

## RESEARCH ARTICLE

# Traditional Chinese Medicine studies for Alzheimer's disease via network pharmacology based on entropy and random walk

Xiaolu Wu<sup>1</sup>, Shujuan Cao<sup>1\*</sup>, Yongming Zou<sup>2</sup>, Fangxiang Wu<sup>3</sup>

**1** School of Mathematical Sciences, Tiangong University, Tianjin, China, **2** Department of Neurology, Tianjin Huanhu Hospital, Tianjin, China, **3** Division of Biomedical Engineering, Department of Mechanical Engineering and Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

\* [sj.cao@163.com](mailto:sj.cao@163.com)**OPEN ACCESS**

**Citation:** Wu X, Cao S, Zou Y, Wu F (2023) Traditional Chinese Medicine studies for Alzheimer's disease via network pharmacology based on entropy and random walk. PLoS ONE 18(11): e0294772. <https://doi.org/10.1371/journal.pone.0294772>

**Editor:** Jinghui Wang, Anhui University of Chinese Medicine, CHINA

**Received:** April 29, 2023

**Accepted:** November 8, 2023

**Published:** November 29, 2023

**Copyright:** © 2023 Wu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting information files](#).

**Funding:** This research was funded The Science Fund of Tianjin Education Commission for Higher Education, grant number 2019KJ025. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

## Abstract

Alzheimer's disease (AD) is a common neurodegenerative disease having complex pathogenesis, approved drugs can only alleviate symptoms of AD for a period of time. Traditional Chinese medicine (TCM) contains multiple active ingredients that can act on multiple targets simultaneously. In this paper, a novel algorithm based on entropy and random walk with the restart of heterogeneous network (RWRHE) is proposed for predicting active ingredients for AD and screening out the effective TCMs for AD. First, Six TCM compounds containing 20 herbs from the AD drug reviews in the CNKI (China National Knowledge Internet) are collected, their active ingredients and targets are retrieved from different databases. Then, comprehensive similarity networks of active ingredients and targets are constructed based on different aspects and entropy weight, respectively. A comprehensive heterogeneous network is constructed by integrating the known active ingredient-target association information and two comprehensive similarity networks. Subsequently, bi-random walks are applied on the heterogeneous network to predict active ingredient-target associations. AD related targets are selected as the seed nodes, a random walk is carried out on the target similarity network to predict the AD-target associations, and the associations of AD-active ingredients are inferred and scored. The effective herbs and compounds for AD are screened out based on their active ingredients' scores. The results measured by machine learning and bioinformatics show that the RWRHE algorithm achieves better prediction accuracy, the top 15 active ingredients may act as multi-target agents in the prevention and treatment of AD, Danshen, Gouteng and Chaihu are recommended as effective TCMs for AD, Yiqitongyutang is recommended as effective compound for AD.

## 1. Introduction

Alzheimer's disease (AD) is a persistent and irreversible neurodegenerative disease, whose main clinical features are progressive memory loss, cognitive decline, functional

independence loss and so on [1]. The pathogenesis of AD is complex, and not fully understood yet. The incidence of AD has been increasing in recent years, which brings a heavy economic burden to the healthcare system around the world. However, current clinical AD drugs approved by the U.S. Food and Drug Administration (FDA) only relieve related symptoms within a certain period, but none of these therapies can effectively halt the development of the disease. Other single-target AD drugs have failed in clinical trials. Therefore, it is essential to identify efficient and low-toxicity AD drugs with multi-targets [2].

Traditional Chinese medicine (TCM) contains multiple active ingredients that can act on multi-targets simultaneously, some active pharmacological compounds of herbs have been proven to be applied to the treatment of many diseases [3]. TCM has a long history of AD, many active ingredients extracted from herbs have fewer side effects and are regarded as potential anti-AD drugs [4]. However, the potential molecular mechanism of TCM for AD is still unclear, which limits further clinical applications. Network pharmacology is integral and systematic by integrating the ideas of pharmacology with network science, systems biology, and bioinformatics, which can be used to screen active ingredients and understand the mechanism of multi-ingredient, multi-target, and multi-pathway of TCM [5]. Therefore, network pharmacology can be used to reveal the associations between AD, active ingredients and targets, which open up a new way to study the mechanism between TCMs and AD.

With the development of high-throughput biomedical data, network-based propagation methods are often used to predict associations among biological components. Random Walk is one of the typical methods based on network propagation. The traditional restart random walk (RWR) algorithm only carries out random walks in a single network [6], many scholars have proposed improved algorithms for RWR on heterogeneous networks. Luo et al. proposed a random walk-based algorithm on the Reliable Heterogeneous Network (RWRHN) to prioritize potential candidate genes for inherited diseases [7]. Li et al. presented a superimposed local random walk algorithm called LRWHLDA to predict the associations between lncRNAs and diseases. Their algorithm overcomes the limitation of lack of known association between nodes [8]. The topological and structural properties of different networks are different. Wang et al. quantified the individual walk length of nodes by using the improved Jaccard index, and proposed an individual bi-random walk algorithm called DR-IBRW for drug repositioning [9]. Luo et al. put forward a bi-random walk algorithm (MBIRW) for drug repositioning, random walks are conducted to predict potential drug-disease associations in the drug similarity network and disease similarity network respectively [10]. The algorithm based on network inference is also applied to the prediction of biological information association. Cheng et al. proposed an algorithm based on network-based inference (NBI), known drug-target association information was used as the initial resource allocation, the final resource allocation information was obtained through two-step propagation to predict drug-target associations [11]. Wang et al. proposed a method based on the guilt-by-association principle, called HGBI for prediction of novel drug-target associations [12]. In addition, KATZ is also a network-based algorithm. Zhu et al. considered the contribution of different walk lengths to the similarity network, and proposed the HMDAKATZ algorithm to predict bacteria-drugs associations [13].

In this paper, we develop a novel algorithm based on entropy and random walk with the restart of heterogeneous network (RWRHE) for predicting active ingredients associated with AD and screening out the effective TCMs for AD. Firstly, the similarity measures of active ingredients are calculated from two aspects, including the chemical structure and the Gaussian interaction profile kernel (GIP kernel) similarity [14]. The similarity measures of targets are calculated from four aspects, including protein sequence, interaction score in String database (release 2021-08) [15], common neighbor, and GIP kernel similarity. Secondly, based on the weight of each similarity measure assigned by information entropy, similarity measures of

active ingredients and targets are integrated into comprehensive similarity measures, respectively. Therefore, a heterogeneous network can be constructed by connecting the comprehensive active ingredient network and target network via known active ingredient-target associations. Next, the active ingredients-target association score matrix is calculated by bi-random walk on heterogeneous networks. Then, AD related targets are selected as the seed nodes, and the random walk is performed on the target similarity network to calculate the AD-target association score vector. Finally, the relationships between AD and active ingredients are predicted and scored, the effects of TCMs for AD are scored and ranked. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis [16] are performed on the top 100 targets by different methods to illustrate the relationship between topological properties and relevant biological functions. Molecular docking is used to assess the ability of top 15 the active ingredients to enter the active pocket of the key enzymes or proteins of AD, the results show that the active ingredients may act as multi-target agents in the prevention and treatment of AD.

## 2. Materials and methods

### 2.1. Datasets

Six TCM compounds from the AD drug reviews in the CNKI (China National Knowledge Internet) are selected, including Dangguishaoyaosan, Yigansan, Yiqitongyutang, Gubenjian-naoye, BKHJ TCM and Yizhi [17–20]. A total of 20 herbs are involved. The active ingredients of these herbs are retrieved from Traditional Chinese Medicine System Pharmacology (TCMSP) database (release 2014–05) [21]. Based on the condition that oral bioavailability (OB) is more than 20% and drug-likeness (DL) is more than 0.1, 583 active ingredients are screened out. The herbs contained in TCM compounds and the active ingredients contained in the herbs are shown in [S1 Table](#).

The datasets used in this study include active ingredients, targets and known active ingredient-target associations, which are collected from two databases: TCMSP and HERB (release 2021–01) [22]. Targets associated with active ingredients are mapped into genes through UniProt database (release 2022–05) [23] for “Homo sapiens” organism. 4313 active ingredient-target associations involving 387 active ingredients and 374 targets are retrieved from TCMSP database. Simultaneously, 5823 active ingredient-target associations involving 453 active ingredients and 642 targets are obtained from HERB database. We take the union of these two datasets as the total dataset, which has 6973 active ingredient-target associations involving 457 active ingredients and 731 targets, as shown in [S2 Table](#).

### 2.2 Similarity measures

**2.2.1 Active ingredient similarity measures.** Let  $R = \{r_1, r_2, \dots, r_{n_r}\}$  be the set of  $n_r$  active ingredients and  $T = \{t_1, t_2, \dots, t_{n_t}\}$  be the set of  $n_t$  targets. The matrix  $A_{rt}$  represents known active ingredient-target associations, and its dimension is  $n_r * n_t$ . The value of  $A_{rt}(i, j)$  is 1 if active ingredient  $i$  and target  $j$  have a known association, otherwise is 0.

The first similarity measure is calculated based on the chemical structure. The SMILES (Simplified Molecular Input Line Entry Specification) describes the chemical structure of active ingredients [24], which could be retrieved from PubChem database (release 2022–06) [25]. For the active ingredients that cannot be retrieved, the SMILES can be obtained from their 2D structure through the Swiss Target Prediction database (release 2019–05) [26]. The Chemical Development Kit is used to compute chemical fingerprints of active ingredients

[27]. The similarity between active ingredients  $r_i$  and  $r_j$  is calculated by the Tanimoto score of the binary fingerprint vector [28], the formula as follows:

$$S_r^f(r_i, r_j) = \frac{f_{r_i} * f_{r_j}}{f_{r_i}^2 + f_{r_j}^2 - f_{r_i} * f_{r_j}} \quad (1)$$

where  $f_{r_i}$  represents the chemical fingerprint vector of the active ingredient  $r_i$ .

Based on the finding that weak similarity between active ingredient pairs provides less information for prediction, the logistic function is used to convert those small similarity values into values close to zero and expand those large similarity values simultaneously [10]. The improved similarity value by the logistic function is redefined as follows:

$$L(S_r^f(r_i, r_j)) = \frac{1}{1 + \exp(c \cdot S_r^f(r_i, r_j) + d)} \quad (2)$$

where  $c$  and  $d$  are the parameters. According to the previous research [29, 30], we set  $d$  as  $\log(9999)$ , which represents  $L(0) = 0.0001$ . We set  $L(0.3) < 0.01$ , which determines  $c$  as  $-15$ . The improved similarity value  $L(S_r^f(r_i, r_j))$  is denoted as  $S_r^l(r_i, r_j)$ .

Based on the assumption that similar active ingredients tend to associate with similar targets, GIP kernel similarity is used to calculate the similarity between active ingredients. The interaction profile  $IP(r_i)$  of active ingredient  $r_i$  is a binary vector representing the known associations between the active ingredient  $r_i$  and targets. The GIP kernel similarity between active ingredient  $r_i$  and active ingredient  $r_j$  is computed as follows:

$$S_r^2(r_i, r_j) = \exp(-\Upsilon_r \left\| IP(r_i) - IP(r_j) \right\|^2) \quad (3)$$

$$\Upsilon_r = \Upsilon_r' / \left( \frac{1}{n_r} \sum_{i=1}^{n_r} \left\| IP(r_i) \right\|^2 \right)$$

where  $IP(r_i)$  denotes the  $i$ -th row of the matrix  $A_{rt}$ .  $\Upsilon_r'$  is set to be 1 according to previous research [14]. The similarity values of active ingredients based on the above two similarity measures are shown in [S3 Table](#).

**2.2.2 Target similarity measures.** The first target similarity measure is calculated based on the protein sequences, which can be retrieved from UniProt database. The Smith-Waterman local alignment algorithm [31] is used to calculate the sequence similarity of targets, and the similarity matrix is denoted as  $S_t^1$ .

The second target similarity measure is based on the interaction confidence scores of the String database. The gene symbols of targets are entered into the String database, and the organisms are selected as "Homo sapiens". Then the interaction scores of target pairs are as elements of the second similarity matrix  $S_t^2$ .

The third target similarity measure is calculated based on the common neighbors of targets. To filter out the target interaction relationships with low confidence, target pairs with interaction scores greater than 0.4 in the String database are regarded as neighbors. The contribution of the common neighbor targets with small degree is greater than that of the common neighbor targets with large degree, each common neighbor target is assigned a weight based on its degree value. The target similarity measure based on the common neighbors is defined as:

$$S_t^3(t_i, t_j) = \sum_{z \in \Gamma(t_i) \cap \Gamma(t_j)} \frac{1}{k(z)} \quad (4)$$

where  $\Gamma(t_i)$  is the set of neighbors of target  $t_i$ ,  $z$  are the common neighbors of targets  $t_i$  and  $t_j$ ,  $k(z)$  is the degree of  $z$ . If two targets have more common neighbors and the degree of common neighbors is small, the similarity between them will be greater than 1, it will be replaced by 0.99.

Based on the assumption that similar targets tend to associate with similar active ingredients, the fourth target similarity measure based on GIP kernel is defined. The interaction profile  $IP(t_i)$  of target  $t_i$  is a binary vector representing the known associations between target  $t_i$  and active ingredients. The GIP kernel similarity between targets  $t_i$  and  $t_j$  is computed as follows:

$$S_t^4(t_i, t_j) = \exp(-\Upsilon_t \|IP(t_i) - IP(t_j)\|^2) \tag{5}$$

$$\Upsilon_t = \Upsilon_t' / \left( \frac{1}{n_t} \sum_{i=1}^{n_t} \|IP(t_i)\|^2 \right)$$

where  $IP(t_i)$  denotes the  $i$ -th column of the matrix  $A_{rt}$ .  $\Upsilon_t'$  is set to be 1 according to previous research [14]. The similarity values of targets based on the above four similarity measures are shown in S4 Table.

**2.2.3 Integrating similarity measures based on entropy.** Different similarity measures contain different similarity information and play different roles in measuring the similarity of node pairs. The information entropy is used to select similarity measures in the previous research [32]. In this study, the weights of different similarity measures are further calculated based on entropy.

For the  $m$ -th similarity matrix for active ingredients, the entropy  $E_i^m$  of the  $i$ -th row is calculated as follows:

$$E_i^m = - \sum_{j=1}^k \frac{s_{ij}}{\sum_{j=1}^k s_{ij}} \log \left( \frac{s_{ij}}{\sum_{j=1}^k s_{ij}} \right) \tag{6}$$

where  $s_{ij}$  represents the similarity value between active ingredients  $i$  and  $j$ .  $k$  indicates the number of active ingredients. We average the entropy of all rows as the final average entropy value. The average entropy of the  $m$ -th similarity matrix is calculated as follows:

$$E_{mean}^m = \frac{\sum_{i=1}^k E_i}{k} \tag{7}$$

The smaller the average entropy of the similarity matrix is, the less random information is delivered by the similarity measure. The similarity measure with small average entropy occupies a significant proportion of the comprehensive similarity measure. The average entropy of each similarity matrix is normalized after taking the reciprocal, then the weight of the  $m$ -th similarity measure is defined as follows:

$$\omega_{rm} = \frac{1/E_{mean}^m}{\sum_n 1/E_{mean}^n} \tag{8}$$

where  $\omega_{rm}$  represents the weight of the  $m$ -th similarity measure of active ingredients. In the same way, the average entropy and weight of each target similarity measure can be obtained.

**Table 1. Entropy and weight of similarity measure.**

		$E_{mean}$	$\omega$
Active ingredient	$S_r^1$	3.9734	0.5986
	$S_r^2$	5.9265	0.4014
Target	$S_t^1$	6.5938	0.2071
	$S_t^2$	4.5929	0.2974
	$S_t^3$	4.9661	0.2750
	$S_t^4$	6.1959	0.2204

<https://doi.org/10.1371/journal.pone.0294772.t001>

The average entropy and corresponding weights of the similarity matrix of two active ingredients and four targets are shown in Table 1.

Finally, the two similarity measures of active ingredients are integrated into a comprehensive active ingredient similarity measure  $S_r$ , and the four similarity measures of targets are integrated into a comprehensive target similarity measure  $S_t$ , which are calculated as follows:

$$\begin{aligned} S_r &= \omega_{r1}S_r^1 + \omega_{r2}S_r^2 \\ S_t &= \omega_{t1}S_t^1 + \omega_{t2}S_t^2 + \omega_{t3}S_t^3 + \omega_{t4}S_t^4 \end{aligned} \quad (9)$$

where  $\omega_{r1}$  and  $\omega_{r2}$  are the weights of active ingredient similarity matrix  $S_r^1$  and  $S_r^2$ , respectively.  $\omega_{t1}$ ,  $\omega_{t2}$ ,  $\omega_{t3}$ ,  $\omega_{t4}$  are the weights of target similarity matrix  $S_t^1$ ,  $S_t^2$ ,  $S_t^3$ ,  $S_t^4$ , respectively. The comprehensive similarity values of active ingredients and targets are shown in S5 Table.

### 2.3 Construction of the heterogeneous network

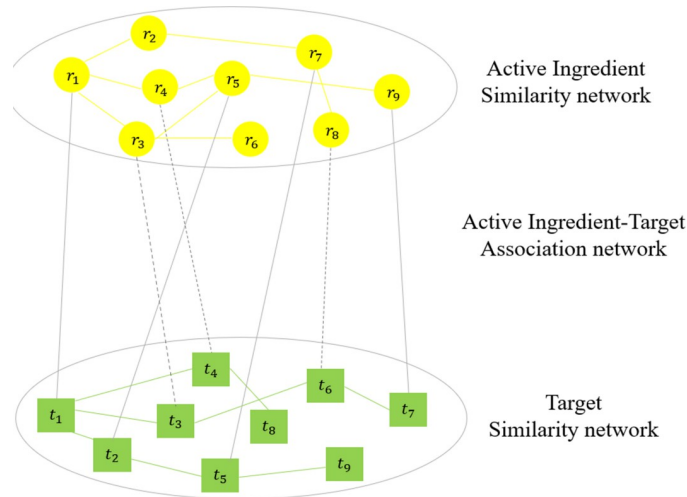
The similarity networks of active ingredients and targets are constructed based on the comprehensive similarity measures of active ingredients and targets.  $R = \{r_1, r_2, \dots, r_{n_r}\}$  is the node set of  $n_r$  active ingredients. The comprehensive similarity  $S_r(r_i, r_j)$  is the weight between active ingredients  $r_i$  and  $r_j$ .  $T = \{t_1, t_2, \dots, t_{n_t}\}$  is the node set of  $n_t$  targets. The comprehensive similarity  $S_t(t_i, t_j)$  is the weight between targets  $t_i$  and  $t_j$ .

In addition, the active ingredient-target bipartite graph  $G(V, E)$  is constructed based on the known active ingredient-target associations.  $V = \{R, T\}$  is the node set containing active ingredient nodes and target nodes.  $E = \{A_{rt}(i, j)\}$  is the edge set. If active ingredient  $i$  and target  $j$  have a known association, the weight of edge between them is 1, otherwise is 0.

The active ingredient-target heterogeneous network can be constructed by integrating the active ingredient similarity network, target similarity network, and the known active ingredient-target association network. The heterogeneous network is illustrated in Fig 1. The yellow circles and green rectangles represent active ingredients and targets, respectively. Solid lines represent known active ingredient-target associations, dashed lines indicate the predicted active ingredient-target associations.

### 2.4 Bi-random walk on the heterogeneous network

On the heterogeneous network, a bi-random walk algorithm is used to predict the active ingredient-target associations score matrix. As the previous research [33], the Laplacian normalization is used to normalize the weight matrix of a network. For the active ingredient similarity



**Fig 1. The active ingredient-target heterogeneous network.** Yellow circles and green rectangles represent active ingredients and targets respectively.

<https://doi.org/10.1371/journal.pone.0294772.g001>

matrix  $S_r$ , the Laplacian normalization is divided into two steps:

$$S'_r(r_i, r_j) = \frac{S_r(r_i, r_j)}{\sqrt{D(r_i, r_i) * D(r_j, r_j)}} \tag{10}$$

$$S''_r(r_i, r_j) = \frac{S'_r(r_i, r_j)}{\sum_j S'_r(r_i, r_j)} \tag{11}$$

where  $D$  is a diagonal matrix, and the elements of the main diagonal are the sum of corresponding rows of the matrix  $S_r$ , i.e  $D(r_i, r_i) = \sum_{k=1}^{n_r} S_r(r_i, r_k)$ . The matrix  $S_r$  after Laplacian normalization is denoted as  $S'_r$ . Likewise, the target similarity matrix  $S_t$  after Laplacian normalization is denoted as  $S'_t$ . In addition, the adjacency matrix  $A_{rt}$  of the active ingredient-target associations is normalized as follows:

$$A'_{rt} = \frac{A_{rt}}{sum(A_{rt})} \tag{12}$$

A random walk begins from an active ingredient, then traverses to other target nodes based on its associated targets. The probabilistic associations between the active ingredient and all targets are obtained. The left random walk is performed on the target similarity network to simulate this process:

$$leftRT_t = \alpha \times RT_{t-1} * S'_t + (1 - \alpha) \times A'_{rt} \tag{13}$$

where  $\alpha = 0.3$  is the restart probability, which controls the probability for the walker staying at the starting node.  $leftRT_t$  is the predicted association matrix between active ingredients and targets in the t-step iteration,  $RT_0 = A'_{rt}$ . The left random walk stops until  $|leftRT_{t+1} - leftRT_t| < 10^{-6}$ , the resulted matrix is denoted  $leftRT$ .

Likewise, a random walk starts from a target node, then traverses other active ingredient nodes based on its known associated active ingredients. Then another active ingredient-target



association matrix is obtained. The right random walk is conducted on the active ingredient similarity network to mimic this process:

$$\text{rightRT}_t = \alpha \times S_r'' \times \text{RT}_{t-1} + (1 - \alpha) \times A_t' \quad (14)$$

where  $\alpha = 0.3$  is the restart probability, the right random walk stops until  $|\text{rightRT}_{t+1} - \text{rightRT}_t| < 10^{-6}$ , the resulted matrix is denoted *rightRT*. The final result is the average of the left and right random walk results, and the final predicted active ingredient-target association matrix is denoted as *RT* as follows:

$$\text{RT} = \frac{\text{leftRT} + \text{rightRT}}{2} \quad (15)$$

The final matrix *RT* is shown in [S6 Table](#).

## 2.5 Prediction of AD-target association

In this study, the targets from the Alzheimer's disease pathway (hsa05010) in the KEGG and the targets contained in the AD mini metabolic network [34] are selected. The intersection of these targets and aforementioned 731 targets contains 72 targets, which form the seed node set. The initial probability vector  $p_0$  of targets is constructed such that equal probability is assigned to each target seed, and probability 0 is assigned to other targets, with the sum of the probabilities equal to 1. Starting from the target seed node, according to the topological property of the target similarity network, random walks are carried out to traverse other targets. Then, the AD-targets association probabilities vector could be predicted as follows:

$$P_t = \alpha * S_t'' * P_{t-1} + (1 - \alpha)P_0 \quad (16)$$

where  $\alpha = 0.3$  is the restart probability.  $P_t$  is the probability vector of the  $t$ -step iteration, and its value represents the association probabilities between targets and AD. The iteration stops until  $|P_{t+1} - P_t| < 10^{-6}$ , the resulted probability vector is denoted as  $P$ , and its values are as shown in [S7 Table](#).

## 2.6 Prediction of AD-active ingredient association

Combined with the final predicted active ingredient-target association matrix *RT* and the AD-targets association vector  $P$ , the association scores of active ingredients and AD are predicted as follows:

$$P_r(r_i) = \sum_j \text{RT}(r_i, t_j) * P(t_j) \quad (17)$$

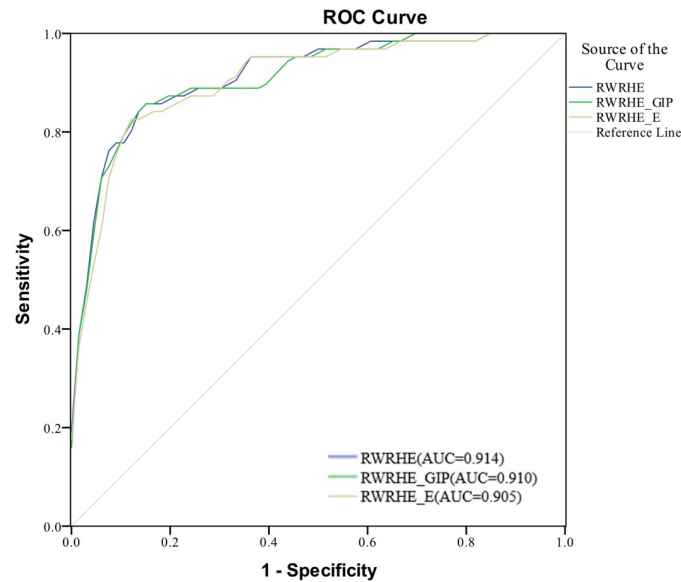
where  $\text{RT}(r_i, t_j)$  represents the final predicted association score of active ingredient  $r_i$  and target  $t_j$ ,  $P(t_j)$  represents the score of the target  $t_j$  associated with AD.  $P_r(r_i)$  represents the score of the active ingredient  $r_i$  associated with AD, as shown in [S8 Table](#).

## 3. Results and discussion

### 3.1 Measuring the effect of similarity by ablation analysis

Different similarity measures in a network can lead to differences in the topology and structure characteristics of the network. Therefore, the iterative process of random walk and the accuracy of prediction results are affected by different measures. Ablation analysis is performed on the comprehensive similarity measure to study the influence of different similarity measures





**Fig 2.** The ROC curves of RWRHE, RWRHE\_GIP and RWRHE\_E.

<https://doi.org/10.1371/journal.pone.0294772.g002>

on the performance of the RWRHE algorithm. We implement two simplified variants of RWRHE.

1. RWRHE\_GIP: By removing GIP kernel similarity, the active ingredient similarity  $S_r^2$  and target similarity  $S_t^4$  based on GIP kernel similarity are removed.
2. RWRHE\_E: By removing entropy, each similarity measure is weighted equally.

The Receiver Operating Characteristic Curve (ROC) reflects the relationship between True Positive Rate (TPR) and False Positive Rate (FPR) at different thresholds. The area under the ROC curve (AUC) is used as an evaluation metric to measure the accuracy of prediction results. “Alzheimer’s Disease” is used as the keyword, 63 active ingredients associated with AD are retrieved from the HERB database as the positive control group. At the same time, the active ingredients related to leukemia, mammary carcinoma, fever, and acute pharyngitis are retrieved, and the active ingredients related to AD are excluded, 66 active ingredients are obtained as the negative control group. The positive and negative control groups are shown in S9 Table. TPR, also known as sensitivity, represents the ratio of active ingredients in positive controls that rank for association with AD above the specified threshold. FPR, also known as 1-specificity, denotes the percentage of active ingredients in negative control group that rank for association with AD above the specified threshold. ROC curves of RWRHE, RWRHE\_GIP, and RWRHE\_E algorithms are shown in Fig 2. The results show that the RWRHE algorithm achieves higher prediction accuracy, and its AUC value is 0.914. The AUC values of RWRHE\_GIP and RWRHE\_E algorithms are 0.910 and 0.905, respectively. The number of nodes and edges in the network corresponding to the RWRHE\_GIP and the RWRHE\_E is the same as that of the RWRHE, but each edge has different weight. The GIP kernel similarity integrates the known active ingredient-target associations information, which makes the prediction results more reliable. The importance of each similarity measure is quantitatively weighted based on entropy, which makes the similarity network more complete. Therefore, integrating GIP kernel similarity and information entropy can improve the performance of prediction results to a certain extent.

### 3.2 Comparison with other methods

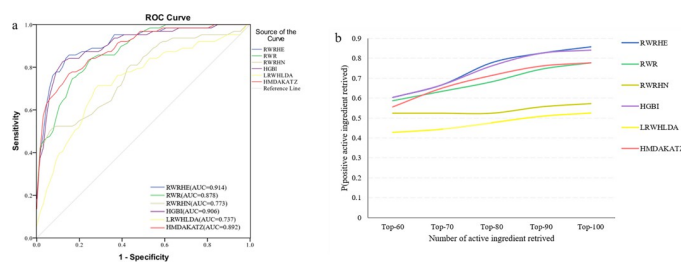
To further evaluate the performance of RWRHE algorithm, other five network-based algorithms RWR [6], RWRHN [7], HGBI [12], LRWLDA [8] and HMDAKATZ [13] are compared with RWRHE. RWR algorithm is implemented on the node similarity network. The walker returns to the seed node with a certain probability in the iterative process. All nodes are ranked according to the final probability. The RWRHN algorithm takes into account the jump probability of nodes in heterogeneous networks when constructing the transition matrix. Active ingredients related to AD are retrieved from the SymMap database (release 2019–01) [35], and the top 50 active ingredients in the inferred evidence score are selected as seed nodes. The RWR and RWRHN algorithms are used to calculate the association scores between AD and all active ingredients. The HGBI algorithm based on heterogeneous graph reasoning infers potential node associations by constructing heterogeneous networks. The LRWLDA algorithm is a superimposed local random walk algorithm. The HMDAKATZ algorithm considers the contribution of different walks to the association probability, and considers that the longer walks tend to contribute less to similarity. We apply HGBI, LRWLDA and HMDAKATZ algorithms to calculate the active ingredient-target association probability matrix, which is compared with the bi-random walk part in this study. The ROC curves of RWRHE and the other five algorithms are shown in Fig 3(a). The results show that the RWRHE algorithm has excellent prediction performance, and its AUC value is 0.914, which is higher than the AUC values of the other five algorithms.

In addition, the ranks of prediction results also play an important role in evaluating the performance of all algorithms. The cumulative distribution function (CDF) of the ranks is employed to compare the performance of different algorithms. CDF refers to the proportion of active ingredients whose ranking exceeds the top- $r$  threshold in the positive control group.  $60 \leq r \leq 100$  are reported as shown in Fig 3(b). The results show that compared with the other five algorithms, the RWRHE algorithm retrieves the largest proportion under the same top- $r$  threshold. The majority of the 63 active ingredients in the positive control group are retrieved in the top 100. Therefore, RWRHE has a good performance in predicting the relationship between AD and active ingredients.

The RWRHE algorithm predicts the association scores and ranks between active ingredients and AD. According to the active ingredients contained in herbs and ranks of active ingredients associated with AD, the effect score of herb  $k$  in treating AD is calculated based on ranks as follows:

$$H_k = \sum_i \frac{1}{\text{rank}(i)} \quad (18)$$

where,  $\text{rank}(i)$  represents the rank of the active ingredient  $i$  associated with AD. Finally, 20



**Fig 3.** (a) The ROC curves of RWRHE, RWR, RWRHN, HGBI, HMDAKATZ and LRWLDA. (b) The ranks of CDF.

<https://doi.org/10.1371/journal.pone.0294772.g003>

**Table 2. Ranking of efficacy of compound in treating AD.**

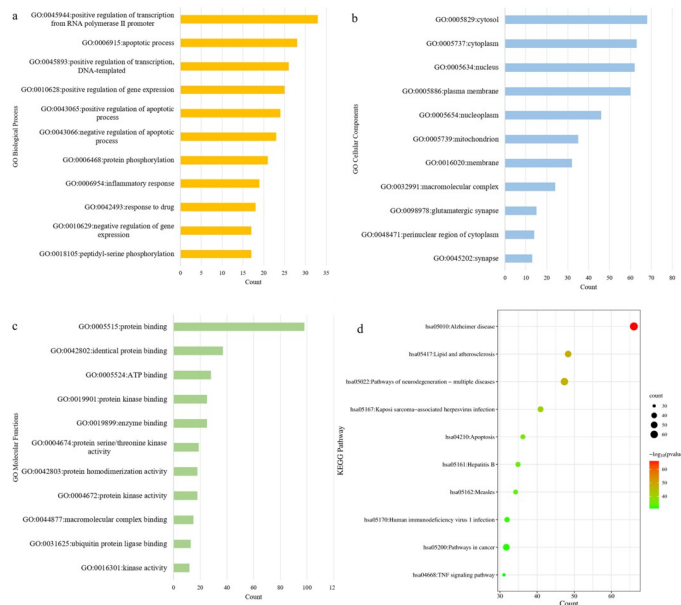
Compound	score	rank
Yiqitongyutang	6.229121413	1
Yigansan	4.196828293	2
BKHJ TCM	2.394613178	3
Gubenjiannaoye	2.28103394	4
Danguishaoyaosan	0.991858956	5
Yizhi	0.165683049	6

<https://doi.org/10.1371/journal.pone.0294772.t002>

herbs are scored and ranked according to the predicted effect scores of herbs for AD, shown in [S10 Table](#), the results show that Danshen have the best effect for AD, followed by Gouteng and Chaihu. Similarly, based on the active ingredients contained in the compound, the efficacy of the compound for AD is scored and ranked according to [Eq \(18\)](#), the results are shown in [Table 2](#). The results show that Yiqitongyutang is the best compound for AD compared with other compounds, which indicates that Yiqitongyutang may have better therapeutic effect for AD.

### 3.3 GO enrichment analysis and KEGG pathway analysis

The top 100 targets associated with AD are regarded as potential targets of AD. To learn more information about TCMs and AD, bioinformatics analysis is used to find relevant information about potential targets of AD. GO enrichment analysis and KEGG pathway analysis are performed to clarify relevant biological information about the potential targets of AD. GO enrichment analysis includes biological process (BP), cell component (CC), and molecular function (MF). To screen out the most significantly enriched biological annotations, the top 11 entries with the lowest P value are selected respectively, as shown in [Fig 4\(a\)–4\(c\)](#). In



**Fig 4. GO enrichment analysis and KEGG pathway analysis of potential targets.** (a): the histogram of GO biological process; (b) the histogram of GO cell component; (c) the histogram of molecular function; (d) the bubble graph of KEGG pathway.

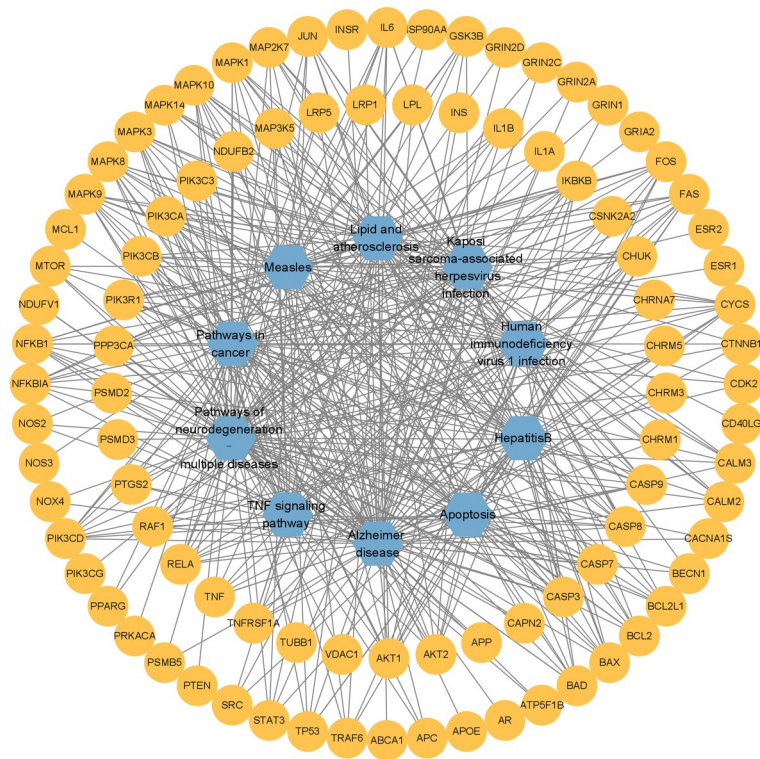
<https://doi.org/10.1371/journal.pone.0294772.g004>

biological process results, the most enriched GO term ‘Positive regulation of transcription from RNA polymerase II promoter’ may be involved in the transcription of the substance related to synaptic connectivity. Synapses play a central role in learning and memory, and disorders of these behaviors can lead to neurological diseases, including AD [36]. The targets are also mainly concentrated on positive and negative regulation of gene expression and apoptotic biological processes, suggesting that apoptosis and gene expression regulation play an important role in the mechanism of AD. In addition, the targets are involved in the biological process of protein phosphorylation. Protein phosphorylation is the key mechanism of AD [37]. Highly phosphorylated tau protein forms neurofibrillary tangles (NFTs), which is one of the main histopathological features of AD [38]. In cell component results, the targets are mainly concentrated on cytoplasm, nucleus, plasma membrane and so on. In molecular functional results, the targets are mainly concentrated on protein kinase binding, protein kinase activity, protein serine/threonine kinase activity and protein kinase activity. Some researchers report that Glycogen synthase kinase-3 $\beta$  (Gsk3 $\beta$ ), Ccyclin-dependent kinase 5 (CDK5) and Microtubule affinity regulating kinase (MARK) may hyper-phosphorylate tau and accelerate the formation of NFTs [39–41]. Cyclin-dependent kinases (Cdks), which are serine/threonine kinases, regulate cell cycle and neuronal differentiation. Cdks pathway may have effects on neuron loss, which is responsible for AD [42].

The involvement of pathways based on potential targets by KEGG pathway enrichment analysis is plentiful. The top 10 pathways with the smallest P value are drawn as bubble graphs in Fig 4(d). The size of the dots in the bubble diagram represents the number of targets enriched in the pathway, and the depth of the color represents the size of the statistical test P-value. The number of these targets enriched in the AD pathway is the highest. In addition, the potential targets are mainly concentrated on neurodegenerative diseases, apoptosis, and cancer pathways.

To further illustrate the relationship between topological properties and relevant biological functions, GO enrichment analysis and KEGG pathway analysis are performed on top 100 targets predicted by the RWRHE\_GIP and RWRHE\_E algorithm, respectively, shown in S11 Table. The results of the two algorithms do not include the GO term ‘Positive regulation of transcription from RNA polymerase II promoter’, the GO term ‘inflammatory response’ and the KEGG pathway ‘Pathways in cancer’. The GO term ‘Positive regulation of transcription from RNA polymerase II promoter’ have contribution to the pathogenesis of AD, which has been discussed above. Numerous studies have revealed the strong contribution of inflammation to AD pathogenesis, A $\beta$  deposition in AD is related to inflammatory response [43–45]. Some cancer-related signaling pathways including FOXO, mTOR, SIRT1, HIF, oxidative stress, inflammation, and metabolism have important roles in regulating aging and AD [46]. The GIP similarity and the weights based on entropy are necessary in integrating networks, which could further display the entries related to AD.

To further explain the mechanism of multi-pathway and multi-target in the treatment of AD with TCMs, a target-pathway network is constructed based on the top 10 pathways of KEGG enrichment analysis. The relationships between the targets and the pathways are intuitively visualized, as shown in Fig 5. The same pathway is connected to different targets, in the meantime, the same target is connected to different pathways, which indicates that they are interactive relationships. Different targets play a synergistic role in the same pathway. And the pathway also regulates gene transcription and effects gene expression. Compared with single-target drugs, TCMs show the advantage of multi-target and multi-pathway.



**Fig 5. The pathway-target network.** The orange circles represent targets, the blue hexagons represent pathways, and the gray lines represent the target-pathway connections.

<https://doi.org/10.1371/journal.pone.0294772.g005>

### 3.4 Molecular docking

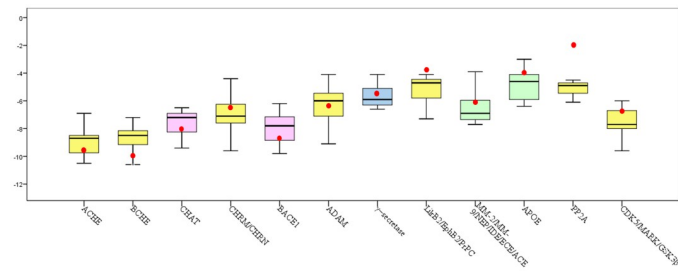
The components of TCM are complex and the targets are diverse. Molecular docking can indicate the binding ability between active ingredients and related targets. The Coach-D server is presented to predict the protein-ligand binding sites and ligand-binding poses by molecular docking [47]. The minimum value of Energy<sup>u</sup> by Coach-D is regarded as the final docking energy to evaluate the binding ability between active ingredient and target.

The panel of ligands is composed of the top 15 active ingredients by RWRHE and several approved acetylcholinesterase (ACHE) inhibitors for AD. Several approved ACHE inhibitors for AD clinically are Galantamine, Donepezil, Rivastigmine, Tacrine and Huperzine A. The first four are approved by FDA, and Huperzine A is approved in China.

76 key enzymes and receptors from AD metabolic network [34] are sorted out to form the panel of proteins, and they are divided into 12 categories according to their main functions in the AD metabolic network, the structure versions of proteins are selected according to better resolution from Protein data bank (PDB), shown in S12 Table.

The minimum values of Energy<sup>u</sup> for these possible complexes formed by ligands and proteins in the panels are calculated by Coach-D. The smaller the minimum value of docking energy is, the stronger the stability of active ingredient-protein site binding is. The docking energies of these five ACHE inhibitors and 12 categories of proteins are computed and the box plots are shown in Fig 6. The docking energies of these five ACHE inhibitors and their common target ACHE are used as the standard for active ingredients. The docking energies of the top 15 active ingredients and 76 proteins in the panel are given by Coach-D, then the





**Fig 6. The box plot of molecular docking.**

<https://doi.org/10.1371/journal.pone.0294772.g006>

average docking values of the top 15 active ingredients and 12 categories of proteins are computed and represented with red dots in Fig 6. The five approved ACHE inhibitors' common target is the enzyme ACHE. BCHE is the isozyme of ACHE, the median docking energies of these five ACHE inhibitors with ACHE/BCHE are -8.7/-8.5, almost all docking values of the top 15 active ingredients and ACHE are less than -8.7 and the average value is -9.73, all docking values of the top 15 active ingredients and BCHE are less than -8.5 and the average value is -10.10. The results indicate that the top 15 active ingredients can bind to ACHE/BCHE and improve the level of acetylcholine (ACH), which is involved in both memory and learning. Similarly, the average docking value of the top 15 active ingredients and BACE1 is -9.45 and less than -8.7, the result shows that the top 15 active ingredients may bind to BACE1, which may reduce the generation of Amyloid- $\beta$  ( $A\beta$ ),  $A\beta$  is the main component of senile plaques. Among the top 15 active ingredients, some active ingredients are confirmed as ACHE or BACE1 inhibitors in previous research through wet experiments. Álvarez-Berbel et al. show that Quercetin and apigenin are characterized as inhibition of the enzymatic activity of ACHE [48]. Beg et al. report that the exposure of AD flies to kaempferol reduces ACHE activity [49]. In addition, Han et al. report that Baicalein exhibits strong BACE1 and ACHE inhibitory properties [50]. Youn et al. show that oleic acid exerts significant noncompetitive inhibitory activity against BACE1 [51]. Also, the docking results show that the top 15 active ingredients may bind to CDK5/MARK/GSK3 $\beta$  and restrain the phosphorylation of tau and reduce the formation of NFTs. It is well known that Memory loss, Senile plaques and NFTs are the main histopathological features of AD. The results show that the active ingredients can enter the active pocket of related targets well and act as multi-target agents in the prevention and treatment of AD.

#### 4. Conclusion

AD is a complex neurodegenerative disease with few approved drugs. TCM has a long history and has a strong clinical basis for more than two thousand years. TCM has the advantage of more active ingredients, multi-target cooperative regulation and lower toxicity. Therefore, the study of new active ingredients from herbs is of great significance for the development of anti-AD drugs.

In this study, a novel random walk algorithm called RWRHE is devised to predict active ingredients and effective TCMs associated with AD based on entropy and random walk with the restart of a heterogeneous network. The comprehensive heterogeneous network is constructed by integrating the known active ingredient-target association network, active ingredient similarity network and target similarity networks based on entropy. The active ingredients and effective TCMs for AD are inferred based on random walks. The results measured by machine learning and bioinformatics show that the RWRHE algorithm achieves better

prediction accuracy. Particularly, the docking energies of the five approved ACHE inhibitors and their common target ACHE are used as the standard for active ingredients, the results show that the top 15 active ingredients may improve the level of Ach, reduce the generation of A $\beta$  and restrain the phosphorylation of tau, which are involved in the main histopathological features of AD. 20 herbs are ranked according to the active ingredients contained in herbs and ranks of active ingredients associated with AD, Danshen, Gouteng and Chaihu are recommended as effective TCMs for AD. Yiqitongyutang is recommended as effective compound for AD in the same way.

This study may provide new directions for building more effective prediction models to identify novel AD drugs. But there are several potential limitations in the current study that could be improved. In future, more validated association data would be incorporated to improve the quality of the heterogeneous network, and more useful and detailed studies would be integrated to improve the prediction ability. Only some predicted active ingredients have been validated for acting on key enzymes of AD in different published papers, we will take experimental validation into account and try to design wet experiments to verify the effectiveness of the predicted active ingredients in future.

## Supporting information

**S1 Table. TCM compounds-herbs-active ingredients.**

(XLSX)

**S2 Table. Active ingredients, targets and known active ingredient-target associations from TCMSP and HERB.**

(XLSX)

**S3 Table. The similarity values of active ingredients based on two similarity measures.**

(XLSX)

**S4 Table. The similarity values of targets based on four similarity measures.**

(XLSX)

**S5 Table. The comprehensive similarity values of active ingredients and targets.**

(XLSX)

**S6 Table. The final active ingredient-target association matrix  $RT$ .**

(XLSX)

**S7 Table. AD-target associations vector  $P$ .**

(XLSX)

**S8 Table. AD-active ingredient association vector  $P_r$ .**

(XLSX)

**S9 Table. The positive and negative control groups.**

(XLSX)

**S10 Table. The predicted effect ranks of herbs in treating AD.**

(XLSX)

**S11 Table. GO enrichment analysis and KEGG pathway analysis on top 100 targets predicted by RWRHE\_GIP and RWRHE\_E.**

(XLSX)



**S12 Table. The panel of proteins.**  
(XLSX)

## Author Contributions

**Methodology:** Xiaolu Wu, Shujuan Cao.

**Supervision:** Shujuan Cao, Yongming Zou, Fangxiang Wu.

**Writing – original draft:** Xiaolu Wu, Shujuan Cao.

**Writing – review & editing:** Fangxiang Wu.

## References

1. Zverova M. Clinical aspects of Alzheimer's disease. *Clinical Biochemistry*. 2019; 72:3–6. <https://doi.org/10.1016/j.clinbiochem.2019.04.015> PMID: 31034802
2. Li L, Zhang L, Yang CC. Multi-Target Strategy and Experimental Studies of Traditional Chinese Medicine for Alzheimer's Disease Therapy. *Current Topics in Medicinal Chemistry*. 2016; 16(5):537–48. <https://doi.org/10.2174/1568026615666150813144003> PMID: 26268330
3. Wu TY, Chen CP, Jinn TR. Traditional Chinese medicines and Alzheimer's disease. *Taiwanese Journal of Obstetrics & Gynecology*. 2011; 50(2):131–5. <https://doi.org/10.1016/j.tjog.2011.04.004> PMID: 21791295
4. Sun ZK, Yang HQ, Chen SD. Traditional Chinese medicine: a promising candidate for the treatment of Alzheimer's disease. *Transl Neurodegener*. 2013 Feb 28; 2(1):6. <https://doi.org/10.1186/2047-9158-2-6> PMID: 23445907
5. Chen J, Chen QL. Current situation and thinking of network pharmacology in the research of traditional Chinese medicine. *Acta Universitatis Traditionis Medicinalis Sinensis Pharmacologiaeque Shanghai*. 2021; 35(05):1–6+13. <https://doi.org/10.16306/j.1008-861x.2021.05.001>
6. Tong HH, Faloutsos C, and Pan JY, Fast random walk with restart and its applications, IEEE International Conference on Data Mining, 2006, . 613–622.
7. Luo JW, Liang SY. Prioritization of potential candidate disease genes by topological similarity of protein-protein interaction network and phenotype data. *Journal of Biomedical Informatics*. 2015; 53:229–36 <https://doi.org/10.1016/j.jbi.2014.11.004> PMID: 25460206
8. Li JC, Zhao HC, Xuan ZW, Yu JW, Feng X, Liao B, et al. A Novel Approach for Potential Human LncRNA-Disease Association Prediction Based on Local Random Walk. *Ieee-Acm Transactions on Computational Biology and Bioinformatics*. 2021; 18(3):1049–59. <https://doi.org/10.1109/TCBB.2019.2934958> PMID: 31425046
9. Wang YH, Guo MZ, Ren YZ, Jia LY, Yu GX. Drug repositioning based on individual bi-random walks on a heterogeneous network. *BMC Bioinformatics*. 2019. <https://doi.org/10.1186/s12859-019-3117-6> PMID: 31874623
10. Luo HM, Wang JX, Li M, Luo JW, Peng XQ, Wu FX, et al. Drug repositioning based on comprehensive similarity measures and Bi-Random walk algorithm. *Bioinformatics*. 2016; 32(17):2664–71. <https://doi.org/10.1093/bioinformatics/btw228> PMID: 27153662
11. Cheng FX, Liu C, Jiang J, Lu WQ, Li WH, Liu GX, et al. Prediction of Drug-Target Interactions and Drug Repositioning via Network-Based Inference. *Plos Computational Biology*. 2012; 8(5)10. <https://doi.org/10.1371/journal.pcbi.1002503> PMID: 22589709
12. Wang W, Yang S, Li J. Drug target predictions based on heterogeneous graph inference. *Pac Symp Biocomput*. 2013:53–64. PMID: 23424111
13. Zhu LZ, Duan GH, Yan C, Wang JX. Prediction of Microbe-drug Associations Based on Chemical Structures and the KATZ Measure. *Current Bioinformatics*. 2021; 16(6):807–19. <https://doi.org/10.2174/1574893616666210204144721>
14. Yan C, Duan GH, Pan Y, Wu FX, Wang JX. DDIGIP: predicting drug-drug interactions based on Gaussian interaction profile kernels. *BMC Bioinformatics*. 2019. <https://doi.org/10.1186/s12859-019-3093-x> PMID: 31874609
15. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021 Jan 8; 49(D1):D605–D612. <https://doi.org/10.1093/nar/gkaa1074> PMID: 33237311

16. Sherman B.T., Hao M., Qiu J., Jiao X., Baseler M.W., Lane H.C., et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*. 23 March 2022.
17. Xie J. Formulating rules of Senile Dementia based on statistical analysis. Nanjing University Of Chinese Medicine.2009.
18. Wang YQ. Effect of Gubenjiannao methods and its formula on and neurotransmitters cytokines expression in brain tissue of senile insomnia rats. Hubei University for Nationalities.2014.
19. Gong J, Qian JY, Wang YX, Qian K, Ming Y. Analysis of medication rules of traditional Chinese medicine in treating Alzheimer's disease. *Journal of Qiqihar Medical University*.2020, 41(16):2040–2042.
20. Xia S, Li QR, Li SM. Network pharmacology-based study on material basis and mechanism of *Alpinia oxyphylla* treatment of Alzheimer's disease. *Chinese Journal of Ethnopharmacology*.2020, 29(22):28–34.
21. Ru J, Li P, Wang J, Zhou W, Li B, Huang C,et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*. 2014 Apr 16; 6:13. <https://doi.org/10.1186/1758-2946-6-13> PMID: 24735618
22. Fang S, Dong L, Liu L, Guo J, Zhao L, Zhang J, et al. HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. *Nucleic Acids Res*. 2021 Jan 8; 49(D1):D1197–D1206. <https://doi.org/10.1093/nar/gkaa1063> PMID: 33264402
23. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021 Jan 8; 49(D1):D480–D489. <https://doi.org/10.1093/nar/gkaa1100> PMID: 33237286
24. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Modeling*, 1988.
25. Kim S, Cheng T, He S, Thiessen PA, Li Q, Gindulyte A, et al. PubChem Protein, Gene, Pathway, and Taxonomy Data Collections: Bridging Biology and Chemistry through Target-Centric Views of PubChem Data. *J Mol Biol*. 2022 Jun 15; 434(11):167514. <https://doi.org/10.1016/j.jmb.2022.167514> PMID: 35227770
26. Antoine Daina, Olivier Michielin, Vincent Zoete. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic acids research*,2019, 47(W1). <https://doi.org/10.1093/nar/gkz382> PMID: 31106366
27. Steinbeck C, Hoppe C, Kuhn S, Floris M, Guha R, Willighagen EL. Recent developments of the chemistry development kit (CDK)- an open-source java library for chemo- and bioinformatics. *Curr Pharm Des* 2006; 12(17): 2111–20. <https://doi.org/10.2174/13816120677585274> PMID: 16796559
28. Tanimoto TT. Elementary mathematical theory of classification and prediction. International Business Machines Corporation, 1958.
29. Chen X, Yan GY. Novel human lncRNA-disease association inference based on lncRNA expression profiles. *Bioinformatics*. 2013; 29(20):2617–24. <https://doi.org/10.1093/bioinformatics/btt426> PMID: 24002109
30. Wen YP, Han GS, Anh VV. Laplacian normalization and bi-random walks on heterogeneous networks for predicting lncRNA-disease associations. *Bmc Systems Biology*. 2018. <https://doi.org/10.1186/s12918-018-0660-0> PMID: 30598088
31. Smith T.F. and Waterman M.S. (1981). Identification of common molecular subsequences. *Journal of molecular biology*. 147(1): p. 195–197. [https://doi.org/10.1016/0022-2836\(81\)90087-5](https://doi.org/10.1016/0022-2836(81)90087-5) PMID: 7265238
32. Yan CK, Wang WX, Zhang G, Wang JL, Patel A. BiRWDDA: A Novel Drug Repositioning Method Based on Multisimilarity Fusion. *J Comput Biol*. 2019 Nov; 26(11):1230–1242. <https://doi.org/10.1089/cmb.2019.0063> PMID: 31140857
33. Zhao ZQ, Han GS, Yu ZG, Li JY. Laplacian normalization and random walk on heterogeneous networks for disease-gene prioritization. *Computational Biology and Chemistry*. 2015; 57:21–8. <https://doi.org/10.1016/j.compbiolchem.2015.02.008> PMID: 25736609
34. Cao SJ, Yu L, Mao JY, Wang Q, Ruan JS. Uncovering the Molecular Mechanism of Actions between Pharmaceuticals and Proteins on the AD Network. *Plos One*. 2015; 10(12)10. <https://doi.org/10.1371/journal.pone.0144387> PMID: 26650760
35. Wu Y, Zhang FL. SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping. *Nucleic acids research*,2019, 47(D1). <https://doi.org/10.1093/nar/gky1021> PMID: 30380087
36. Chen J, Qi Y, Liu CF, Lu JM, Shi J, Shi Y. MicroRNA expression data analysis to identify key miRNAs associated with Alzheimer's disease. *Journal of Gene Medicine*. 2018; 20(6)10. <https://doi.org/10.1002/jgm.3014> PMID: 29543360

37. Oliveira J, Costa M, de Almeida MSC, da Cruz E Silva OAB, Henriques AG. Protein Phosphorylation is a Key Mechanism in Alzheimer's Disease. *J Alzheimers Dis*. 2017; 58(4):953–978. <https://doi.org/10.3233/JAD-170176> PMID: 28527217
38. Oliveira JM, Henriques AG, Martins F, Rebelo S, Silva O. Amyloid-beta Modulates Both A beta PP and Tau Phosphorylation. *Journal of Alzheimers Disease*. 2015; 45(2):495–507. <https://doi.org/10.3233/jad-142664> PMID: 25589714
39. Drewes G, Ebnet A, Preuss U, Mandelkow EM, Mandelkow E. MARK, a novel family of protein kinases that phosphorylate microtubule-associated proteins and trigger microtubule disruption. *Cell*. 1997; 89: 297–308. [https://doi.org/10.1016/s0092-8674\(00\)80208-1](https://doi.org/10.1016/s0092-8674(00)80208-1) PMID: 9108484
40. Baum L, Hansen L, Masliah E, Saitoh T. Glycogen synthase kinase 3 alteration in Alzheimer disease is related to neurofibrillary tangle formation. *Mol. Chem. Neuropathol*. 1996; 29: 253–261. <https://doi.org/10.1007/BF02815006> PMID: 8971700
41. Patrick GN, Zukerberg L, Nikolic M, de la Monte S, Dikkes P, Tsai LH. Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature*. 1999 Dec 9; 402(6762):615–22. <https://doi.org/10.1038/45159> PMID: 10604467
42. Monaco EA 3rd, Vallano ML. Role of protein kinases in neurodegenerative disease: cyclin-dependent kinases in Alzheimer's disease. *Front Biosci*. 2005 Jan 1; 10:143–59. <https://doi.org/10.2741/1516> PMID: 15574357
43. Zotova E, Bharambe V, Cheaveau M, Morgan W, Holmes C, Harris S, et al. Inflammatory components in human Alzheimer's disease and after active amyloid-beta(42) immunization. *Brain*. 2013; 136:2677–96. <https://doi.org/10.1093/brain/awt210> PMID: 23943781
44. Heneka MT, Carson MJ, El Khoury J, Landreth GE, Brosseron F, Feinstein DL, et al. Neuroinflammation in Alzheimer's disease. *Lancet Neurol*. 2015; 14:388–405. [https://doi.org/10.1016/S1474-4422\(15\)70016-5](https://doi.org/10.1016/S1474-4422(15)70016-5) PMID: 25792098
45. Dansokho C, Heneka MT. Neuroinflammatory responses in Alzheimer's disease. *Journal of Neural Transmission*. 2018; 125(5):771–9. <https://doi.org/10.1007/s00702-017-1831-7> PMID: 29273951
46. Guo JP, Cheng J, North BJ, Wei WY. Functional analyses of major cancer-related signaling pathways in Alzheimer's disease etiology. *Biochimica Et Biophysica Acta-Reviews on Cancer*. 2017; 1868(2):341–58. <https://doi.org/10.1016/j.bbcan.2017.07.001> PMID: 28694093
47. Wu Q, Peng Z, Zhang Y, Yang J. COACH-D: improved protein-ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Res*. 2018 Jul 2; 46(W1):W438–W442. <https://doi.org/10.1093/nar/gky439> PMID: 29846643
48. Álvarez-Berbel I, Espargaró A, Viayna A, Caballero AB, Busquets MA, Gámez P, et al. Three to Tango: Inhibitory Effect of Quercetin and Apigenin on Acetylcholinesterase, Amyloid- $\beta$  Aggregation and Acetylcholinesterase-Amyloid Interaction. *Pharmaceutics*. 2022 Oct 30; 14(11):2342. <https://doi.org/10.3390/pharmaceutics14112342> PMID: 36365159
49. Beg T, Jyoti S, Naz F, Rahul, Ali F, Ali SK, et al. Protective Effect of Kaempferol on the Transgenic *Drosophila* Model of Alzheimer's Disease. *Cns & Neurological Disorders-Drug Targets*. 2018; 17(6):421–9. <https://doi.org/10.2174/1871527317666180508123050> PMID: 29745345
50. Han J, Ji Y, Youn K, Lim G, Lee J, Kim DH, et al. Baicalein as a Potential Inhibitor against BACE1 and AChE: Mechanistic Comprehension through In Vitro and Computational Approaches. *Nutrients*. 2019 Nov 7; 11(11):2694. <https://doi.org/10.3390/nu11112694> PMID: 31703329
51. Youn K, Yun EY, Lee J, Kim JY, Hwang JS, Jeong WS, et al. Oleic acid and linoleic acid from *Tenebrio molitor* larvae inhibit BACE1 activity in vitro: molecular docking studies. *J Med Food*. 2014 Feb; 17(2):284–9. <https://doi.org/10.1089/jmf.2013.2968> PMID: 24548007