Journal of Pharmaceutical Analysis 14 (2024) 100914

Contents lists available at ScienceDirect

Journal of Pharmaceutical Analysis

journal homepage: www.elsevier.com/locate/jpa

Original article

Model informed precision medicine of Chinese herbal medicines formulas—A multi-scale mechanistic intelligent model

Yuanyuan Qian [a, c, d, 1], Xiting Wang [b, 1], Lulu Cai [a, c, d, f, 1], Jiangxue Han [a, c, d, f, 1], Zhu Huang [a, c, d], Yahui Lou [a, c, d], Bingyue Zhang [a, c, d], Yanjie Wang [a, c, d], Xiaoning Sun [e], Yan Zhang [e], Aisong Zhu [a, c, d, *]

[a] Zhejiang Key Laboratory of Blood-Stasis-Toxin Syndrome, Zhejiang Chinese Medical University, Hangzhou, 310000, China
[b] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China
[c] Zhejiang Engineering Research Center for "Preventive Treatment" Smart Health of Traditional Chinese Medicine, Zhejiang Chinese Medical University, Hangzhou, 310000, China
[d] College of Basic Medical Sciences, Zhejiang Chinese Medical University, Hangzhou, 310000, China
[e] Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang, 110032, China
[f] Academy of Chinese Medical Science, Zhejiang Chinese Medical University, Hangzhou, 310000, China

ARTICLE INFO

Article history:
Received 20 August 2023
Received in revised form
21 November 2023
Accepted 7 December 2023
Available online 9 December 2023

Keywords:
Chinese herbal medicine formulas
Precision medicine
Mathematical modeling
Systems biology
Coronary heart disease
Depression
Ischemia-reperfusion
Inflammation

ABSTRACT

Recent trends suggest that Chinese herbal medicine formulas (CHM formulas) are promising treatments for complex diseases. To characterize the precise syndromes, precise diseases and precise targets of the precise targets between complex diseases and CHM formulas, we developed an artificial intelligence-based quantitative predictive algorithm (DeepTCM). DeepTCM has gone through multilevel model calibration and validation against a comprehensive set of herb and disease data so that it accurately captures the complex cellular signaling, molecular and theoretical levels of traditional Chinese medicine (TCM). As an example, our model simulated the optimal CHM formulas for the treatment of coronary heart disease (CHD) with depression, and through model sensitivity analysis, we calculated the balanced scoring of the formulas. Furthermore, we constructed a biological knowledge graph representing interactions by associating herb-target and gene-disease interactions. Finally, we experimentally confirmed the therapeutic effect and pharmacological mechanism of a novel model-predicted intervention in humans and mice. This novel multiscale model opened up a new avenue to combine "disease syndrome" and "macro micro" system modeling to facilitate translational research in CHM formulas.

1. Introduction

Traditional Chinese medicine (TCM) has a history of treating diseases for over two thousand years, and an increasing number of Asian and Western countries have recognized its efficacy. Chinese herbal medicine formulas (CHM formulas) have long been used for the treatment of complex diseases and their complications, such as angina pectoris, myocardial infarction and stroke [1,2]. However, there are two problems that limit the development of CHM formulas to achieve precision medicine [1]. On the one hand, TCM relies heavily on rich experience; for example, different doctors prescribe different CHM formulas for the same disease, so there is a lack of efficacy evaluation criteria. On the other hand, many CHM formulas, which contain 2 or more herbs and have different compatibility rules, lack specific pharmacological mechanisms and active ingredients, which pose challenges for Western researchers or physicians. Therefore, there is an urgent need to establish a judgment evaluation model to improve the interpretability of CHM formulas.

With the development of deep learning (DL) and TCM network pharmacology, researchers have realized the use of artificial intelligence (AI) methods to mine the knowledge of CHM formulas and the use of databases to speculate the targets and pathways of herbs. A multistage analysis method that integrates propensity case matching, complex network analysis, and herb set enrichment analysis was proposed to identify effective herb prescriptions for particular diseases (e.g., insomnia) [3]. A recent study proposed a recommendation system (FordNet) by fusing phenotype and molecular information, which promoted the shift of TCM research

* Corresponding author. College of Basic Medical Sciences, Zhejiang Chinese Medical University, Hangzhou, 310000, China.
E-mail address: aisong.zhu@zcmu.edu.cn (A. Zhu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jpha.2023.12.004

patterns from “experience based, macro” to “data based, macro micro combined” patterns [4]. However, there is still no adequate consideration of the unique characteristics of TCM theory, principles, formulas, and Chinese medicinal substances and a lack of systematic models to realize the possibility of precision medicine with CHM formulas. To achieve the goal of aim, we should solve three problems. First, we need to precisely recognize CHM formula syndromes. Second, we need to precisely select different diseases for which CHM formulas are suitable. Third, we need to confirm the precise targets of complex diseases for CHM formulas.

The diagnostic and therapeutic methods of TCM are based on the differentiation of syndromes and the use of CHM formulas. Therefore, taking sufficient data from CHM formulas into account will help to comprehensively and deeply understand the mechanism of drug synergy to effectively identify clinically relevant herbal combinations [5–7]. In addition, given that most diseases are associated with characteristic changes in gene expression profiles, such changes are used as indicators reflecting the underlying mechanisms of diseases [8–13]. We built a DL framework for the integration of CHM formulas and Western medicine, which combines “disease syndrome” and “macro micro”, DeepTCM. In this algorithm, we developed an efficacy generation model based on the unified language model (ULM) that enables DeepTCM to infer the macroscopic efficacy of CHM formulas and compare it with human knowledge. In addition, we constructed a biological knowledge graph representing interactions by associating herb targets and genes. Through 2 scenarios, we developed a scoring function that combines the correlation between CHM formula macro efficacy and disease score to quantify the regulatory link between CHM formulas and complex diseases (the precise syndromes, diseases and targets).

Using DeepTCM, we screened for potential therapeutic interventions and the most relevant targets for CHM formula-complex diseases (e.g., coronary heart disease (CHD) and CHD with depression). The results revealed that DeepTCM attained 78.1% agreement scores that were manually corrected in the tenfold validation. Furthermore, we evaluated the relative efficacies and mechanisms of Huangjin shuangshen jiawei (HJSSJW), an empirical CHM formula, proposed by DeepTCM analysis with diseases in terms of improvement in cardiac function and depression recovery according to the disease score. Finally, we experimentally validated HJSSJW by demonstrating its ability to improve cardiac function and depression recovery in mice and patients. In short, this study presents DeepTCM as the first endeavor to use a quantitative predictive DL-based algorithm to simulate and estimate the scientificity and rationality of CHM formulas in preclinical settings. It can serve as a useful platform to provide novel systems-level insights into precision medicine for CHM formulas and complex diseases.

2. Methods

2.1. Construction and evaluation of the DeepTCM algorithm

The DeepTCM algorithm consists of three parts: a CHM formula efficacy generation model (CHM-EG), CHM formulas and a disease ontology score model (CHM-DOS), and a CHM formula-disease score function.

2.1.1. Dataset collection and processing

We constructed a large-scale TCM corpus (TCM-CPS) derived from TCM textbooks, guidelines and works (Table S1). A high-quality CHM formula dataset was created to facilitate mapping Chinese herb composition to the efficacy description of CHM formulas. The benchmark dataset was derived from highly recognized industry-standard data, including the Chinese Pharmacopoeia (2020 Edition), National Essential Medicines Catalog (2018 Edition), Compilation of National Standard for Traditional Chinese Medicines, National Planning Textbook for Colleges and Universities of TCM (Tenth Edition), National Basic Medical Insurance Medicine List (2021 Edition), and other clinical medical guidelines.

We also integrated bioinformatics and systems pharmacology approaches to observe the underlying molecular mechanisms of CHM formulas. The target set for a given CHM formula was collected from HERB [5], a comprehensive TCM database. The expression dataset was obtained from the Gene Expression Omnibus (GEO) repository with the accession number (http://www.ncbi.nlm.nih.gov/geo).

2.1.2. Construction of the CHM-EG

First, we applied the word embedding method to preprocess the data. Given a sentence of TCM knowledge as an input sequence $k = \{k_1, k_2, \ldots, k_n\}$, the tokenization method was performed to translate k to $x = \{x_0, x_1, \ldots, x_n, x_{n+1}\}$. Here, $\{x_1, \ldots, x_n\}$ is the corresponding set of tokenized k according to tokenization vocabulary, and x_0 is defined as an added special start-of-sequence (SOS) token; therefore x_{n+1} represents a special end-of-sequence (EOS) token. Subsequently, three types of word embedding were generated for each token x_i according to x, including token, position, and segment embedding. The Transformer was applied as a backbone network to encode contextual information. The sequence x was first packed together into $\boldsymbol{H}^0 = [x_0, x_1, \ldots, x_n, x_{n+1}]$. Then, an L-layer transformer was used to encode the input:

$$\boldsymbol{H}^l = Transformer_l\left(\boldsymbol{H}^{l-1}\right),$$

where $l \in [1, L]$, and $\boldsymbol{H}^L = [h_1^L, \ldots, h_{|x|}^L]$. Here, the hidden vector h_i^L was used as the contextualized representation of the input token x.

Then, we applied bidirectional encoder representations from transformers (BERT) as the basic pretrained language representation model for the downstream TCM text understanding and generation tasks. Here, the BERT-based Chinese pretrained language model, NEZHA-base [6], was applied to initialize the model weight. Thus, the base model used a similar architecture with NEZHA-base, which employed a 12-layer transformer with 768 hidden sizes and 12 attention heads. The activation function is GeLU. The maximum length of the input sequence was 512. The token masking probability was 15%. The learning rate was set to 1e−5, and Adam was applied as an optimization algorithm. The basic model was optimized by minimizing the cross-entropy loss computed using the predicted and original tokens.

Finally, we defined CHM formula efficacy prediction as a sequence-to-sequence generation task, namely, efficacy generation, considering that the doctor infers the efficacy of the CHM formulas in a knowledge-driven generation way. Next, we constructed the CHM-EG according to the UNILM approach [7]. The self-attention masks technique was incorporated into the CHM-EG models, thus enabling the original basic model to perform sequence generation. For the l-th Transformer layer, the output of a self-attention head $\mathbf{A}_l$ is computed via:

$$\begin{cases} \mathbf{Q} = \mathbf{H}^{l-1}\mathbf{W}_l^Q, \mathbf{K} = \mathbf{H}^{l-1}\mathbf{W}_l^K \\ \mathbf{A}_l = softmax\left(\dfrac{QK^T}{\sqrt{d_k}} + \mathbf{M}\right)\left(\mathbf{H}^{l-1}\mathbf{V}_l\right) \end{cases},$$

where $\mathbf{H}^{l-1} \in \mathbb{R}^{|x| \times d_h}$ was linearly projected to a triple of queries, keys and values using parameter matrices $\mathbf{W}_l^Q, \mathbf{W}_l^K, \mathbf{V}_l \in \mathbb{R}^{d_h \times d_k}$, respectively, and the mask matrix $\mathbf{M} \in \mathbb{R}^{|x| \times |x|}$ determines whether a pair of tokens could be attended to each other. The mask matrices $\mathbf{M}$ were

used to control what context a token can attend to when computing its contextualized representation. Correspondingly, a mask cross-entropy loss method was also proposed for model loss calculation. The specific fine-tuning procedure also refers to UNILM, which likewise used self-attention masks as pretraining, with the objective of maximizing the likelihood of masked tokens in a given context. Ultimately, we used a beam search with a beam size of one to generate an efficacy description from the CHM formulas.

2.1.3. Construction of CHM-DOS

For the disease ontology score CHM-DOS algorithm, all the targets of the CHM formulas were initially retrieved from the HERB database, and disease enrichment analysis was performed on the targets. DisGeNet [8] was applied as the disease ontology database in this study, and the enrichment score after normalization was taken as the DOS score.

2.1.4. Construction of the CHM-disease score function

The CHM-disease score function was proposed to balance the results from the above approaches, including natural language inference and network analysis, formulated as follows:

$$\text{CHM}-\text{disease score}=\frac{1}{N}\sum_{i=0}^{N}\rho D\left(1+\frac{C}{D}+\mu C\right),$$

where D is the DOS score, C is the semantic consistency (SC) score, ρ and μ are weighted values, and N is the number of test samples (CHM-disease pairs). Here, ρ and μ were set as 0.5 and 2, respectively.

2.1.5. Construction of minimal semantic unit segmentation operation

Thereafter, we disassembled the entire CHM formula efficacy description into a single set of nonseparable subefficacy. We then proposed a SC score to evaluate the hit ratio (fraction of coverage) between the CHM formula efficacy generated by CHM-EG and the true efficacy. In addition, we assessed the CHM-EG performance in two dimensions, namely, semantic ambiguity (SA) and the semantic deficiency (SD) between true and generated CHM formula efficacy. The SC score and SD score were formulated as follows:

$$\begin{cases} SC\ score=\dfrac{P_E\cap T_E}{T_E} \\ SD\ score=\dfrac{T_E-(P_E\cap T_E)}{T_E} \\ SA\ score=\dfrac{P_E-(P_E\cap T_E)}{P_E} \end{cases},$$

where $\{P_{e_1},P_{e_2}\ldots P_{e_m}\}\in P_E$, $\{T_{e_1},T_{e_2}\ldots T_{e_n}\}\in T_E$, P_E is the set of subefficacy generated by CHM-EG and T_E is the set of true (original) subefficacy of particular CHM formulas. Then, manual correction processes were performed when evaluating the performance of CHM-EG. The average score on all 10-fold validation datasets was taken as the model performance score for the CHM-EG.

2.2. Bioinformatics analysis

The upset algorithm was constructed to detect the shared and specific targets among Chinese herbs in CHM formulas. Then, the resultant network was visualized using Cytoscape (https://cytoscape.org/), in which each node represents one herb or target. After eliminating the overlapping targets from the target set, the targets were uploaded to Metascape [9] and converted into human Entrez gene identities (IDs). Then, the enrichment analysis of DisGeNet was performed.

The raw gene expression data of CHD and depression were extracted from GSE98583 and GSE98793, respectively. By using the Bioconductor package Limma [10], differentially expressed genes (DEGs) between the two groups were identified. Genes with |log2-fold-change (FC)| > 1 and $P < 0.05$ were regarded as statistically significant DEGs. Subsequently, based on the online analysis service of the STRING database with a threshold of interaction score >0.4, the protein protein interaction (PPI) network was constructed. Then, a plug-in Cytoscape software, CytoHubba [11], was used to screen the top key genes in PPI networks through the maximal clique centrality (MCC) methods.

2.3. Bioinformatics analysis of CIBERSORT and WGCNA

Having performed the CIBERSORT [12] algorithm on normalized gene expression matrices, we aimed to evaluate the relative proportions of 22 types of infiltrating immune cells. Significant immune cells between disease and healthy control samples were screened with the threshold Wilcoxon test at $P < 0.05$ and plotted by R. Subsequently, by applying the weighted gene coexpression network analysis (WGCNA) [13] package in R, coexpression networks were built to identify co-expression modules. The median absolute deviation (MAD) algorithm was applied to identify the top 10,000 genes with high expression variance for WGCNA. The “pickSoftThreshold” algorithm was used to determine the “soft” thresholding power (β) in constructing the biologically significant scale-free network. We further calculated the gene significance and module membership (MM) to relate modules to clinical traits.

2.4. Preparation method for CHM formulas (HJSSJW and ShexiangBaoxin Wan)

Sanjiu Pharmaceutical Co., Ltd. (Shenzhen, China) rationally designed the process route, production equipment and quality control methods, and established a set of process research modes in line with the characteristics of compound preparations of HJSSJW granule. In the clinical trial, the dosage of HJSSJW was 2.63 g/kg/day. In the mouse experiments, HJSSJW contained a low (10.74 g/kg/day), middle (21.47 g/kg/day, which matches the recommended dose in the clinic) and high dose (42.94 g/kg/day). The blank control group and the model group were given equal doses of sterile drinking water. The HJSSJW granule was dissolved in sterile water, and ultrasonic treatment was conducted for 40 min. Before high-performance liquid chromatography coupled with linear ion trap-Orbitrap tandem mass spectrometry (UHPLC-LTQ-Orbitrap MS) detection and analysis, the drug was filtered through a 0.22 μM filter membrane.

With a Venusil MP C_{18} (250 mm × 4.6 mm, 5 μm) (Aglient, Beijing, China) chromatographic column, acetonitrile (A)−water (B) was used as the mobile phase, gradient elution (0−40 min, 0%−30% A; 40−50 min, 30%−80% A), flow rate was 1.0 mL/min, detection wavelength was 210 nm, column temperature was 30 °C, injection volume was 10 μL, electric spray ion source (ESI+), high resolution detection (FTMS), resolution of 60,000 FWHM and data dependence scanning (IDA) were used to collect dat, ion source temperature was 350 °C, sheath flow rate was 35 arb, auxiliary air velocity was 10 arb, blowback flow rate was 0 arb, spray voltage was 4.0 kV, capillary temperature was 300 °C, data acquisition mode was positive ion mode, and scanning range of primary mass spectrometry parention was *m/z* 100−1,000. HCD mode was adopted for secondary mass spectrometry, dynamic background subtraction was turned on, and collision voltage was 45 V.

By consulting the literature in the database, the chemical components of the whole prescription and each single medicine were summarized, and the components of each single medicine were summarized, including molecular formula, molecular structure formula, relative molecular weight, molecular ions, source of medicinal materials, and Chinese and English names. We established a database dedicated to HJSSJW granules and compared and analyzed the data collected by HPLC-LTQ/Orbitrap-MS/MS to quickly identify known compounds.

Shanghai Hutchison Pharmaceuticals Co., Ltd. (Shanghai, China) producted ShexiangBaoxin Wan (SXBXW). In the mouse experiments, the dosage of SXBXW was 253.4 mg/kg/day. The blank control group and the model group were given equal doses of sterile drinking water. The SXBXW was dissolved in sterile water and subjected to ultrasonic treatment for 40 min.

2.5. Human serum samples

We designed a multicenter, randomized, double-blinded, placebo-controlled clinical trial to explore the safety and efficacy of HJSSJW granules in patients with CHD after percutaneous coronary intervention (PCI). We chose participants who had depression between 2020 and 2021 in the Affiliated Hospital of Liaoning University of Traditional Chinese Medicine. All protocols and procedures followed the rules of Standard Protocol Items Recommendations for Interventional Trials (SPIRIT) and were approved by the Review Board (Clinical Trial Registration: https://www.chictr.org.cn/; identifier: ChiCTR2000037233.). A total of 318 participants were enrolled in the study and randomly assigned to the HJSSJW group or the placebo group in a 1:1 ratio for 12 weeks of treatment and 24 weeks of follow-up. All patients underwent standard Western medicine treatment. The clinical characteristics of the individuals mentioned in this study are shown in Table S2. After 12 weeks, 40 participants' serum samples were collected after centrifugation at 3,000 rpm for 15 min at 4 °C and kept at −80 °C before analysis. All samples were only used for research objectives, and informed consent forms were signed by donors or their families per the regulations in China overseeing the collection of informed consent documents. All procedures were performed in compliance with relevant laws and institutional guidelines and the appropriate institutional committees approved them.

2.6. Animal preparation

All mouse experiments and protocols were performed in accordance with the requirements of the Animal Care and Use Committee of Zhejiang Chinese Medical University and the Institutional Animal Care and Use Committee at the Center for Animal Experiments of Zhejiang Chinese Medical University (Approval No.: 20210802-09). Animals were under humane care with avoidance of animal abuse. C57BL/6J mice and ApoE$^{-/-}$ mice were purchased from the Shanghai SLAC Laboratory Animal Co., Ltd (Shanghai, China). Mice were maintained on a 12 h light/dark cycle from 8 am to 8 pm at a comfortable temperature (25 °C) with water and food. Animals were allocated to each treatment group on a random basis, which was carried out in accordance with standard laboratory procedures of randomization and blinding. There were 4–5 mice in each group. The mice were anesthetized by isoflurane followed by cervical dislocation. Mouse body and heart tissues were weighed, and the tissues were instantly frozen in liquid nitrogen. Blood was collected in a BD Microtainer (Franklin Lakes, NJ, USA), and serum was collected after centrifugation at 3,000 rpm for 15 min at 4 °C. All procedures were performed in compliance with relevant laws and institutional guidelines, and the appropriate institutional committees approved them.

2.7. Mouse models of myocardial ischemia-reperfusion (I/R)

The I/R model was induced in C57BL/6J male mice aged 8 weeks (24–25 g). The mice that underwent a left lateral thoracotomy were anesthetized by intraperitoneal injection of pentobarbital sodium and connected to Micro-vent 1 (Hallowell, Pittsfield, MA, USA). The hearts of the mice were exposed to the air. We used a 7-0 Prolene suture to ligate the left anterior descending (LAD) coronary artery. The occlusion was removed after 30 min to allow myocardial reperfusion. After 30 min of myocardial infarction, the slipknot was released, and myocardial reperfusion commenced for 3 h, 6 h, 12 h, 24 h, 3 day and 7 day, respectively. The sham group underwent the same procedures without LAD ligation. To evaluate the efficacy of HJSSJW, mice were given a gavage of HJSSJW in the early 1 week. The cardiac functional test was subjected to echocardiographic examination one day before sacrifice using an Ultrahigh resolution small animal ultrasound imaging system vevo 1,100 (VisualSonics, Toronto, Canada) with a 30 MHz transducer following the manufacturer's instructions.

A “multi-hit” MI/R model was induced in ApoE$^{-/-}$ male mice aged 8 weeks (24–28 g). Mice were fed a normal diet (ND; Trophic, Nantong, China) or high-fat diet (HFD; Trophic, Nantong, China) for 12 weeks. To evaluate the efficacy of HJSSJW, mice were given a gavage of HJSSJW in the last 3 weeks. After 9 weeks of HFD, mice were subjected to 6 stressors at random for 3 weeks (two stressors per day from Day 1 to Day 21; the stressors included bedding deprivation overnight, restraint stress (1 h), food and water deprivation (12 h), wet bedding (12 h), and tail clamp (5 min)). After behavioral tests, the mice were subjected to a myocardial I/R model in accordance with a previous study.

2.8. Behavioral tests

2.8.1. Forced swimming test (FST)

The mice were gently placed in a clear glass cylinder (height: 19 cm; diameter: 14 cm) containing with fresh water (height: 14 cm from the bottom; temperature: 23 ± 2 °C). The test lasted 6 min and was videotaped. The immobility time in the last 4 min was calculated by three experienced observers. Despair behavior involved floating or only slight movement to maintain balance.

2.8.2. Tail suspension test (TST)

The mice were suspended by tail on a station (50 cm above the floor) for 6 min with a short adhesive tape. The tape was placed approximately 1 cm from the tip of the tail and connected to a box. The processes were videotaped, and the immobility time in the last 4 min was scored by three experienced observers. The cumulative immobility time was recorded.

2.8.3. Sucrose preference test (SPT)

The mice were provided and acclimatized to two identical drinking bottles (one bottle contained water, and the other contained 1.5% sucrose solution) for 48 h. Each bottle was switched at 24 h. For the next 24 h, the mice were deprived of drinking water. The SPT was performed by providing the mice once again with water and 1.5% sucrose solution during which the position of each bottle was switched after 12 h. The weight of the water and the sucrose solution were measured. The ratio of sucrose solution was calculated as 100% × ((weight (sucrose solution))/(weight (sucrose solution + water (weight))).

2.8.4. Open field test (OFT)

The mice were gently placed in one corner of a square Plexiglas box (50 cm × 50 cm × 50 cm) separately. A camera recorded the routes of mice for 6 min, and the Etho Vision video tracking system smart 3.0 (RWD, Shenzhen, China) analyzed the total distances traveled during the test. The box was cleaned with 75% ethanol after each test to remove olfactory cues.

2.9. RNA preparation for quantitative real-time PCR (RT-PCR)

Total RNA was extracted from heart and hippocampal tissues using a SteadyPure Universal RNA Extraction Kit (Agbio, Changsha, China) according to the manufacturer's instructions. cDNA was synthesized by the Evo M-MLV RT Premix for qPCR Kit (Agbio). RT-qPCRs were performed using UltraSYBR Mixture (CWBio, Taizhou, China) on a LightCycler 384 Real-Time System (Roche, Basle, Switzerland). The primers are shown in Table S3. Gene expression was calculated by the 2−ΔΔCT method and normalized to β-actin.

2.10. Western blot analysis

The heart and hippocampus samples were lysed in radioimmunoprecipitation assay (RIPA) lysis buffer with protease inhibitor cocktail (Solarbio, Beijing, China) and phosphatase inhibitors (Sigma, Shanghai, China) by a tissue homogenizer. The lysates were collected and centrifuged at 12,000 *g* for 15 min. The protein concentration was determined using a BCA protein assay kit (Beyotime, Shanghai, China). Protein was denatured for 5 min at 100 °C with 5 × Laemmli sample buffer. Equal amounts of protein (10 μg) were separated by 8%−12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto polyvinylidene fluoride (PVDF) membranes. The membranes were incubated with 5% skim milk powder buffer for 60 min at room temperature and incubated with primary antibodies overnight at 4 °C. After incubation with primary antibodies, the membranes were incubated for 1 h with the corresponding secondary antibodies. Chemiluminescence was detected with ECL Western blot detection kits (Biosharp, Hefei, China) with a Bio-Rad image system (Hercules, CA, USA). The intensity of the bands was analyzed by ImageJ software. The primary antibodies are shown in Table S4. The secondary antibodies were as follows: horseradish peroxidase (HRP)-conjugated secondary anti-rabbit (1:5000, *V/V*) and HRP-conjugated secondary anti-mouse (1:5000, *V/V*).

2.11. Histological and immunofluorescence staining

Mice were perfused with saline for 3 min, and heart, liver and lung tissues were collected. The tissues were fixed in 4% paraformaldehyde at room temperature and embedded in paraffin after one week. For staining, the tissues were serially sectioned at 4 μm thickness. Standard hematoxylin−eosin staining (H&E), Masson's trichrome and Sirius red staining were performed on these sections to evaluate histopathology and collagen deposition, respectively. To assess lipid droplets in the liver, liver samples were embedded in optimal cutting temperature (OCT) compound. The sections of 8 μm thickness were stained with oil red O staining. Histopathological images were acquired with a NanoZoomer2.0RS digital pathology grade slide scanner (Hamamatsu, Shizuoka, Japan). For immunofluorescence analysis, the sections were blocked with 10% bovine serum albumin at room temperature for 1 h and incubated with primary antibodies, including anti-MMP9 (1:200, *V/V*) or anti-P65 (1:200, *V/V*), overnight, followed by incubation with the corresponding secondary antibody (1:500, *V/V*) for 1 h. Finally, the nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) (Sigma, Shanghai, China). The immunofluorescence images were captured on an OLYMPUS VS120 Virtual Slide Microscope (Olympus, Tokyo, Japan).

2.12. Biochemical and ELISA

The concentrations of creatine kinase (CK), total cholesterol (TC), triglyceride (TG), low-density lipoprotein (LDL) and high-density lipoprotein (HDL) in the study were detected by diagnostic reagent kits (Nanjing Jiancheng Bioengineering Inc., Nanjing, China) following the guidance of the manufacturer. The serum concentrations of troponin T (cTnI) and 5-hydroxytryptamine (5-HT) were examined via enzyme-linked immunosorbent assay (ELISA) following the manufacturer's instructions.

2.13. Statistical analysis

Prism software (GraphPad Prism 8) was used for statistical analysis. Two-tailed Student's *t*-test was used for comparisons between the two groups. $P < 0.05$ was considered statistically significant. All the data were exported into Adobe Illustrator CS5 for the preparation of figures and are shown as the mean ± standard error of the mean (SEM).

3. Results

3.1. Construction of the DeepTCM algorithm

The DeepTCM algorithm consists of three submodules. The first submodule was the text generation model based on CHM-EG (Fig. 1A), which generated the CHM formula efficacy for a given list of Chinese herbs. The CHM-DOS applied bioinformatics methods to perform disease enrichment formula analysis for the targets of CHM formulas (Fig. 1B), which was then used as the correlation score between the CHM formulas and the diseases. The third sub-module, CHM-score, was a balanced scoring function to integrate the scores of CHM-EG and CHM-DOS (Fig. 1C). The code involved in the DeepTCM algorithm is available at the https://github.com/AITCM/DeepTCM.

3.2. Construction of the benchmark dataset

In this study, we preliminarily collected 6,011 CHM formulas, normalized the name of the Chinese herb according to the Chinese pharmacopoeia, and normalized the synonyms of the CHM formula efficacy. CHM formulas containing herbal extracts and compounds were removed, deduplication operations were performed on data with the same Chinese herb and efficacy information, and 2,184 entries were retained as the dataset for the task. We conducted statistical analysis on the benchmark dataset. The frequency analysis revealed that the majority of CHM formulas (94.4%) contained fewer than 20 Chinese herbs, with those containing fewer herbs being predominant (accounting for 64.9%) (Fig. 2A). We used the Jaccard coefficient to analyze the similarity between prescriptions, and the results showed that the resemblance was concentrated in the region below 0.3, accounting for 30% (Fig. 2B). On the other hand, the majority of prescriptions (83.2%) were concentrated in the area below 15, and the similarity was mainly focused on sites below 0.4 (Figs. 2C and D). The efficacy of CHFs is primarily distributed within the range of 0−0.4, with a higher concentration in the segment below 0.25 (Fig. 2E). This indicates that efficacy descriptions were quite different, and the connotation was relatively rich in semantics. We also found that

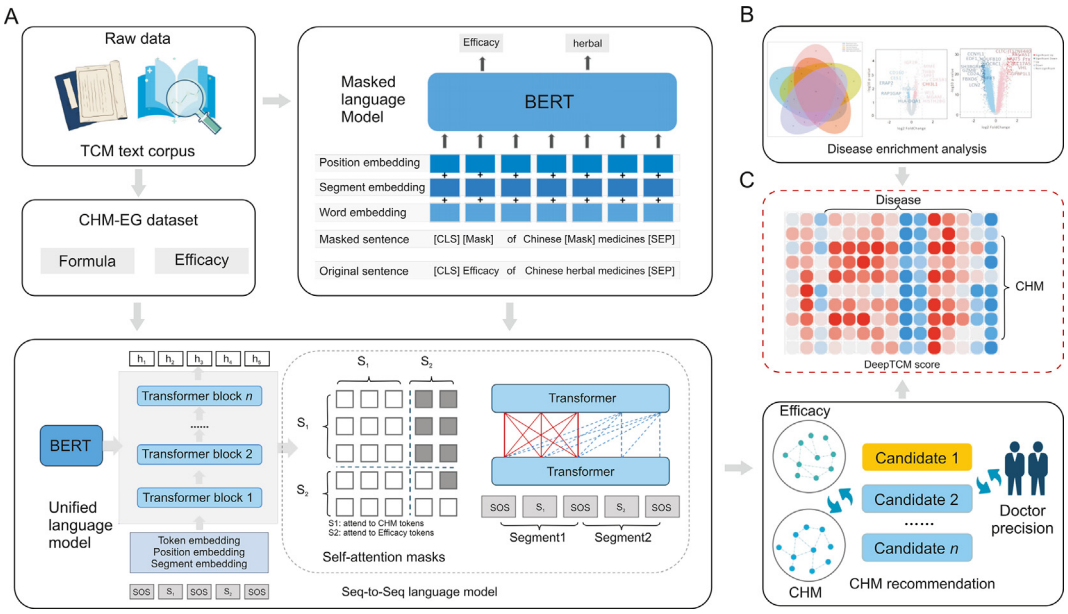

Fig. 1. A schematic representation of the deep learning model underlying an artificial intelligence-based quantitative predictive algorithm (DeepTCM). The large-scale bidirectional encoder representations from transformers-based pre-trained language model and fine-tuned it using traditional Chinese medicine medical texts and a language random mask strategy. We took Chinese herbal medicine (CHM) formulas information as the model input and applied the unified language model to perform word vector embedding and transformer-based feature extraction on the CHM formulas. We constructed a masked cross-entropy loss for the generative model as the loss function, and also introduced the BeamSearch method in the model generation stage for generating efficacy information of candidate CHM formulas. (A) CHM formula efficacy generation model (CHM-EG). (B) CHM formulas and a disease ontology score model (CHM-DOS). (C) CHM-diseases score. TCM: traditional Chinese medicine; BERT: bidirectional encoder representations from transformers.

some Chinese herbs were widely used in many cases of CHM formulas; for instance, licorice was priced at 637 CHM formulas, followed by Angelica (422), Rehmannia (356), and Poria (309) (Fig. 2F).

3.3. Evaluation of DeepTCM

The SC score reflected the fit or hits between effects, so the higher the score, the better. In contrast, the SA score reflected the conceptual conflict between efficacy, and the SD score reflected the lack of information about the generated efficacy, so a lower score was better. We subsequently performed manual corrections based on the scoring function. We then performed a tenfold cross-validation of the model by applying the three scores. CHM-EG achieved an average SC score of 0.781, an average SA score of 0.189, and an average SD score of 0.240 (Fig. 3A). In terms of the scoring trend, the consistency scores are mainly distributed in the high score interval, indicating that the model has good hit ability. In addition, we built common deep learning models to compare with the CHM-EG algorithm, namely, the Seq2Seq and BART models. The Seq2Seq model achieved an average SC score of 0.708, an average SA score of 0.267, and an average SD score of 0.293. The BART model achieved an average SC score of 0.729, an average SA score of 0.236, and an average SD score of 0.269.

Then, we applied DeepTCM to screen CHM formulas for CHD and depression diseases. CHD affects >11.39 million people in China and has become one of the leading causes of death globally [14,15], while ischemic heart disease (IHD) accounts for 16% of the world's total deaths [16]. Recently, an increasing number of studies have indicated that depression is common and is an independent risk factor for CHD [17]. It was reported that CHD patients have a 4.3% higher rate of depression than the general public. Depression is a highly prevalent risk factor for the pathogenesis of CHD [18]. Furthermore, depression not only severely affects quality of life but also increases the risk of major adverse cardiovascular events (MACEs) [19]. Anticardiovascular medical treatment combined with antidepression medical treatment or cognitive behavioral therapy is mainly used to treat this complex disease [20,21]. Nevertheless, these therapies have not shown satisfactory results, among which the side effects of antidepression medical treatment are inevitable in long-term treatment and there is some controversy regarding whether they improve the prognosis of patients with CHD [22,23]. As a supplementary medical treatment, TCM is attracting increasing attention worldwide, and its efficacy needs to be evaluated using modern methods [24]. According to TCM guidelines and expert experience related to cardiovascular and cerebrovascular diseases, we selected 15 candidate CHM formulas (Table S5) for estimation of correlation. We applied DeepTCM to calculate the association of candidate CHM formulas with diseases of interest and drew a cluster score heatmap (Fig. 3B).

Finally, we used the Mfuzz method to analyze the trend characteristics of the comprehensive association score between candidate CHM formulas and diseases. The results showed that the CHM Formulas F14 and F6 had a higher score trend (Figs. 3C and D). We compared the scores of each candidate CHM formula in vascular disease (Fig. 3E) and depression (Fig. 3F). The results showed that F14 (HJSSJW) had a high score and served as the hit CHM.

3.4. Integrated bioinformatics analysis of HJSSJW in CHD with depression

HJSSJW consists of 13 Chinese medicinal herbs that “activate blood circulation, detoxify and alleviate depression”. Toxicological studies have not revealed any significant side effects. The active components included ferulic acid, catechin, adenosine, tanshinoneIIA, kaempferol, vanillic acid, butylidenephthalide, caffeic acid, senkyunolide A, qinanone G, tryptophan, harpagide, *etc* (Fig. S1 and Table S6).

Our study showed that HJSSJW supported the collection of 515 target records after deleting the overlapping data. We performed an

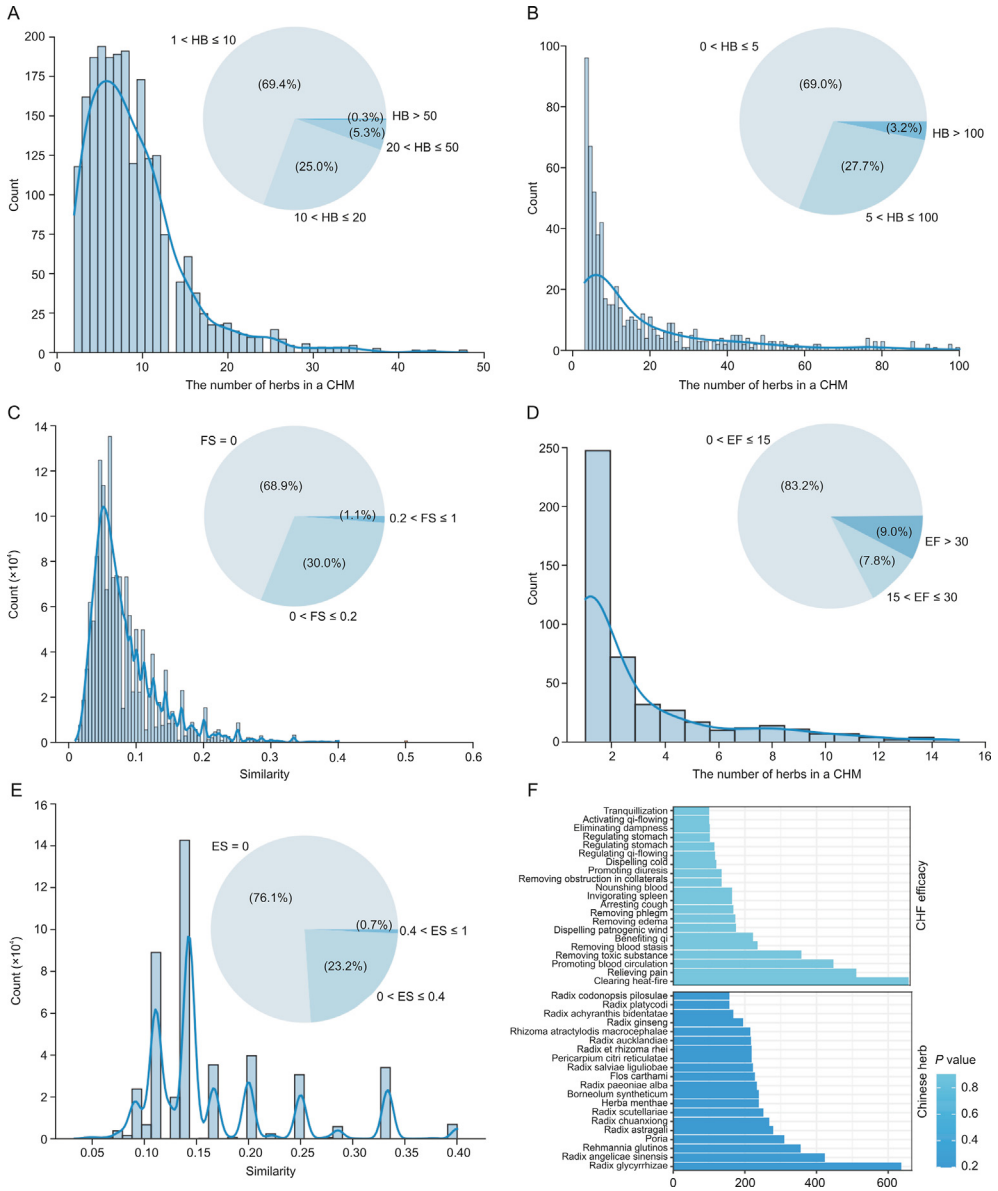

Fig. 2. Analysis of benchmark dataset. (A) Distribution of herb numbers in Chinese herbal medicine (CHM) formulas. (B) Distribution of herb frequency in CHM formulas. (C) Distribution of similarity among CHM formulas. (D) Distribution of efficacy frequency of CHM formulas. (E) Distribution of efficacy similarity among CHM formulas. (F) Top 20 most commonly used herbs and efficacy of CHM formulas. SC: semantic consistency; SA: semantic ambiguity; SD: semantic deficiency; HB: the number of herbs in a CHM; FS: the similarity between CHM; EF: the The number of efficacy in a CHM; ES: the similarity between efficacy.

overlap analysis of the set of targets for each herb (Fig. S2A). We then constructed the "herb-target" network by Cytoscape (Fig. S2B). The top 20 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) terms and Pattern genes (Figs. S2C–E). Subsequently, we performed an enrichment analysis of HJSSJW targets for all DisGeNet diseases, and the top 20 enriched diseases. The targets regulated by the prescription had high enrichment scores in five vascular and nervous system diseases, including "vascular diseases", "myocardial ischemia", "reperfusion injury", "major depressive disorder" and "depressive symptoms" (Fig. S2F).

We applied Venn analysis to observe five disease-related overlapping targets. The results showed 14 common targets in these diseases (Figs. 4A and B), including *interleukin* (*Il*)*-6* (*Il-6*), *angiotension converting enzyme* (*Ace*), *interferon gamma* (*Ifnγ*), *Il-1α*, *epidermal growth factor* (*Egf*), *Il-10*, *vascular endothelial growth factor receptor 2* (*Kdr*), *nitric oxide synthase 3* (*Nos3*), *catenin* (*Cat*), *IL-1β*, *C-X-C motif chemokine ligand 10* (*Cxcl10*), *chemokine (C–C motif) ligand 2* (*Ccl2*), *prostaglandin endoperoxide synthase 2* (*Ptgs2*) and *tumor necrosis factor* (*Tnf*). We then applied the PPI network analysis algorithm based on the STRING database and MCC based on cytoHubba to observe the interaction pattern of these targets. Hence, we observed the pair association of a single target with five diseases and drew the a cluster heatmap, where *Il-6*, *Il-1β* and *Tnf* showed a higher score trend (Fig. 4C).

"Myocardial ischemia" and "reperfusion injury" were the main symptoms and high-risk types of CHD [25]. "Major depressive disorder" and "depressive symptoms" belong to the category of depression. We hypothesized that HJSSJW mainly acted on CHD with depression. We established a scale-free (scale-free $R^2 = 0.90$) coexpression network with soft-thresholding powers $\beta = 16$ and $\beta = 7$ for CHD and depression, respectively (Figs. S3A and B). Then,

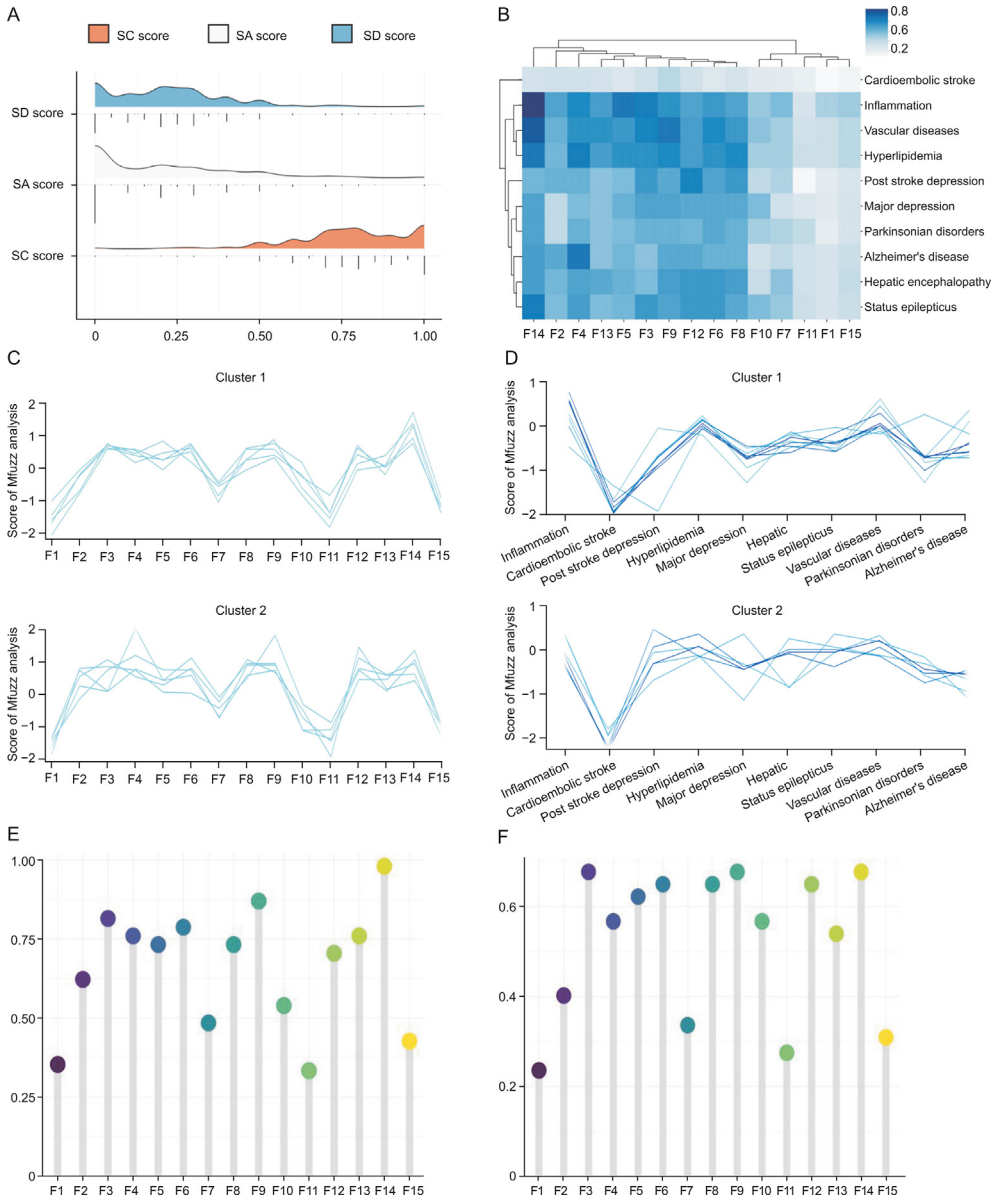

Fig. 3. Training results of the deep learning model underlying DeepTCM. (A) Distribution of predictive evaluation in 10-fold cross-validation. (B) DeepTCM scores of 15 Chinese herbal medicine (CHM) formulas in treating diseases associated with complex diseases. (C, D) Mfuzz analysis of DeepTCM scores variations of 15 CHM formulas for complex diseases. (E) DeepTCM scores of 15 candidate CHM formulas for vascular diseases. (F) DeepTCM scores of 15 candidate CHM formulas for depression.

the initial modules and merged modules are displayed under the clustering tree (Figs. S3C and D). The relationship between the modules and clinical characteristics (disease state and health management) was investigated. The results indicated that the green–yellow module was negatively correlated with healthy controls (r = -0.49, P = 0.04) in CHD and the black module (r = -0.35, $P = 4 \times 10^{-4}$) in depression (Figs. S3E and F). Finally, we performed differential gene analyses for CHD and depression. We calculated the log2-fold change (log2FC) value and P value for each gene in CHD and depression and plotted volcano plots (Figs. S3G and H). Due to the fewer DEGs in CHD under this threshold, we adjusted the threshold of log2FC value to $|\log2FC| \geq 0.5$. Thus, a total of 59 DEGs for CHD were identified (Figs. S3I and J).

Then, we constructed a PPI fusion network of HJSSJW and multiple key disease genes (Fig. S4). The MCC algorithm was used to identify the key hub genes of CHD with depression and HJSSJW (Fig. 4D). The results showed that *Il-6*, *Il-1β*, *Il-10* and so on were the top-ranked key genes (Fig. 4E), suggesting that HJSSJW may regulate and treat CHD with depression through these key targets.

3.5. HJSSJW is associated with major cardiovascular events in CHD with depression patients

PCI is the most important therapeutic method for CHD and aims to reconstitute the insufficient myocardial blood supply due to severe coronary stenosis or vascular occlusion [26]. Recently, an increasing number of studies have indicated that depression following PCI is common and is an independent risk factor for CHD [18,19,27]. A prospective study found that 20.7% of patients arranged for cardiac coronary artery bypass graft surgery had high depression scores, and that the rate reached 28% at six months after surgery [28]. To determine the clinical relevance of HJSSJW, we enrolled 40 patients with CHD after the PCI procedure with

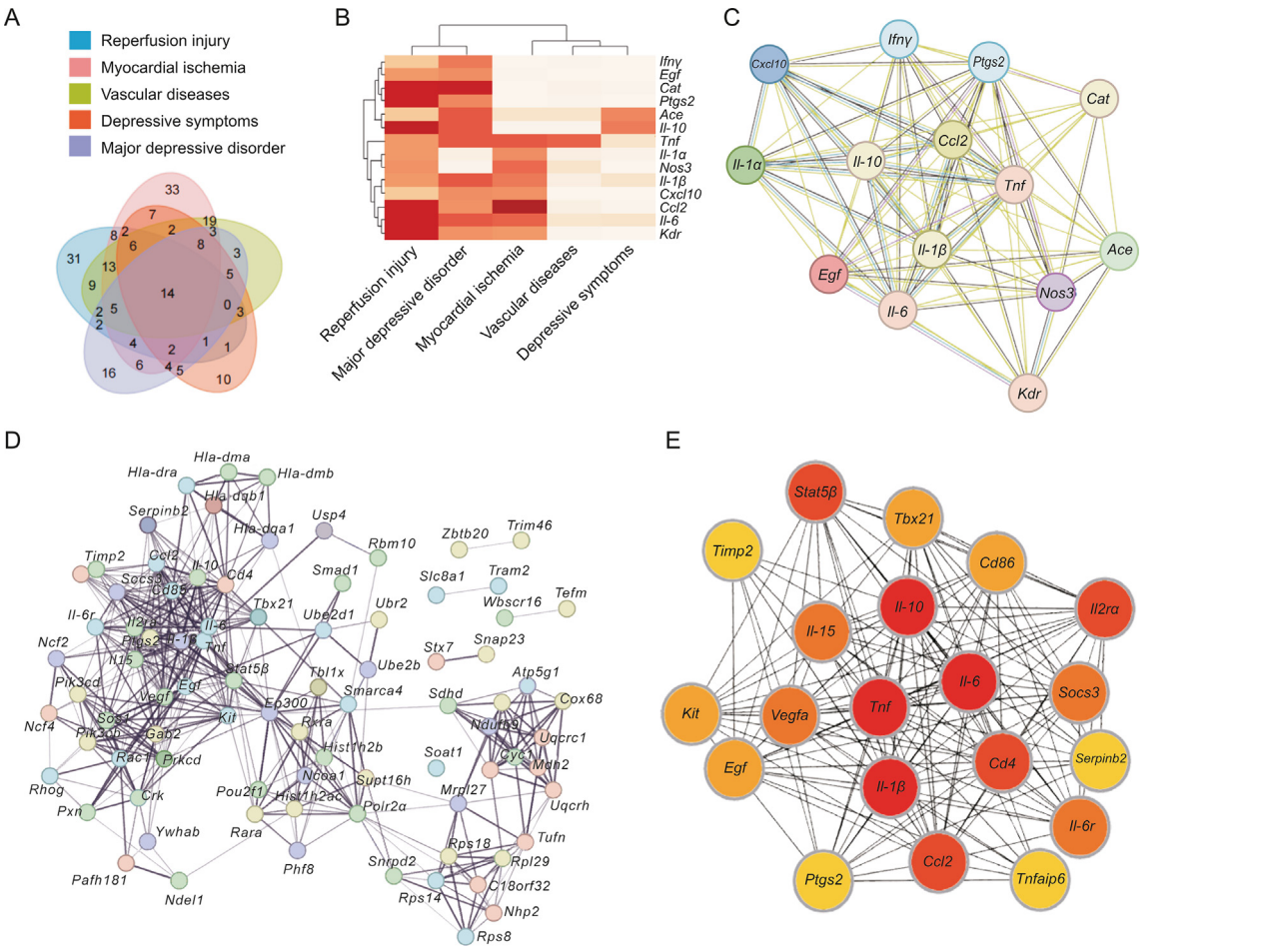

Fig. 4. Systematic pharmacological analysis of Huangjin shuangshen jiawei (HJSSJW) and coronary heart disease (CHD) with depression. (A) Veen diagram of 5 enriched diseases in DisGeNET. (B) Protein protein interaction (PPI) network and maximal clique centrality (MCC) analysis of the overlap targets among 5 enriched diseases. (C) Gene-disease correlation between overlap targets and 5 enriched diseases. (D) The most important diseases genes. (E) The MCC algorithm was applied to pinpoint the essential targets in the PPI network between HJSSJW and CHD with depression.

depression. Sequential serum samples at 12 weeks after HJSSJW or placebo were collected (Fig. 5A). Serum TC, TG and HDL levels were not significantly different between placebo and HJSSJW patients (Figs. S5A–C). However, the serum LDL level was significantly decreased in the HJSSJW group compared to the placebo group (Fig. S5D). In addition, the blood profile demonstrated significant changes in serum serum creatine phosphokinase isoenzyme (CK-MB) levels (Fig. 5B). The 5-HT level and number of HAMD 24 in HJSSJW group were significantly less than those in the other groups (Figs. 5C and D). Next, we investigated the association of serum inflammatory cytokines in 40 patients. We found that TNF-α was not significantly different between placebo and HJSSJW (Fig. S6) but decreased IL-6 in patients with HJSSJW (Fig. 5E).

3.6. HJSSJW protected against clinical symptoms in a mouse model of I/R

To better explore the efficacy and mechanism of HJSSJW. A formula was used to convert the therapeutic dose in humans to the therapeutic dose in mice. I/R injury is a major cause of adverse outcomes of revascularization after CHD [29]. To further investigate the dynamic pathological alterations of I/R for HJSSJW, time-series analysis was performed in hearts after 30 min of ischemia regimen followed by reperfusion (duration 3 h and 1, 7 days) by performing intragastric administration [30] of formulations one week in advance (Fig. 6A). Cardiac function and adverse cardiac remodeling (as indicated by cardiac fibrosis) were evaluated on day 7. We found that HJSSJW significantly reduced the relative heart weight (Fig. 6B). The cardiac function in HJSSJW mice was significantly preserved as evidenced by improving the left ventricular ejection fraction (LVEF) and lowering serum myocardial enzymes of cardiac injury (Figs. 6C–E). Furthermore, histopathology of cardiac tissue on day 7 after reperfusion revealed a significantly smaller disorderly arrangement of smooth muscle fibers and elastic fibers (Fig. 6F), while RT-PCR analysis showed lower degrees of functional gene expression (Fig. 6G).

Cardiac fibrosis secondary to CHD is known to promote cardiac dysfunction [31]. To test whether HJSSJW affects cardiac fibrosis, we used Masson's trichrome stain and sirius red stain (Fig. 6H). We found that HJSSJW obviously reduced fibrosis and collagen accumulation. HJSSJW also obviously decreased fibrosis genes such as *Tgfb1* and *Col1* expression (Fig. 6I). Matrix metalloprotein 9 (MMP9) was the key factor associated with cardiac fibrosis [32], and we observed that cardiac tissue at day 7 after reperfusion was significantly decreased from HJSSJW (Figs. 6J and K).

To further test whether HJSSJW affects inflammation in the I/R model, we measured the expression of inflammatory genes in the heart. HJSSJW significantly reduced inflammatory cytokine genes such as *Il-6* and *Il-1β* in the I/R model (Fig. 6L). Moreover, quite

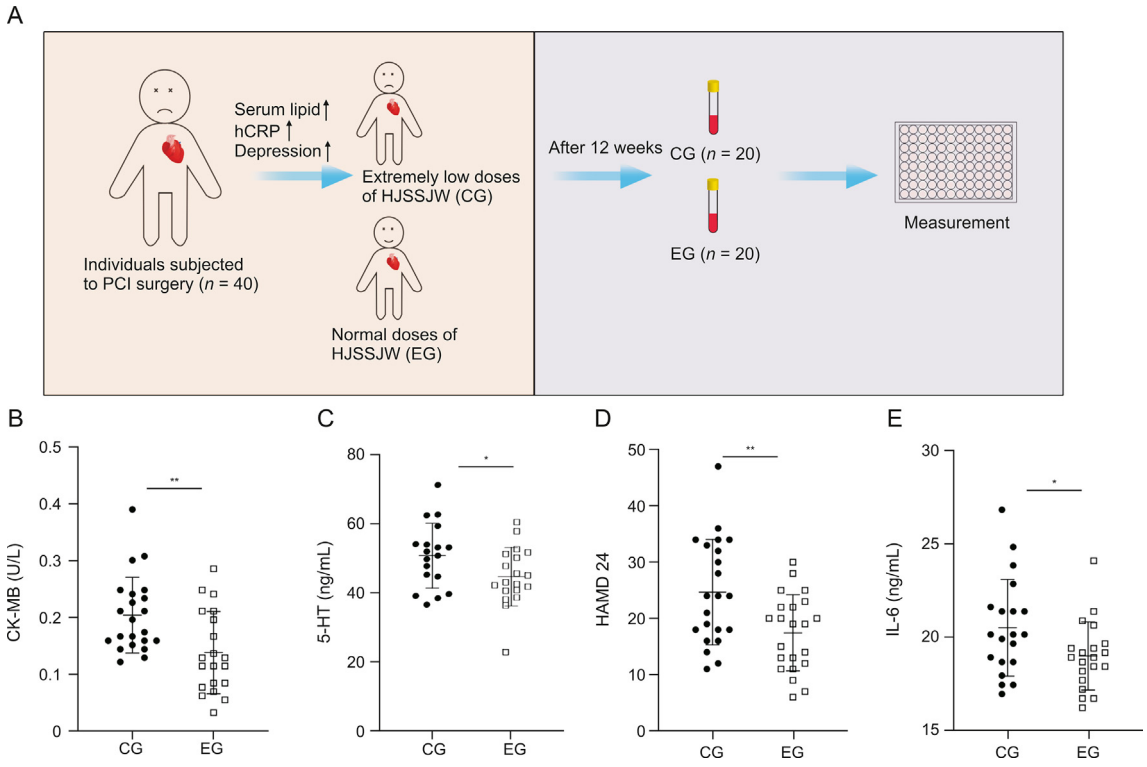

Fig. 5. Huangjin shuangshen jiawei (HJSSJW) controls major cardiovascular events in coronary heart disease (CHD) with depression patients. (A) Flow diagrams for assessing the serum lipid or cardiac function or inflammation levels and its prognostic value for CHD with depression in patients with CHD receiving percutaneous coronary intervention (PCI). (B) Serum creatine phosphokinase isoenzyme (CK-MB) levels in patients with CHD receiving PCI. (C) Serum 5-hydroxytryptamine (5-HT) levels in patients with CHD receiving PCI. (D) The score of depression in patients with CHD receiving PCI. (E) Serum interleukin 6 (IL-6) in patients with CHD receiving PCI. All *P* values were determined by one-way analysis of variance. $^{*}P < 0.05$, $^{**}P < 0.01$ compared with the control group. CG: extremely low doses of HJSSJW; EG: nomal doses of HJSSJW.

interestingly, the expression of IL-6 and IL-1β rather than TNF-α protein in the heart was reduced by HJSSJW, with a peak on day 7 (Fig. S7).

To explore the molecular mechanism underlying the impact of HJSSJW on the heart, we performed network pharmacology analysis of HJSSJW. The enriched pathways were related to the IL-6 signaling pathway, transforming growth factor (TGF)-β signaling pathway, positive regulation of IL-6 production and inhibitor of nuclear factor kappa-B (IκB) kinase/nuclear factor kappa-B (NF-κB) signaling (Fig. S2). Then, we used Western blot analysis and observed that IκB protein expression in cardiac tissue on day 1 after reperfusion was significantly decreased and that phosphorylated NF-κB protein expression in cardiac tissue on day 7 after reperfusion was significantly decreased in HJSSJW mice compared with I/R mice (Fig. S8A), suggesting that HJSSJW reduced NF-κB signaling activation in heart tissue in the I/R model. Immunohistochemistry showed HJSSJW suppression of NF-κB nuclear translocation (Fig. S8B), showing that NF-κB was the main target of HJSSJW.

Through I/R experiments, we found that the middle dose of HJSSJW had the best efficacy and could improve mouse survival (Fig. S9). Therefore, we selected a clinically effective dose for the next experiment.

To better explore the efficacy and confirm the feasibility of DeepTCM, we compared the therapeutic effect between a higher-scoring formulation (F14, HJSSJW) and a lower-scoring formulation (F1, SXBXW) in a mouse model of I/R. In the mouse experiments, the dosages of SXBXW and HJSSJW matched the recommended doses in the clinic. We found that HJSSJW significantly reduced the relative heart weight and heart volume compared with SXBXW (Figs. S10A and B). The cardiac function in HJSSJW mice was significantly preserved, as evidenced by improving LVEF and lowering serum myocardial enzymes of cardiac injury compared with SXBXW (Figs. S10C–E). Furthermore, histopathology revealed a significantly smaller disorderly arrangement of smooth muscle fibers and elastic fibers in the HJSSJW group than in the SXBXW group (Fig. S10F), while RT-PCR analysis showed lower degrees of functional gene expression (Fig. S10G). To compare the efficacy of cardiac fibrosis, we used Masson's trichrome stain and Sirius red stain (Fig. S10H). We found HJSSJW obviously reduced fibrosis and collagen accumulation compared with SXBXW. HJSSJW also obviously decreased fibrosis genes such as *Tgfb1* and *Col1* expression in the heart compared with SXBXW (Fig. S10I). We also observed MMP9 expression and found that HJSSJW decreased MMP9 versus SXBXW (Figs. S10J and K). Interestingly, we observed that when treated with HJSSJW, the mouse survival rates achieved a significant increase when compared to SXBXW treatment in the model of I/R (Fig. S10L). All results suggested that HJSSJW is better than SXBXW in protecting against clinical symptoms in a mouse model of I/R. The comparison of the predicted results with experimental data confirms the feasibility of DeepTCM.

3.7. HJSSJW protected against clinical symptoms in a mouse model of CHD with depression

To determine whether HJSSJW has clinically relevant effects, we examined it in a mouse model of CHD with depression. APOE$^{-/-}$ male mice were subjected to a "two-hit" model induced by a combination of HFD and CUMS, followed by I/R (Fig. 7A). After 4 weeks of HJSSJW, we found that it caused reductions in the rate of body weight gain compared with the model group (Fig. S11A), while food intake had no reductive effect (Fig. S11B). Interestingly,

we observed that when HJSSJW treatment was administered, mouse survival rates were significantly increased when compared with vehicle treatment in the model of CHD with depression (Fig. S11C). Then, the staining of fat and liver showed that HJSSJW treatment eliminated lipid accumulation (Figs. S11D and E), decreased the levels of TC, TG and LDL, and increased the levels of HDL (Fig. S11F–I).

To determine whether HJSSJW treatment affects cardiac dysfunction and pathological alterations, we measured a series of indices and found that HJSSJW caused significant reductions in the heart weight ratio, LV mass and relative myocardial enzymes compared with the model group (Figs. 7B–D). Furthermore, for the heart failure phenotypes that recapitulated the clinical symptoms of CHD with depression, including cardiac contractile dysfunction, LVEF and LVFS, HJSSJW had strong therapeutic effects in the heart (Figs. 7E and S12). The histopathology of cardiac tissue at day 7 after reperfusion revealed a significantly smaller disorderly arrangement of smooth muscle fibers and elastic fibers in the HJSSJW group than in the CHD with depression group (Fig. 7F). The ameliorating effect of HJSSJW on heart functional genes (Figs. 7G–I), fibrosis (Figs. 7J–M) and the inflammatory response (Figs. 7N and S13) after reperfusion was largely reversed by CHD with depression. Moreover, these were significantly increased IκB protein expression and phosphorylated NF-κB protein expression in cardiac tissue on day 7 after reperfusion (Fig. S14A) and prevention of NF-κB nuclear translocation (Fig. S14B), suggesting that HJSSJW reduced NF-κB signaling activation in CHD with depression.

3.8. HJSSJW protected against depression in a mouse model of CHD with depression

Numerous studies have shown that HFD induces depression [33,34]. To study the comorbidity of CHD and associated behavioral phenotypes in an animal model, we employed several behavioral assays, including the FST, TST, SPT and OFT (Fig. 8A). The chronic HFD and CUMS treatment mice showed profound depression-like responses in the FST, TST and SPT, and HJSSJW could remit these phenomena (Figs. 8B–D). Moreover, HJSSJW mice spent significantly more time in the center zone in the OFT than HFD- and CUMS-treated mice (Figs. 8E and F). It can be concluded from the results that depression-like responses developed along with chronic HFD and CUMS treatment, and HJSSJW possessed therapeutic effects.

The hippocampus functions as a learning, memory and other cognitive brain area. Previous studies have suggested that the hippocampus is also a contributing factor to depression [35,36]. These processes involve changes in ionized calcium binding adaptor molecule-1 (Iba-1), postsynaptic density 95 (PSD-95) [37], activity-regulated cytoskeleton-associated (Arc) [38] and glial fibrillary acidic protein (GFAP) protein levels. HJSSJW suppressed Iba-1 and GFAP protein levels and increased PSD-95 and Arc protein levels in the HFD and CUMS treatment groups. All of these results suggested that HJSSJW improved CHD with depression in mice (Figs. 8G and H).

We also tested the effect of HJSSJW on CHD with depression mouse-associated cytokines. According to the results of cytokines measured by the RT-PCR and Western blot in the hippocampal tissue, the mRNA and protein levels of the inflammatory factors *IL-6* and *TNF-α* in the control and HJSSJW groups were significantly lower than those in the model group (Figs. 8I and S15).

4. Discussion

This study proposed a quantitative predictive algorithm DeepTCM, which facilitated the calculation of the correlation between CHM formulas and complex diseases, offered multiscale results support at both the molecular and theoretical levels of TCM, and achieved the goal of clinical precision medicine. We then evaluated HJSSJW for cardiovascular disease using the DeepTCM and experimental testing in mouse and human disease models. Validation suggested that HJSSJW confers disease-relevant impacts against I/R and CHD with depression. We have also demonstrated that the activation of the IL-6 -induced NF-κB signaling pathway is a critical target for the development of agents to combat CHD.

Bioinformatics and DL methods have demonstrated potential in identifying synergistic drug combinations for diseases [39] and significantly improving the precision of combination therapies [40]. However, several issues still challenge the current application of systematic computational methods in the study of CHM formulas [3,41].

Different proportions of herbs have different syndromes, and a dominant systematic effect emerges. For example, when the six Chinese herbs Semen Persicae, Flos Carthami, Radix Rehmanniae Preparata, Radix Angelicae Sinensis, Rhizoma Chuanxiong, and Radix Paeoniae Rubra are combined, the systematic CHM formula efficacy of “blood-activating and stasis-resolving” emerges [42]. In our study, we constructed a large-scale TCM-CPS and created a high-quality benchmark dataset to facilitate mapping Chinese herb composition to efficacy descriptions of CHM formulas. The study involved the testing of 15 prescriptions. The results indicated the same syndrome as the drug instructions. The experimental CHM formula HJSSJW exhibited a therapeutic effect characterized by “activating blood circulation, detoxifying and alleviating depression”.

Model-guided drug development involves building a scoring system that quantifies the physiological, pharmacological, and disease process information obtained in the process of drug development through modeling [43]. To overcome this difficulty, we utilized two synergy metrics, prior biological knowledge and biostatistical inference, to predict complex herbal medicine and multi disease association studies, which were applied to high-throughput screening of active CHM formulas. The results showed that HJSSJW had the highest score and was the best CHM formula for CHD with depression. The main compound composition of HJSSJW has the efficacy of reducing lipids, expanding the coronary artery, anticoagulating, preventing thrombosis, and exerting antioxidation anti-inflammatory and neuroprotective actions [44–46]. As the basic prescription in HJSSJW, Si-Miao-Yong-An (SMYA) decoction could significantly promote isoproterenol-induced heart failure [47]. Furthermore, *hyperici perforati herba* and *Aquilariae lignum resinatum* had considerable efficacy and safety with depression [48,49].

In our study on the pharmacological mechanism of HJSSJW, we employed integrated bioinformatics analysis to unravel the connection between HJSSJW and CHD with depression. By employing systematic pharmacology and a PPI network-based MCC algorithm, we successfully identified a set of potentially crucial targets of HJSSJW in regulating CHD and depression, including IL-6, IL-1β, and others, which was similar to the clinical results [50]. The outcomes of these integrative analyses suggest that HJSSJW may regulate CHD with depression by targeting these crucial proteins. This provides a good foundation for further experimental verification. In addition, our data indicated that IL-6 and IL-1β were overexpressed in clinical research and the mouse model, while HJSSJW could decrease their expression levels. To explore the molecular mechanism underlying the impact of HJSSJW on the heart, we also performed integrated bioinformatics analysis of HJSSJW. The enriched pathways were related to the IL-6 signaling pathway, TGF-β signaling pathway, positive regulation of IL-6 production and IκB kinase/NF-κB signaling. It has been reported that the IκB kinase/NF-κB pathway is enriched in inflammatory, cardiovascular and neurodegenerative diseases [51]. NF-κB has been found to alleviate heart failure with I/R and depression by

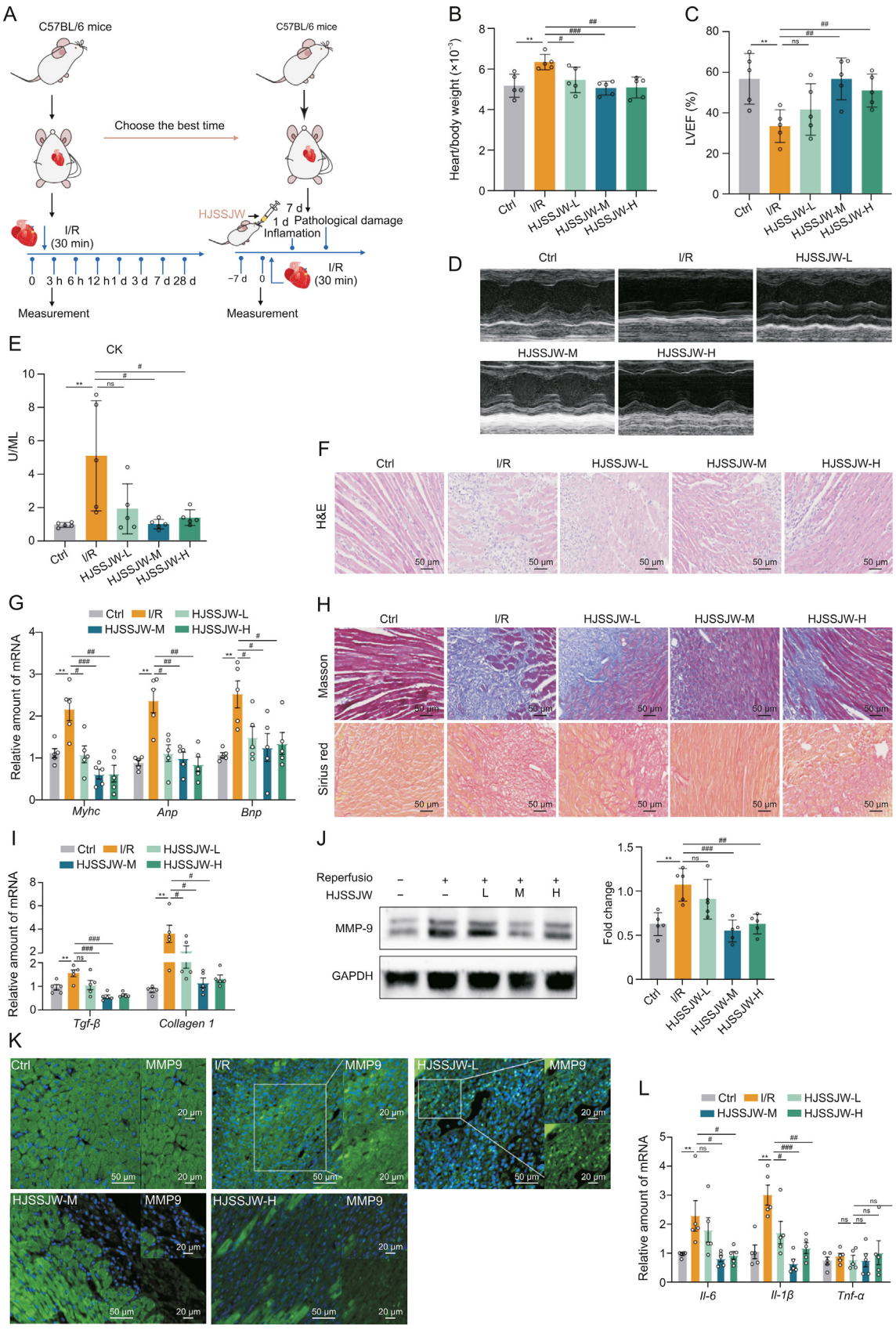

Fig. 6. The anti-coronary heart disease (CHD) prescription Huangjin shuangshen jiawei (HJSSJW) prevented ischemia reperfusion (I/R). (A) Experimental processes for the I/R in C57BL/6 male mice. $n = 5$ mice per group. Mice were intragastric administration water or HJSSJW one week before I/R. (B) Heart weight/body weight ratio. (C) Left ventricular ejection fraction (LVEF). (D) Image. (E) Serum creatine kinase (CK). (F) Hematoxylin-eosin staining (H&E) staining of cardiac interstitial disorders. (G) mRNA quantification of *myosin heavy chain* (*Myhc*), *atrial natriuretic peptide (Anp)* and *brain natriuretic peptide (Bnp)*. (H) Trichrome staining. Red: cardiomyocytes; blue: fibrosis. Sirius red staining. Red: collagen of

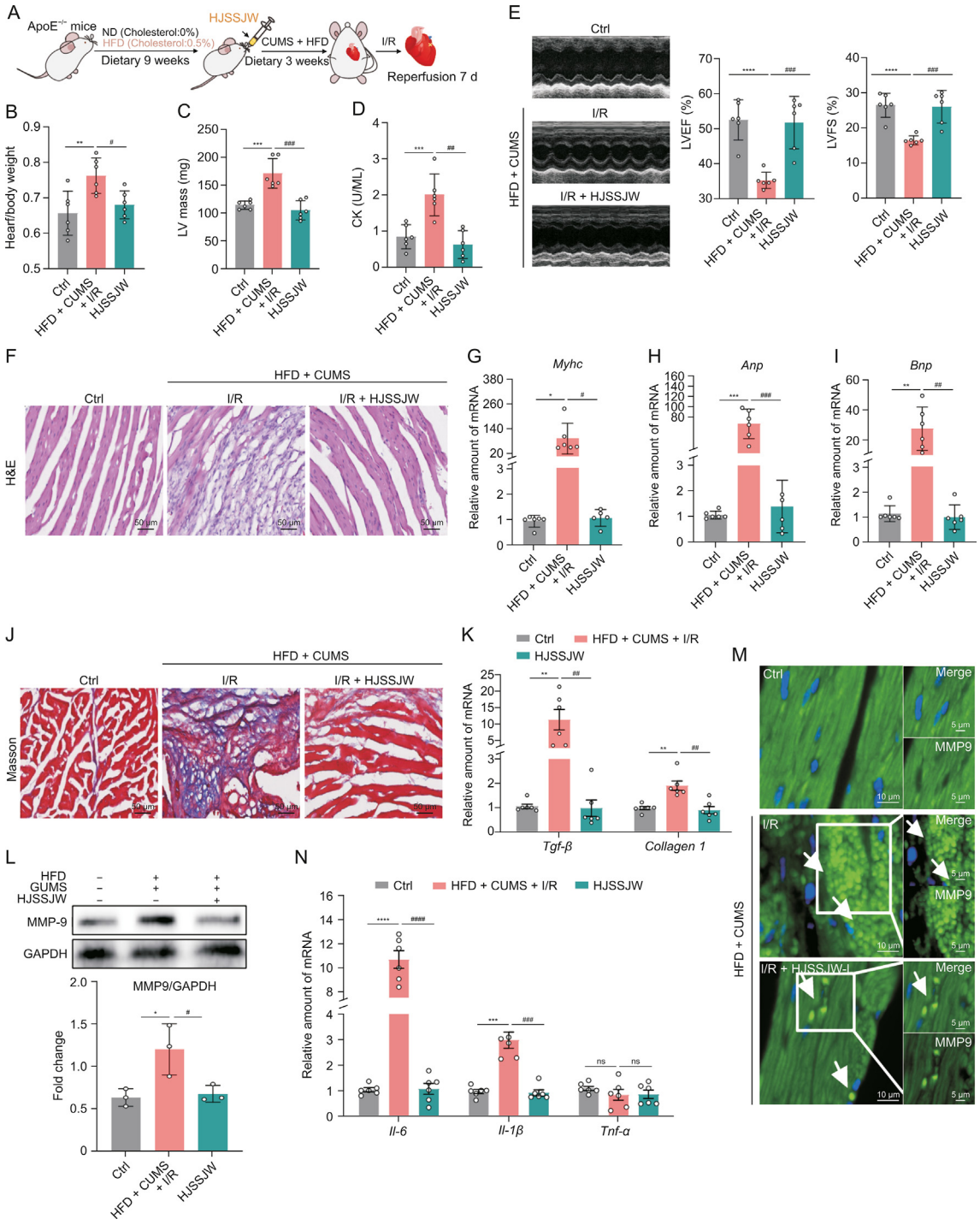

Fig. 7. Huangjin shuangshen jiawei (HJSSJW) reversed cardiac dysfunction, fibrosis, and inflammation in mice model of coronary heart disease (CHD) with depression. (A) Experimental processes in ApoE$^{-/-}$ male mice. $n = 6$ mice per group. Mice were intragastric administration water or HJSSJW from 10 week. (B) Heart weight/body weight ratio. (C) Left ventricular (LV) mass. (D) Serum creatine kinase (CK). (E) Echocardiographic measurement of cardiac functional and structural parameters. (F) Hematoxylin-eosin staining (H&E) staining of cardiac interstitial disorders. (G) mRNA quantification of *myosin heavy chain (Myhc)*. (H) mRNA quantification of *atrial natriuretic peptide (ANP)*. (I) mRNA quantification of *brain natriuretic peptide (BNP)*. (J) Trichrome staining. Red: cardiomyocytes; blue: fibrosis. (K) mRNA quantification of *transforming growth factor (TGF)-β* and *Collagen 1*. (L) Western blot analysis of matrix metalloproteinase-9 (MMP9) protein. (M) Immunohistochemistry for MMP9. (N) mRNA quantification of *interleukin 6 (IL-6)*, *interleukin 1β (IL-1β)*, *tumor necrosis factorα (TNF-α)*. All *P* values were determined by one-way analysis of variance. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ compared with the saline group. $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$ compared with the model group. ns: not significance. ND: normal diet; HFD: high fit diet; CUMS: chronic mild unpredictable stress.

cardiac interstitial fibrosis. (I) mRNA quantification of *transforming growth factor (TGF)-β* and *Collagen 1*. (J) Western blot analysis of matrix metalloproteinase-9 (MMP9) protein. (K) Immunohistochemistry for MMP9. (M) mRNA quantification of *interleukin 6 (Il-6)*, *interleukin 1β (Il-1β)*, *tumor necrosis factorα (Tnf-α)*. All *P* values were determined by one-way analysis of variance. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared with the saline group. $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$ compared with the I/R group. ns: not significance. GAPDH: glyceraldehyde-3-phosphate dehydrogenase.

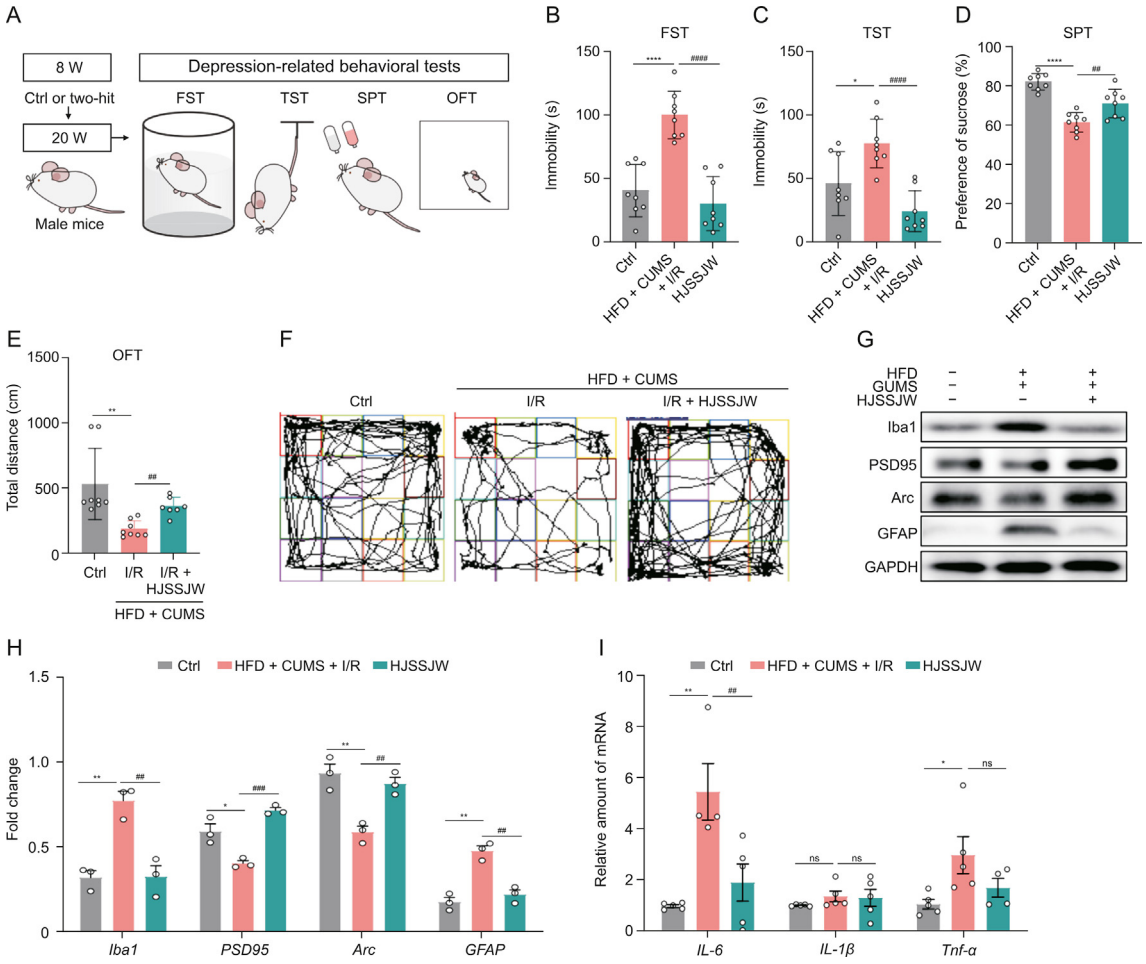

Fig. 8. Huangjin shuangshen jiawei (HJSSJW) reduced hippocampus inflammation and reverses depression in mice model of coronary heart disease (CHD) with depression. (A) Experimental paradigms and behavioral tests in ApoE$^{-/-}$ male mice before operation. (B) Forced swimming test (FST). (C) Tail suspension test (TST). (D) Sucrose preference test (SPT). (E) Open field test (OFT). (F) Activity trajectory diagram of OFT. (G) Western blot analysis of adaptor molecule-1 (Iba-1), postsynaptic density 95 (PSD-95), activity-regulated cytoskeleton-associated (Arc), and glial fibrillary acidic protein (GFAP) in the ApoE$^{-/-}$ male mice hippocampus after operation. (H) Quantitative analyses. (I) mRNA quantification of *IL-6*, *IL-1β*, *TNF-α* in the ApoE$^{-/-}$ male mice hippocampus after operation. All *P* values were determined by one-way analysis of variance. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$ compared with the Saline group. $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$, $^{\#\#\#\#}P < 0.0001$ compared with the model group. ns: not significance. GAPDH: glyceraldehyde-3-phosphate dehydrogenase.

inhibiting inflammasome formation [52,53]. Another study observed increased the NF-κB levels in patients with coronary artery disease and depression, suggesting that NF-κB pathway may be involved in the pathophysiology of comorbid CHD with depression [54]. Our data also confirmed that HJSSJW mediates CHD and CHD with depression through the IκB kinase/NF-κB pathway. Through the above analysis, we believe the potential of DeepTCM to reveal the association between the efficacy of CHM formulas and complex diseases. We showcase the power of combining computational biology and biological experiments in discovering and understanding the mechanism of action of HJSSJW in modulating CHD with depression.

In the future, we will improve the model in three aspects. First, according to the pathological evolution law of complex diseases, we will gain a deeper understanding of the mechanisms and variability of diseases and treatments, as well as their translation to clinical outcomes, to update high-quality benchmark datasets for CHM formulas and diseases. Second, the study of the "dose effect" is also an important topic in the process of reconstructing the herbal system. We will explore the optimal dosage range of CHM formulas and explain the law of dosage and effect of CHM formulas through clinical research, combined with pharmacokinetics (PK), pharmacodynamics (PD) and other modern research techniques. Third, we will incorporate single-cell transcriptome data to address the question of the mechanism of action of the CHM formulas against a cell-specific phenotype of diseases.

5. Conclusion

In summary, the proposed DeepTCM algorithm can be a powerful tool for studying complex herbal medicine and multi-disease association studies, which can be applied to high-throughput screening of active CHM formulas. Bioinformatics methods such as diseases and pathway enrichment analysis are then embedded into the model to obtain the correlation between CHM formulas and diseases at the micro level and to discover the critical regulatory targets, biological pathways and processes. In the future, single-cell multiomics approaches can be further integrated to gain a more comprehensive understanding of the mechanisms underlying drugs and diseases from both macro and micro

perspectives. Since the dosage of the same herb in different CHM formulas varies, quantitative prediction becomes a more challenging task.

CRediT author statement

Yuanyuan Qian: Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - Original draft preparation, Reviewing and Editing, Visualization; **Xiting Wang:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - Reviewing and Editing, Visualization; **Lulu Cai:** Methodology, Validation, Formal analysis; **Jiangxue Han:** Methodology, Formal analysis; **Zhu Huang:** Methodology; Writing - Reviewing and Editing; **Yahui Lou:** Methodology, Writing - Reviewing and Editing; **Binyue Zhang:** Methodology; **Yanjie Wang:** Formal analysis, Reviewing and Editing; **Xiaoning Sun:** Methodology; **Yan Zhang:** Methodology; **Aisong Zhu:** Resources, Writing - Reviewing and Editing, Supervision, Funding. acquisition.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No.: 82174246), the National Key R&D Program of China (Grant No.: 2019YFC1708701), the Postdoctoral Innovation Talent Support Program (Grant No.: BX20220329). We appreciate the great help/technical support/ experimental support from the Public Platform of Medical Research Center, Academy of Chinese Medical Science, Zhejiang Chinese Medical University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpha.2023.12.004.

References

[1] C. Wang, M. Niimi, T. Watanabe, et al., Treatment of atherosclerosis by traditional Chinese medicine: Questions and quandaries, Atherosclerosis 277 (2018) 136–144.
[2] R. Zhang, X. Zhu, H. Bai, et al., Network pharmacology databases for traditional Chinese medicine: Review and assessment, Front. Pharmacol. 10 (2019), 123.
[3] K. Yang, R.S. Zhang, L.Y. He, et al., Multistage analysis method for detection of effective herb prescription from clinical data, Front. Med. 12 (2018) 206–217.
[4] W. Zhou, K. Yang, J. Zeng, et al., FordNet: Recommending traditional Chinese medicine formula via deep neural network integrating phenotype and molecule, Pharmacol. Res. 173 (2021), 105752.
[5] S. Fang, L. Dong, L. Liu, et al., HERB: A high-throughput experiment- and reference-guided database of traditional Chinese medicine, Nucleic Acids Res. 49 (2021) D1197–D1206.
[6] J. Wei, X. Ren, X. Li, et al., NEZHA: Neural contextualized representation for Chinese language understanding, (2019): arXiv:1909.00204.
[7] L. Dong, N. Yang, W.H. Wang, et al., Unified language model pre-training for natural language understanding and generation, (2019): arXiv.1905.03197.
[8] J. Piñero, À. Bravo, N. Queralt-Rosinach, et al., DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants, Nucleic Acids Res. 45 (2017) D833–D839.
[9] Y. Zhou, B. Zhou, L. Pache, et al., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, Nat. Commun. 10 (2019), 1523.
[10] M.E. Ritchie, B. Phipson, D. Wu, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015), e47.
[11] C.H. Chin, S.H. Chen, H.H. Wu, et al., cytoHubba: Identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (2014), S11.
[12] A.M. Newman, C.L. Liu, M.R. Green, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457.
[13] P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis, BMC Bioinform. 9 (2008), 559.
[14] Writing Committee of the Report on Cardiovascular H, Diseases in C, Report on Cardiovascular Health and Diseases in China 2021: An Updated Summary, B Biomedical and Environmental Sciences 35 (2022) 573–603.
[15] V. Vaccarino, L. Badimon, J.D. Bremner, et al., Depression and coronary heart disease: 2018 position paper of the ESC working group on coronary pathophysiology and microcirculation, Eur. Heart J. 41 (2020) 1687–1696.
[16] World Health Organization, The top 10 causes of death. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. (Accessed 20 December 2022).
[17] R.W. Yeh, H. Tamez, E.A. Secemsky, et al., Depression and angina among patients undergoing chronic total occlusion percutaneous coronary intervention: The OPEN-CTO registry, JACC Cardiovasc. Interv. 12 (2019) 651–658.
[18] R.M. Carney, K.E. Freedland, Depression and coronary heart disease, Nat. Rev. Cardiol. 14 (2017) 145–155.
[19] A.L. Beatty, J.A. Doll, D.W. Schopfer, et al., Cardiac rehabilitation participation and mortality after percutaneous coronary intervention: Insights from the veterans affairs clinical assessment, reporting, and tracking program, J. Am. Heart Assoc. 7 (2018), e010010.
[20] K. Kotseva, D. Wood, D. De Bacquer, et al., Determinants of participation and risk factor control according to attendance in cardiac rehabilitation programmes in coronary patients in Europe: EUROASPIRE IV survey, Eur. J. Prev. Cardiol. 25 (2018) 1242–1251.
[21] S.J. Olsen, H. Schirmer, T. Wilsgaard, et al., Cardiac rehabilitation and symptoms of anxiety and depression after percutaneous coronary intervention, Eur. J. Prev. Cardiol. 25 (2018) 1017–1025.
[22] C.J. Spindelegger, K. Papageorgiou, R. Grohmann, et al., Cardiovascular adverse reactions during antidepressant treatment: A drug surveillance report of German-speaking countries between 1993 and 2010, Int. J. Neuropsychopharmacol. 18 (2014), pyu080.
[23] M.K. Jha, A. Qamar, M. Vaduganathan, et al., Screening and management of depression in patients with cardiovascular disease: JACC state-of-the-art review, J. Am. Coll. Cardiol. 73 (2019) 1827–1845.
[24] Y. Wang, J. Xu, J. Yang, et al., Effects of Guanxinshutong capsules as complementary treatment in patients with chronic heart failure: Study protocol for a randomized controlled trial, Front. Pharmacol. 11 (2021), 571106.
[25] F.J. Neumann, M. Sousa-Uva, 'ten commandments' for the 2018 ESC/EACTS guidelines on myocardial revascularization, Eur Heart J. 39 (2018), 3759.
[26] T. Doenst, A. Haverich, P. Serruys, et al., PCI and CABG for treating stable coronary artery disease: JACC review topic of the week, J. Am. Coll. Cardiol. 73 (2019) 964–976.
[27] R.W. Yeh, H. Tamez, E.A. Secemsky, et al., Depression and angina among patients undergoing chronic total occlusion percutaneous coronary intervention: The OPEN-CTO registry, JACC Cardiovasc. Interv. 12 (2019) 651–658.
[28] B. Korbmacher, S. Ulbrich, H. Dalyanoglu, et al., Perioperative and long-term development of anxiety and depression in CABG patients, Thorac, Cardiovasc. Surg. 61 (2013) 676–681.
[29] D.M. Yellon, D.J. Hausenloy, Myocardial reperfusion injury, N Engl J Med. 357 (2007) 1121–1135.
[30] Y. Li, B. Chen, X. Yang, et al., S100a8/a9 signaling causes mitochondrial dysfunction and cardiomyocyte death in response to ischemic/reperfusion injury, Circulation 140 (2019) 751–764.
[31] A. González, E.B. Schelbert, J. Díez, et al., Myocardial interstitial fibrosis in heart failure: Biological and translational perspectives, J. Am. Coll. Cardiol. 71 (2018) 1696–1706.
[32] S. Chandra, K.C. Ehrlich, M. Lacey, et al., Epigenetics and expression of key genes associated with cardiac fibrosis: *NLRP3, MMP2, MMP9, CCN$_2$/CTGF* and *AGT*, Epigenomics 13 (2021) 219–234.
[33] M. Ogrodnik, Y. Zhu, L.G.P. Langhi, et al., Obesity-induced cellular senescence drives anxiety and impairs neurogenesis, Cell Metab. 29 (2019) 1061–1077.e8.
[34] X. Xie, H. Yang, J.J. An, et al., Activation of anxiogenic circuits instigates resistance to diet-induced obesity via increased energy expenditure, Cell Metab. 29 (2019) 917–931.e4.
[35] N.V. Gulyaeva, Functional neurochemistry of the ventral and dorsal hippocampus: Stress, depression, dementia and remote hippocampal damage, Neurochem. Res. 44 (2019) 1306–1322.
[36] W. Liu, T. Ge, Y. Leng, et al., The role of neural plasticity in depression: From hippocampus to prefrontal cortex, Neural Plast. 2017 (2017), 6871089.
[37] B. Compans, C. Camus, E. Kallergi, et al., NMDAR-dependent long-term depression is associated with increased short term plasticity through autophagy mediated loss of PSD-95, Nat. Commun. 12 (2021), 2849.
[38] M.C.V. Shandilya, A. Gautam, Hippocampal arc induces decay of object recognition memory in male mice, Neuroscience 431 (2020) 193–204.
[39] W. Jin, J.M. Stokes, R.T. Eastman, et al., Deep learning identifies synergistic drug combinations for treating COVID-19, Proc. Natl. Acad. Sci. USA 118 (2021), e2105070118.
[40] X. Li, E.K. Dowling, G. Yan, et al., Precision combination therapies based on recurrent oncogenic coalterations, Cancer Discov. 12 (2022) 1542–1559.
[41] D. Bu, Y. Xia, J. Zhang, et al., FangNet: Mining herb hidden knowledge from TCM clinical effective formulas using structure network algorithm, Comput. Struct. Biotechnol. J. 19 (2020) 62–71.
[42] X. Wang, T. Wang, Y. Wang, et al., Research progress on classical traditional Chinese medicine Taohong Siwu Decoction in the treatment of coronary heart disease, Biomedecine Pharmacother. 152 (2022), 113249.

[43] B. Ravikumar, T. Aittokallio, Improving the efficacy-safety balance of polypharmacology in multi-target drug discovery, Expert Opin. Drug Discov. 13 (2018) 179–192.
[44] D. Li, Y. Rui, S. Guo, et al., Ferulic acid: A review of its pharmacology, pharmacokinetics and derivatives, Life Sci. 284 (2021), 119921.
[45] M. Fan, W. Yang, M. He, et al., Occurrence, synthesis and biological activity of 2-(2-phenyethyl)chromones, Eur. J. Med. Chem. 237 (2022), 114397.
[46] I. Bernatova, Biological activities of (-)-epicatechin and (-)-epicatechin-containing foods: Focus on cardiovascular and neuropsychological health, Biotechnol. Adv. 36 (2018) 666–681.
[47] M. Liao, Q. Xie, Y. Zhao, et al., Main active components of Si-Miao-Yong-An Decoction (SMYAD) attenuate autophagy and apoptosis via the PDE5A-AKT and TLR4-NOX4 pathways in isoproterenol (ISO)-induced heart failure models, Pharmacol. Res. 176 (2022), 106077.
[48] Q.X. Ng, N. Venkatanarayanan, C.Y. Ho, Clinical use of Hypericum perforatum (St John's wort) in depression: A meta-analysis, J. Affect. Disord. 210 (2017) 211–221.
[49] H. Li, Y. Li, X. Zhang, et al., The combination of *Aquilaria sinensis (lour.) gilg* and *Aucklandia costus falc. volatile oils exerts* antidepressant effects in a CUMS-induced rat model by regulating the HPA axis and levels of neurotransmitters, Front. Pharmacol. 11 (2021), 614413.
[50] G.M. Khandaker, V. Zuber, J.M.B. Rees, et al., Shared mechanisms between coronary heart disease and depression: Findings from a large UK general population-based cohort, Mol. Psychiatry 25 (2020) 1477–1486.
[51] S. Singh, T.G. Singh, Role of nuclear factor kappa B (NF-κB) signalling in neurodegenerative diseases: An mechanistic approach, Curr. Neuropharmacol. 18 (2020) 918–935.
[52] Y. Dai, S. Wang, S. Chang, et al., M2 macrophage-derived exosomes carry microRNA-148a to alleviate myocardial ischemia/reperfusion injury via inhibiting TXNIP and the TLR4/NF-κB/NLRP3 inflammasome signaling pathway, J. Mol. Cell. Cardiol. 142 (2020) 65–79.
[53] Q. Wang, H. Bi, H. Huang, et al., Electroacupuncture prevents the depression-like behavior by inhibiting the NF-κB/NLRP3 inflammatory pathway in hippocampus of mice subjected to chronic mild stress, Neuropsychobiology 81 (2022) 237–245.
[54] W. Ma, D. Shen, J. Liu, et al., Statin function as an anti-inflammation therapy for depression in patients with coronary artery disease by downregulating interleukin-1β, J. Cardiovasc. Pharmacol. 67 (2016) 129–135.