Biomedicine & Pharmacotherapy 78 (2016) 272–279

Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Original article

Drug discovery of neurodegenerative disease through network pharmacology approach in herbs

Zhipeng Ke[a], Xinzhuang Zhang[a], Zeyu Cao[a], Yue Ding[a], Na Li[a], Liang Cao[a], Tuanjie Wang[a], Chenfeng Zhang[a], Gang Ding[a], Zhenzhong Wang[a], Xiaojie Xu[b,*], Wei Xiao[a,**]

[a] National Key Laboratory of Pharmaceutical New Technology for Chinese Medicine, Kanion Pharmaceutical Corporation, Lianyungang, China
[b] Beijing National Laboratory for Molecular Sciences (BNLMS), State Key Laboratory of Rare Earth Materials Chemistry and Applications, College of Chemistry and Molecular Engineering, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 26 August 2015
Received in revised form 8 January 2016
Accepted 13 January 2016

Keywords:
Neurodegenerative diseases
Network pharmacology
Drug discovery
Virtual screening

ABSTRACT

Neurodegenerative diseases, referring to as the progressive loss of structure and function of neurons, constitute one of the major challenges of modern medicine. Traditional Chinese herbs have been used as a major preventive and therapeutic strategy against disease for thousands years. The numerous species of medicinal herbs and Traditional Chinese Medicine (TCM) compound formulas in nervous system disease therapy make it a large chemical resource library for drug discovery. In this work, we collected 7362 kinds of herbs and 58,147 Traditional Chinese medicinal compounds (Tcmcs). The predicted active compounds in herbs have good oral bioavailability and central nervous system (CNS) permeability. The molecular docking and network analysis were employed to analyze the effects of herbs on neurodegenerative diseases. In order to evaluate the predicted efficacy of herbs, automated text mining was utilized to exhaustively search in PubMed by some related keywords. After that, receiver operator characteristic (ROC) curves was used to estimate the accuracy of predictions. Our study suggested that most herbs were distributed in family of Asteraceae, Fabaceae, Lamiaceae and Apocynaceae. The predictive model yielded good sensitivity and specificity with the AUC values above 0.800. At last, 504 kinds of herbs were obtained by using the optimal cutoff values in ROC curves. These 504 herbs would be the most potential herb resources for neurodegenerative diseases treatment. This study would give us an opportunity to use these herbs as a chemical resource library for drug discovery of anti-neurodegenerative disease.

1. Introduction

Neurodegenerative disease (ND) is an umbrella term for illnesses caused by the progressive neurons loss and associated functions. It is characterized by a more or less severe loss of certain cognitive functions including learning and memory [1,2]. Of these, Alzheimer's disease (AD), Parkinson's disease (PD) and Huntington's disease (HD) are the most common, which affect millions of people worldwide [3]. The pathogenesis of ND is very complicated, involved in misfolded proteins, protofibril formation, ubiquitin–proteasome dysfunction, oxidative and nitrosative stress, mitochondrial injury, *etc.* [4]. Drugs are rarely approved due to the clinical failure of lead compounds [5]. Up to now, there are only a few drugs for ND, such as AchEIs, NMDA receptor antagonist for the cognitive manifestations [6] and dopamine agonists [7]. Nonetheless, the approved drugs were limited to treat ND in practice, because of the presence of drug adverse [8–11] and the complicated mechanism of ND [12–15]. Accordingly, conventional drug-discovery approaches basing on 'one gene, one drug, one disease' philosophy may not offer the best pathway towards the therapeutics for ND [16].

Recent tendency in drug discovery is the rational design of multi-target therapeutics [16,17]. Multi-target therapeutics with higher efficacy and lower toxicity, can simultaneously adjust multiple links of the complex disease network [18,19]. By virtue of their component diversity, higher multi-target activity and lower toxicity, herbs become an important source for developing multi-target drugs. Recently, herbal medicines have received considerable attention as potential treatments for ND [20]. Researches showed that some common herbs have been utilized for treating ND, such as *Cynodon dactylon, Ginkgo biloba, Camellia sinensis, Radix polygalae* and *Acorus tatarinowii* [21,22]. The multiple components

* Corresponding author.
** Corresponding author.
E-mail addresses: xiaojxu@pku.edu.cn (X. Xu), xw_kanion@163.com (W. Xiao).

http://dx.doi.org/10.1016/j.biopha.2016.01.021
0753-3322/

of herbs could hit multiple targets and exert effects of mutual reinforcement [23,24]. In clinical practice, herbal medicines have been proven to be effective in improving the complex conditions, caused by complicated pathogenesis. This accumulated experience showed herbal medicines were promising for ND treatment. In this work, network pharmacology method was utilized to study the effective compounds of anti-neurodegenerative diseases herbs in the context of interactions between compounds and disease-related targets. Network pharmacology is a novel tool for drug discovery, providing insights into a comprehensive cellular and organism behavior at systems level [25]. This research could provide an opportunity for the treatment and prevention of ND in the near future.

2. Materials and methods

2.1. Data sets

Adatabase, Traditional Chinese Medicine Network Pharmacology Intelligent Information Platform (TCMN), was constructed, which contains 7362 kinds of herbs and 58147 Tcmcs, Herbs and TCMs were collected from Zhonghua Bencao [26] and were checked by two taxonomy databases: Global Biodiversity Information Facility (http://data.gbif.org) and The Plant List (http://www.theplantlist.org/). Taxonomies of herbs were also downloaded from the two online databases. Structures of Tcmcs were obtained from the Universal Natural Products Database (UNPD) [27], Chinese Natural Product Database (CNPD) [28] and Reaxys by the scientific names of herbs. Targets of AD, PD, HD and other NDs were collected from Therapeutic Target Database (TTD, http://bidd.nus.edu.sg/group/ttd/TTD.asp) and DrugBank (http://www.drugbank.ca/). In this work, UniProtKB [29] of the targets were obtained using the International Classification of Disease ICD-10-CM codes (HD:G10, PD:G20, AD:G30 and other dementia:F1-F3) from TTD and diseases name (dementia, Alzheimer, Parkinson and Huntington) form DrugBank. The crystal or NMR structures of targets were downloaded from RCSB Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) by UniProtKB, and the PDB must meet the following criteria: (1) Organism: Homo sapiens. (2) Resolution$\leq$2.5 Å. (3) The crystal structure must contain ligand. The targets were then utilized to predict the anti-neurodegenerative diseases effects of Tcmcs.

2.2. Molecular docking and phylogenetic tree construction

Virtual screening has been successful in prioritizing large chemical libraries to identify experimentally active compounds. Methodologies used in virtual screening such as molecular docking and scoring have advanced to the point where they can rapidly and accurately identify lead compounds in addition to predicting native binding conformations [30]. In this study, virtual screening methods were employed to predict the efficacy of herbs and Tcmcs. Targets collected from TTD and DrugBank were used in virtual screening. Discovery Studio was utilized to prepare protein targets by deletion of the hetero atoms and addition of hydrogen atoms. After that, the autodock 4.01 [31] in DOVIS 2.0 [32] was employed to perform the molecular docking. For each docking, the binding site was set as a40 $\times$ 40 $\times$ 40 Å cube centered on the occupied space of the original ligand and the spacing of energy grid points was defined as 0.375 Å. Default settings were used for all the other parameters. The compound which docking score was higher than the original ligand was believed to have an effect on the target. In order to select all the compounds with good activity, a scoring threshold was used. In this study, the compounds which docking scores above 5.00 were chosen for further predication of actions about herbs and targets. In order to study the distribution of herbs at the family level, we used the family name of species to construct a phylogenetic tree by an iTOL tool [33]. The tree file was autogenerated in iTOL by the taxonomy IDs which were got from Taxonomy Database (http://www.ncbi.nlm.nih.gov/taxonomy) by using the scientific name of herb.

2.3. Physicochemical properties analysis

The molecular structure is associated with the physicochemical properties. Thus, molecular descriptors was used to analyze the pharmacokinetics behavior of a compound in the early stages of drug discovery [34]. In this study, molecule descriptors were calculated by Discovery Studio (version 2.5). Six physicochemical parameters, that is AlogP, Molecular Weight (MW), polar surface area (PSA), the number of rotatable bonds (RBs) and the number of H-bond donors (HBDs) and acceptors (HBAs), were chosen to predict the oral bioavailability of the predicted active compounds.

2.4. Prediction of herbs and model validation

We used a docking score weighted model to predict the coefficient between herbs (PCH) and targets (PCT) which were calculated as follows:

$$\mathrm{PCH} = \sum_{i=1}^{n} \mathrm{Score}_i \quad \text{(a)}$$

$$\mathrm{PCT} = \sum_{i=1}^{m} \mathrm{Score}_i \quad \text{(b)}$$

where Score_i was the docking score of compound i to the target, n was the number of effective compounds contained in the herb, and m was the number of targets. In order to validate the above predictions, the positive samples were constructed by known activities of herbs, which were gathered by literature mining in PubMed (http://www.ncbi.nlm.nih.gov/pubmed). A strategy named Arrowsmith was utilized to identify biologically meaningful links between any two sets of Parkinson's disease, Alzheimer disease, Huntington disease, Dementia and Neurodegenerative disease in PubMed, and six keywords were identified at last. In detail; we searched literatures in the abstract and title from PubMed by using the terms: "dementia"; "neural stem cell"; "Alzheimer disease"; "Parkinson's disease"; "Huntington disease" and "Neurodegenerative disease "combined with the scientific name of herbs. The active herbs; which were reported by two or more literatures; were defined as positive samples. In this study; receiver operator characteristic (ROC) curves which were performed by IBM SPSS Statistics (IBM Corp., NY, Armonk) [35] were used to estimate the accuracy of predictions. Sensitivity; specificity; the areas under the ROC curve (AUC) and optimal cutoff value values were calculated from ROC curves.

2.5. Target-Herb-TCM network construction

NDs are associated with a multifactorial etiology. Despite significant advances in the study of the molecular mechanisms altered in the development and progression of NDs, effective therapeutic agents are still rare [36–38]. TCM has been practiced for thousands of years and been the primary means for maintaining health as well as preventing and treating human diseases [39]. The analysis of the target, herb and TCM may provide a practical therapeutic strategy against the disease. In this study, the most promising predictions were utilized for Target-Herb-TCM (THT) network construction. The THT network was constructed by Cytoscape 3.2 [40] and analyzed by network analysis plugin [41].