the degree of absolute enrichment of a gene set in each sample within a given dataset.

For the given sample S, the expression values of gene G of size NG were firstly rank-normalized. Then, the empirical cumulative distribution functions (ECDF) of the genes were used to calculate an enrichment score of $ES\ (G, S)$ by a sum (integration) of the difference between a weighted ECDF of the genes in the signature P_G^{ω} and the ECDF of the remaining genes P_{NG} as follows, where α is set to 0.25.

$$ES(G,S) \ = \ \sum_{i=1}^{N} [P_G^{\omega}(G,S,i) \ - \ P_{NG}(G,S,i)] \tag{1}$$

$$P_G^{\omega}(G,S,i) \ = \ \sum_{rj \in G, j \leq i} \frac{|rj|^{\alpha}}{\sum_{rj \in G} |rj|^{\alpha}} \tag{2}$$

$$P_{NG}(G,S,i) \ = \ \sum_{rj \notin G, j \leq i} \frac{i}{(N - N_G)} \tag{3}$$

This procedure differs from the classic GSEA procedure in that the gene list is ranked by absolute expression and the enrichment score is obtained by an integration of the difference between the ECDFs. For the GDSC dataset, 228 resultant mRNA pathways, 201 methyl pathways, and 227 CNV pathways were extracted. In total, 656 multi-omics pathways were obtained. In addition, 223 mRNA pathways, 220 methyl pathways, and 228 CNV pathways were extracted for the CCLE datasets with a total of 671 multi-omics pathways. Finally, these pathway activity patterns were used to calculate drug–pathway associations and the drug pathway level similarity.

3.3. Predict Drug Response Activity through Recommendation System Based on Multi-Omics Pathway Activity Profiles

Here, we modified the CADRRes framework to predict drug response. The previous report indicated that the CADRRes algorithm outperformed other existing methods including elastic net regression and random permutations [13]. The first step in CaDRReS is to define the cell line features. In contrast to the original calculation with gene expression information, here we calculated cell line features based on the activity of the multi-omics pathway, i.e., by using Pearson's correlation to compute every pair of cell lines, while utilizing the multi-omics pathway activity profiles. In total, we obtained 904 and 402 cell line features for GDSC and CCLE, respectively.

Models were pre-trained and tested independently on both CCLE and GDSC to avoid biases toward either of the datasets [55,56]. Then, the matrix factorization was used to train drug sensitivity models, which were computed based on the equation:

$$\hat{S_{ui}} = \mu + b_i^Q + b_u^P + q_i \cdot p_u = \mu + b_i^Q + b_u^P + q_i (x_u W_p)^T \tag{4}$$

where $\hat{S_{ui}}$ is the computed sensitivity score of cell line u to drug i; μ is the general mean drug response; b_i^Q and b_u^P are bias terms for drug i and cell line u, respectively; $q_i, p_u \in R^f$ are vectors for drug i and cell line u in the f-dimensional latent space; and $W_P \in R^{d \times f}$ is a transformation matrix that projects cell line features $x_u \in R^d$ onto the latent space.

The value of f was set at 10. The source code of CaDRReS can be downloaded at https://github.com/CSB5/CaDRReS. Here, we used the CaDRReS_train_and_test.py file to predict GDSC and CCLE drug response activity data.

Two indicators were used to evaluate the drug response predicted by the recommendation system. The normalized discounted cumulative gain (NDCG) is a generally accepted score for comparing recommendations. It was computed as follows:

$$NDCG(\hat{r}, s) = \frac{DCG(\hat{r}, s)}{DCG(r, s)} \tag{5}$$

$$DCG(\hat{r}, s) = \sum_i \frac{2_i^s - 1}{\log_2 \hat{i} + 1} \tag{6}$$

where $\hat{r}$ is the output rank of drugs tested on a cell line, s is the observed drug activity values, and r is the known rank of drugs based on the drug activity values. The numerator in the DCG is designed to give more weight to drugs with higher sensitivity scores, while the denominator takes precedence over predicting drugs with a higher rank. In addition, the "sum of squared error" loss function was defined as:

$$L(\theta) = \frac{1}{2|K|} \sum_u \sum_i e_{ui}^2 \tag{7}$$

$$e_{ui} = s_{ui} - \hat{s_{ui}} \tag{8}$$

where s_{ui} and $\hat{s_{ui}}$ are the observed and predicted activity for cell line u using drug i, respectively; $\theta = \{b_i, b_u, W_p, q_i\}$; and $|K|$ is the number of drug activity values in the dataset.

3.4. Calculate the Functional Similarity of Drug–Drug Pairs Based on Multi-Omics Pathways Profiles

We first calculated the Pearson correlation between the multi-omics pathway activity profiles and the drug activity profiles predicted by the recommendation system across cell lines in GDSC and CCLE datasets, respectively:

$$r = \frac{\sum_{u=1}^n (D_{u,i} - \overline{D_i})(P_{u,j} - \overline{P_j})}{\sqrt{\sum_{u=1}^n (D_{u,i} - \overline{D_i})^2} \sqrt{\sum_{u=1}^n (P_{u,j} - \overline{P_j})^2}} \tag{9}$$

where $D_{u,i}$ is the predicted drug activity profiles with drug i and cell lines u, and $P_{u,j}$ is the multi-omics pathway activity profiles in method 3.2 with pathway j. Then, for both drug i and pathway j, the correlation based on the same cell lines as u can be computed. The Pearson correlation coefficient reflects the association between the multi-omics pathway activity and the drug response activity.

Next, based on the above multi-omics pathway–drug correlation matrix, the Pearson correlation coefficient of each pair of drugs is calculated, which reflects the functional similarity of each pair of drugs at the level of the multi-omics pathways and describes the functional similarity between the multi-omics pathway activity patterns, in which drugs with higher functional similarity have a higher probability of being involved in related biological pathways and treating similar diseases.

3.5. Molecular Structural Similarities and Drug Target Similarities

In order to make a strict comparison with the other similarity of drugs, we calculated the two-dimensional structural similarity of the drugs and the similarity of the protein-protein interaction (PPI) distance of the drug targets. All similarity measures were finally normalized to be in the range (0, 1). Here, the two-dimensional structure data of the drug comes from the SMILES format of the PubChem database, which was downloaded from Pubchem [57,58]. The similarity score between two drug molecules is calculated according to the two-dimensional Tanimoto coefficient score [59], based on R packages "RxnSim" [60].

It extracts structural features and is then defined as the size of the intersection divided by the size of the union of the feature sets. The structural features determine absorption, distribution, metabolism, excretion, and toxicity properties, which ultimately affect the pharmacological activity of the drug molecule.

The target data come from The Drug Gene Interaction Database [61] and the PPI data were downloaded from the STRING database [62]. The distances between each pair of drug targets were computed using an all-pairs shortest paths algorithm, found on the human PPI network for drugs, associated with more than one gene, and the minimal distance between the associated genes was used. Then, the shortest distances were transformed to similarity values using the formula described in Perlman et al. [63]: $S(p, p') = Ae^{-bD(p, p')}$, where $S(p, p')$ is the similarity value between two proteins and $D(p, p')$ is the shortest path between them in the PPI network. Then, *A, b* was set to 0.9 and 1 according to Perlman et al. [63]. Self-similarity was assigned a value of 1.

4. Conclusions

The long and expensive drug discovery process needs to develop its machine learning method for drug repurposing. We integrate a variety of transcriptomic resources, methylation, and CNV data into a recommendation framework. Not only can it accurately predict drug response, but also provide an interpretable mechanism on drug effects at the same time. This is particularly useful in personal drug prescriptions in cancer care and/or combating drug resistance. In addition, accumulated evidence revealed that some cancers share common mechanisms and molecular pathogenesis with neurodegenerative diseases, such as Alzheimer's disease (AD) [64,65]. Thus, this framework can be utilized as a computational tool for generating mechanistic hypotheses and drug replacement for both cancer and AD when the drug activity and multi-omics data are comprehensive.

Supplementary Materials: Table S1: Candidate drug repurposing pairs and their functional structural and target similarities. Candidate drug repurposing pairs are selected with a cutoff function similarity of >0.95.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Not available.

References

1. Zahan, T.; Das, P.K.; Akter, S.F.; Habib, R.; Rahman, H.; Karim, R.; Islam, F. Therapy resistance in cancers: Phenotypic, metabolic, epigenetic and tumour microenvironmental perspectives. *Anti-Cancer Agents Med. Chem.* **2020**, *20*, 2190–2206. [CrossRef]
2. DiMasi, J.A.; Grabowski, H.G.; Hansen, R.W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **2016**, *47*, 20–33. [CrossRef] [PubMed]
3. Ashburn, T.T.; Thor, K.B. Drug repositioning: Identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov.* **2004**, *3*, 673–683. [CrossRef] [PubMed]
4. Dudley, J.T.; Deshpande, T.; Butte, A.J. Exploiting drug-disease relationships for computational drug repositioning. *Briefings Bioinform.* **2011**, *12*, 303–311. [CrossRef]
5. Pijl, H.; Ohashi, S.; Matsuda, M.; Miyazaki, Y.; Mahankali, A.; Kumar, V.; Pipek, R.; Iozzo, P.; Lancaster, J.L.; Cincotta, A.H.; et al. Bromocriptine: A novel approach to the treatment of type 2 diabetes. *Diabetes Care* **2000**, *23*, 1154–1161. [CrossRef] [PubMed]

6. Zhao, K.; So, H.-C. Using drug expression profiles and machine learning approach for drug repurposing. *Methods Mol. Biol.* **2019**, *1903*, 219–237. [CrossRef] [PubMed]

7. Zhou, X.; Dai, E.; Song, Q.; Ma, X.; Meng, Q.; Jiang, Y.; Jiang, W. In silico drug repositioning based on drug-miRNA associations. *Brief. Bioinform.* **2020**, *21*, 498–510. [CrossRef] [PubMed]

8. Peyvandipour, A.; Saberian, N.; Shafi, A.; Donato, M.; Draghici, S. A novel computational approach for drug repurposing using systems biology. *Bioinformatics* **2018**, *34*, 2817–2825. [CrossRef]

9. Jiang, L.; Chen, Q.; Bei, M.; Shao, M.; Xu, J. Characterizing the tumor RBP-ncRNA circuits by integrating transcriptomics, interactomics and clinical data. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 5235–5245. [CrossRef]

10. Mehrabad, E.M.; Hassanzadeh, R.; Eslahchi, C. PMLPR: A novel method for predicting subcellular localization based on recommender systems. *Sci. Rep.* **2018**, *8*, 12006. [CrossRef]

11. Corrado, G.; Tebaldi, T.; Costa, F.; Frasconi, P.; Passerini, A. RNAcommender: Genome-wide recommendation of RNA–protein interactions. *Bioinformatics* **2016**, *32*, 3627–3634. [CrossRef] [PubMed]

12. Wang, A.; Lim, H.; Cheng, S.-Y.; Xie, L. ANTENNA, a multi-rank, multi-layered recommender system for inferring reliable drug-gene-disease associations: Repurposing diazoxide as a targeted anti-cancer therapy. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *15*, 1960–1967. [CrossRef]

13. Suphavilai, C.; Bertrand, D.; Nagarajan, N. Predicting cancer drug response using a recommender system. *Bioinformatics* **2018**, *34*, 3907–3914. [CrossRef] [PubMed]

14. Wu, J.; Powell, F.; Larsen, N.A.; Lai, Z.; Byth, K.F.; Read, J.; Gu, R.-F.; Roth, M.; Toader, D.; Saeh, J.C.; et al. Mechanism and in vitro pharmacology of TAK1 inhibition by (5Z)-7-Oxozeaenol. *ACS Chem. Biol.* **2013**, *8*, 643–650. [CrossRef]

15. Munster, P.N.; Aggarwal, R.; Hong, D.; Schellens, J.H.; Van Der Noll, R.; Specht, J.M.; Witteveen, P.O.; Werner, T.L.; Dees, E.C.; Bergsland, E.K.; et al. First-in-human phase I study of GSK2126458, an oral pan-class I phosphatidylinositol-3-kinase inhibitor, in patients with advanced solid tumor malignancies. *Clin. Cancer Res.* **2016**, *22*, 1932–1939. [CrossRef] [PubMed]

16. Williams, C.K.; Li, J.-L.; Murga, M.; Harris, A.L.; Tosato, G. Up-regulation of the Notch ligand Delta-like 4 inhibits VEGF-induced endothelial cell function. *Blood* **2006**, *107*, 931–939. [CrossRef]

17. Wells, S.A., Jr.; Robinson, B.G.; Gagel, R.F.; Dralle, H.; Fagin, J.A.; Santoro, M.; Baudin, E.; Elisei, R.; Jarzab, B.; Vasselli, J.R.; et al. Vandetanib in patients with locally advanced or metastatic medullary thyroid cancer: A randomized, double-blind phase III trial. *J. Clin. Oncol.* **2012**, *30*, 134–141. [CrossRef]

18. Hidalgo, M. Erlotinib: Preclinical investigations. *Oncology* **2003**, *17*, 11–16.

19. Chitnis, M.M.; Yuen, J.S.; Protheroe, A.S.; Pollak, M.; Macaulay, V.M. The type 1 insulin-like growth factor receptor pathway. *Clin. Cancer Res.* **2008**, *14*, 6364–6370. [CrossRef]

20. Pollak, M. The insulin and insulin-like growth factor receptor family in neoplasia: An update. *Nat. Rev. Cancer* **2012**, *12*, 159–169. [CrossRef]

21. Gao, J.; Chang, Y.S.; Jallal, B.; Viner, J. Targeting the insulin-like growth factor axis for the development of novel therapeutics in oncology. *Cancer Res.* **2012**, *72*, 3–12. [CrossRef]

22. King, H.; Aleksic, T.; Haluska, P.; Macaulay, V.M. Can we unlock the potential of IGF-1R inhibition in cancer therapy? *Cancer Treat. Rev.* **2014**, *40*, 1096–1105. [CrossRef]

23. Parker, A.S.; Cheville, J.C.; Janney, C.A.; Cerhan, J.R. High expression levels of insulin-like growth factor–I receptor predict poor survival among women with clear-cell renal cell carcinomas. *Hum. Pathol.* **2002**, *33*, 801–805. [CrossRef]

24. Spentzos, D.; Cannistra, S.A.; Grall, F.; Levine, D.A.; Pillay, K.; Libermann, T.A.; Mantzoros, C.S. IGF axis gene expression patterns are prognostic of survival in epithelial ovarian cancer. *Endocr.-Relat. Cancer* **2007**, *14*, 781–790. [CrossRef]

25. Dale, O.T.; Aleksic, T.; Shah, K.A.; Han, C.; Mehanna, H.; Rapozo, D.C.; Sheard, J.D.H.; Goodyear, P.; Upile, N.S.; Robinson, M.; et al. IGF-1R expression is associated with HPV-negative status and adverse survival in head and neck squamous cell cancer. *Carcinogenesis* **2015**, *36*, 648–655. [CrossRef] [PubMed]

26. Kim, J.-S.; Kim, E.S.; Liu, D.; Lee, J.J.; Solis, L.; Behrens, C.; Lippman, S.M.; Hong, W.K.; Wistuba, I.I.; Lee, H.-Y. Prognostic implications of tumoral expression of insulin like growth factors 1 and 2 in patients with non–small-cell lung cancer. *Clin. Lung Cancer* **2014**, *15*, 213–221. [CrossRef] [PubMed]

27. Ulanet, D.B.; Ludwig, D.L.; Kahn, C.R.; Hanahan, D. Insulin receptor functionally enhances multistage tumor progression and conveys intrinsic resistance to IGF-1R targeted therapy. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10791–10798. [CrossRef] [PubMed]

28. Belfiore, A.; Frasca, F.; Pandini, G.; Sciacca, L.; Vigneri, R. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr. Rev.* **2009**, *30*, 586–623. [CrossRef]

29. Buck, E.; Gokhale, P.C.; Koujak, S.; Brown, E.; Eyzaguirre, A.; Tao, N.; Rosenfeld-Franklin, M.; Lerner, L.; Chiu, M.I.; Wild, R.; et al. Compensatory insulin receptor (IR) activation on inhibition of insulin-like growth factor-1 receptor (IGF-1R): Rationale for cotargeting IGF-1R and IR in cancer. *Mol. Cancer Ther.* **2010**, *9*, 2652–2664. [CrossRef]

30. Janssen, J.A.; Varewijck, A.J. IGF-IR targeted therapy: Past, present and future. *Front. Endocrinol.* **2014**, *5*, 224. [CrossRef]

31. Ji, Q.-S.; Mulvihill, M.J.; Rosenfeld-Franklin, M.; Cooke, A.; Feng, L.; Mak, G.; O'Connor, M.; Yao, Y.; Pirritt, C.; Buck, E.; et al. A novel, potent, and selective insulin-like growth factor-I receptor kinase inhibitor blocks insulin-like growth factor-I receptor signaling in vitro and inhibits insulin-like growth factor-I receptor–dependent tumor growth in vivo. *Mol. Cancer Ther.* **2007**, *6*, 2158–2167. [CrossRef] [PubMed]

32. Mulvihill, M.J.; Cooke, A.; Rosenfeld-Franklin, M.; Buck, E.; Foreman, K.; Landfair, D.; O'Connor, M.; Pirritt, C.; Sun, Y.; Yao, Y.; et al. Discovery of OSI-906: A selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor. *Future Med. Chem.* **2009**, *1*, 1153–1171. [CrossRef] [PubMed]

33. Jones, R.L.; Kim, E.S.; Nava-Parada, P.; Alam, S.; Johnson, F.M.; Stephens, A.W.; Simantov, R.; Poondru, S.; Gedrich, R.; Lippman, S.M.; et al. Phase I study of intermittent oral dosing of the insulin-like growth factor-1 and insulin receptors inhibitor OSI-906 in patients with advanced solid tumors. *Clin. Cancer Res.* **2015**, *21*, 693–700. [CrossRef] [PubMed]

34. Puzanov, I.; Lindsay, C.R.; Goff, L.; Sosman, J.; Gilbert, J.; Berlin, J.; Poondru, S.; Simantov, R.; Gedrich, R.; Stephens, A.; et al. A phase I study of continuous oral dosing of OSI-906, a dual inhibitor of insulin-like growth factor-1 and insulin receptors, in patients with advanced solid tumors. *Clin. Cancer Res.* **2015**, *21*, 701–711. [CrossRef]

35. Fassnacht, M.; Berruti, A.; Baudin, E.; Demeure, M.J.; Gilbert, J.; Haak, H.; Kroiss, M.; Quinn, D.I.; Hesseltine, E.; Ronchi, C.L.; et al. Linsitinib (OSI-906) versus placebo for patients with locally advanced or metastatic adrenocortical carcinoma: A double-blind, randomised, phase 3 study. *Lancet Oncol.* **2015**, *16*, 426–435. [CrossRef]

36. van der Veeken, J.; Oliveira, S.; Schiffelers, R.M.; Storm, G.; van Bergen En Henegouwen, P.M.; Roovers, R.C. Crosstalk between epidermal growth factor receptor- and insulin-like growth factor-1 receptor signaling: Implications for cancer therapy. *Curr. Cancer Drug Targets* **2009**, *9*, 748–760. [CrossRef]

37. Qi, H.W.; Shen, Z.; Fan, L.H. Combined inhibition of insulin-like growth factor-1 receptor enhances the effects of gefitinib in a human non-small cell lung cancer resistant cell line. *Exp. Ther. Med.* **2011**, *2*, 1091–1095. [CrossRef]

38. Suda, K.; Mizuuchi, H.; Sato, K.; Takemoto, T.; Iwasaki, T.; Mitsudomi, T. The insulin-like growth factor 1 receptor causes acquired resistance to erlotinib in lung cancer cells with the wild-type epidermal growth factor receptor. *Int. J. Cancer* **2014**, *135*, 1002–1006. [CrossRef]

39. Camirand, A.; Zakikhani, M.; Young, F.; Pollak, M. Inhibition of insulin-like growth factor-1 receptor signaling enhances growth-inhibitory and proapoptotic effects of gefitinib (Iressa) in human breast cancer cells. *Breast Cancer Res.* **2005**, *7*, R570. [CrossRef]

40. Jones, H.; Gee, J.M.W.; Barrow, D.; Tonge, D.; Holloway, B.; Nicholson, R. Inhibition of insulin receptor isoform-A signalling restores sensitivity to gefitinib in previously de novo resistant colon cancer cells. *Br. J. Cancer* **2006**, *95*, 172–180. [CrossRef]

41. Urtasun, N.; Vidal-Pla, A.; Pérez-Torras, S.; Mazo, A. Human pancreatic cancer stem cells are sensitive to dual inhibition of IGF-IR and ErbB receptors. *BMC Cancer* **2015**, *15*, 223. [CrossRef]

42. Zanella, E.R.; Galimi, F.; Sassi, F.; Migliardi, G.; Cottino, F.; Leto, S.M.; Lupo, B.; Erriquez, J.; Isella, C.; Comoglio, P.M.; et al. IGF2 is an actionable target that identifies a distinct subpopulation of colorectal cancer patients with marginal response to anti-EGFR therapies. *Sci. Transl. Med.* **2015**, *7*, 272ra12. [CrossRef]

43. Iorio, F.; Knijnenburg, T.A.; Vis, D.J.; Bignell, G.R.; Menden, M.P.; Schubert, M.; Aben, N.; Gonçalves, E.; Barthorpe, S.; Lightfoot, H.; et al. A landscape of pharmacogenomic interactions in cancer. *Cell* **2016**, *166*, 740–754. [CrossRef] [PubMed]

44. Ghandi, M.; Huang, F.W.; Jané-Valbuena, J.; Kryukov, G.V.; Lo, C.C.; McDonald, E.R., 3rd; Barretina, J.; Gelfand, E.T.; Bielski, C.M.; Li, H.; et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **2019**, *569*, 503–508. [CrossRef]

45. Gautier, L.; Cope, L.; Bolstad, B.M.; Irizarry, R.A. Affy—Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **2004**, *20*, 307–315. [CrossRef] [PubMed]

46. Greenman, C.D.; Bignell, G.; Butler, A.; Edkins, S.; Hinton, J.; Beare, D.; Swamy, S.; Santarius, T.; Chen, L.; Widaa, S.; et al. PICNIC: An algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics* **2010**, *11*, 164–175. [CrossRef]

47. 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. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012**, *483*, 603–607. [CrossRef]

48. The Cancer Cell Line Encyclopedia Consortium; The Genomics of Drug Sensitivity in Cancer Consortium. Pharmacogenomic agreement between two cancer cell line data sets. *Nature* **2015**, *528*, 84–87. [CrossRef]

49. The International HapMap Consortium. The international hapmap project. *Nature* **2003**, *426*, 789–796. [CrossRef] [PubMed]

50. Venkatraman, E.S.; Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **2007**, *23*, 657–663. [CrossRef] [PubMed]

51. Di, J.; Zheng, B.; Kong, Q.; Jiang, Y.; Liu, S.; Yang, Y.; Han, X.; Sheng, Y.; Zhang, Y.; Cheng, L.; et al. Prioritization of can-didate cancer drugs based on a drug functional similarity network constructed by integrating pathway activities and drug activities. *Mol. Oncol.* **2019**, *13*, 2259–2277. [CrossRef] [PubMed]

52. Hänzelmann, S.; Castelo, R.; Guinney, J. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinform.* **2013**, *14*, 7. [CrossRef]

53. Barbie, D.A.; Tamayo, P.; Boehm, J.S.; Kim, S.Y.; Moody, S.E.; Dunn, I.F.; Schinzel, A.C.; Sandy, P.; Meylan, E.; Scholl, C.; et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **2009**, *462*, 108–112. [CrossRef]

54. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef]

55. Haibe-Kains, B.; El-Hachem, N.; Birkbak, N.J.; Jin, A.C.; Beck, A.H.; Aerts, H.J.; Quackenbush, J. Inconsistency in large pharmacogenomic studies. *Nature* **2013**, *504*, 389–393. [CrossRef] [PubMed]

56. Haverty, P.M.; Lin, E.; Tan, J.; Yu, Y.; Lam, B.; Lianoglou, S.; Neve, R.M.; Martin, S.; Settleman, J.; Yauch, R.L.; et al. Reproducible pharmacogenomic profiling of cancer cell line panels. *Nature* **2016**, *533*, 333–337. [CrossRef]

57. Weininger, D. SMILES, a chemical language and information system. 1—Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36. [CrossRef]

58. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **2021**, *49*, D1388–D1395. [CrossRef]

59. Tanimoto, T. IBM Internal Report 17th Nov; DJ Rogers and TT Tanimoto. *Science* **1957**, *132*, 1115–1118.

60. Giri, V.; Sivakumar, T.V.; Cho, K.M.; Kim, T.Y.; Bhaduri, A. RxnSim: A tool to compare biochemical reactions. *Bioinformatics* **2015**, *31*, 3712–3724. [CrossRef]

61. Freshour, S.L.; Kiwala, S.; Cotto, K.C.; Coffman, A.C.; McMichael, J.F.; Song, J.J.; Griffith, M.; Griffith, O.L.; Wagner, A.H. Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res.* **2021**, *49*, D1144–D1151. [CrossRef]

62. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef]

63. Perlman, L.; Gottlieb, A.; Atias, N.; Ruppin, E.; Sharan, R. Combining drug and gene similarity measures for drug-target elucidation. *J. Comput. Biol.* **2011**, *18*, 133–145. [CrossRef]

64. Zabłocka, A.; Kazana, W.; Sochocka, M.; Stańczykiewicz, B.; Janusz, M.; Leszek, J.; Orzechowska, B. Inverse correlation between Alzheimer's disease and cancer: Short overview. *Mol. Neurobiol.* **2021**, *58*, 6335–6349. [CrossRef]

65. Chen, D.; Hao, S.; Xu, J. Revisiting the relationship between Alzheimer's disease and cancer with a circRNA perspective. *Front. Cell Dev. Biol.* **2021**, *9*, 647197. [CrossRef]