

Identification of High-Frequency Drug Combinations for Perimenopausal Syndrome and Associated Conditions via Multilevel Association Rule Mining and Network Pharmacology

Hao Wang

Wangjing Hospital of China Academy of Chinese Medical Sciences

Haoran Li

Beijing University of Chinese Medicine

Yixuan Wang

Beijing University of Chinese Medicine

Ruolan Zhang

Beijing University of Chinese Medicine

Meixi Wang

Beijing University of Chinese Medicine

Suyang Ma

Beijing University of Chinese Medicine

Xinrui Men

Beijing University of Chinese Medicine

Haoxiang Li

Beijing University of Chinese Medicine

Yuelin Huang

Beijing University of Chinese Medicine

Liqun Yu

Wangjing Hospital of China Academy of Chinese Medical Sciences

Hao Li

wb6867@163.com

Wangjing Hospital of China Academy of Chinese Medical Sciences

Article

Keywords: Perimenopausal syndrome, Traditional medicine, Association rule mining, Network pharmacology, Molecular docking, Molecular dynamics simulation

Posted Date: September 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10567069/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Identification of High-Frequency Drug Combinations for Perimenopausal Syndrome and Associated Conditions via Multilevel Association Rule Mining and Network Pharmacology

Hao Wang^{1†}, Haoran Li^{2†}, Yixuan Wang², Ruolan Zhang², Meixi Wang², Suyang Ma², Xinrui Men², Haoxiang Li², Yuelin Huang², Liquun Yu¹, Hao Li^{1*}

¹Wangjing Hospital, China Academy of Chinese Medical Sciences, Beijing, China

²Beijing University of Chinese Medicine, Beijing, China

* **Correspondence:**

Hao Li

wb6867@163.com

*The two authors share the correspondence

[†]The first two authors contributed equally to this work and share first authorship

Keywords: Perimenopausal syndrome; Traditional medicine; Association rule mining; Network pharmacology; Molecular docking; Molecular dynamics simulation

Abstract

Objective To integrate association rule mining (ARM) with network pharmacology to identify high-frequency herbal combinations for perimenopausal syndrome (PMS) and its associated anxiety, depression, and insomnia, and to explore their mechanisms. **Method** Traditional medical formulas, herbs, compounds, genes, and phenotypes were collected from multiple public databases (KTKP, TCM-ID, KampoDB, etc.). The Apriori algorithm was used for herb–phenotype ARM to identify core herbal combinations, validated by herb–phenotype proximity ($Z_h \leq -0.15$). Compound–phenotype ARM and network analysis were then conducted to identify key compounds, confirmed by literature. A PPI network was built to screen core genes, followed by molecular docking and 100 ns molecular dynamics simulations to validate binding stability. **Result** ARM identified six core herbal combinations for PMS, two for anxiety, two for insomnia, and two for depression (e.g., “*Angelica sinensis*–*Ligusticum chuanxiong*–*Rehmannia glutinosa*” for PMS). All Z_h values were below -0.15 . Six compounds (reduced glutathione, D-tryptophan, linalool, GABA, folic acid, resveratrol) and three core genes (ESR1, MYC, APP) were obtained. All compounds showed strong binding affinity (< -6.0 kcal/mol) and stable dynamics over 100 ns. **Conclusion** By integrating ARM with network pharmacology, this study successfully identified high-frequency herbal combinations and core compounds for PMS and its comorbidities, supporting the “treating different diseases with the same approach” principle and providing data-driven candidates for experimental validation and drug development.

1 Introduction

Perimenopausal syndrome (PMS) refers to a range of physiological and psychological symptoms in women resulting from fluctuations or declines in sex hormone levels during the menopausal transition^[1]. With increasing life expectancy, the menopausal transition and postmenopausal period have become the longest and most significant stages in a woman’s life.

Epidemiological data indicate that the number of perimenopausal women in China has reached 120 million, representing the largest such population among all developing countries, accounting for 10% of the country's total population and 23% of the global perimenopausal female population^[2]. The quality of life of women affected by PMS is significantly impaired, placing a substantial burden on both families and society. Patients with PMS frequently present with neuropsychiatric symptoms such as anxiety, depression, and insomnia, which often co-occur and further contribute to the complexity of the condition^[3].

Currently, the management of PMS in modern medicine primarily relies on hormone replacement therapy and anxiolytic and antidepressant medications; however, long-term use may be associated with certain adverse effects and contraindications^[4]. In contrast, traditional medicine (TM), which emphasizes holistic regulation and treatment based on pattern differentiation, offers unique advantages in managing complex, multifactorial conditions such as PMS^[5]. According to traditional Chinese medicine (TCM) theory, the fundamental pathogenesis of PMS lies in kidney essence deficiency and gradual exhaustion of Tian Gui, often involving the liver, heart, and other viscera, leading to blood deficiency with liver depression and mental restlessness^[6,7]. Therefore, identifying effective herbal combinations from ancient texts and formulas is of great significance for guiding clinical practice and elucidating the scientific basis of TM.

Therefore, this study employed a highly interpretable analytical framework that integrates association rule mining (ARM) with network pharmacology to identify potentially effective combinations of herbal medicines and compounds in traditional medicine. ARM, as a rule-based analytical approach^[8], is capable of processing large-scale data with high interpretability, offering a potential solution to the limitations of machine learning methods. To this end, we systematically collected diverse datasets covering traditional Chinese medicine formulas, herbal ingredients, compounds, target proteins, diseases, and their interrelationships. Using these data, ARM was applied to screen candidate herbal combinations or compounds with potential therapeutic activity, while network pharmacology was employed to provide supporting evidence at the molecular level for the selected candidates. Molecular docking and molecular dynamics simulations were further conducted to validate the efficacy of the predicted herbal combinations and compounds. The framework proposed in this study not only facilitates drug discovery from traditional Chinese medicine but also advances translational medicine research by elucidating the potential pharmacological mechanisms and associated biological pathways of candidate substances.

2 Materials and methods

2.1 Materials

Information on traditional medicine prescriptions, herbal materials, compounds, genes, phenotypes, and their interactions from Korea, China, and Japan was collected from the Korean Traditional Knowledge Portal (KTKP, <https://koreantk.com/>), the Traditional Chinese Medicine Information Database (TCM-ID), KampoDB (<https://wakanmoview.inm.u-toyama.ac.jp/kampo/>), and the COlleCtion of Open NatUral producTs (COCONUT, <https://coconut.naturalproducts.net/>) database.

Compound–phenotype and compound–gene associations were obtained from DrugBank (<https://www.drugbank.ca/>)^[9], the Therapeutic Target Database (TTD, <https://idrblab.org/ttd/>)^[10], the Comparative Toxicogenomics Database (CTD, <https://ctdbase.org/>)^[11], BindingDB (<https://www.bindingdb.org/rwd/bind/index.jsp>)^[12], the Search Tool for Interacting Chemicals

(STITCH, <https://stitch.embl.de>)^[13], and FooDB (<https://foodb.ca/>).

Gene–phenotype associations were collected from the CTD, while herb–compound associations were retrieved from the Collective Molecular Activities of Useful Plants (CMAUP, <https://bidd2.nus.edu.sg/CMAUP/>)^[14].

A total of 16,260 prescriptions, 17,519 phenotypes, 5,187 herbal medicines, 43,454 compounds, and 25,344 genes were compiled. In addition, a human protein–protein interaction network comprising 26,771 nodes and 856,635 edges was retrieved from the Biological General Repository for Interaction Datasets (BioGRID, v.4.4.218, <https://thebiogrid.org>)^[15].

Targets for perimenopausal syndrome (menopause), anxiety, depression, and insomnia were retrieved from the GeneCards database (<https://www.genecards.org/>)^[16] and the OMIM database (<https://www.ncbi.nlm.nih.gov/omim/>)^[17] using the keywords “perimenopausal syndrome,” “menopausal syndrome,” “climacteric,” and “menopause.” Gene sets obtained from the two databases were merged, and duplicate targets were removed to obtain the final target set for each disease.

2.2 Method overview

The proposed integrative analysis framework in this study consists of four main steps: (i) herb level analysis, to identify potentially therapeutic herbal combinations and evaluate their associations with phenotypes; (ii) compound level analysis, to screen for effective compounds and validate consistency through intersection analysis; (iii) structural network analysis, to identify key action targets based on network topological features; and (iv) molecular docking and molecular dynamics simulation, to validate the binding characteristics between candidate compounds and targets at the molecular level (Fig 1).

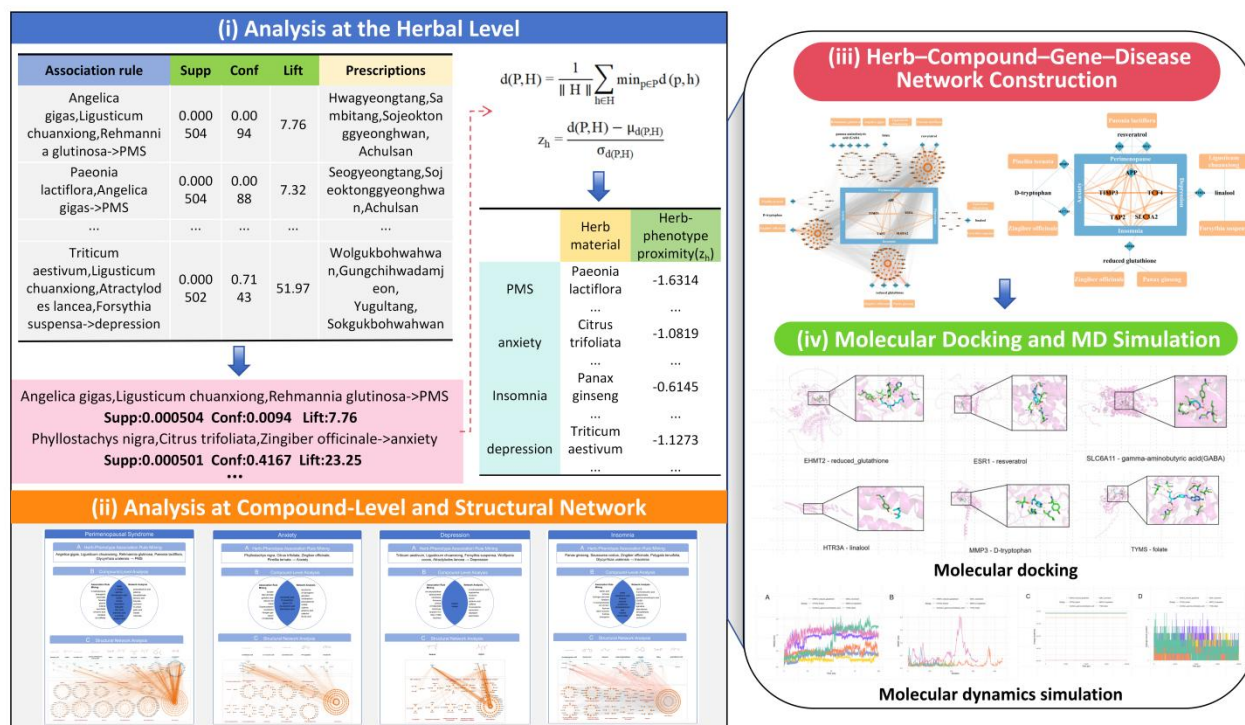


Figure 1. Overview of the integrated analysis pipeline. The schematic diagram illustrates the four main steps of

the proposed framework (see Methods for details). Representative association rules, prescriptions, herb–phenotype proximity values, compound–target docking pairs, and molecular dynamics simulation results are shown for perimenopausal syndrome (PMS), anxiety, depression, and insomnia. The multi-layer network visualizes the interactions among herbs, compounds, core targets, and phenotypes.

2.3 Analysis at the Herbal Level

2.3.1 Herbal-Phenotype Association Rule Mining

Based on the hypothesis that certain herbal combinations frequently appearing in prescriptions for specific phenotypes may possess therapeutic potential against those phenotypes, the Apriori algorithm was employed to identify optimal herbal combinations from 16,260 prescriptions involving 5,279 herbs for the treatment of PMS, anxiety, depression, and insomnia. The rules were expressed in the form “herb A, herb B, ... → phenotype” and evaluated using three metrics: support, confidence, and lift. The minimum support threshold was set to 0.0005.

2.3.2 Herbal-Phenotypic Network Analysis

To validate the reliability of the ARM results, a network proximity approach based on shortest distance was employed to calculate the relative proximity (Z_h) between the targets of the inferred herbal combinations and the phenotype-associated gene sets within the human protein – protein interaction network^[18]. Protein and gene interaction data related to PMS, anxiety, depression, and insomnia were collected from BioGRID to construct the human protein – protein interaction network. The herb – phenotype proximity Z_h , representing the relative proximity between the herbal combination and the corresponding phenotype using the antecedent of the selected association rule, was calculated as follows:

$$d(P, H) = \frac{1}{\|H\|} \sum_{h \in H} \min_{p \in P} d(p, h)$$

$$Z_h = \frac{d(P, H) - \mu_{d(P, H)}}{\sigma_{d(P, H)}}$$

where P represents the phenotype-associated gene set, H denotes the herbal target set, and $d(P, H)$ is the shortest path length between nodes in P and H . A reference distribution of expected distances (μ and σ) was obtained through 1,000 random permutation tests. A Z_h value ≤ -0.15 was defined as indicating a significant association between the herbal combination and the phenotype.

2.4 Analysis at the Compound Level

2.4.1 Compound-Phenotype Association Rule Mining

A transaction dataset comprising compounds from herbal medicines and their associated phenotype information was constructed. The Apriori algorithm was then applied to mine association rules in the form of “compound → phenotype,” with a minimum support threshold set to 0.008.

2.4.2 Compound–phenotype network analysis

Using the same method described in Section 2.2.2, the network proximity (Z_c) between the targets of compounds present in the core herbal combinations and disease-related targets was

calculated, and compounds with $Z_c \leq -0.15$ were selected. Candidate compounds were obtained by taking the intersection of the results from compound–phenotype ARM and compound–phenotype network analysis. Subsequently, drug-likeness assessment (Lipinski’s rule of five) was performed using the SwissADME platform^[19], combined with extensive literature review, to ultimately determine the core compounds for further in-depth analysis in this study.

$$d(P, C) = \frac{1}{\|C\|} \sum_{c \in C} \min_{p \in P} d(p, c)$$

$$z_c = \frac{d(P, C) - \mu_d(P, C)}{\sigma_d(P, C)}$$

2.5 Structural network analysis

To identify key biological targets, herbal compound-related and herb-related targets (genes) associated with PMS, anxiety, depression, and insomnia were retrieved from databases. KEGG enrichment analysis of disease-related target genes was performed using clusterProfiler (v4.10.1)^[20], with KEGG pathway annotations obtained from the KEGG database^[21]. Statistical significance was determined using a threshold of $\text{adj.P} < 0.05$. Cytoscape was employed to visualize the herb compound-related targets, candidate compounds, and genes associated with PMS, anxiety, depression, and insomnia.

2.6 Core gene screening and network construction

The intersecting genes between the targets of the core compounds and disease-related targets were imported into the STRING database to construct a PPI network, with a highest confidence threshold (>0.9) applied to obtain high-reliability interactions. The resulting data were imported into Cytoscape, and multiple topological algorithms, including MCC, Degree, BottleNeck, Closeness, and Radiality from the CytoHubba plugin, were employed to identify the top five genes, the intersection of which was defined as the core genes^[22,23]. Cytoscape was then used to integrate the core herbs, core compounds, core genes, and their corresponding disease associations to construct a multi-level network diagram, intuitively illustrating the multi-target synergistic action pattern of the herbal combinations.

2.7 Molecular docking and molecular dynamics simulation validation

2.7.1 Molecular docking

The three-dimensional structures of the core compounds were downloaded from the PubChem database, and the protein structures of the core genes were obtained from the AlphaFold database. Molecular docking simulations were performed using the CB-Dock2 online platform, and binding energies (kcal/mol) were calculated. A binding energy < -5.0 kcal/mol was generally considered to indicate favorable binding activity. The conformation with the lowest binding energy was selected for visualization analysis using PyMOL software.

2.7.2 Molecular dynamics simulation

To evaluate the stability and dynamic behavior of the optimal docking complexes, the six

complexes with the lowest binding energies were selected for 100 ns molecular dynamics simulations using GROMACS 2024.2 software^[24]. The AMBER99SB-ILDN force field and the TIP3P water model were employed. Parameters such as root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) were calculated to assess conformational stability, flexibility changes, and binding mode reliability of the complexes during the simulation.

3 Results

3.1 Herb-level analysis results

3.1.1 Herb–phenotype association rule mining (ARM)

To identify high-frequency herbal combinations potentially associated with PMS and its related conditions (anxiety, depression, and insomnia), the Apriori algorithm was applied to perform herb–phenotype association rule mining based on 16,260 prescriptions and 628 herbs. The results showed that the combination with the strongest association with PMS was “*Angelica sinensis*–*Ligusticum chuanxiong*–*Rehmannia glutinosa*” (Danggui–Chuanxiong–Dihuang); for anxiety, the strongest association was observed for “*Bambusa textilis*–*Citrus aurantium*–*Zingiber officinale*” (Zhuru–Zhiqiao–Shengjiang) and “*Zingiber officinale*–*Citrus trifoliata*–*Phyllostachys nigra*–*Pinellia ternata*” (Shengjiang–Zhiqiao–Zhuru–Banxia); for insomnia, the combination “*Panax ginseng*–*Aucklandia lappa*–*Zingiber officinale*–*Polygala tenuifolia*” (Renshen–Muxiang–Shengjiang–Yuanzhi) and “*Saussurea costus*–*Glycyrrhiza uralensis*–*Zingiber officinale*–*Polygala tenuifolia*” (Muxiang–Gancao–Shengjiang–YuanZhi) showed the strongest association; and for depression, the combination “*Triticum aestivum*–*Ligusticum chuanxiong*–*Forsythia suspensa*–*Poria cocos*” (Xiaomai–Chuanxiong–Lianqiao–Fuling) was the most strongly associated. The lift values of all rules were substantially greater than 1, indicating significant positive associations between these herbal combinations and the corresponding phenotypes. Using a minimum support threshold of 0.0005 and in combination with literature review, the top six recommended combinations for perimenopause were merged into a key herbal combination set, while the top two combinations for anxiety, insomnia, and depression were respectively merged into key herbal combination sets for subsequent analysis. Prescriptions containing each herbal combination were extracted (Tab 1).

Tab 1 Inferred association rules and related prescriptions for phenotypes such as PMS, anxiety, depression, and insomnia

Association rule	Support	Confidence	Lift	Prescriptions
<i>Angelica gigas</i> , <i>Ligusticum chuanxiong</i> , <i>Rehmannia glutinosa</i> ->perimenopause	0.000504	0.0094	7.76	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
<i>Paeonia lactiflora</i> , <i>Angelica gigas</i> ->perimenopause	0.000504	0.0088	7.32	Seogyongtang,Sojeoktonggyeonghwan,Achulsan
<i>Ligusticum chuanxiong</i> , <i>Rehmannia glutinosa</i> ->perimenopause	0.000504	0.0078	6.42	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
<i>Angelica gigas</i> , <i>Rehmannia glutinosa</i> ->perimenopause	0.000504	0.0050	4.15	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan

Association rule	Support	Confidence	Lift	Prescriptions
Angelica gigas,Ligusticum chuanxiong->perimenopause	0.000504	0.0049	4.04	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
Angelica gigas,Glycyrrhiza uralensis->perimenopause	0.000504	0.0035	2.87	Hwagyeongtang,Seogyongtang,Sambitang,Achulsan
Phyllostachys nigra,Citrus trifoliata,Zingiber officinale->anxiety	0.000501	0.4167	23.25	Samhoondamtang,Gwibiondamtang,Cheongsimondamtang,Ondamtang,Gamiondamtang
Zingiber officinale,Citrus trifoliata,Phyllostachys nigra,Pinellia ternata->anxiety	0.000501	0.4167	23.25	Gwibiondamtang,Cheongsimondamtang,Ondamtang,Samhoondamtang,Gamiondamtang
Panax ginseng,Saussurea costus,Zingiber officinale,Polygala tenuifolia->Insomnia	0.000599	0.7500	36.84	Maehaekgigitagyeongheombang1,Gamiyukbitang,Gamigwibitang,Gamiondamtang,Gwibitang,Yeongsingwibitang
Saussurea costus,Glycyrrhiza uralensis,Zingiber officinale,Polygala tenuifolia->Insomnia	0.000599	0.7500	36.84	Maehaekgigitagyeongheombang1,Gamiyukbitang,Gamigwibitang,Gamiondamtang,Gwibitang,Yeongsingwibitang
Triticum aestivum,Ligusticum chuanxiong,Forsythia suspensa,Wolfiporia cocos->depression	0.000502	0.8333	60.63	Wolgukbohwhawan,Gungchihwadamjeon,Yugultang,Sokgukbohwhawan
Triticum aestivum,Ligusticum chuanxiong,Atractylodes lancea,Forsythia suspensa->depression	0.000502	0.7143	51.97	Wolgukbohwhawan,Gungchihwadamjeon,Yugultang,Sokgukbohwhawan

204

205 3.1.2 Herb–phenotype network proximity validation

206 To validate the reliability of the ARM results, a network proximity-based approach was
207 employed to calculate the relative proximity (Z_h) between the target sets of the inferred herbal
208 combinations and the disease-associated gene sets within the human protein–protein interaction
209 network. The results showed that all screened herbal combinations exhibited Z_h values below -0.15
210 (Tab 2), indicating significant associations between these combinations and the corresponding
211 phenotypes.

212 Furthermore, the herb–disease proximity was calculated for the top 10 herb–phenotype
213 association rules derived from ARM, while 10 randomly generated herbal combinations were used as
214 a reference to compute the corresponding herb–disease proximity (Fig 2). The results demonstrated
215 that the herbal combinations identified by ARM exhibited a lower median Z_h score.

216 Tab 2 Analysis Results of Herbal Levels for PMS, Anxiety, Depression and Insomnia

	Herb material	Herb-phenotype proximity(z_h)
perimenopause	<i>Paeonia lactiflora</i>	-1.6314
	<i>Glycyrrhiza uralensis</i>	-1.1484
	<i>Angelica gigas</i>	-1.0369
	<i>Rehmannia glutinosa</i>	-0.7336
	<i>Ligusticum chuanxiong</i>	-1.3285
anxiety	<i>Citrus trifoliata</i>	-1.0819
	<i>Zingiber officinale</i>	-1.1426
	<i>Pinellia ternata</i>	-1.0222
	<i>Phyllostachys nigra</i>	-1.3660
Insomnia	<i>Panax ginseng</i>	-0.6145
	<i>Glycyrrhiza uralensis</i>	-0.4715
	<i>Zingiber officinale</i>	-0.3817
	<i>Saussurea costus</i>	-0.6732
	<i>Polygala tenuifolia</i>	-0.9548
depression	<i>Triticum aestivum</i>	-1.1273
	<i>Atractylodes lancea</i>	-0.7372
	<i>Wolfiporia cocos</i>	-1.0334
	<i>Forsythia suspensa</i>	-1.0827
	<i>Ligusticum chuanxiong</i>	-1.0455

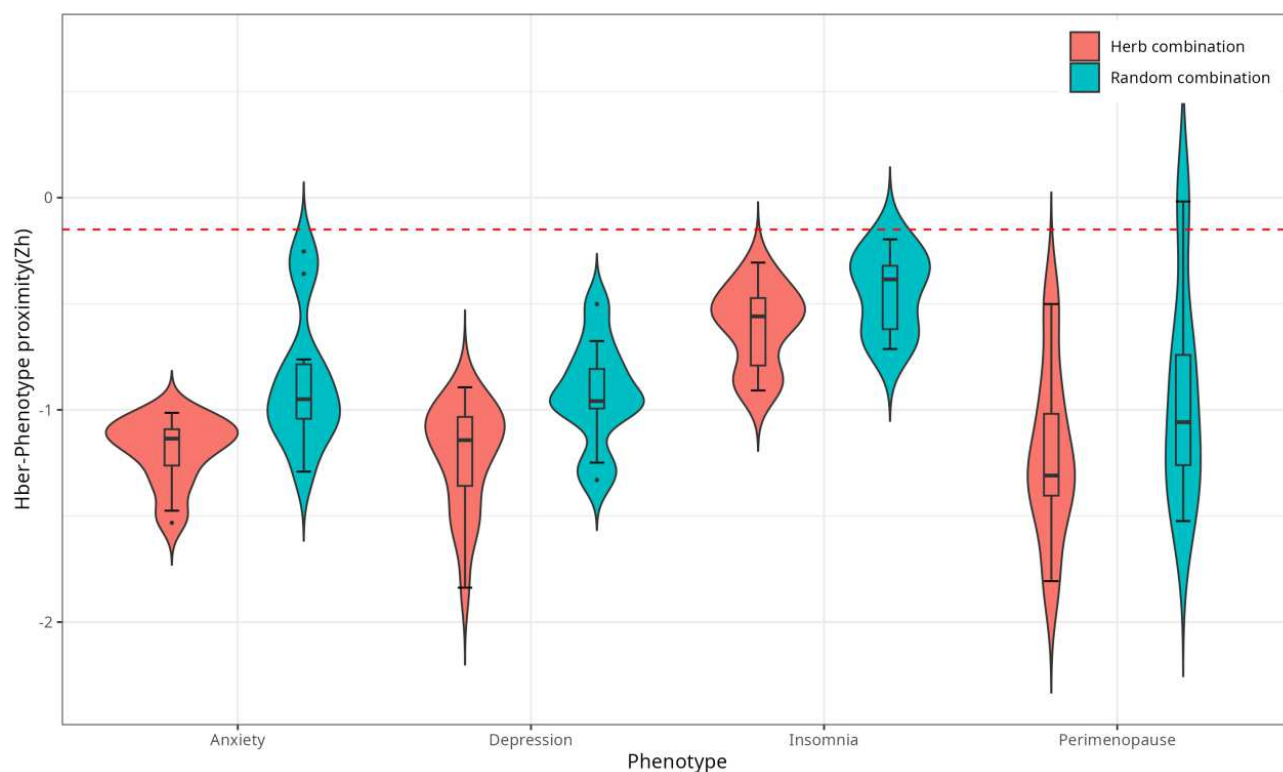


Fig 2 Extracts the compound-phenotype ARM and the compounds identified in the compound-phenotype network analysis. The x-axis represents diseases, the y-axis represents the proximity of the drug to the disease, the red part represents the drugs selected by the ARM rule, and the green part represents the randomly selected drug combinations.

3.2 Compound-level analysis results and screening of core compounds

Based on the herb-level analysis, compounds present in the inferred herbal combinations were further investigated. Compound-phenotype association rule mining (minimum support threshold of 0.008) and compound-phenotype network proximity analysis (threshold $Z_c \leq -0.15$) were performed to identify compounds associated with each phenotype, and the results from the two approaches were intersected to obtain candidate compounds (Fig 3). The results showed that the numbers of candidate compounds for PMS, anxiety, depression, and insomnia were 10, 5, 7, and 2, respectively. Combined with an extensive literature review^[25-30], six compounds were ultimately selected as core compounds for further analysis in this study: reduced glutathione, D-tryptophan, linalool, gamma-aminobutyric acid (GABA), folate, and resveratrol.

3.3 Structural network analysis results

Cytoscape was used to visualize herb compound-related targets, candidate compounds, and genes associated with PMS, anxiety, depression, and insomnia (Fig 3).

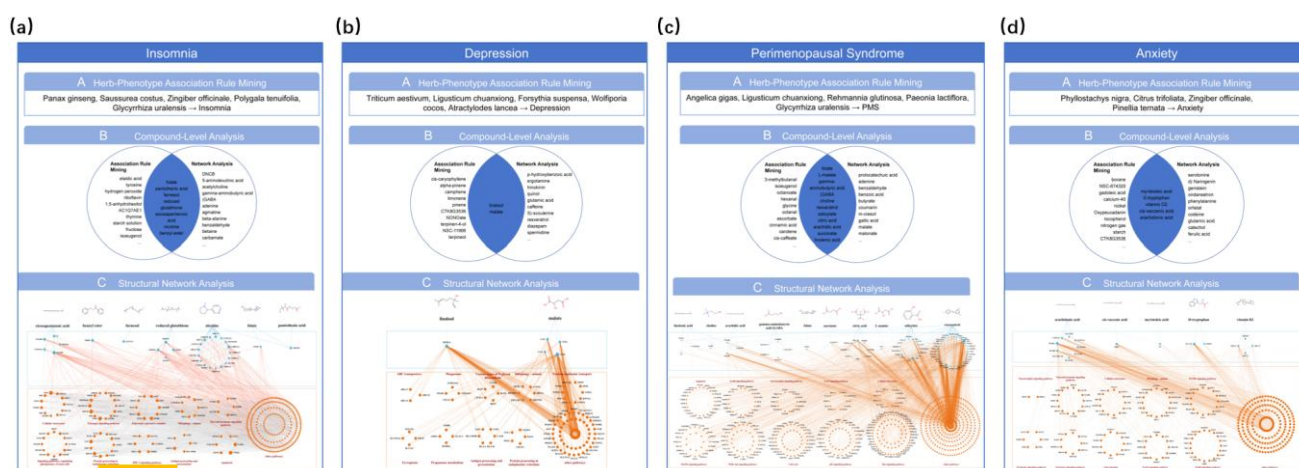


Fig 3. Integrated herb-compound-phenotype analysis for perimenopausal syndrome (PMS), anxiety, depression and insomnia. Note: (a), Analysis results of PMS, including herb-phenotype association rules, Venn diagram of candidate compounds from two mining strategies, and compound-target-disease structural network; (b), Corresponding analytical results for anxiety; (c), Matching analytical results for depression; (d), Parallel analytical results for insomnia. All panels share three modules: herb association rule mining, candidate compound intersection, and network visualization. Separate single diagrams for each disease are shown in Supplementary Figures S1–S4.

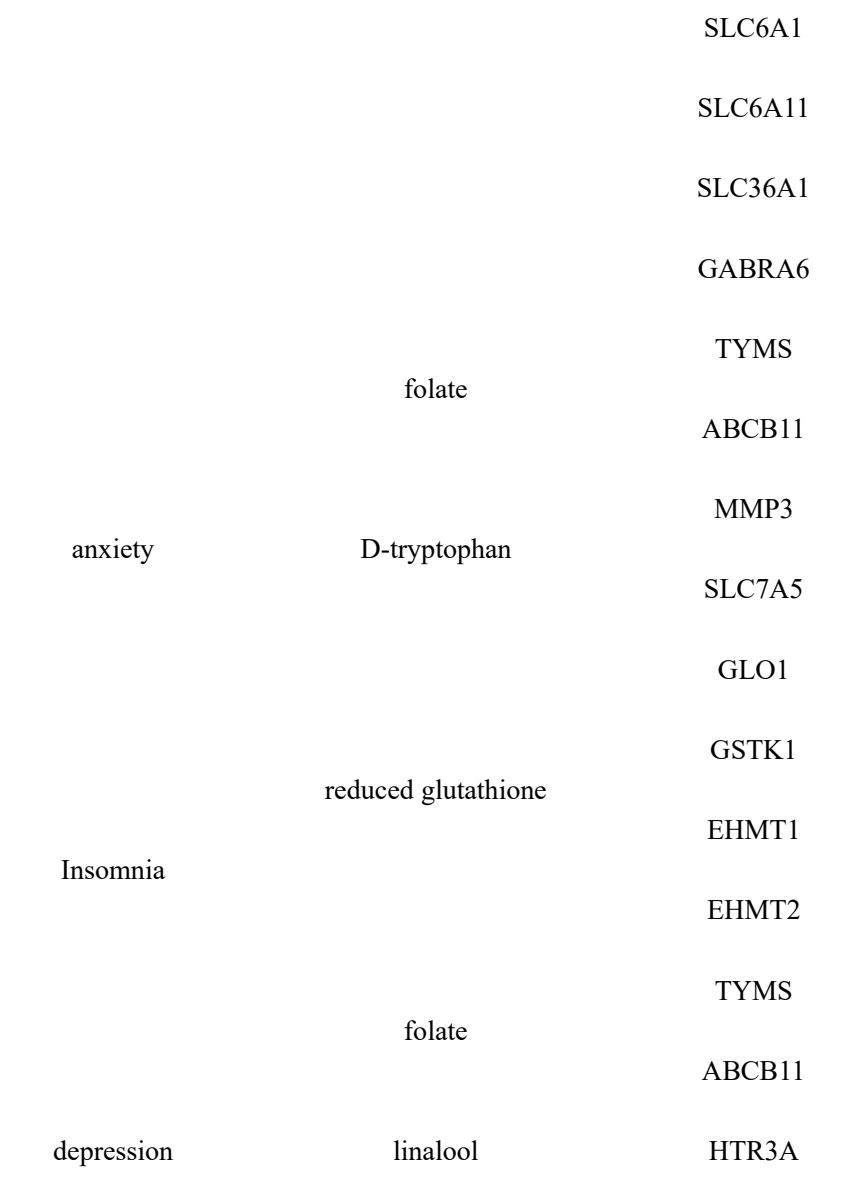
3.4 Core target screening and network construction

3.4.1 Core gene screening

To identify key targets through which the core compounds act on each disease, the overlapping genes between the targets of each core compound and the corresponding disease targets were obtained, and a PPI network was constructed. Five algorithms (MCC, Degree, BottleNeck, Closeness, and Radiality) from the CytoHubba plugin were used to identify the top five genes for each disease, and the intersections were defined as the core genes (Tab 3). Specifically, 47 overlapping genes were identified between the targets of resveratrol and PMS-related targets, and after screening with the five algorithms, three core genes (MYC, ESR1, and APP) were identified as shared core genes. All five targets of GABA (GABBR2, SLC6A1, SLC6A11, SLC36A1, and GABRA6) were identified as core genes. The core genes for folate were TYMS and ABCB11; for D-tryptophan, MMP3 and SLC7A5; for reduced glutathione, GLO1, GSTK1, EHMT1, and EHMT2; and for linalool, HTR3A.

Tab 3 Structural Network Analysis of PMS, Anxiety, Depression, and Insomnia and Search Results of Biological Pathways

disease	compound	herb target
perimenopause	resveratrol	MYC
		ESR1
		APP
	gamma-aminobutyric acid (GABA)	GABBR2



256

257 3.4.2 Construction of the “herb–core compound–core gene–disease” network

258 To systematically illustrate the complex regulatory relationships among the core herbs, core
 259 compounds, and core genes, a multi-level interaction network was constructed using Cytoscape (Fig
 260 4). The network comprised 9 core herbs, 6 core compounds, 16 drug targets, and 614 disease-related
 261 genes. The network intuitively revealed that core compounds (e.g., resveratrol and GABA) can act on
 262 multiple disease targets simultaneously, reflecting the “multi-component, multi-target” synergistic
 263 characteristics of traditional Chinese medicine. Notably, the APP gene was identified not only as a
 264 target of resveratrol but also as a common disease-related gene shared by PMS, anxiety, depression,
 265 and insomnia, suggesting that it may serve as a key hub linking these conditions.

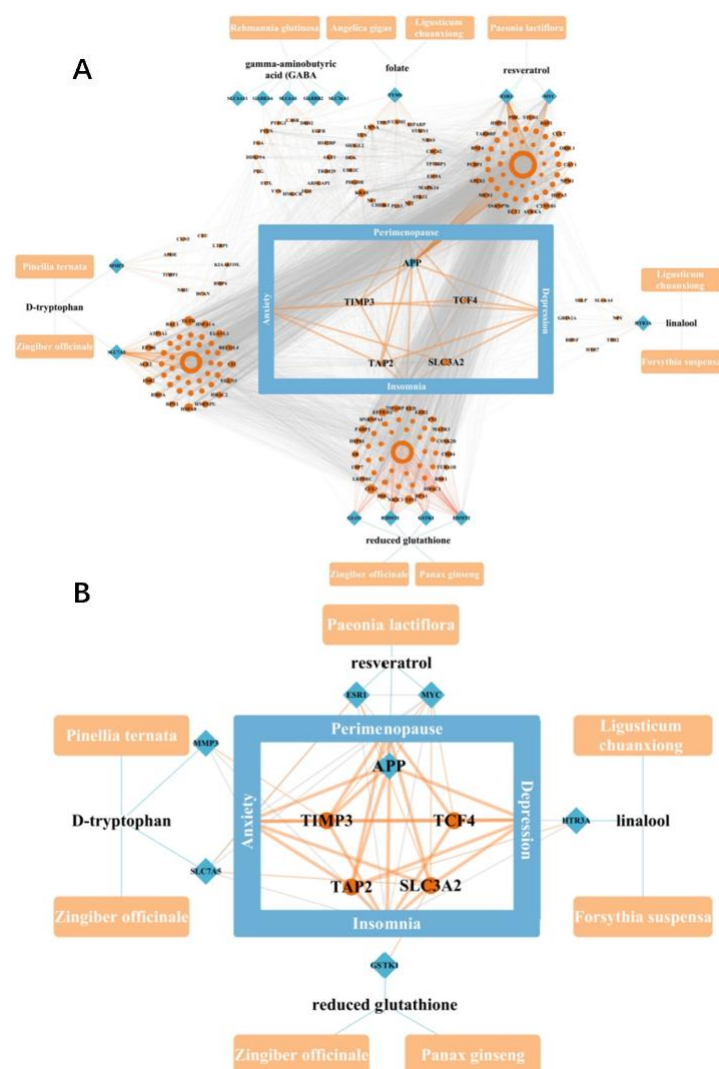


Figure 4. Herb - compound - core gene - disease network analysis. Note: A, Global herb-compound-target-disease network. Rectangles with orange background denote herbal medicines, connected nodes represent bioactive compounds, blue rhombuses stand for drug targets, orange circles indicate disease targets, and rectangles with blue background represent disorders. B, Simplified network focusing on shared hub genes across perimenopausal syndrome, anxiety, depression and insomnia. The core shared targets of all four phenotypes are displayed in the central region, among which APP acts as both a drug target and a disease target. Separate individual diagrams are provided in Supplementary Figures S5–S6.

3.5 Molecular docking and molecular dynamics simulation validation

3.5.1 Molecular docking results

To evaluate the binding potential between the core compounds and the corresponding core genes, molecular docking simulations were performed using the CB-Dock2 platform. A lower binding energy indicated more stable ligand–receptor binding. As shown in Tab S1, the binding energies of all six core compounds with their respective core proteins were as follows: folate-TYMS (−9.1 kcal/mol), resveratrol-ESR1 (−7.8 kcal/mol), D-tryptophan-MMP3 (−7.7 kcal/mol), reduced glutathione-EHMT2 (−6.8 kcal/mol), linalool-HTR3A (−6.1 kcal/mol), gamma-aminobutyric acid (GABA)-SLC6A11 (−4.3 kcal/mol), suggesting strong binding affinity^[18]. Among them, folate exhibited the lowest binding energy with TYMS. The conformation with the lowest binding energy

for each complex was selected and visualized using PyMOL (Fig 5A).

3.5.2 Molecular dynamics simulation results

To evaluate the stability of the aforementioned optimal docking complexes under dynamic conditions, 100 ns molecular dynamics simulations were performed for the six complexes. The stability of the complexes was analyzed by calculating parameters including root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and total energy (Fig 5B).

RMSD analysis (Fig 5Ba) showed that all complexes exhibited small fluctuations in RMSD values throughout the simulation, with none exceeding 2 nm, indicating that the systems reached equilibrium. Among them, the ESR1 – resveratrol complex maintained consistently low RMSD values with the smallest fluctuations, indicating the highest structural stability; the TYMS – folate complex showed a slight increase in RMSD during the later stage of the simulation but remained generally stable.

RMSF analysis (Fig 5Bb) revealed a prominent RMSF peak near residue 200 of HTR3A, suggesting high flexibility in this region, which may represent a conformational change induced by ligand binding or the location of the binding site. In contrast, proteins such as ESR1 and MMP3 exhibited generally low RMSF values, indicating minimal conformational changes and relatively rigid structures throughout the simulation.

Total energy analysis (Fig 5Bc) showed that the total energy of all systems remained at low levels with stable trends and no drastic fluctuations, confirming good simulation stability and high reliability of the results.

Hydrogen bond count analysis (Fig 5Bd) showed that the TYMS – folate and EHMT2 – reduced glutathione complexes formed a relatively high number of hydrogen bonds with fluctuations during most of the simulation, suggesting possible dynamic processes of binding and dissociation or transitions between binding sites. In contrast, complexes such as ESR1 – resveratrol and MMP3 – D-tryptophan exhibited fewer hydrogen bonds with relatively stable patterns over most of the simulation, indicating relatively fixed binding modes.

Collectively, the molecular dynamics simulation results demonstrated favorable binding stability for all six core compounds with their respective target proteins, further validating the reliability of the molecular docking findings.

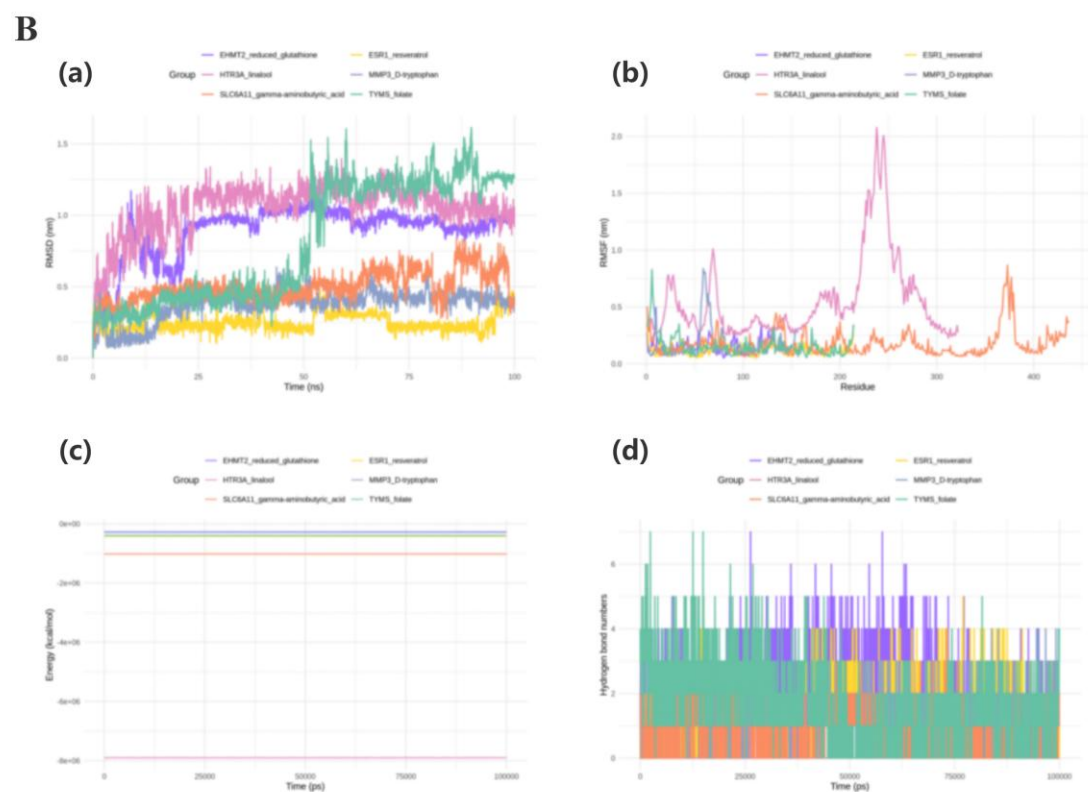
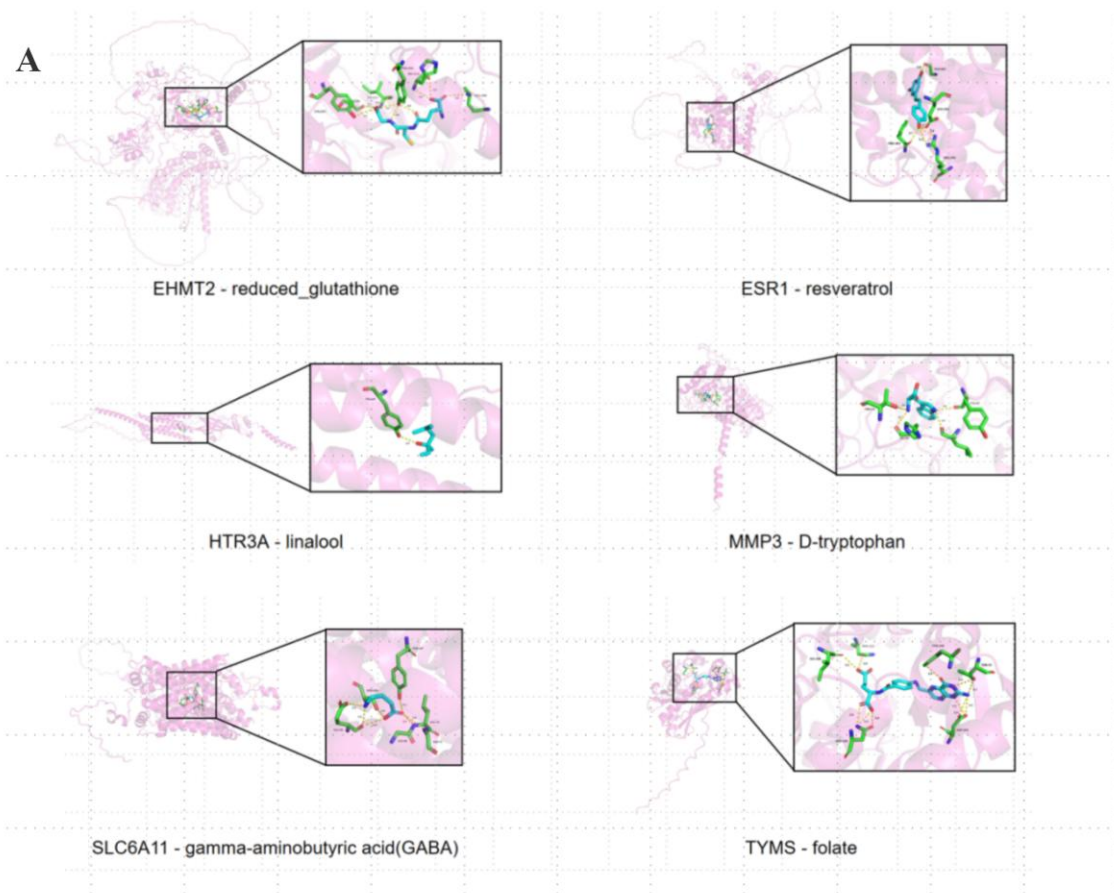


Fig. 5 Molecular docking visualization and molecular dynamics simulation. Note: B(a), Root-mean-square deviation (RMSD) as a function of time. The x-axis represents time, and the y-axis represents RMSD; smaller fluctuations indicate more stable protein–ligand binding. B(b), The x-axis represents the amino acid residue number, and the y-axis represents root-mean-square fluctuation (RMSF); larger fluctuations indicate greater flexibility of protein amino acid residues. B(c), Energy changes during molecular dynamics simulation. The x-axis represents time, and the y-axis represents total energy; lower energy and smaller fluctuations indicate greater stability. B(d), Changes in the number of hydrogen bonds formed between the compound and the protein over time. The x-axis represents time, and the y-axis represents the number of hydrogen bonds; a higher number of hydrogen bonds indicates more stable binding.

4 Discussion

This study adapted and extended the multi-level association rule mining and network pharmacology integration framework proposed by Yu et al.^[31] to systematically identify potential high-frequency herbal combinations for the treatment of perimenopausal syndrome (PMS) and its commonly associated conditions (anxiety, depression, and insomnia). Traditional Chinese medicine (TCM) formulas, characterized by their multi-component and multi-target features, offer unique advantages in managing complex psychosomatic disorders such as PMS^[32,33]; however, the underlying principles governing herb compatibility are often rooted in empirical experience and lack data-driven systematic interpretation. By mining high-frequency herb pairs through algorithmic approaches, this study not only provides a scientific basis for clinical practice but also offers new perspectives for understanding the dialectical logic of “treating different diseases with the same approach” and “treating the same disease with different approaches” in TCM.

4.1 Interpretation of high-frequency herbal combinations from the perspective of TCM theory

Through association rule mining, this study identified a series of herbal combinations strongly associated with the core symptoms of PMS as well as with associated mood and sleep disturbances. For PMS itself, the combination “*Angelica sinensis*–*Ligusticum chuanxiong*–*Rehmannia glutinosa*” (Danggui–Chuanxiong–Dihuang) emerged with the highest frequency, closely aligning with the core structure of the classic formula Siwu Tang (Four Substances Decoction, composed of Danggui, Chuanxiong, Baishao, and Shudihuang). According to TCM theory, the fundamental pathogenesis of PMS lies in kidney essence deficiency and gradual exhaustion of Tian Gui, often involving the liver, heart, and other viscera, leading to blood deficiency with liver depression and mental restlessness^[34,35]. Siwu Tang is a representative formula for nourishing and invigorating blood, and its core herb group “Danggui–Chuanxiong–Dihuang” embodies the therapeutic principles of nourishing blood, activating blood circulation, and tonifying kidney yin^[36], and is commonly used to alleviate core symptoms such as hot flashes and night sweats. Moreover, the frequent occurrence of sub-combinations such as “Baishao–Danggui” and “Chuanxiong–Dihuang” further supports the synergistic roles of this core herb group in harmonizing qi and blood and nourishing the liver and kidney.

More notably, this study uncovered specific herb pairs associated with common comorbid conditions. For perimenopausal anxiety, the high-frequency combinations “*Bambusa textilis*–*Citrus aurantium*–*Zingiber officinale*” (Zhuru–Zhiqiao–Shengjiang) and “*Pinellia ternata*–*Zingiber officinale*–*Bambusa textilis*–*Citrus aurantium*” (Banxia–Shengjiang–Zhuru–Zhiqiao) reflect the therapeutic strategies of regulating qi, resolving phlegm, harmonizing the stomach, and calming the mind. In TCM, anxiety and irritability are often attributed to liver qi stagnation and phlegm-heat disturbing the heart^[37,38]; such combinations may alleviate tension by modulating autonomic nervous

function and neurotransmitter balance^[39]. For perimenopausal depression, combinations such as “*Triticum aestivum*–*Ligusticum chuanxiong*–*Forsythia suspensa*–*Poria cocos*” (Xiaomai–Chuanxiong–Lianqiao–Fuling) and “*Triticum aestivum*–*Ligusticum chuanxiong*–*Atractylodes lancea*–*Forsythia suspensa*” (Xiaomai–Chuanxiong–Cangzhu–Lianqiao) suggest a therapeutic strategy that simultaneously emphasizes “soothing the liver, relieving depression, strengthening the spleen, and calming the heart” along with “clearing heat, resolving phlegm, invigorating blood, and opening stagnation.” Within these combinations, the pairing of Xiaomai (the principal herb in Ganmai Dazao Tang)^[40] and Chuanxiong (which invigorates blood and circulates qi)^[41] appears particularly critical. For insomnia, combinations such as “*Panax ginseng*–*Polygala tenuifolia*–*Aucklandia lappa*–*Zingiber officinale*” (Renshen–Yuanzhi–Muxiang–Shengjiang) and “*Aucklandia lappa*–*Glycyrrhiza uralensis*–*Zingiber officinale*–*Polygala tenuifolia*” (Muxiang–Gancao–Shengjiang–Yuanzhi) highlight the therapeutic approaches of “supplementing qi, nourishing the middle, communicating heart and kidney, regulating qi, and calming the mind.” Yuanzhi calms the mind and enhances cognition, Renshen supplements qi and generates fluids, and Muxiang regulates qi and invigorates the spleen, collectively targeting the insomnia pathogenesis associated with dual deficiency of heart and spleen and disharmony between heart and kidney^[42].

4.2 Pharmacological mechanisms of the core compounds

Through network pharmacology and literature review, this study identified six core compounds: resveratrol, GABA, folate, D-tryptophan, reduced glutathione, and linalool. These compounds may exert synergistic effects in the treatment of PMS and its associated conditions.

Resveratrol is a natural polyphenolic compound widely found in plants such as grapes and *Polygonum cuspidatum*. Molecular docking results showed that resveratrol exhibits high binding affinity for the estrogen receptor ESR1, and molecular dynamics simulations further confirmed the structural stability of this complex throughout the 100 ns simulation. Multiple studies have demonstrated that resveratrol can improve pain perception, bone metabolism, menopausal symptoms, and overall health status in postmenopausal women^[43,44]. As a key molecule in the estrogen signaling pathway, ESR1 plays a central role in the pathogenesis of PMS^[45]. Resveratrol may partially substitute for estrogen by modulating ESR1 activity, thereby alleviating symptoms such as hot flashes and mood swings caused by estrogen fluctuations.

γ -Aminobutyric acid (GABA) is the primary inhibitory neurotransmitter in the central nervous system and plays a critical role in regulating mood, sleep, and anxiety^[46,47]. This study found that GABA acts on multiple targets, including GABBR2, SLC6A1, and SLC6A11, which are involved in GABA synthesis, transport, and receptor binding^[48]. Dysfunction of the GABAergic system is closely associated with anxiety and insomnia^[26]. The high-frequency herbal combinations identified in this study may exert sedative and calming effects by influencing GABAergic neurotransmission.

D-Tryptophan is the D-isomer of tryptophan, which can be converted in the body to serotonin and subsequently participates in melatonin synthesis^[27]. Melatonin is a key neurohormone regulating sleep–wake rhythms. This study found that D-tryptophan targets MMP3 and SLC7A5, suggesting that it may improve sleep and mood by modulating amino acid metabolism and neurotransmitter balance. This finding is consistent with previous studies indicating that the tryptophan metabolic pathway plays an important role in the pathogenesis of depression and insomnia^[49].

Reduced glutathione is a major antioxidant in the body that protects cells from oxidative stress damage^[50]. Oxidative stress is recognized as an important pathological mechanism underlying various diseases, including PMS, anxiety, and depression^[28]. This study found that reduced

glutathione targets genes such as GLO1, GSTK1, EHMT1, and EHMT2, which are involved in oxidative stress responses and epigenetic regulation^[51,52]. Linalool is a natural monoterpene compound with anxiolytic and sedative effects, and its target HTR3A is the 5-hydroxytryptamine receptor 3A subunit, which participates in mood regulation and anxiety responses^[29]. Folate is involved in one-carbon metabolism and is essential for neurotransmitter synthesis and DNA methylation; folate deficiency has been associated with an increased risk of depression^[30].

4.3 Molecular mechanisms of core genes

Through PPI network analysis and multiple topological algorithms, this study identified core genes including MYC, ESR1, and APP. APP (amyloid precursor protein) is a key gene in Alzheimer's disease^[53], but its role in PMS and related conditions has garnered increasing attention in recent years^[54]. This study found that APP serves not only as a target of resveratrol but also as a common disease-related gene shared by PMS, anxiety, depression, and insomnia, suggesting that it may serve as a key hub linking these conditions. Previous studies have shown that sleep disturbances may induce the production of pro-inflammatory cytokines, leading to abnormal APP metabolism and amyloid accumulation, thereby increasing the risk of neurodegenerative diseases^[55]. Fluctuations in estrogen levels in patients with PMS may affect APP expression and processing, consequently influencing neural function.

MYC, an important transcription factor, is involved in cell proliferation, apoptosis, and metabolic regulation. Recent studies have suggested that MYC may participate in regulating circadian rhythms and sleep homeostasis^[56]. ESR1 is estrogen receptor alpha, a core molecule in the estrogen signaling pathway, and its role in the pathogenesis of PMS has been widely recognized^[57]. The multi-level "herb-core compound-core gene-disease" network constructed in this study intuitively reveals that core compounds (e.g., resveratrol and GABA) can act on multiple disease targets simultaneously, reflecting the "multi-component, multi-target" synergistic characteristics of TCM.

4.4 Limitations of the study

This study has several limitations. First, association rule mining is highly dependent on the comprehensiveness and accuracy of the databases used, and potential data bias may exist, with some potentially effective herbal combinations possibly not being captured. Second, this study is currently at the stage of data mining and mechanistic hypothesis generation; the efficacy, optimal compatibility ratios, and specific mechanisms of action of the identified herb pairs require further validation through in vitro and in vivo pharmacological experiments as well as clinical studies.

5 Conclusion

In summary, by integrating multi-level association rule mining with network pharmacology strategies, this study identified high-frequency herbal combinations and core compounds for the treatment of PMS and its associated conditions (anxiety, depression, and insomnia) from large-scale data and preliminarily elucidated their potential mechanisms of action. This study not only provides data-driven support for understanding the scientific basis of the "treating different diseases with the same approach" principle in traditional medicine but also offers candidate molecules for the modern treatment and drug development of PMS. Future research should further conduct in vitro and in vivo experimental validation to explore the multi-level regulatory network of "symptom-herb pair-target-pathway," thereby advancing the precision and scientific rigor of TCM formula-based treatment for complex diseases.

Key Points

- Multi-level ARM and network pharmacology screen six high-frequency herbal combinations for PMS and its comorbidities.
- Six core compounds and three core genes are validated.
- Molecular docking and 100 ns MD simulations confirm stable compound–target binding.
- Data-driven evidence supports the “treating different diseases with the same approach” principle and provides candidate compounds for drug development.

6 Author Contributions

Conceptualization: Hao Wang.
Data curation: Haoran Li.
Formal analysis: Yixuan Wang.
Funding acquisition: Liquan Yu, Hao Li.
Methodology: Ruolan Zhang.
Project administration: Meixi Wang.
Software: Suyang Ma, Xinrui Men.
Supervision: Hao Li.
Validation: Haoxiang Li, Yuelin Huang.
Visualization: Yixuan Wang.
Writing – original draft: Hao Wang, Haoran Li.
Writing – review & editing: Liquan Yu, Hao Li.

7 Funding

This work was supported by the Nursery cultivation program of Beijing Traditional Chinese Medicine Science and Technology Development Fund [grant number BJZYQN-2023-12]; the High Innovation Program – Summit Youth Talents Spring Blossom Project in 2025 [grant number G202534225]; and the Beijing Medical Health Collaborative Project on High-Quality Research and Development of Traditional Chinese Medicine [grant number zyygzl-2025-017].

8 Data Availability

All data used in this study are publicly available from the databases cited in the Materials and Methods section. The specific database sources and URLs are as follows: KTKP (<https://koreantk.com/>), TCM-ID (accessed via internal database; no public URL), KampoDB (<https://wakanmoview.inm.u-toyama.ac.jp/kampo/>), COCONUT (<https://coconut.naturalproducts.net/>), DrugBank (<https://www.drugbank.ca/>), TTD (<https://idrblab.org/ttd/>), CTD (<https://ctdbase.org/>), BindingDB (<https://www.bindingdb.org/>), STITCH (<https://stitch.embl.de/>), FooDB (<https://foodb.ca/>), CMAUP (<https://bidd2.nus.edu.sg/CMAUP/>), BioGRID (<https://thebiogrid.org/>), GeneCards (<https://www.genecards.org/>), and OMIM (<https://www.ncbi.nlm.nih.gov/omim>). All data were accessed between January 2024 and December 2024. The processed datasets and custom analysis scripts (R/Python) are available from the corresponding author upon reasonable request.

9 Conflicts of Interest

The authors declare that there are no conflicts of interest.

10 References

- [1] Zhang X ,Luo L ,Wang C , et al.Research progress on Chaihu Shugan San in treating perimenopausal syndrome: A review.[J].Medicine,2024,103(52):e41044.DOI:10.1097/MD.00000000000041044
- [2] Chen W, Chen M, Tang H, et al. Advances in diagnosis and treatment of perimenopausal syndrome[J]. Open Life Sciences,2023,18(1):20220754-20220754.DOI:10.1515/BIOL-2022-0754.
- [3] Min SH, Yang Q, Docherty SL, et al. Symptom Clusters and Key Symptoms Among Midlife Perimenopausal and Postmenopausal Women With and Without Metabolic Syndrome[J].Nursing Research,2022, 71(4):E28-E38.DOI:10.1097/NNR.0000000000000591.
- [4] Tang BX, Meng QY, Xie C, et al. Ziyin Bushen Decoction Alleviates Perimenopausal Syndrome in Rats by Enhancing Estradiol Production.[J].Evidence-based complementary and alternative medicine : eCAM,2020.(2020):8895809-8895809.DOI:10.1155/2020/8895809.
- [5] Xiang L ,Ziqi L ,Jie L , et al.Network pharmacology approaches for research of Traditional Chinese Medicines.[J].Chinese journal of natural medicines,2023,21(5):323-332.DOI:10.1016/S1875-5364(23)60429-7.
- [6] Junjie L ,Caiming W ,Wen'na L , et al.An Effective Treatment of Perimenopausal Syndrome by Combining Two Traditional Prescriptions of Chinese Botanical Drugs[J].Frontiers in Pharmacology,12.(2021): 744409-744409.DOI:10.3389/FPHAR.2021.744409.
- [7] Xueqin C ,Caiming W ,Wen'na L , et al.The Metabolomic Rationale for Treating Perimenopausal Syndrome as Kidney Deficiency[J].Evidence-Based Complementary and Alternative Medicine,2020,2020:8568353.DOI:10.1155/2020/8568353.
- [8] Yu H ,Choi K ,Kim Y J , et al.Multi-level association rule mining and network pharmacology to identify the polypharmacological effects of herbal materials and compounds in traditional medicine.[J].Briefings in bioinformatics,2025,26(4):bbaf328.DOI:10.1093/BIB/BBAF328.
- [9] Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018.[J].Nucleic acids research,2018,46(D1):D1074-D1082.DOI:10.1093/nar/gkx1037.
- [10] Yunxia W ,Song Z ,Fengcheng L , et al.Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics.[J].Nucleic acids research,2020,48(D1):D1031-D1041.DOI:10.1093/nar/gkz981.
- [11] Davis AP, Grondin CJ, Johnson RJ, et al. Comparative Toxicogenomics Database (CTD): update 2021. [J].Nucleic acids research,2020,49(D1):D1138-D1143.DOI:10.1093/nar/gkaa891.
- [12] Gilson MK, Liu T, Baitaluk M, et al. BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology.[J].Nucleic acids research,2016,44(D1):D1045-53.DOI:10.1093/nar/gkv1072.
- [13] Damian S ,Alberto S ,Christian M V , et al.STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data.[J].Nucleic acids research,2016,44(D1):D380-D384.DOI:10.1093/nar/gkv1277.
- [14] Xian Z, Peng Z, Yali W, et al.CMAUP: a database of collective molecular activities of useful plants. [J].Nucleic acids research,2019,47(D1):D1118-D1127.DOI:10.1093/nar/gky965.
- [15] Oughtred R, Stark C, Breitkreutz BJ, et al. The BioGRID interaction database: 2019 update.[J].Nucleic acids research,2019,47(D1):D529-D541.DOI:10.1093/nar/gky1079.
- [16] Safran M, Dalah I, Alexander J, et al. GeneCards Version 3: the human gene integrator.[J].Database : the journal of biological databases and curation,2010,2010baq020.DOI:10.1093/database/baq020.

- 522 [17] Amberger JS, Bocchini CA, Schiettecatte F, et al. OMIM.org: Online Mendelian Inheritance in Man
523 (OMIM®), an online catalog of human genes and genetic disorders.[J].Nucleic acids research,2015,43
524 (Database issue):D789-98.DOI:10.1093/nar/gku1205.
- 525 [18] Yang BZ, Xiang B, Wang T, et al. Neurogenetic underpinnings of nicotine use severity: Integrating th
526 e brain transcriptomes and GWAS variants via network approaches[J].Psychiatry Research,2024,3341158
527 15.DOI:10.1016/J.PSYCHRES.2024.115815.
- 528 [19] Antoine D ,Olivier M ,Vincent Z .SwissADME: a free web tool to evaluate pharmacokinetics, drug-lik
529 eness and medicinal chemistry friendliness of small molecules.[J].Scientific reports,2017,7(1-4):42717.D
530 OI:10.1038/srep42717
- 531 [20] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among
532 gene clusters.[J].Omics : a journal of integrative biology,2012,16(5):284-7.DOI:10.1089/omi.2011.0118.
- 533 [21] Minoru K, Miho F ,Yoko S, et al. KEGG for taxonomy-based analysis of pathways and genomes.[J].
534 Nucleic acids research,2022,51(D1):DOI:10.1093/NAR/GKAC963
- 535 [22] Zhou H ,Huertas J ,Maristany J M , et al. Multiscale structure of chromatin condensates explains pha
536 se separation and material properties.[J].Science (New York, N.Y.),2025,390(6777):eadv6588.DOI:10.112
537 6/SCIENCE.ADV6588.
- 538 [23] Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from compl
539 ex interactome.[J].BMC systems biology,2014,8 Suppl 4(S-4):S11.DOI:10.1186/1752-0509-8-S4-S11.
- 540 [24] Páll S, Zhmurov A, Bauer P, et al. Heterogeneous parallelization and acceleration of molecular dynam
541 ics simulations in GROMACS[J].JOURNAL OF CHEMICAL PHYSICS,2020,153(13):134110.DOI:10.10
542 63/5.0018516.
- 543 [25] Xixi J, Yanan S, Zhengqian L, et al.Melatonin regulates the circadian rhythm to ameliorate postoperat
544 ive sleep disorder and neurobehavioral abnormalities in aged mice.[J].CNS neuroscience & therapeutics.
545 2024;30:e14436.DOI:10.1111/CNS.14436.
- 546 [26] Rockstrom MD, Chen L, Taishi P, et al. Tumor necrosis factor alpha in sleep regulation[J].Sleep Med
547 icine Reviews,2018,4069-78.DOI:10.1016/j.smrv.2017.10.005.
- 548 [27] Jenkins T, Nguyen J, Polglaze K, et al. Influence of Tryptophan and Serotonin on Mood and Cogniti
549 on with a Possible Role of the Gut-Brain Axis[J].Nutrients,2016,8(1):56.DOI:10.3390/nu8010056.
- 550 [28] Liu N ,Liu Y ,Wang Y , et al.Oxidative cell death in the central nervous system: mechanisms and th
551 erapeutic strategies[J].Frontiers in Cell and Developmental Biology,13.(2025):1562344-1562344.DOI:10.3
552 389/FCELL.2025.1562344.
- 553 [29] Shinhui L, Woog J, Sanung E, et al. Molecular Regulation of Betulinic Acid on $\alpha\beta 4$ Nicotinic Acet
554 ylcholine Receptors[J].Molecules,2021,26(9):2659-2659.DOI:10.3390/MOLECULES26092659.
- 555 [30] Ghilotti F, Bellocco R, Trolle Lagerros Y, et al. Relationship between sleep characteristics and marker
556 s of inflammation in Swedish women from the general population.[J].Journal of sleep research,2020,30
557 (2):e13093.DOI:10.1111/jsr.13093.
- 558 [31] Yu H, Choi K, Kim JY, et al. Multi-level association rule mining and network pharmacology to ident
559 ify the polypharmacological effects of herbal materials and compounds in traditional medicine.[J].Briefi
560 ngs in bioinformatics,2025,26(4):bbaf328.DOI:10.1093/BIB/BBAF328.
- 561 [32] Yuan H ,Ma Q ,Cui H, et al.How Can Synergism of Traditional Medicines Benefit from Network Ph
562 armacology?[J].Molecules,2017,22(7):1135-1135.DOI:10.3390/molecules22071135
- 563 [33] Cheuka M P ,Mayoka G ,Mutai P , et al.The Role of Natural Products in Drug Discovery and Devel
564 opment against Neglected Tropical Diseases[J].Molecules,2016,22(1):58.DOI:10.3390/molecules22010058.
- 565 [34] Lan J, Wu C, Liang W, et al. An Effective Treatment of Perimenopausal Syndrome by Combining T
566 wo Traditional Prescriptions of Chinese Botanical Drugs[J].Frontiers in Pharmacology,2021,12744409-74
567 4409.DOI:10.3389/FPHAR.2021.744409.
- 568 [35] Chen X, Wu C, Liang W, et al. The Metabolomic Rationale for Treating Perimenopausal Syndrome a

- s Kidney Deficiency[J].Evidence-Based Complementary and Alternative Medicine,2020,2020:8568353.D
OI:10.1155/2020/8568353.
- [36] Dan X, Ming L, Ying L, et al.An Analysis of the Clinical Medication Rules of Traditional Chinese
Medicine for Polycystic Ovary Syndrome Based on Data Mining[J].Evidence-Based Complementary and
Alternative Medicine,2023,20236198001-6198001.DOI:10.1155/2023/6198001.
- [37] Katherine F, Jonathan K, Rubbia S M, et al.Anxiety and Depressive Symptoms and Disorders in Chil
dren and Adolescents With Migraine: A Systematic Review and Meta-analysis.[J].JAMA pediatrics,202
2,DOI:10.1001/JAMAPEDIATRICS.2022.3940.
- [38] O'Leary KB, Khan JS. Pharmacotherapy for Anxiety Disorders.[J].The Psychiatric clinics of North Am
erica,2024,47(4):689-709.DOI:10.1016/J.PSC.2024.04.012.
- [39] Kim TW, Jeong JH, Hong SC. The impact of sleep and circadian disturbance on hormones and meta
bolism.[J].International journal of endocrinology,2015;2015:591729.DOI:10.1155/2015/591729.
- [40] Wang X. Summary of Professor Zhao Jianjun's experience in treating insomnia based on data mining
system.[D].Changchun University of Chinese Medicine,2020.DOI:10.26980/d.cnki.gcczc.2020.000302
- [41] Yuzhi Z ,Tian L ,Shiwei Z, et al.Study on antidepressant mechanism of Radix Bupleuri-Radix Paeonia
e Alba herb pair by metabonomics combined with 1H nuclear magnetic resonance and ultra-high-perfo
rmance liquid chromatography-tandem mass spectrometry detection technology.[J].The Journal of pharm
acy and pharmacology,2021,73(9):1262-1273.DOI:10.1093/JPP/RGAB061.
- [42] Xulong H ,Huajuan J ,Jin P , et al.Study on the potential mechanism, therapeutic drugs and prescripti
ons of insomnia based on bioinformatics and molecular docking[J].Computers in Biology and Medicin
e,149.(2022):106001-106001.DOI:10.1016/J.COMPBIOMED.2022.106001.
- [43] Cardinali DP. Melatonin as a chronobiotic/cytoprotector: its role in healthy aging. [J].Biological Rhyth
m Research,2019;50(1):28-45.DOI:10.1080/09291016.2018.1491200.
- [44] Amini A ,Soureshjani H S ,Sherwin M C , et al.The Effects of Resveratrol on Menopausal Cardio-m
etabolic Changes: A Systematic Review.[J].Cardiovascular & hematological disorders drug targets,2025,
114 - 130.DOI:10.2174/011871529X398468250910225946.
- [45] Xiang X ,Jiang W ,Palasuberniama P , et al.Estrogen receptor-autophagy axis protects the heart, aorta
and kidneys during perimenopausal aging: Evidence from human cohort and mouse experiments.[J].Li
fe sciences,2026,387124183.DOI:10.1016/J.LFS.2025.124183.
- [46] Li Y ,Wang Y ,Li L , et al.Astrocytic H3 receptors regulate anxiety through GABA signaling.[J].Neur
on,114.(2026):1-16.DOI:10.1016/J.NEURON.2026.01.020.
- [47] Dai M ,Hong H ,Mao Y , et al.Thalamic GABA+ levels are negatively associated with neuropsychiatr
ic symptoms in patients with insomnia[J].Frontiers in Human Neuroscience,20.(2026):1750271-1750271.
DOI:10.3389/FNHUM.2026.1750271.
- [48] Kang Q J ,Delahanty A .Endoplasmic reticulum protein retention and disturbed proteostasis is a comm
on pathology for a subset of autism: evidence from mutations in GABAA receptors and GABA transp
orter 1[J].Frontiers in Molecular Neuroscience,18.(2026):1671331-1671331.DOI:10.3389/FNMOL.2025.16
71331.
- [49] Karbownik MS, Kręczyńska J, Wiktorowska-Owczarek A, et al. Decrease in salivary serotonin in resp
onse to probiotic supplementation with *Saccharomyces boulardii* in healthy volunteers under psychologi
cal stress: secondary analysis of a randomized, double-blind, placebo-controlled trial.[J].Frontiers in En
docrinology,2022,12:800023-800023.DOI:10.3389/FENDO.2021.800023.
- [50] Alom S ,Ali F ,Nongkhaw T F , et al.Glutathione at the Crossroads of Redox Homeostasis and Gluc
ose Metabolism in Diabetes Mellitus[J].Indian Journal of Clinical Biochemistry,2026,(prepublish):1-17.D
OI:10.1007/S12291-026-01409-5.
- [51] Lee J ,Lee J ,Tae H I , et al.Gut microbiota modulation of epigenetic target EHMT2: Lactaseibacill
us rhamnosus Fb7-311 regulated renal cell carcinoma apoptosis and metastasis[J].Experimental & Mole
cular Medicine,2026,(prepublish):1-16.DOI:10.1038/S12276-026-01659-6.

- [52] Ryu Y T ,Tae H I ,Han S T , et al.Epigenetic alterations of TP53INP1 by EHMT2 regulate the cell cycle in gastric cancer.[J].Experimental hematology & oncology,2024,13(1):86.DOI:10.1186/S40164-024-00554-Y.
- [53] Zhu X ,Lai X ,Deng J , et al.Naturally occurring antibodies against ASC reduced in Alzheimer's disease and alleviating AD-type pathology in APP/PS1 mouse.[J].Experimental neurology,2026,115728.DOI:10.1016/J.EXPNEUROL.2026.115728.
- [54] Li L X ,Lin H Z ,Chen R S , et al.Tiaogeng decoction improves mild cognitive impairment in menopausal APP/PS1 mice through the ERs/NF- κ b/AQP1 signaling pathway[J].Phytomedicine,138.(2025):156391-156391.DOI:10.1016/J.PHYMED.2025.156391.
- [55] Irwin MR, Vitiello MV. Implications of sleep disturbance and inflammation for Alzheimer's disease dementia.[J].The Lancet Neurology,2019,18(3):296-306.DOI:10.1016/S1474-4422(18)30450-2.
- [56]]Su H ,Zhou Q ,Guo L , et al.Lanostane triterpenoids from Ganoderma luteomarginatum and their cytotoxicity against four human cancer cell lines[J].Phytochemistry,2018,15689-95.DOI:10.1016/j.phytochem.2018.09.003
- [57] Gardini ES, Fiacco S, Mermone L, et al. Sleep and methylation of estrogen receptor genes, ESR1 and GPER, in healthy middle-aged and older women: findings from the women 40+ healthy aging study. [J].Nature and science of sleep,2020;12:525-536.DOI:10.2147/NSS.S256102.

11 Author biography

Hao Wang (born 1989), M.D., focusing on integrated traditional Chinese and Western medicine for the prevention and treatment of perimenopausal syndrome and other gynecological diseases. Affiliation: Wangjing Hospital, China Academy of Chinese Medical Sciences.

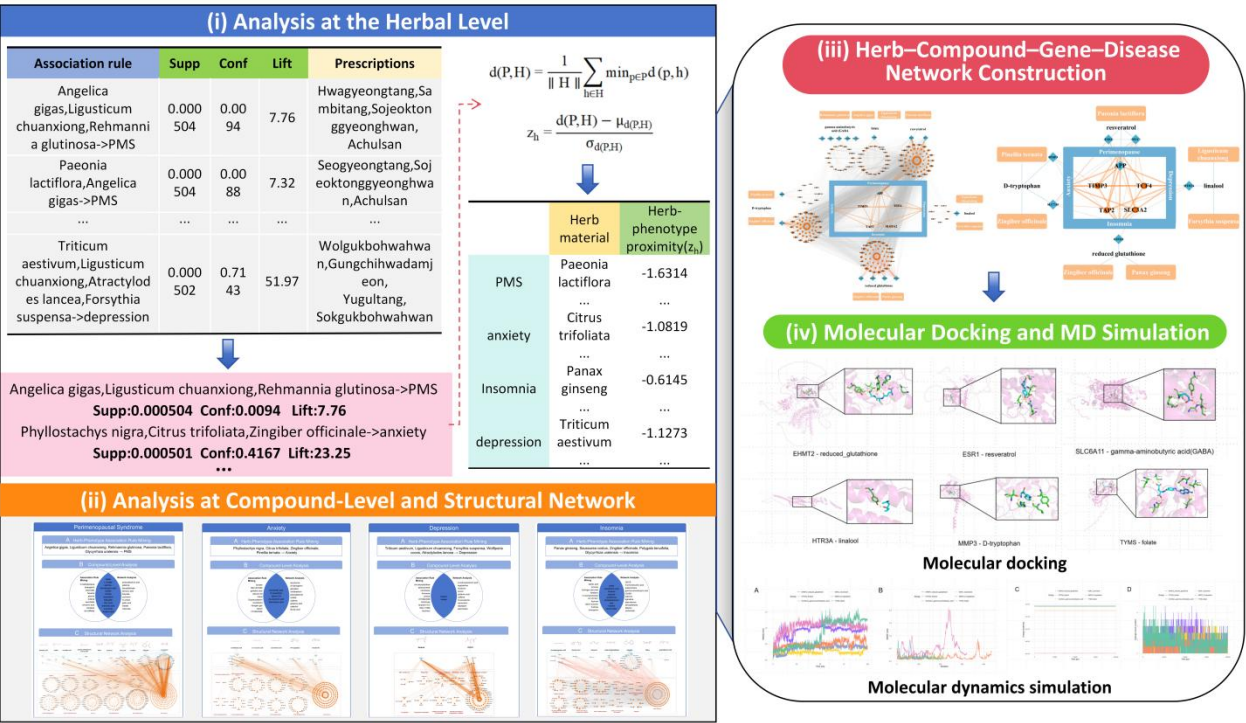


Figure 1. Overview of the integrated analysis pipeline. The schematic diagram illustrates the four main steps of the proposed framework (see Methods for details). Representative association rules, prescriptions, herb-phenotype proximity values, compound-target docking pairs, and molecular dynamics simulation results are shown for perimenopausal syndrome (PMS), anxiety, depression, and insomnia. The multi-layer network visualizes the interactions among herbs, compounds, core targets, and phenotypes.

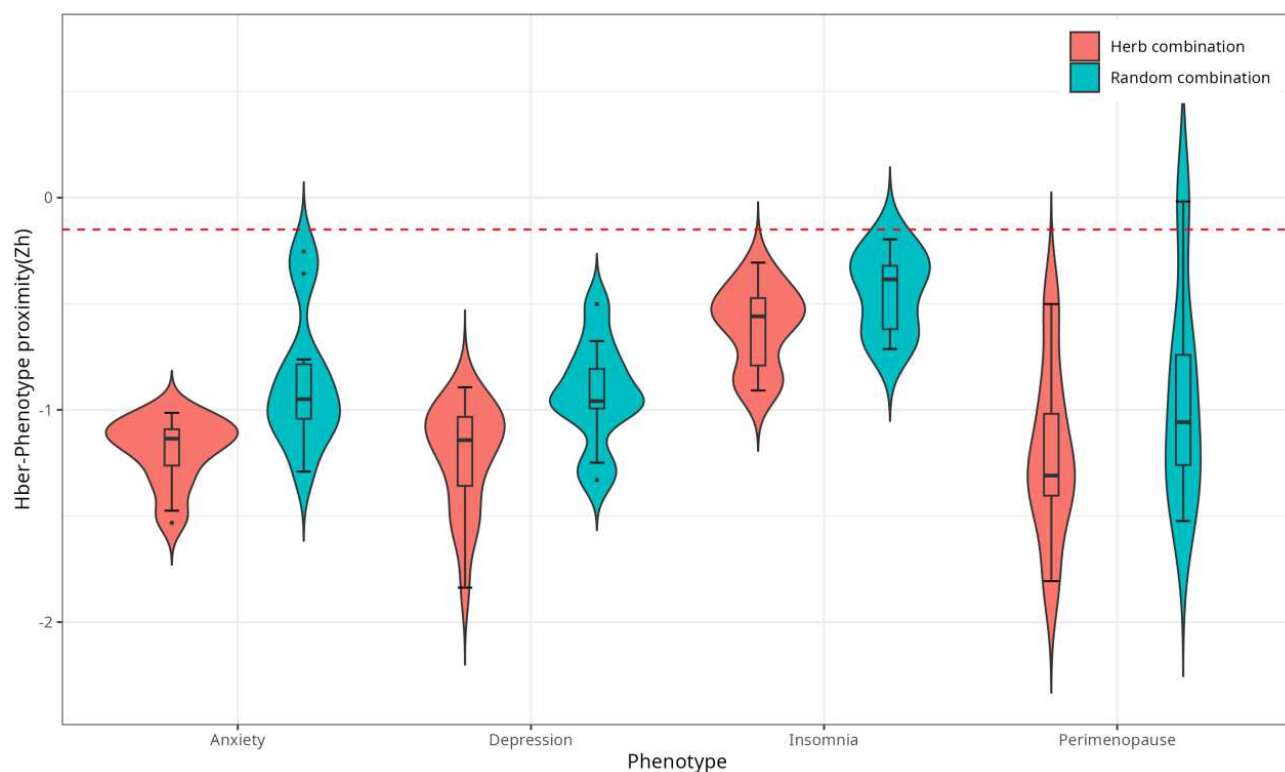


Fig 2 Extracts the compound-phenotype ARM and the compounds identified in the compound-phenotype network analysis. The x-axis represents diseases, the y-axis represents the proximity of the drug to the disease, the red part represents the drugs selected by the ARM rule, and the green part represents the randomly selected drug combinations.

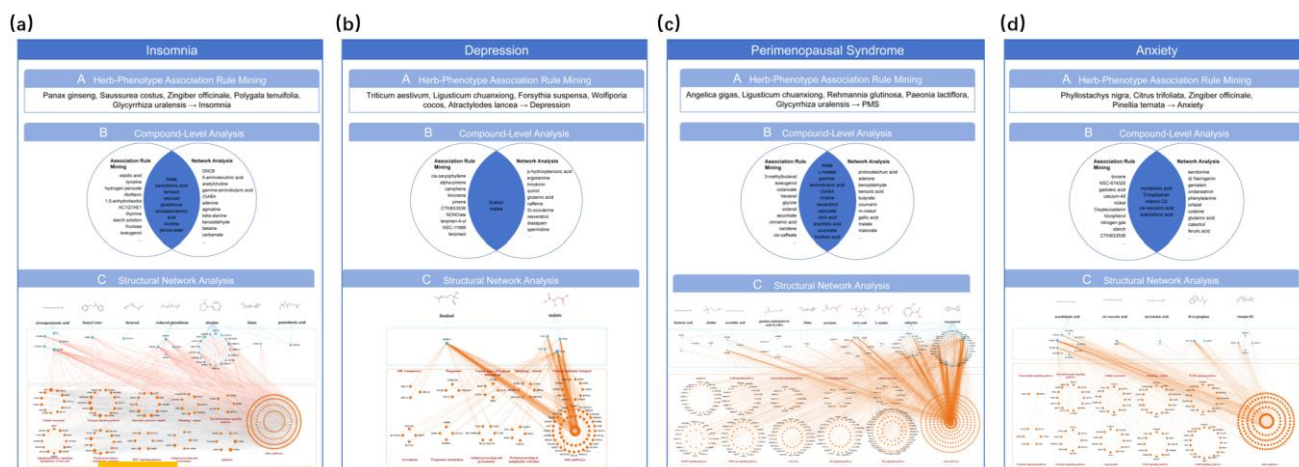


Fig 3. Integrated herb-compound-phenotype analysis for perimenopausal syndrome (PMS), anxiety, depression and insomnia. Note: (a), Analysis results of PMS, including herb-phenotype association rules, Venn diagram of candidate compounds from two mining strategies, and compound-target-disease structural network; (b), Corresponding analytical results for anxiety; (c), Matching analytical results for depression; (d), Parallel analytical results for insomnia. All panels share three modules: herb association rule mining, candidate compound intersection, and network visualization. Separate single diagrams for each disease are shown in Supplementary Figures S1–S4.

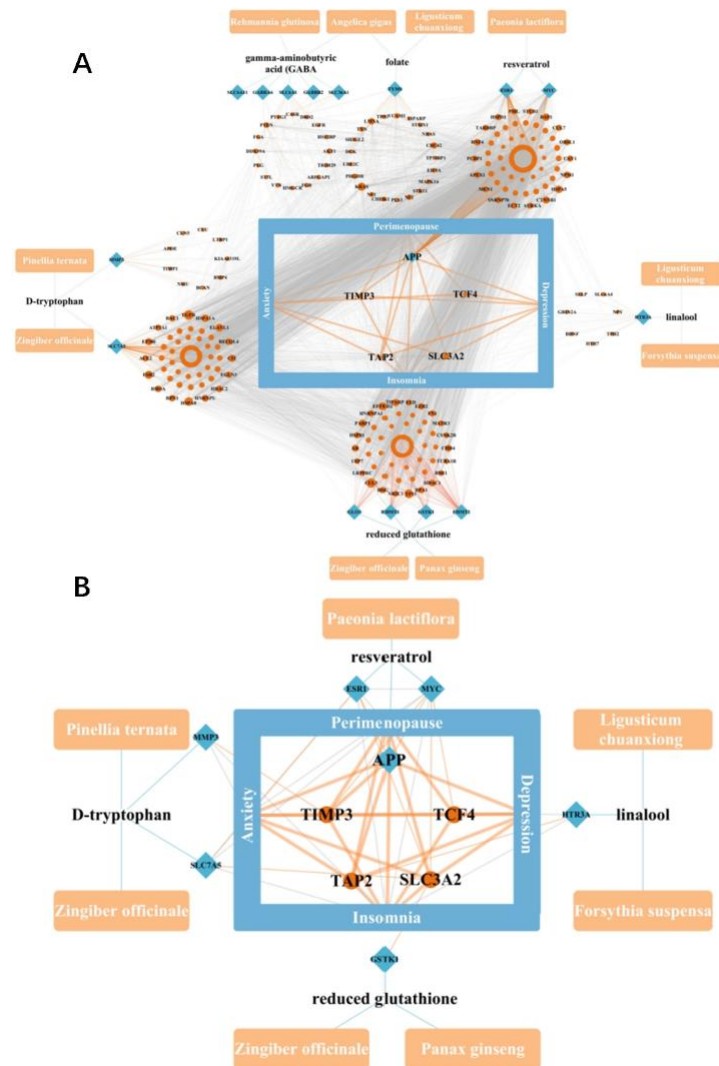


Figure 4. Herb - compound - core gene - disease network analysis. Note: A, Global herb-compound-target-disease network. Rectangles with orange background denote herbal medicines, connected nodes represent bioactive compounds, blue rhombuses stand for drug targets, orange circles indicate disease targets, and rectangles with blue background represent disorders. B, Simplified network focusing on shared hub genes across perimenopausal syndrome, anxiety, depression and insomnia. The core shared targets of all four phenotypes are displayed in the central region, among which APP acts as both a drug target and a disease target. Separate individual diagrams are provided in Supplementary Figures S5–S6.

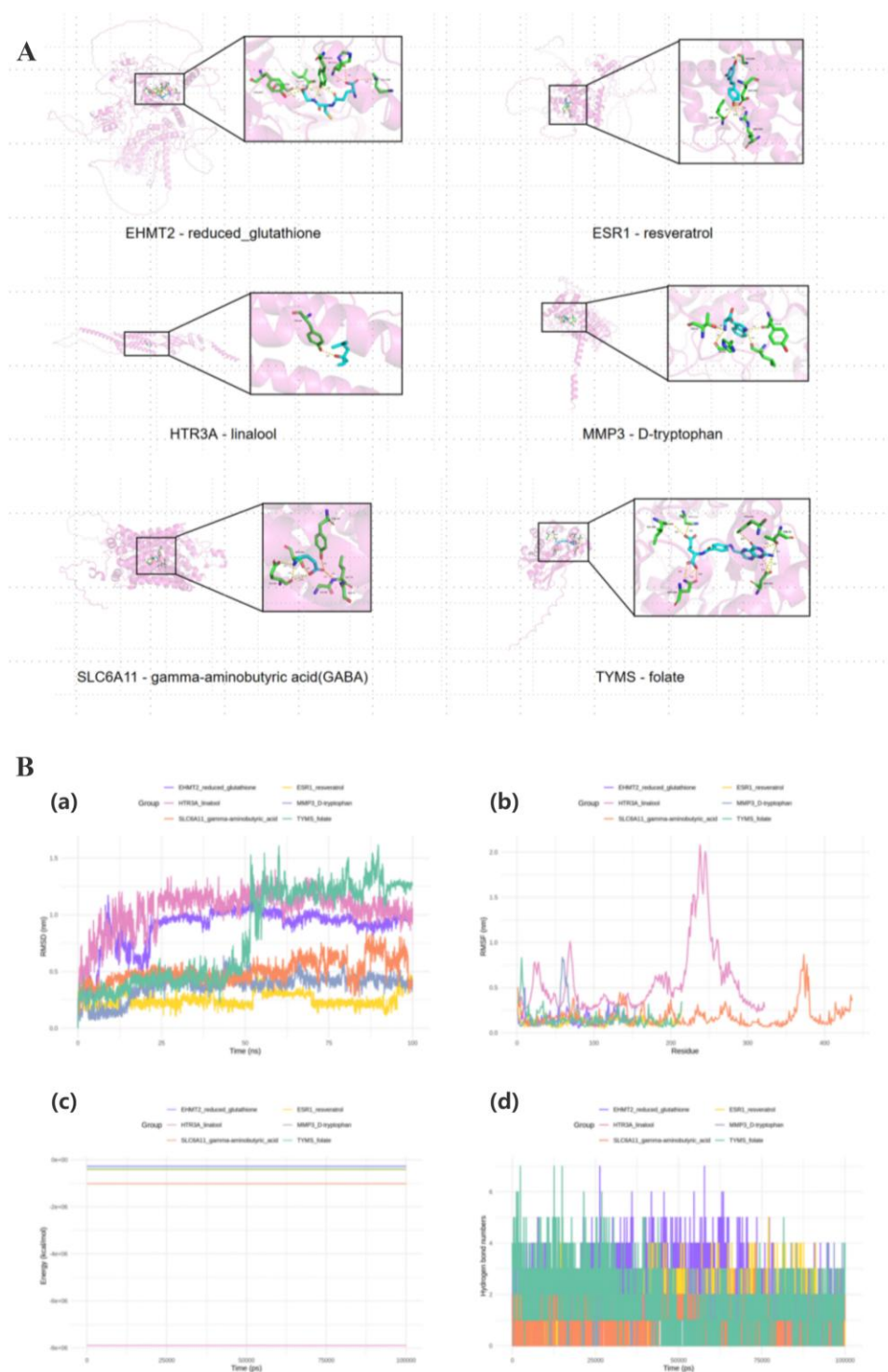


Fig. 5 Molecular docking visualization and molecular dynamics simulation. Note: B(a), Root-mean-square deviation (RMSD) as a function of time. The x-axis represents time, and the y-axis represents RMSD; smaller fluctuations indicate more stable protein–ligand binding. B(b), The x-axis represents the amino acid residue number, and the y-axis represents root-mean-square fluctuation (RMSF); larger fluctuations indicate greater flexibility of protein amino acid residues. B(c), Energy changes during molecular dynamics simulation. The x-axis represents time, and the y-axis represents total energy; lower energy and smaller fluctuations indicate greater stability. B(d), Changes in the number of hydrogen bonds formed between the compound and the protein over time. The x-axis represents time, and the y-axis represents the number of hydrogen bonds; a higher number of hydrogen bonds indicates more stable binding.

693
694
695

Tab 1 Inferred association rules and related prescriptions for phenotypes such as PMS, anxiety, depression, and insomnia

Association rule	Support	Confidence	Lift	Prescriptions
Angelica gigas,Ligusticum chuanxiong,Rehmannia glutinosa->perimenopause	0.000504	0.0094	7.76	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
Paeonia lactiflora,Angelica gigas->perimenopause	0.000504	0.0088	7.32	Seogyeeongtang,Sojeoktonggyeonghwan,Achulsan
Ligusticum chuanxiong,Rehmannia glutinosa->perimenopause	0.000504	0.0078	6.42	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
Angelica gigas,Rehmannia glutinosa->perimenopause	0.000504	0.0050	4.15	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
Angelica gigas,Ligusticum chuanxiong->perimenopause	0.000504	0.0049	4.04	Hwagyeongtang,Sambitang,Sojeoktonggyeong hwan,Achulsan
Angelica gigas,Glycyrrhiza uralensis->perimenopause	0.000504	0.0035	2.87	Hwagyeongtang,Seogyeeongtang,Sambitang,Achulsan
Phyllostachys nigra,Citrus trifoliata,Zingiber officinale->anxiety	0.000501	0.4167	23.25	Samhoondamtang,Gwibiondamtang,Cheongsimondamtang,Ondamtang,Gamiondamtang
Zingiber officinale,Citrus trifoliata,Phyllostachys nigra,Pinellia ternata->anxiety	0.000501	0.4167	23.25	Gwibiondamtang,Cheongsimondamtang,Ondamtang,Samhoondamtang,Gamiondamtang
Panax ginseng,Saussurea costus,Zingiber officinale,Polygala tenuifolia->Insomnia	0.000599	0.7500	36.84	Maehackgigitagyeongheombang1,Gamiyukbitang,Gamigwibitang,Gamiondamtang,Gwibitang,Yeongsingwibitang
Saussurea costus,Glycyrrhiza uralensis,Zingiber officinale,Polygala tenuifolia->Insomnia	0.000599	0.7500	36.84	Maehackgigitagyeongheombang1,Gamiyukbitang,Gamigwibitang,Gamiondamtang,Gwibitang,Yeongsingwibitang
Triticum aestivum,Ligusticum chuanxiong,Forsythia suspensa,Wolfiporia cocos->depression	0.000502	0.8333	60.63	Wolgukbohwhawan,Gungchihwadamjeon,Yugultang,Sokgukbohwhawan
Triticum aestivum,Ligusticum chuanxiong,Atractylodes lancea,Forsythia suspensa->depression	0.000502	0.7143	51.97	Wolgukbohwhawan,Gungchihwadamjeon,Yugultang,Sokgukbohwhawan

696

697

Tab 2 Analysis Results of Herbal Levels for PMS, Anxiety, Depression and Insomnia

	Herb material	Herb-phenotype proximity(z_h)
perimenopause	Paeonia lactiflora	-1.6314
	Glycyrrhiza uralensis	-1.1484
	Angelica gigas	-1.0369
	Rehmannia glutinosa	-0.7336
	Ligusticum chuanxiong	-1.3285
anxiety	Citrus trifoliata	-1.0819
	Zingiber officinale	-1.1426
	Pinellia ternata	-1.0222
	Phyllostachys nigra	-1.3660
	Panax ginseng	-0.6145
Insomnia	Glycyrrhiza uralensis	-0.4715
	Zingiber officinale	-0.3817
	Saussurea costus	-0.6732
	Polygala tenuifolia	-0.9548
	Triticum aestivum	-1.1273
depression	Atractylodes lancea	-0.7372
	Wolfiporia cocos	-1.0334
	Forsythia suspensa	-1.0827
	Ligusticum chuanxiong	-1.0455

698

699

700
701

Tab 3 Structural Network Analysis of PMS, Anxiety, Depression, and Insomnia and Search Results of Biological Pathways

disease	compound	herb target
perimenopause	resveratrol	MYC
		ESR1
		APP
		GABBR2
		SLC6A1
	gamma-aminobutyric acid (GABA)	SLC6A11
		SLC36A1
		GABRA6
		TYMS
		ABCB11
anxiety	D-tryptophan	MMP3
		SLC7A5
	reduced glutathione	GLO1
		GSTK1
Insomnia	folate	EHMT1
		EHMT2
	folate	TYMS
		ABCB11
depression	linalool	HTR3A

702

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigS1S6TabS1.docx](#)