



Application of Ferulic Acid for Alzheimer's Disease: Combination of Text Mining and Experimental Validation

Guilin Meng^{1,2†}, Xiulin Meng^{3†}, Xiaoye Ma¹, Gengping Zhang⁴, Xiaolin Hu⁵, Aiping Jin¹, Yanxin Zhao^{1*} and Xueyuan Liu^{1*}

¹Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai, China, ²School of Computer Science and Informatics, Indiana University, Bloomington, IN, United States, ³Houma People's Hospital, Linfen, China, ⁴Library of Tongji University, Shanghai, China, ⁵School of Life Sciences, Tsinghua University, Beijing, China

OPEN ACCESS

Edited by:

Ghulam Md Ashraf,
King Abdulaziz University,
Saudi Arabia

Reviewed by:

Tarique Khan,
Buck Institute for Research on Aging,
United States
Luigia Trabace,
University of Foggia, Italy

*Correspondence:

Yanxin Zhao
287594350@qq.com
Xueyuan Liu
1510922@tongji.edu.cn

[†]Co-first authors.

Received: 02 February 2018

Accepted: 11 May 2018

Published: 29 May 2018

Citation:

Meng G, Meng X, Ma X, Zhang G,
Hu X, Jin A, Zhao Y and Liu X
(2018) Application of Ferulic Acid for
Alzheimer's Disease: Combination of
Text Mining and
Experimental Validation.
Front. Neuroinform. 12:31.
doi: 10.3389/fninf.2018.00031

Alzheimer's disease (AD) is an increasing concern in human health. Despite significant research, highly effective drugs to treat AD are lacking. The present study describes the text mining process to identify drug candidates from a traditional Chinese medicine (TCM) database, along with associated protein target mechanisms. We carried out text mining to identify literatures that referenced both AD and TCM and focused on identifying compounds and protein targets of interest. After targeting one potential TCM candidate, corresponding protein-protein interaction (PPI) networks were assembled in STRING to decipher the most possible mechanism of action. This was followed by validation using Western blot and co-immunoprecipitation in an AD cell model. The text mining strategy using a vast amount of AD-related literature and the TCM database identified curcumin, whose major component was ferulic acid (FA). This was used as a key candidate compound for further study. Using the top calculated interaction score in STRING, BACE1 and MMP2 were implicated in the activity of FA in AD. Exposure of SHSY5Y-APP cells to FA resulted in the decrease in expression levels of BACE-1 and APP, while the expression of MMP-2 and MMP-9 increased in a dose-dependent manner. This suggests that FA induced BACE1 and MMP2 pathways maybe novel potential mechanisms involved in AD. The text mining of literature and TCM database related to AD suggested FA as a promising TCM ingredient for the treatment of AD. Potential mechanisms interconnected and integrated with A β aggregation inhibition and extracellular matrix remodeling underlying the activity of FA were identified using *in vitro* studies.

Keywords: Alzheimer disease, BACE1, curcumin, ferulic acid, MMP2, STRING, text mining

INTRODUCTION

Alzheimer's disease (AD) is a chronic neurodegenerative disease that usually progresses from short memory loss to dementia, and accounts for 50%–70% of dementia cases (Burns and Iliffe, 2009). According to the World Alzheimer Report (Prince, 2015), 46.8 million people worldwide are living with dementia, and this number is estimated to reach 131.5 million by 2050, which will result in an

increasing burden on society and families. In addition, the cost of long-term care, home services, and non-professional caregivers is greater than the cost of direct medical care (Bullock, 2004; Winblad et al., 2016; Yokoyama et al., 2016).

Despite enormous financial and research investments, appropriate interventions to prevent the progress of AD are lacking (Iqbal and Grundke-Iqbal, 2011; Selkoe, 2013). Based on the failure of a number of novel AD drugs, investigators are increasingly convinced that AD is not a single but rather a multifactorial disease (Iqbal et al., 2013), and hence, drugs that target one node on the classical pathway have little effect on the AD disease network. Since AD is a multifactorial disease, drugs that modulate systemic or multiple targets are of interest.

Traditional Chinese medicine (TCM) compositions usually exert systemic impact and can be a source of drug repositioning efforts (Wang et al., 2011). TCM treatments are natural herbs discovered by the ancient Chinese and evolved through at least 3000 years of clinical practice. TCM is gaining increasing attention with the emergence of integrative and personalized medicine, characterized by pattern differentiation on individual variance and treatments based on natural herbal synergism (Wang and Wei, 2009). With the growing popularity and promising approach of TCM applicability, the ever-increasing demand for understanding the pharmacological mechanisms and potential drug efficacy are the major issues that need to be addressed.

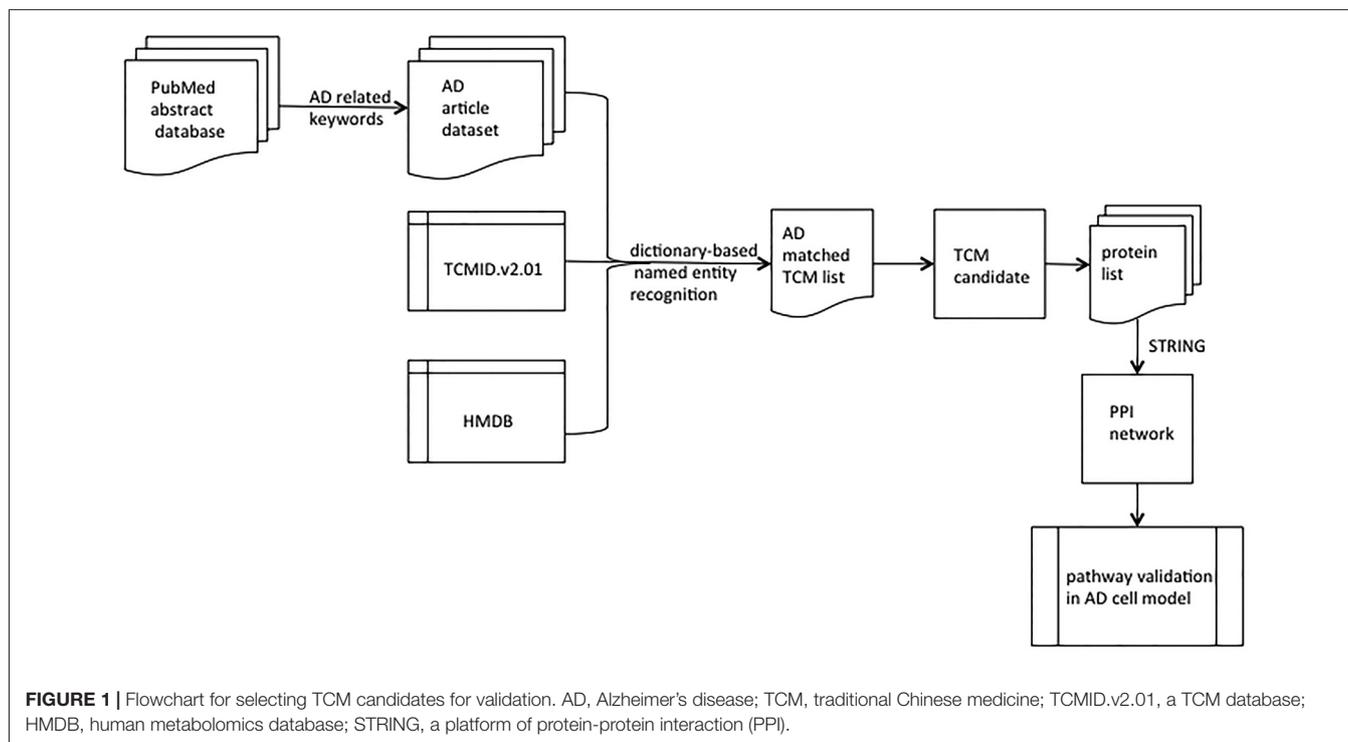
In this study, we sought to shed light on TCM for AD. What typical TCM treatments could be effective for AD, and what are the underlying target-based mechanisms? How can we

integrate systemic and target-based understandings of the disease and treatments (Cho et al., 2006)? With the overwhelming amount of biomedical knowledge recorded in texts, text mining is essential for identifying, extracting, managing, integrating and exploiting this information to discover new, hidden, or unsuspected information. Text mining is a computer-based discovery of new, previously unknown information, which automatically extracts information from different written resources (Ding et al., 2013), drawing on information retrieval, statistics, and computational linguistics. It has considerable potential for drug target discovery and re-labeling of existing drugs. Some typical proven drug repositioning cases are available for text mining, such as the beneficial effect of estrogen on human memory discovered by Smalheiser and Swanson (1996), thalidomide for treating acute pancreatitis extracted by Weeber et al. (2003), and the association of migraine with AMPA receptors identified using Litlinker (Yetisgen-Yildiz and Pratt, 2006).

Herein, we report an approach for finding an appropriate TCM for AD through the utilization of text-mining from literature database, exploring the underlying therapeutic mechanisms followed by searching for protein-protein interactions (PPI) using the STRING platform, and finally using the SHSY5Y-APP AD cell line model for validation.

MATERIALS AND METHODS

Our first aim was to select a TCM candidate from the extensive literature collection. The study workflow is shown in **Figure 1**.



Data Collection and Extraction to Find a TCM Candidate for AD

First, we assembled an AD literature dataset by retrieving articles from PubMed using AD-related keywords: “Alzheimer or Mild cognitive impairment or Dementia or Significant memory concern or Subjective memory complaint,” a resource for extracting and defining TCM candidates. The TCM database, TCMID.v2.01 (Chen et al., 2006), was then utilized, which included names, stitch_id, PubChem_id, synonyms, formula, SMILES strings, and the source of the involved chemicals. The TCM terminologies, mentioned in the abstracts of the retrieved articles were then extracted using dictionary-based named entity recognition (i.e., simple word matching) provided by LingPipe¹. Using this method, the selection of the TCM for AD was retrieved in a relatively short duration.

Furthermore, we matched the PubChem ID of TCM and the origin ontology in the human metabolome database (HMDB; Southan et al., 2013) to determine the optimal TCM candidates. After focusing on a possible lead TCM candidate, we retrieved articles in the AD dataset, and extracted protein names that co-occurred with the candidate TCM using the dictionary-based entity recognition from NCBI protein list². This gave us with a list of possible proteins affected by the TCM candidate.

Inferring Possible PPI (Li et al., 2017) Networks of the TCM Candidate Using STRING

STRING presents a specific and productive functional relationship between two proteins into a combined interaction confidence score, which is derived from the co-expression score, experimentally determined interaction score, and the automated text mining score. In this system, the automated text mining score is higher than or approximately equal to the experimentally determined interaction score since it is integrated from these scores. However, if the text mining score was lower than the experimentally determined interaction, there were two possibilities: (1) the experimentally determined interaction score was a false positive; (2) only a few studies are available related to these two proteins; however, experimental validation could be conducted.

We deposited the protein list mentioned above in the multiple protein column in the search webpage of STRING³ and acquired the PPI network after deleting results with co-expression scores >0 (already validated PPI), or experimentally determined interaction score \times automated text mining score = 0 (little relevance). In the next step, we rearranged the PPI network according to text mining scores and obtained the top proteins in the network, which were most likely related to candidate mechanisms in AD.

¹<http://alias-i.com/lingpipe/index.html>

²<https://www.ncbi.nlm.nih.gov/protein/>

³<https://string-db.org/>

Validation of Protein Expression and PPI in an AD Cell Line Model

In order to generate strong evidence not only in the data level, we validated the possible mechanisms mined from STRING using AD cell lines.

Cell Culture and Treatment With TCM Candidate for AD

SHSY5Y-APP cells, a classic cell line for AD research, were a kind gift from Shanghai Jiao Tong University. The cells were cultured in MEM supplemented with 10% heat-inactivated fetal bovine serum (FBS), 100 units/mL penicillin, and 100 μ g/mL streptomycin (Invitrogen, Carlsbad, CA, USA) at 95% humidity, 37°C, and 5% CO₂ in an incubator. The cells were passaged by trypsinization every 2–3 days. The SHSY5Y-APP cells were treated with different doses of the TCM candidate for 24 h. In the existing researches, the effective ferulic acid (FA) concentrations vary from 10 nM to 1 mM without toxic reactions in a variety of cell lines, in accordance with the point that FA is highly safe for daily and long-term consumption (Thakkar et al., 2015; Sompong et al., 2017; Zhang et al., 2017). In line with a previous study using the same cell line (Cui et al., 2013), micromolar (μ M) was chosen as the unit for FA concentration and upgraded in steps of 0 μ M, 15 μ M, 30 μ M and 60 μ M.

Co-immunoprecipitation Assays (co-IP)

Cell lysates were centrifuged (10,000 \times g) at 4°C for 15 min. Proteins were then immunoprecipitated with the relevant antibodies to determine interactions. The precleared Protein A/G Plus-Agarose beads (Merck KGaA, Darmstadt, Germany) were incubated with the immunocomplexes for 2 h and washed four times with phosphate-buffered saline. The immunoprecipitates were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; Merck KGaA, Darmstadt, Germany), followed by transfer to polyvinylidene difluoride (PVDF) membrane (Amresco, OH, USA). The antibody-antigen complexes were visualized using the UPV software according to the manufacturer's instructions. The immunoreactive bands were quantified to confirm the appropriate levels of proteins.

Western Blot

Preparation of Protein Samples

After TCM candidate exposure for 24 h, SHSY5Y-APP cells were washed with pre-cooled 4°C PBS, and then the wash solution was discarded. The above procedure was repeated twice. PMSF was added to lyse the cells on ice with frequent shaking for 30 min. After lysis, the cells were scraped with a clean scraper, and then the cell debris and lysate were transferred and centrifuged at 12,000 rpm for 5 min at 4°C. The supernatant after centrifugation was stored at –20°C.

Determination of Protein Concentration

The standard BCA assay procedure was done as previously described (Huang et al., 2010). After blocking, the membranes were probed with the following primary antibodies (Cell

Signaling, Beverly, MA, USA) using different dilutions: rabbit anti-MMP2 (92 kDa, Abcam ab92539, 1:2000), rabbit anti-MMP9 (92 kDa, Abcam ab38898, 1:2000), rabbit anti-BACE1 (68 kDa, Abcam ab183612, 1:1000), rabbit anti-APP (87 kDa, Abcam ab15272, 1:600), and mouse anti-beta-actin (42 kDa, Boster, BM0627, 1:200). All experiments were performed at least three times.

Electrophoresis

We prepared the 12% separation gel, 10% separation gel and 5% concentration gel. The prepared protein sample and the maker were added to 40 μ g. After the sample was added, constant 80 V electrophoresis was performed until the bromophenol blue indicator was linear at the junction of the concentrated gel and the separation gel, and the pressure was changed to constant 120 V. This process took about 1.5 h. Next, we removed the gel and the target band according to the Marker. The PVDF membrane was soaked in methanol for several seconds and soaked in the electroporation buffer together with the filter paper. The transfer membrane conditions were as below: β -actin 200 mA 90 min, BACE1 200 mA 120 min, APP, MMP2 and MMP9 250 mA 120 min.

Immunoblotting and Analysis

The PVDF membrane was soaked in TBST containing 5% skimmed milk powder and shaken at room temperature for 2 h. We mixed the ECL reagent with the stable peroxidase solution in a ratio of 1:1, added the solution onto the PVDF membrane. X-ray film was placed in the solution, flushed, dried, scanned, and finally analyzed grayscale value with BandScan 5.0 software (NIH, USA). Statistical analysis was performed using SPSS 20.0 software (SPSS, Chicago, IL, USA). Quantitative data are presented as mean \pm standard deviation (SD) of triplicates in an independent experiment that was repeated three times. Data were compared using Student's unpaired *t*-test for direct

comparison between two-groups and the Tukey-Kramer test after a significant one-way analysis of variance (ANOVA), and *F*-test for multiple-group comparisons. *P* < 0.05 was considered as statistically significant.

RESULTS

Text Mining Using AD Literature and the TCM Database

We retrieved 195,882 articles from PubMed using AD-related keywords and assembled an AD article dataset.

After matching TCMID.v2.01 ingredients to the AD article database, we extracted a list of AD-related TCM ingredients with PubChem IDs, which was checked for origin ontology in HMDB.

We ranked the TCM ingredients by the number of mentions and focused on the top 20 frequent terms after deleting common words, such as "protein," "glucose," "amino acid," and others (Table 1).

Next, we checked the origin ontology of all the 20 components. In Table 1, a total of 12 endogenous ingredients, including Tau, Acetylcholine, Dopamine, Melatonin, Glutathione, Aspartate, Serotonin, Tyrosine, Serine, Levodopa, Estradiol and Creatine were selected. These endogenous ingredients could not only be absorbed from the environment but also produced and synthesized within the organism or system. Cholinesterase does not have an origin result in HMDB, and hence, our list was narrowed down to Glutamate, Choline, Nicotine, Scopolamine, Curcumin, Methionine and Physostigmine, of which Glutamate, Choline, Methionine and Physostigmine could be extracted from a broad list of drug and food options, in which includes Nicotine, Scopolamine and Curcumin. Literature suggested that Scopolamine induced retrograde amnesia, or an inability to recall events prior to its administration (Coletti et al., 2014), and hence, it was

TABLE 1 | TCM ingredients' AD-matched list in PubMed.

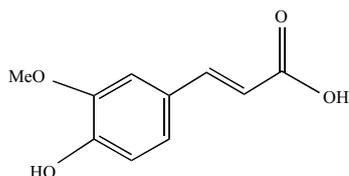
No.	Frequency	TCM ingredients	PubChem ID	Origin ontology in HMDB
1	41288	Tau	156615	Endogenous
2	5431	Glutamate	5128032	Bacteria and beans
3	3609	Cholinesterase	4460501	Not available
4	3557	Acetylcholine	5315629	Endogenous
5	3453	Dopamine	681	Endogenous
6	2824	Choline	305	Many plants and animal organs
7	1999	Nicotine	89594	Tobacco
8	1973	Melatonin	896	Endogenous
9	1898	Glutathione	124886	Drug metabolite and Endogenous
10	1722	Aspartate	5460541	Endogenous
11	1718	Serotonin	5202	Endogenous
12	1593	Tyrosine	6057	Endogenous
13	1534	Scopolamine	5184	Solanaceae
14	1432	Serine	5951	Endogenous
15	1372	Curcumin	969516	<i>Curcuma longa</i>
16	1089	Levodopa	6047	Endogenous
17	939	Estradiol	5757	Drug, Endogenous, Food
18	931	Methionine	6137	Drug metabolite and Food
19	788	Creatine	586	Endogenous
20	785	Physostigmine	5983	Drug

TCM, traditional Chinese medicine; HMDB, human metabolome database.

deleted from our list. Compared to the double-edged function of Nicotine, Curcumin has not been shown to cause any toxicity despite its daily consumption for centuries in Asian countries (Maheshwari et al., 2006). Thus, we first focused on Curcumin in the candidate list, an obvious TCM component that is extracted from *Curcuma longa*, a common plant in China.

In order to confirm its effect, we extracted all the sentences that contained “curcumin” from the AD article database. From the 107 retrieved sentences, one sentence inferred that Curcumin was suitable for treating AD; whereas, FA appeared in the same sentence with Curcumin at a high frequency. Three representative sentences are shown in **Table 2**.

Curcumin and FA share some similarities, and as a major metabolite of curcumin, FA has better bioavailability and metabolic stability than curcumin, thus rendering it as a better candidate (Badavath et al., 2016). Thus, we re-assigned our TCM target from curcumin to FA. FA also denoted as 3-(4-hydroxy-3-methoxyphenyl)-2-propenoic acid, has the following chemical structure:



After selecting FA as our TCM target, the next step was to understand the possible mechanism of FA in AD. Given the apparent complexity of the FA mechanism network, understanding its involvement in the underlying AD pathological pathways was a challenge.

We retrieved 178,725 articles using FA-related keywords “Curcumin, or FA, or Sodium Ferulate” in PubMed. From these articles, we extracted a list of proteins that was co-mentioned with FA, using the dictionary-based entity recognition. This resulted in 178 proteins that were ranked by the number of times mentioned with links to sentence sources in PubMed. After deleting a large number of false positives using the auto stop list (Fenner, 2008) of drug abbreviations, experimental test abbreviations, cell lines, synonyms of other genes, and common serum proteins, we reduced the list of proteins to 20, which are listed below:

APOB, BACE1, BCL2, CCNB1, CCND1, ERBB2, GAPDH, GSR, HMOX1, MMP2, MYB, NOS1, PCNA, PEA15, PIK3CA, PPARA, PTGS2, RAF1, TXN and VEGFA.

Potential PPI Network in STRING

Next, we entered these proteins into STRING in order to obtain direct as well as indirect protein associations (**Figure 2**).

The PPI scores were also exported into **Table 3**. The BACE1 and MMP2 combined score ranked on top among the interactions, however it had a lower automated text mining score than in the experimentally determined interaction score. We selected BACE1-MMP2 interaction as the target PPI. The edges connecting BACE1 and MMP2 (**Figure 1**) are and , which indicates that BACE1 and MMP2 may interact with each other. However, when we searched for *in-silico* evidence, the two words occurred in the full-text of some experimental articles, albeit without any direct correlation, such that the experimentally determined interaction was a false positive with a high validation possibility.

Novel Hypothesis for the FA Related Mechanism in AD

Based on the above results, we hypothesized that BACE1 and MMP2 were closely linked to the mechanism of FA. The two possibilities are as follows: these two proteins interacted directly, which could be validated by co-IP; in addition to proteolytic cleaving of the amyloid precursor protein (APP), the extracellular matrix proteins may also have a role in the AD pathological pathways, and these two pathways were always concurrent in AD.

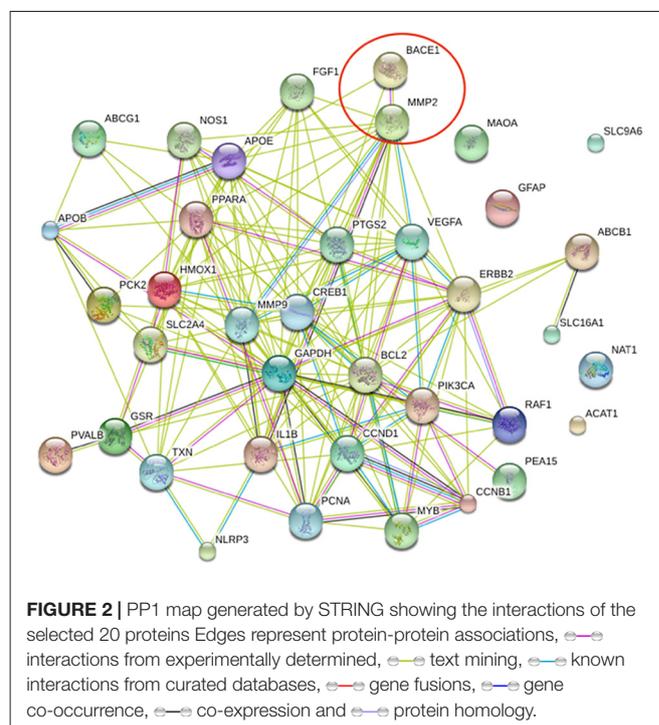


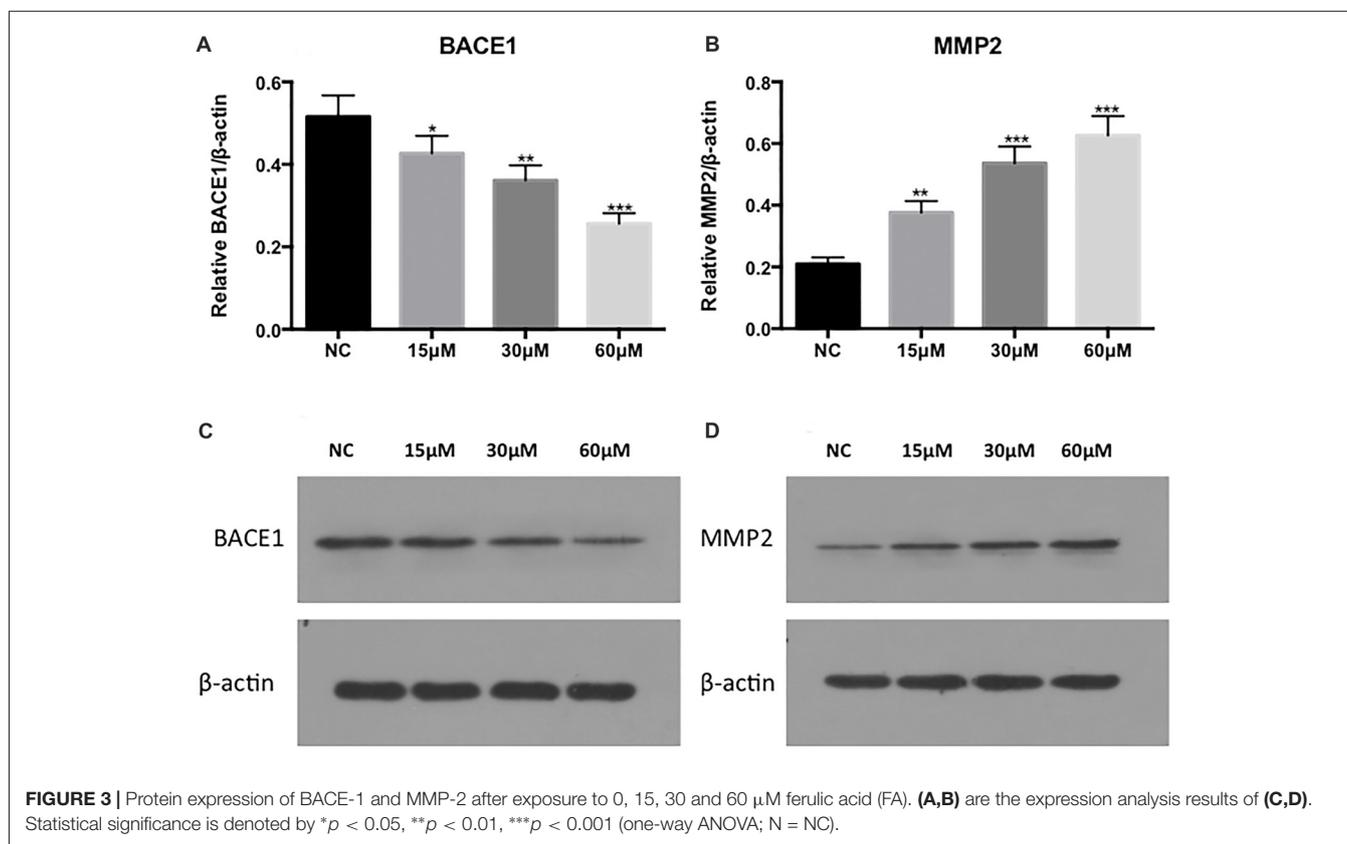
FIGURE 2 | PPI map generated by STRING showing the interactions of the selected 20 proteins. Edges represent protein-protein associations, interactions from experimentally determined, text mining, known interactions from curated databases, gene fusions, gene co-occurrence, co-expression and protein homology.

TABLE 2 | Sentences retrieved from the literature based on the entities' biomedical researches.

PMID	Sentences retrieved from literature
16387689	Because it can modulate the expression of these targets, curcumin is now being used to treat cancer, arthritis, diabetes, Crohn's disease, cardiovascular diseases, osteoporosis, Alzheimer's disease, psoriasis and other pathologies.
17127365	Food supplementation with curcumin and ferulic acid is considered a nutritional approach to reduce oxidative damage and amyloid pathology in Alzheimer's disease.
26592858	Ferulic acid has structural similarity with curcumin, which is known for its monoamine oxidase (MAO) inhibitory activity.

TABLE 3 | Ranking scores of protein interactions in STRING.

Node 1	Node 2	Co-expression	Experimentally determined interaction	Automated text mining	Combined score
BACE1	MMP2	0	0.534	0.149	0.586
PPARA	APOB	0.053	0	0.578	0.583
CCND1	MMP2	0	0	0.575	0.575
NOS1	BCL2	0.048	0	0.566	0.569
APOB	GAPDH	0	0.179	0.491	0.564
BCL2	PPARA	0	0.377	0.304	0.548
MYB	GAPDH	0.061	0	0.517	0.527
RAF1	GAPDH	0.059	0	0.503	0.512
TXN	BCL2	0	0	0.507	0.507
PCNA	PTGS2	0	0	0.507	0.507
PCNA	TXN	0.2	0.102	0.367	0.506
PCNA	MMP2	0	0	0.506	0.506
PPARA	GAPDH	0	0.042	0.501	0.501
NOS1	GAPDH	0.084	0	0.476	0.5
ERBB2	RAF1	0	0.292	0.76	0.487
BACE1	GAPDH	0.053	0	0.465	0.472
ERBB2	CCNB1	0	0	0.454	0.454
BCL2	GSR	0.048	0	0.446	0.45
NOS1	MMP2	0.055	0	0.442	0.45
TXN	VEGFA	0	0	0.442	0.442
PCNA	MYB	0.058	0	0.426	0.436
VEGFA	CCNB1	0	0	0.432	0.432
TXN	APOB	0	0	0.417	0.416
MYB	PIK3CA	0.055	0.104	0.349	0.401



The SHSY5Y-APP cell lines were passaged every 2–3 days by trypsinization, and treated with 0, 15, 30 and 60 μM FA (Yuanmu Tech, China) for 24 h. After FA

exposure for 24 h, the BACE-1 expression decreased and MMP-2 expression increased in a dose-dependent manner (Figure 3).

We next tested the expression of APP, another dominant protein in the A β aggregation pathway, which is positively correlated with BACE1. After exposure to FA, the proteolytic cleavage of APP and APP enzymolysis is decreased, thereby improving the AD process. In addition, we tested MMP9, another protein in the extracellular matrix pathway. We observed that extracellular matrix protein expression was increased after FA exposure and contributes to the pathological process of AD (Figure 4).

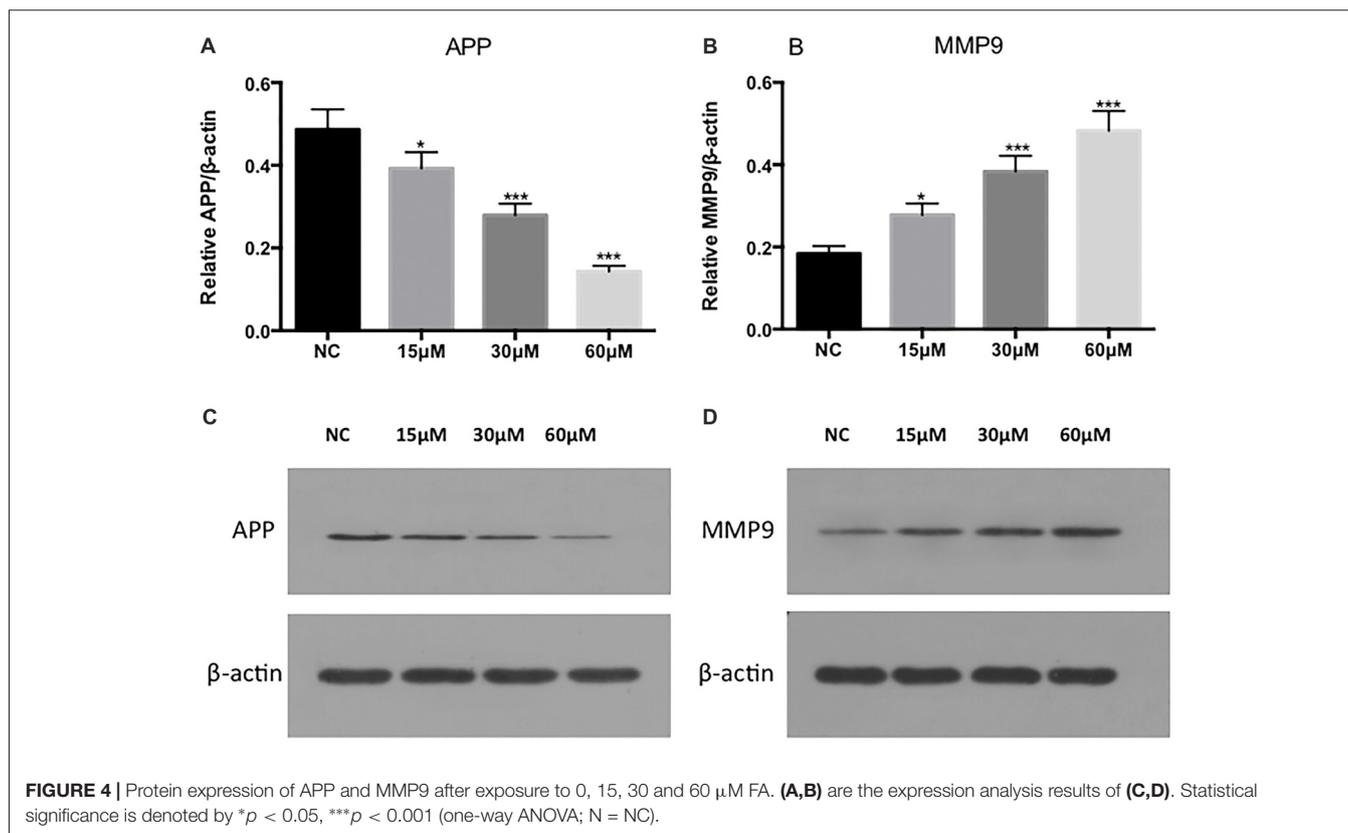
DISCUSSION

TCM Candidate Selection for AD Using Data Mining

In this study, we used text mining to select a TCM candidate for subsequent validation. FA was selected and *in vitro* validation was performed to understand the potential mechanisms involved in AD. Furthermore, the current study used the information-medicine integrated system to map TCM for AD research. In addition, text mining was coupled to the experimental validation to assess the drug selection outcomes. This offset the information gap and maximized the utilization of existing knowledge to select the optimal TCM candidate to study.

Drug discovery for AD is no longer a game of chance or just limited to the availability of new technology. Societal expectations about drug efficacy are rising; thus, early-stage

drug discovery necessitates accessible, standardized data sets to generate a complete scenario of the physiological function and disease relevance. Some pioneering studies have focused on drug repurposing, such as systematic “omics” data mining of genome-wide association studies (GWAS), HMDB, epigenomics and proteomics data (Zhang et al., 2016; Pimplikar, 2017). These studies suggested drugs that were applicable for other diseases having novel anti-AD indications. These attempts were very logical however these studies did not consider TCM as a source of information for AD research. TCM is a good source for drug discovery. The uniqueness of the TCM system is based on the philosophical logic underlying daily practices (Ho et al., 2011), which was accumulated over thousands of years of empirical studies and provides a unique view of the relationships between the human body and the universe (Gu and Chen, 2014). Therefore, a better understanding of TCM and key learning from the past with appropriate strategies for the future is essential to make a significant difference. These theories render the proposed approach useful in identifying novel relationships between diseases and drugs that have a high probability of being physiologically effective. On the other hand, existing TCM drug mining primarily focuses on the assessment of ancient classic literature, with less analysis of herbal components (May et al., 2014; Pae et al., 2016), and thus may affect knowledge dissemination. Our study is the first to combine well-known TCM database with text mining approaches. This led us to select FA as a lead candidate for experimental validation for AD.



MMP2-BACE1 Mechanisms of FA in AD

Our findings suggested that FA might be a promising multi-targeted TCM with a therapeutic potential for AD (Jung et al., 2016). We evaluated the APP and BACE1 inhibitory activities, which inhibit Ab aggregation; in addition, the matrix clearance properties of MMP-2 and MMP-9 implicated FA was actively involved in the alteration of matrix proteins and that it played a major role in *in vitro* extracellular matrix remodeling. As shown in a previous *in situ* proximity ligation assay (*in situ* PLA), which is a new technique to monitor PPI with high specificity and sensitivity, it was found that APP, MMP2 and MMP9 all interacted with TGFB1, and the interaction of MMP2 and BACE1 was also positive (Chen et al., 2014). Furthermore, from the Human Protein Reference Database (HPRD) in the STRING platform (Higashi and Miyazaki, 2003), the COOH-terminal parts of APP were found to interact with the extracellular matrix and highly selectively inhibit MMP2, in which the decapeptide region of APP was likely an active site-directed inhibitor toward MMP2.

The pathways of PPI at the molecular level include cellular transduction and biological function. Hence, the two pathways of A β aggregation inhibition and extracellular matrix remodeling were interconnected and integrated to the biological function-signaling map for AD. The results of these analyses might have potential application in exploring FA mechanism because they can be used as rational targets to inhibit the function of pathways essential to AD. In the multi-targeted AD model, APP cleavage, inhibition of A β deposition, and extracellular matrix remodeling are co-operative interactions involved in AD pathology, which could be attractive therapeutics with respect to pharmacokinetics and pharmacodynamics when compared to a specific highly specific single target molecule. These results highlight the prospective beneficial effects of FA as a therapeutic agent against AD pathology.

One limitation of this study was that no animal model was validated, and the experimental validation in the AD cell model was not sufficient to make a conclusive statement regarding the potential efficacy and benefit of FA. However, we found evidence in previous study of FA's protective effects on different animal models of intra-cerebroventricular (i.c.v.) injection of A β 1–42 in mice and APP/PS1 mutant transgenic

mice (Jung et al., 2016), which shows potent anti-oxidant and anti-inflammatory activities. Future studies should discuss the in-depth mechanisms of FA together with the physiological data to evaluate FA efficacy and involvement in an AD animal model. These include: What are the safety implications of the different doses of FA for AD? What biomarkers exist for FA metabolites? In addition, understanding the role of FA within the system, the pathways and networks of the different protein interactions are invaluable.

CONCLUSION

In summary, we demonstrate that the combination of text mining and professional medical knowledge is an effective approach for finding new mechanisms underlying the clinical therapeutics for AD. Equipped with this data, the clinical scientist can obtain information in a short period of time without searching large volumes of articles. Moreover, using *in vitro* studies for validation, the data-driven results were based on not only a hypothesis but also true novel findings of potential mechanisms interconnected and integrated by A β aggregation inhibition and extracellular matrix remodeling underlying the activity of FA. The present study strongly supported text mining of the ever-increasing volume of literature and TCM database as a drug repositioning approach for elucidating FA as a promising TCM ingredient for treating AD.

AUTHOR CONTRIBUTIONS

GM, YZ and XMeng designed the study. XMa and AJ performed experiments and prepared figures. XMa, GZ and XH analyzed the data. GM and XMeng wrote and discussed all sections of the manuscript. All authors reviewed and approved the manuscript.

FUNDING

This study was supported by the International Exchange Program for Graduate Students, Tongji University (2016020033), the National Natural Science Foundation of China (81571033, 81771131) and the Shanghai Science and Technology Committee (17411950100, 17411950101).

REFERENCES

- Badavath, V. N., Baysal, I., Uçar, G., Mondal, S. K., Sinha, B. N., and Jayaprakash, V. (2016). Monoamine oxidase inhibitory activity of ferulic acid amides: curcumin-based design and synthesis. *Arch. Pharm.* 349, 9–19. doi: 10.1002/ardp.201500317
- Bullock, R. (2004). The needs of the caregiver in the long-term treatment of Alzheimer disease. *Alzheimer Dis. Assoc. Disord.* 18, S17–S23. doi: 10.1097/01.wad.0000127493.65032.9a
- Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *BMJ* 338:b158. doi: 10.1136/bmj.b158
- Chen, T. C., Lin, K. T., Chen, C. H., Lee, S. A., Lee, P. Y., Liu, Y. W., et al. (2014). Using an *in situ* proximity ligation assay to systematically profile endogenous protein-protein interactions in a pathway network. *J. Proteome Res.* 13, 5339–5346. doi: 10.1021/pr5002737
- Chen, X., Zhou, H., Liu, Y. B., Wang, J. F., Li, H., Ung, C. Y., et al. (2006). Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation. *Br. J. Pharmacol.* 149, 1092–1103. doi: 10.1038/sj.bjp.0706945
- Cho, C. R., Labow, M., Reinhardt, M., van Oostrum, J., and Peitsch, M. C. (2006). The application of systems biology to drug discovery. *Curr. Opin. Chem. Biol.* 10, 294–302. doi: 10.1016/j.cbpa.2006.06.025
- Colettis, N. C., Snitcofsky, M., Kornisiuk, E. E., Gonzalez, E. N., Quillfeldt, J. A., and Jerusalinsky, D. A. (2014). Amnesia of inhibitory avoidance by scopolamine is overcome by previous open-field exposure. *Learn. Mem.* 21, 634–645. doi: 10.1101/lm.036210.114
- Cui, L., Zhang, Y., Cao, H., Wang, Y., Teng, T., Ma, G., et al. (2013). Ferulic acid inhibits the transition of amyloid- β 2 monomers to oligomers but accelerates the transition from oligomers to fibrils. *J. Alzheimers Dis.* 37, 19–28. doi: 10.3233/JAD-130164

- Ding, Y., Song, M., Han, J., Yu, Q., Yan, E., Lin, L., et al. (2013). Entitymetrics: measuring the impact of entities. *PLoS One* 8:e71416. doi: 10.1371/journal.pone.0071416
- Fenner, M. (2008). Duplication: stop favouring applicant with longest list. *Nature* 452:29. doi: 10.1038/452029a
- Gu, P., and Chen, H. (2014). Modern bioinformatics meets traditional Chinese medicine. *Brief. Bioinform.* 15, 984–1003. doi: 10.1093/bib/bbt063
- Higashi, S., and Miyazaki, K. (2003). Identification of a region of β -amyloid precursor protein essential for its gelatinase A inhibitory activity. *J. Biol. Chem.* 278, 14020–14028. doi: 10.1074/jbc.M212264200
- Ho, Y. S., So, K. F., and Chang, R. C. (2011). Drug discovery from Chinese medicine against neurodegeneration in Alzheimer's and vascular dementia. *Chin. Med.* 6:15. doi: 10.1186/1749-8546-6-15
- Huang, T., Long, M., and Huo, B. (2010). Competitive binding to cuprous ions of protein and BCA in the bicinchoninic acid protein assay. *Open Biomed. Eng. J.* 4, 271–278. doi: 10.2174/1874120701004010271
- Iqbal, K., Flory, M., and Soininen, H. (2013). Clinical symptoms and symptom signatures of Alzheimer's disease subgroups. *J. Alzheimers Dis.* 37, 475–481. doi: 10.3233/JAD-130899
- Iqbal, K., and Grundke-Iqbal, I. (2011). Opportunities and challenges in developing Alzheimer disease therapeutics. *Acta Neuropathol.* 122, 543–549. doi: 10.1007/s00401-011-0878-z
- Jung, J. S., Yan, J. J., Li, H. M., Sultan, M. T., Yu, J., Lee, H. S., et al. (2016). Protective effects of a dimeric derivative of ferulic acid in animal models of Alzheimer's disease. *Eur. J. Pharmacol.* 782, 30–34. doi: 10.1016/j.ejphar.2016.04.047
- Li, T., Wernersson, R., Hansen, R. B., Horn, H., Mercer, J., Slodkovic, G., et al. (2017). A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. Methods* 14, 61–64. doi: 10.1038/nmeth.4083
- Maheshwari, R. K., Singh, A. K., Gaddipati, J., and Srimal, R. C. (2006). Multiple biological activities of curcumin: a short review. *Life Sci.* 78, 2081–2087. doi: 10.1016/j.lfs.2005.12.007
- May, B. H., Zhang, A., Lu, Y., Lu, C., and Xue, C. C. (2014). The systematic assessment of traditional evidence from the premodern Chinese medical literature: a text-mining approach. *J. Altern. Complement. Med.* 20, 937–942. doi: 10.1089/acm.2013.0372
- Pae, S. B., Yun, B. C., Han, Y. K., Choi, B. T., Shin, H. K., and Baek, J. U. (2016). Cognitive-enhancing herbal formulae in Korean medicine: identification of candidates by text mining and literature review. *J. Altern. Complement. Med.* 22, 413–418. doi: 10.1089/acm.2015.0257
- Pimplikar, S. W. (2017). Multi-omics and Alzheimer's disease: a slower but surer path to an efficacious therapy? *Am. J. Physiol. Cell Physiol.* 313, C1–C2. doi: 10.1152/ajpcell.00109.2017
- Prince, M. J. (2015). World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. Available online at: <http://www.alz.co.uk/research/world-report-2015>
- Selkoe, D. J. (2013). The therapeutics of Alzheimer's disease: where we stand where we are heading. *Ann. Neurol.* 74, 328–336. doi: 10.1002/ana.24001
- Smalheiser, N. R., and Swanson, D. R. (1996). Linking estrogen to Alzheimer's disease: an informatics approach. *Neurology* 47, 809–810. doi: 10.1212/WNL.47.3.809
- Sompong, W., Cheng, H., and Adisakwattana, S. (2017). Ferulic acid prevents methylglyoxal-induced protein glycation, DNA damage, and apoptosis in pancreatic β -cells. *J. Physiol. Biochem.* 73, 121–131. doi: 10.1007/s13105-016-0531-3
- Southan, C., Sitzmann, M., and Muresan, S. (2013). Comparing the chemical structure and protein content of ChEMBL, DrugBank, human metabolome database and the therapeutic target database. *Mol. Inform.* 32, 881–897. doi: 10.1002/minf.201300103
- Thakkar, A., Chenreddy, S., Wang, J., and Prabhu, S. (2015). Ferulic acid combined with aspirin demonstrates chemopreventive potential towards pancreatic cancer when delivered using chitosan-coated solid-lipid nanoparticles. *Cell Biosci.* 5:46. doi: 10.1186/s13578-015-0041-y
- Wang, H., Ding, Y., Tang, J., Dong, X., He, B., Qiu, J., et al. (2011). Finding complex biological relationships in recent PubMed articles using Bio-LDA. *PLoS One* 6:e17243. doi: 10.1371/journal.pone.0017243
- Wang, J. F., and Wei, D. Q. (2009). Role of structural bioinformatics and traditional Chinese medicine databases in pharmacogenomics. *Pharmacogenomics* 10, 1213–1215. doi: 10.2217/pgs.09.81
- Weeber, M., Vos, R., Klein, H., De Jong-Van Den Berg, L. T., Aronson, A. R., and Molema, G. (2003). Generating hypotheses by discovering implicit associations in the literature: a case report of a search for new potential therapeutic uses for thalidomide. *J. Am. Med. Inform. Assoc.* 10, 252–259. doi: 10.1197/jamia.m1158
- Winblad, B., Amouyel, P., Andrieu, S., Ballard, C., Brayne, C., Brodaty, H., et al. (2016). Defeating Alzheimer's disease and other dementias: a priority for European science and society. *Lancet Neurol.* 15, 455–532. doi: 10.1016/S1474-4422(16)00062-4
- Yettigen-Yildiz, M., and Pratt, W. (2006). Using statistical and knowledge-based approaches for literature-based discovery. *J. Biomed. Inform.* 39, 600–611. doi: 10.1016/j.jbi.2005.11.010
- Yokoyama, J. S., Wang, Y., Schork, A. J., Thompson, W. K., Karch, C. M., Cruchaga, C., et al. (2016). Association between genetic traits for immune-mediated diseases and Alzheimer disease. *JAMA Neurol.* 73, 691–697. doi: 10.1001/jamaneurol.2016.0150
- Zhang, M., Schmitt-Ulms, G., Sato, C., Xi, Z., Zhang, Y., Zhou, Y., et al. (2016). Drug repositioning for Alzheimer's disease based on systematic 'omics' data mining. *PLoS One* 11:e0168812. doi: 10.1371/journal.pone.0168812
- Zhang, X. D., Wu, Q., and Yang, S. H. (2017). Ferulic acid promoting apoptosis in human osteosarcoma cell lines. *Pak. J. Med. Sci.* 33, 127–131. doi: 10.12669/pjms.331.12066

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Meng, Meng, Ma, Zhang, Hu, Jin, Zhao and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.