Herb-CMap: Novel Hot Diffusion Algorithm to Identify Bioactive Ingredients from Herbal Medicine Based on Gene Perturbation Profiling

Yinyin Wang[a, #], Yihang Sui[a, #], Jiaqi Yao[a], Jiang Hong[a], Qimeng Tian[a], Yun Tang[b], Jing Tang[c], Ninghua Tan[a, *]

Affiliation:

[a] Department of TCMs Pharmaceuticals, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, 211198, PR China

[c] Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, FI-00014, Finland

[b] Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai, 200237, China

* Corresponding authors.

E-mail addresses: nhtan@cpu.edu.cn (Ninghua Tan)

# These authors made equal contributions to this work.

Abstract

Herbal medicine, especially Traditional Chinese medicine (TCM), is a valuable resource of natural products for drug discovery with obvious therapeutic effects. However, the mechanisms of action (MOAs) of herbal medicine are often unknown due to limited target information and the complexity of multiple ingredients and targets. This study developed a Herb-CMap algorithm to prioritize active ingredients and targets within herbal medicine by integrating transcriptomics-based gene perturbation data with a random walk algorithm. This methodology bridges the gap between gene perturbation of herbal medicine and its therapeutic target. Using the Suhuang antitussive capsule (Suhuang) for treating cough variant asthma (CVA) as a case study, we identify and experimentally verify that quercetin and luteolin directly interact with Il17a, Pik3cb, Pik3cd, Akt1, and Tnf. These interactions inhibit the IL-17 signaling pathway and inactivate PI3K, Akt, and NF-κB, preventing lung inflammation and treating CVA. Our findings demonstrate the potential of the Herb-CMap methodology to provide insights into the molecular MOAs of herbal medicine, thereby advancing drug discovery from herbal medicine.

Keyword: Herbal medicine; TCM, Gene perturbation; Random walk algorithm

Herb-CMap: Identify bioactive ingredients from herbal medicine through the Gene Perturbation by random walk algorithm

Abstract

Herbal medicine, especially Traditional Chinese medicine (TCM), is a valuable resource of natural products for drug discovery with obvious therapeutic effects. However, the mechanisms of action (MOAs) of herbal medicine are often unknown due to limited target information and the complexity of multiple ingredients and targets. This study developed a Herb-CMap algorithm to prioritize active ingredients and targets within herbal medicine by integrating transcriptomics-based gene perturbation data with a random walk algorithm. This methodology bridges the gap between gene perturbation of herbal medicine and their therapeutic target. Using the Suhuang antitussive capsule (Suhuang) for treating cough variant asthma (CVA) as a case study, we identify and experimentally verify that quercetin and luteolin directly interact with Il17a, Pik3cb, Pik3cd, Akt1, and Tnf. These interactions inhibit the IL-17 signaling pathway and inactivate PI3K, Akt, and NF-κB, preventing lung inflammation and treating CVA. Our findings demonstrate the potential of the Herb-CMap methodology to provide insights into the molecular MOAs of herbal medicine, thereby advancing the drug discovery from herbal medicine.

Keyword: Herbal medicine, TCM prescriptions; MOAs; Gene perturbation; Random walk algorithm

1. Introduction

Traditional Chinese medicine (TCM) has been extensively used in China and other countries, presenting exclusive advantages in drug discovery as significant resources. TCM prescriptions adopt the strategy of multiple ingredients, targets, and pathways to achieve optimal therapeutic effects [1]. However, besides complex ingredients, there could also be many drug-drug interactions in one TCM prescription, making it challenging to clarify the molecular MOAs of TCM prescriptions *via* wet experiments or human experience.

With the development of computational methods, various cutting-edge techniques have been applied to discover the MOAs of herbal medicine, such as molecular docking, machine learning methods [2,3], and network-based methods [4,5]. In particular, network-based methods can extract internal relations according to network topological characteristics, thus revealing hidden relationships, such as drug-disease associations and drug-drug interactions [6]. The premise behind network-based methods is the observation that genes with similar phenotypes and related functions are located close to each other in the protein-protein interaction (PPI) network. However, in these traditional strategies for network analysis in herbal medicine, the interactions between two ingredients and two herbs are usually ignored [7]. Recently, Cheng et al. proposed a novel methodology for analyzing drug combinations and drug-disease associations by calculating the overall network distance between the subnetwork of their targets in PPI [8]. We applied the network distance model to identify herb pairs and quantify their interactions and showed that the network distance algorithm can better capture the biological interactions between herb ingredients [9].

As we know, one drug could create effects by binding to certain therapeutical targets and thus causing downstream cascade reactions. Despite the advantages of network methods in herbal medicine, the lack of reliable target information on ingredients is one of the most challenging problems[10]. Due to the development of high-throughput sequencing and system biology, multi-omics data are often

integrated into computational target predicting methods, such as molecular docking [11], machine learning methods [12], and network-based methods [13], to improve prediction accuracy [14], especially the network perturbation method. For example, the random walk algorithm can be used to prioritize the critical drug targets of the compound based on the topological characteristics of the horizontal disturbance of the gene transcriptome among the PPI network [15]. Xiao et al., for instance, used differentially expressed genes to establish the weighted gene co-expression network to obtain the hub genes[16]. Another network perturbation called local radiality was developed to prioritize the targets by integrating topological data and the perturbed gene information [17]. However, as the application domain of these methods was designed to determine drug-drug or drug-disease associations, they have rarely been applied to analyze the complex systems of herbal medicine. Nevertheless, multi-omics integration methods have achieved considerable success in the study of herbal medicine[18], for instance, interpreting TCM syndrome classifications by integrating proteomic and metabolomic data[19], exploring the therapeutic effects of bioactivity ingredients in herbs by integrating lipidomic and transcriptomic data[20], and identifying diagnostic markers of herbal medicine based on network pharmacological analysis of TCM metabolomic data[21]. These exciting findings show the potential of integrating multi-omics data to explore the MOAs of herbal medicine.

Cough variant asthma (CVA) is a type of asthma with bronchial hyperreactive eosinophilic airway inflammation and airway remodeling. Severe and frequent coughs will damage the respiratory tract and aggravate respiratory infection, which has affected patients' quality of life and caused a considerable burden to society [22,23]. CVA has the features of paroxysmal, spasmodic, and repetitive coughs, which corresponds to the “wind cough” syndrome in TCM theory. The Suhuang antitussive capsule (Suhuang) is one of the TCM prescriptions used to treat the “wind cough” syndrome. Suhuang, which consists of nine herbs, *Ephedrae Herba* (MAHUANG), *Perillae Folium* (ZISUYE), *Pheretima* (DILONG), *Cicadae Periostracum* (CHANTUI), *Arctii Fructus* (NIUBANGZI), *Schisandrae Chinensis Fructus* (WUWEIZI), *Peucedani Radix* (QIANHU), *Eriobotryae Folium* (PIPAYE), and *Perillae*

117 *Fructus* (ZISUZI), is mainly used to treat cough, throat itching, and other diseases
118 caused by "wind evil invading lung", especially for CVA and post-infectious cough
119 (PIC). Although Suhuang has significant clinical therapeutic effects, its
120 mechanisms of action remain unclear. Our laboratory has made progress in
121 previous studies on the MOAs of Suhuang for CVA [19,24-26]. However, despite this
122 progress, the main active ingredients, their specific therapeutic targets, and the
123 MOAs have still not been systematically clarified. Therefore, more comprehensive
124 research is needed to explain the efficacy of Suhuang for treating CVA.

125 In this study, we have introduced a novel network perturbation method known as
126 Herb-CMap. This approach is designed to effectively identify therapeutic targets
127 and bioactive ingredients within Traditional Chinese Medicine (TCM)
128 prescriptions. Herb-CMap achieves this by integrating transcriptional gene
129 perturbation data with Protein-Protein Interaction (PPI) networks. The schematic
130 representation of our study can be observed in Fig. 1 and Fig. S1, with the following
131 key steps: Initially, we extracted three crucial types of network graphs from
132 publicly available databases, including the TCM-ingredient-target, disease-gene
133 associations, and PPI networks. Subsequently, we employed a random walk
134 algorithm within the PPI network to enhance transcriptional changes. This
135 process assigned scores to each target. A higher perturbation score indicates a
136 significant proximity of the target to the perturbed genes. We also further
137 conducted molecular docking simulations and wet experiments to validate our
138 prioritized targets and their corresponding ingredients. We firmly believe that the
139 Herb-CMap methodology introduced in our study will provide invaluable insights
140 into unraveling the MOAs underlying complex herbal medicine by connecting the
141 gene perturbation of herbal medicine to its key therapeutical targets.

142 In this study, we have introduced an innovative network perturbation method,
143 Herb-CMap. This method is specifically crafted to adeptly identify therapeutic
144 targets and bioactive ingredients within herbal medicine. Herb-CMap
145 accomplishes this task by seamlessly integrating transcriptional gene
146 perturbation data with Protein-Protein Interaction (PPI) networks. The schematic

147 overview of our study is presented in Fig. 1 and Fig. S1, illustrating the following
148 pivotal steps: Initially, we extracted three essential types of network graphs from
149 publicly accessible databases, encompassing the herb-ingredient-target network,
150 disease-gene associations network, and PPI networks； subsequently, we applied
151 a random walk algorithm within the PPI network to amplify transcriptional
152 changes. This procedure assigned scores to each target, with a higher perturbation
153 score signifying a more significant proximity of the target to the perturbed genes;
154 to further validate our prioritized targets and their corresponding ingredients, we
155 use Suhuang for CVA as a case study to identify key target and active ingredients.
156 These findings were verified by both molecular docking simulations and wet
157 experiments, suggesting this approach effectively connects the gene perturbation
158 of herbal medicine to their crucial therapeutic targets, thereby enhancing our
159 understanding of their therapeutic efficacy. We strongly believe that the Herb-
160 CMap methodology introduced in our study will provide invaluable insights into
161 unraveling the Mechanisms of Action (MOAs) that underlie complex herbal
162 medicine.

163 **2. Materials and methods**

164 *2.1 Data collection*

165 The relevant ingredients were collected from six TCM databases, including the
166 Encyclopedia of Traditional Chinese Medicine (ETCM) [27], the Traditional Chinese
167 Medicine Systems Pharmacology (TCMSP) [28], the Traditional Chinese Medicine
168 Integrated Database (TCMID) [29], the Traditional Chinese Medicine on Immuno-
169 Oncology (TCMIO) database [30], the Traditional Chinese Medicine Information
170 Database (TCM-ID) [31] and Medical Subject Headings (MESH). Only compounds
171 that could be found in the PubChem database [32] were kept for further study.
172 Ingredient-target relationships from the STITCH database [33] and combined scores >
173 700 were kept for further study. Subsequently, the properties of components were
174 calculated by ADME-lab [34]. Disease-related vital targets were also collected. For
175 example, CVA, or "cough variant asthma," was utilized as a keyword to collect CVA-

176 related genes from the literature. Unlike associated genes from GeneCards [35],
177 OMIM [36], and DisGeNET [37], only those reported as target genes were kept for
178 further study.

179 The RNA-seq of a TCM prescription can be collected from GEO or manually
180 prepared by experiments. The differentially expressed genes (DEGs) analysis was
181 performed by Deseq2 package in R for further network perturbation calculation.

182 *2.2 DEG-weighted PPI network construction*

183 Protein-protein interactions (PPIs) were extracted from the STRING database [38].
184 The obtained targets were standardized to Entreze ID by the MyGene package in
185 Python.

186 First, the binary network was constructed by PPIs with matrix A_{N*N} with N genes;
187 $A_{i,j}$ is the connection between gene *i* and gene *j* with $A_{i,j} = 1$ if *i* and *j* are connected
188 else 0.

189 We obtained the log2 foldchange value of all the genes as a vector $F = (f_1, f_2, \ldots f_N)$
190 with f_i as the log2 foldchange value of a specific gene *i*. The DEG-weighted PPI
191 network (as matrix B) can be established by taking the |log2 foldchange| value of
192 genes as the weight of nodes and the average weight of two nodes as the weight of
193 edges.

194 *2.3 Proximity-weighted PPI network construction*

195 Functionally closely related targets tend to be closer to each other in the PPI
196 network, and the shortest path in the network can be used to quantify the
197 interactions of two targets in the DEG-weighted PPI network.

198 A proximity-weighted PPI network was constructed to quantify the interactions
199 between targets of herb, disease genes, and DEGs. The proximity network for DEGs
200 and disease genes can be defined as matrix $C_{M\times K}$. M is the set of DEGs, and K is the
201 set of the targets of TCM prescriptions. We define the distance between any node

DEG i in M and node j in K as C_{ij}:

$$C_{ij} = wsp(i,j) \quad (1)$$

where wsp is the Shortest_path function and it is calculated by the NetworkX package in Python. The pairwise weighted shortest path is calculated as a proximity-weighted PPI network.

Next, we construct our DEG-target association network from matrix $C_{M\times K}$ via Z score transformation according to the distance value. Z-scored C_{ij} is calculated for each value in C, and the *p*-value is expressed as Z:

$$\overline{C} = \frac{\sum_{i\in M}\sum_{j\in K}(C_{ij})}{\|M\|\|K\|} \quad (2)$$

$$Z - Score(C_{ij}) = \frac{C_{ij}-\overline{C}}{\sigma(C)} \quad (3)$$

Only target-target pairs with *p*-value < 0.025 were kept connected, and the final transformed proximity-weighted PPI network is defined as D.

$$D_{ij} = C_{ij}, if\ p\ value < 0.025\ else\ D_{ij} = 0 \quad (4)$$

2.4 Prioritization of the CVA treatment genes using the local radiality network perturbation method

The union potential targets from multiple ingredients were collected in the TCM prescriptions study. The hypothesis was that the critical targets of active ingredients would activate downstream pathways in a cascade, thus regulating the neighbor targets in the PPI network. On the other hand, the therapeutic disease-related genes would also be significantly regulated by the TCM prescription. They would thus be closely related to the DEGs in the PPI network.

Radiality is a well-known centrality measure describing the level of node reachability of a node via different shortest paths of the network. It was acknowledged that deregulated genes were closer to known targets than most proteins in the network. Based on this hypothesis, local radiality (LR) (Isik et al., 2015) was applied to prioritize the critical disease-associated genes by combining

perturbed genes and functional interaction network information. The LR score of node *i* in the network D is calculated as follows:

$$LR(i) = \frac{\sum_{dg \in DG} |sp(i, dg, G)|}{|DG|} \quad (5)$$

Here, the function |sp| calculates the weighted shortest path that connects the gene dg and the node n in D; |DG| indicates the total number of connected genes. The LR utilizes drug perturbation data (i.e., connected genes) and topological information (i.e., shortest path distance). Thus, it implements the hypothesis about the proximity of DEGs, targets, and disease genes. In our study, LR (n) represents the critical score of disease genes for causing gene perturbation of TCM prescription/herbal medicine.

2.5 Prioritization of the active ingredients and action targets using the random walk method

It was reported that a drug's targets and DEGs were typically closely connected in PPI within three steps [39]. Here, network perturbation methods are applied to infer the critical targets by aggravating the gene perturbation of all neighborhood DEGs to determine the active ingredients acting on these crucial targets and pathways.

We used a random walk with restart (RWR) algorithm called heat diffusion to screen an essential target query that interacted with DEGs closely. The RWR network perturbation algorithm is described as follows: Diffusion applies heat to each target as the heat source in the target-DEG network, and the heat flows through connecting edges to adjacent targets or DEGs. It then produces a list of targets ranked by the heat they accumulated. A target with many connections will have a higher rank, and an isolated target will tend to have a low rank, which has screened the key targets. The specific calculation formula is the following:

$$d = h * \exp(-Lt) \quad (6)$$

$$L = D - A \quad (7)$$

Where “h” is a vector representing the original query or signal. “d” is another

vector, which is the result of the diffusion process. The degree matrix D is a diagonal matrix where each diagonal element corresponds to the degree (number of connections) of a node in the graph. The adjacency matrix A encodes the connectivity between nodes. It represents how the original signal has spread or diffused over the network. L is the Laplacian matrix of the network graph and is defined as the difference between the degree matrix (D) and the graph's adjacency matrix (A). The expression “exp(*)” is the matrix exponential . “t” is a scalar parameter representing the total time of diffusion. It controls how far and quickly the original signal spreads over the network. A larger value of t will allow the signal to spread further, while a smaller value will limit its spread. Here, we conduct a diffusion process on a graph target-DEG network, where “h” represents the transcript gene perturbation values of nodes, and the Laplacian matrix is used to model the diffusion behavior. The parameter 't' controls the extent and speed of the diffusion. The result of this process is the vector 'd,' which represents the gene perturbation score of targets and DEGs after diffusion. We use a default value of t = 0.1 [40].

2.6 Case study: Identifying the active ingredients and MOAs of Suhuang for CVA

Male Wistar rats were intraperitoneally injected with ovalbumin (OVA) to establish the CVA rat model [26]. The rats from the experimental group were intragastrically administrated 7 g/kg of Suhuang crude drug for two weeks. Subsequently, the rats were sacrificed by euthanasia. The lungs from the control, CVA, and CVA+Suhuang groups were collected and subjected to transcriptome sequencing.

The Deseq2 package in R. Gene obtained the differentially expressed genes (DEGs) analysis set enrichment analysis (GSEA) based on Go Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) datasets to understand these DEGs' biological functions and related pathways. RcisTarget transcription factor (TF) analysis was conducted to find the critical genes to determine the key DEGs' specific regulated mechanisms. Moreover, the Connectivity Map (CMap) was

284 adopted to excavate the functional connections between CVA and the genetic
285 perturbation of Suhuang [41]. In order to study whether the gene expression of DEGs
286 in different tissues is consistent with the meridian classifications of Suhuang in
287 TCM theory, the gene expression of DEGs at the transcription level was extracted
288 from the Human Protein Atlas databases. More detail about the CVA model and
289 average RNA data analysis for Suhuang can be seen in **Supplementary Note 1**.
290 Finally, the Herb-CMap method was used to prioritize Suhuang's crucial genes and
291 targets for CVA.

292 *2.7 Verifying the critical targets and active ingredients*

293 *Verification of targets and pathways:* RAW264.7 cells were plated in 6-well plates
294 at a density of 2 × 10^6 cells and pretreated with different concentrations of
295 Suhuang (0.1 μg/ml, 0.2 μg/ml, 0.4 μg/ml), and then stimulated with 1 μg/ml LPS
296 for six h. Total RNA was isolated from cells using Trizol reagent (Takara, Otsu,
297 Japan) according to the manufacturer's protocol. RNAs were reverse transcribed
298 to cDNAs for quantitative PCR. The primer sequences were as follows:

299 **Table 1. Sequences of the primers used for real-time PCR.**

Primer name	Forward primer sequences	Reverse primer sequences
Mouse IL-17A	CTGATGCTGTTGCTGCTGCTG	ACACGCTGAGCTTTGAGGGATG

300 *Validation of active ingredient and target interaction:* We used AutoDock Vina for
301 molecular docking [42]. According to the scores of MCODE [43], the core targets of
302 Suhuang for treating CVA were selected. Then, their crystal structures were
303 downloaded from the RCSB Protein Data Bank (PDB) [44]. Next, these crystal
304 structures were modified by PyMol to remove the original ligands and water
305 molecules and then saved in the PDBQT format. The three-dimensional structures
306 of related compounds were downloaded from PubChem. Then, the bond energy of
307 the compound ligand molecules was processed via Chem3D 15.0 to reduce
308 intermolecular repulsion and saved in the PDB format. Finally, the three-

309 dimensional crystal structures of target proteins and corresponding compound
310 ligands were introduced into AutoDock Vina.

311 *Verification of active ingredients:* RAW264.7 cells were plated in 96-well plates at
312 a density of 1×10^4 cells. RAW264.7 cells were treated with LPS (1 μg/mL) and
313 different concentrations of luteolin, quercetin, and Dexamethasone (2.5, 5, 10, 20,
314 40, 80 μM). Then, the cells were incubated for 24 h, and the content of NO in cell
315 culture supernatant was determined by using the Griess reaction. The NO
316 inhibition rate of different drugs was calculated according to the following formula:

317 $$Inhibitory\ ratio(\%) = \frac{C_{drug} - C_{blank}}{C_{LPS} - C_{blank}} \times 100$$

318 where C_{blank} is the content of NO in unprocessed cell culture, C_{LPS} is the content of
319 NO in LPS-induced cell culture, and C_{drug} is the content of NO in cell culture treated
320 with LPS and different concentrations of drugs. The IC_{50} values for luteolin,
321 quercetin, and Dexamethasone were determined on three replicates of 6
322 concentrations. Statistical analysis and the IC_{50} values were calculated using
323 GraphPad Prism version 8.0.2 software from GraphPad Software Incorporated.

324 **3. Results**

325 *3.1 Data statistics on the herb-ingredient-target network of Suhuang*

326 In total, 1,377 ingredients were collected from six TCM databases. The overlapping
327 ingredients among these six databases are shown in Fig. 2A. The number of
328 ingredients varied among different databases, with 849 ingredients from TCMIO,
329 followed by TCMSP, and only 250 ingredients from ETCM. Furthermore, the TCMSP
330 and TCMIO databases shared the most common compounds (n = 406). Primarily,
331 we found only 14 compounds appearing in all six databases, suggesting the
332 necessity of data integration to address the data complexity.

333 The ingredients in Suhuang are derived from nine herbs, with $N_{MAHUANG}$ = 452,
334 $N_{WUWEIZI}$ = 323, N_{PIPAYE} = 240, N_{QIANHU} = 152, N_{ZISUYE} = 144, N_{ZIZUZI} = 135, $N_{NIUBANGZI}$ =
335 89, $N_{CHANTUI}$ = 15, and N_{DILONG} = 1(Fig. 2B). MAHUANG shared the most common

336 components with PIPAYE and WUWEIZI (n = 21 and 21 separately). Except for two
337 animal drugs, DILONG and CHANTUI, the other seven herbs have a high percentage
338 of compounds with good OB and DL ranging from around 80% to 95% (Fig. 2C).
339 As shown in Fig. 2D, 60.6% of ingredients were predicted to bind with less than
340 five targets of the compounds. Only 63 ingredients owned more than ten targets.

341 *3.2 DEG analysis and their tissue distribution*

342 To grasp the gene perturbation of Suhuang for CVA, RNA-seq data were tested in
343 three groups: control, CVA, and Suhuang. Compared to the control group, a total of
344 589 genes were identified as differentially expressed genes (DEGs, *p*.adjust < 0.05
345 and |log2 foldchange| > 2) for the CVA group, with 465 upregulated and 124
346 downregulated (Fig. 3A). Furthermore, a total of 790 genes were identified as
347 DEGs when comparing the Suhuang group to the CVA group, with 300 upregulated
348 and 490 downregulated (Fig. 3B). In total, 338 common DEGs were identified
349 between CVA/Control and Suhuang/CVA (Fig. 3C), suggesting that these
350 abnormal DEGs in the disease condition returned to the normal level after the
351 treatment of Suhuang (Fig. 3D).

352 Fig. 4A shows the distribution of TCM properties of the nine herbs in Suhuang,
353 including four natures, five flavors, 12 kinds of meridian classifications, and six
354 main efficacies. Regarding the nature of Suhuang, it has warm and cold medicinal
355 characteristics. Regarding the flavors, pungent was the most frequent. In TCM
356 theory, the pungent flavor tended to affect the surface and upper parts of the
357 human body. Furthermore, CVA is a respiratory disease. Accordingly, all herbs in
358 Suhuang were classified to the lung meridian except DILONG. These herbs,
359 especially MAHUANG and CHANTUI, have been reported to have extensive effects,
360 such as refrigeration, analgesia, relieving wheeling, immunological enhancements,
361 immuno-suppression, and hypoglycemia.

362 Considering that the meridian classifications in the TCM theory might provide a
363 reasonable interpretation of Suhuang's therapeutic effects on CVA, we further

364 investigated whether the tissue distributions of these DEGs at the transcription
365 level were consistent with the meridian distribution of Suhuang from TCM theory.
366 As shown in Fig. 4B, the lung was one of the top 10 highest averaged expressed
367 tissues, and the top 10 tissues were not significantly different. On the other hand,
368 except for the top 10 genes, the expression of DEGs in normal human lung tissue
369 was significantly higher than the other 53 tissues. Interestingly, the expression of
370 the DEGs in the liver and urinary bladder tissues was the highest, which was
371 consistent with the liver meridian of CHANGTUI and DILONG and the bladder
372 meridian of MAHUANG and DILONG (Fig. 4A). Taken together, we verified that the
373 results of gene perturbation of Suhuang were in line with the meridian
374 classifications in TCM theory and that Suhuang for treating CVA was closely related
375 to the functional regulation in the lung tissue.

376 *3.3 Identifying CVA-related pathways* via *pathway enrichment analysis*

377 Three enrichment analysis methods, GO, KEGG, and GESA, were performed to
378 systematically grasp these DEGs' biological functions. We also utilized the CMap
379 and TF analysis to identify the underlying molecular mechanisms further.

380 In total, 263 GO-biological processes (GO: BP) terms and 26 KEGG pathways were
381 enriched significantly by GO and KEGG analysis, respectively (p.adjust value < 0.05,
382 Table S1-2). We also mapped the DEGs onto GSEA analysis for GO: BP gene sets,
383 and the EnrichmentMap plugin in Cytoscape was utilized to identify the
384 hierarchical relationships of these pathways (Table S3). Consequently, the top
385 enriched GO: BP terms for Suhuang were related to inflammatory cytokines in the
386 immune system, such as response to interleukin-1, response to interferon-gamma,
387 and negative regulation of the immune system process (Fig. 3E-F). Therefore, it
388 appears that Suhuang might relieve CVA by extensively regulating the pathways
389 associated with the immune system. According to KEGG pathway enrichment
390 analysis, half of the top 20 significantly enriched pathways were associated with
391 CVA, such as the Asthma pathway, IL-17 signaling pathway, TNF signaling pathway,
392 and Toll-like receptor signaling pathway (Fig. 5A).

393 Meanwhile, we mapped the DEGs onto the CMap database to find similar gene
394 signatures. The top and bottom 10 MOAs, respectively, were illustrated in Fig. 5B.
395 The MOAs in Suhuang for treating CVA were positively associated with inhibiting
396 PI3K and negatively associated with activating NF-κB, suggesting that Suhuang
397 may treat CVA through inhibiting PI3K and NF-κB signaling pathway (Fig. 5B, Fig.
398 S2).

399 Interestingly, both KEGG and GSEA showed that the IL-17 signaling pathway was
400 a significant enrichment pathway in Suhuang for treating CVA (Fig. 5A and C, Table
401 S4). However, it was studied that the IL-17 signaling pathway was closely related
402 to Th1 airway inflammation in asthma, while CVA was generally recognized as Th2
403 airway inflammation [45-47]. Therefore, we inferred that Suhuang may have other
404 therapeutic effects on Th1 asthma (Fig. 5D). The NF-κB pathway from GSEA and
405 CMap analysis was also significantly enriched (Fig. 5E and F). In brief, multiple
406 pathway enrichment methods indicated that Suhuang may regulate these
407 pathways to treat CVA in model rats.

408 According to transcription factor (TF) analysis, 86 motifs were significantly
409 enriched. The result for the top three TFs based on motifs was listed in Fig. 5G.
410 RELA was an essential member of the NF-κB family, with 29 DEGs enriched in this
411 related motif. BCL3, with 31 DEGs enriched in this related motif, was highly similar
412 to the IκB family of NF-κB inhibitors [48,49]. These two TFs participated in many
413 cellular responses, such as cell proliferation and transformation, apoptosis,
414 inflammation, and immune response, by binding NF-κB or other TFs and
415 regulating target genes (Fig. 5H). These results were consistent with our previous
416 findings about the critical role of the NF-κB signaling pathway in Suhuang for
417 inhibiting non-resolving inflammation and treating acute lung diseases. Our
418 previous work found that Suhuang could relieve acute lung injury by inhibiting the
419 NF-κB signaling pathway [19].

3.4 Prioritization of the therapeutic genes using the local radiality network perturbation method

Generally speaking, the weighted shortest path (S) between DEGs and CVA genes was negatively correlated with the corresponding |log2foldchange|, suggesting that the more significantly the DEGs expressed, the more functionally close it was to CVA genes (Fig. 6A) in PPI. In addition, DEGs could be classified into three clusters by hierarchical clustering analysis according to the weighted shortest path among 54 disease genes. The weighted shortest path between the CVA genes and DEGs from cluster Ⅱ (average $|\log2\ \text{foldchange}|_{II}$=4.086, average S_{II} = 1.977) was significantly lower than that of the other two cluster DEGs (average $S_{I,\ III}$ = 5.025 and 3.172, average $|\log2\ \text{foldchange}|_{I,\ III}$=2.50,3.10) (Fig. 6B, Fig. 6C), suggesting the importance of the DEGs from cluster Ⅱ. Therefore, further pathway enrichment analysis was performed based on DEGs of cluster Ⅱ. We found that the enriched pathways were similar to those from all the DEGs, suggesting that cluster Ⅱ DEGs could characterize Suhuang perturbation and could be considered as signature genes (Fig. 6D). These results indicated that the weighted proximity network based on DEGs could reveal the vital therapeutic genes of Suhuang for treating CVA.

Then, the LR of network perturbation was leveraged to quantify the importance of disease genes. As shown in Figure 6G, *Tnf* obtained the highest importance score, 0.47, followed by *Tpsb2* (0.48) and *Mmp12* (0.48). *Tnf* was enriched in the IL-17 signaling pathway and was a crucial gene. *Tpsb2* has been implicated as a mediator in the pathogenesis of asthma and other allergic and inflammatory disorders [50]. *Mmp12*, which encodes a member of the peptidase M10 family of matrix metalloproteinases (MMPs), may play a role in aneurysm formation, and mutations in this gene are associated with lung function and chronic obstructive pulmonary disease (COPD) [51].

Il17a (score = 1.00) was significantly enriched in the IL-17 signaling pathway, and *Pik3cd* (score = 0.92), enriched in the PI3K-Akt signaling pathway, was also

449 prioritized using our network perturbation method.

450 Taken together, 15 genes involved in IL-17 and PI3K-Akt signaling pathways were
451 prioritized using network perturbation methods.

452 *3.5 Prioritization of the critical targets and corresponding ingredients using the*
453 *random walk with restart network perturbation method*

454 First, 928,485 target-DEG interactions (T-Gs) were calculated as proximity matrix
455 C_{MxN}. Furthermore, 19,009 T-Gs were identified as significant interactions and kept
456 for the proximity network construction ($z < 0$, $p < 0.025$), covering 694 targets (Fig.
457 S4A). We introduced these target genes into the KEGG pathway enrichment
458 analysis. IL-17, PI3K-Akt, and NF-κB signaling pathways were significantly
459 enriched ($p < 0.01$) (Table S1).

460 With target-DEG interactions, the random walk algorithm was performed to find
461 the key targets. Fig. 7A illustrates the top 20 targets with the highest heat values.
462 Primarily, we found that *Il13*, *Rela* enriched in the IL-17 signaling pathway and
463 *Ywhab*, *Rela*, *Pik3cb* enriched in the PI3K-Akt signaling pathway are among the top
464 20 targets (Fig. 7A), suggesting Suhuang may influence these two pathways to
465 relieve CVA.

466 Interestingly, 499 of 694 targets were not significantly changed at the
467 transcription level and were kept in our proximity network. More importantly,
468 according to the heat values from network perturbation, 16 of the top 20 targets
469 were not DEGs, such as *Rela*, *Ywhab*, and *Pik3cb* (Fig. 7A). These results suggested
470 that our network inference methods could reveal potential targets that were not
471 significantly expressed at the transcription level. We also investigated the
472 functions of these 499 targets via KEGG pathway enrichment. We found that the
473 PI3K-Akt and IL-17 signaling pathways were significantly enriched again (Fig.
474 S4B). Forty-five targets were enriched in the PI3K-Akt signaling pathway, and 23
475 were enriched in the IL-17 signaling pathway.

476 *3.6 Experimental validation of the active ingredients and corresponding targets.*

477 To identify active ingredients, 23 and 45 targets enriched in IL-17 and PI3K-Akt
478 signaling pathways were selected as critical targets. Correspondingly, molecular
479 docking was performed on the 17 and 21 active compounds predicted to interact
480 with these targets (Fig. 7B-C), respectively. The enriched targets and
481 corresponding ingredients, as well as binding interactions, were shown in Fig. 7D.
482 In particular, quercetin represented low binding energy with *Il17a*, *Pik3cb*, *Pik3cd*,
483 and *Akt1* as -7.4, -7.6, -8.8, -6.9 kcal/mol, respectively. In addition, luteolin also
484 showed stable docking results with *Tnf* and *Akt1,* whose binding energy values
485 were -8.7 and -6.8 kcal/mol, respectively. The above results illustrated that
486 quercetin and luteolin could alleviate lung inflammation by regulating IL-17 and
487 PI3K-Akt signaling pathways and might be the main active compounds of Suhuang
488 for treating CVA.

489 More importantly, according to Real-time PCR, we found that the mRNA expression
490 of *Il17a* was reduced by Suhuang significantly (Fig. 8A). Quercetin and luteolin
491 showed lower IC_{50} than positive control drug Dexamethasone (Fig. 8B).

492 Finally, we propose a novel MOA of Suhuang for treating CVA (Fig. 8C). We found
493 that PI3K-Akt, NF-κB, and IL-17 signaling pathways were involved in the
494 pathophysiological process of CVA. On the one hand, Phosphoinositide 3-kinases
495 (PI3Ks) are involved in the immune response. The lipid products of PI3K activate
496 Akt and then release NF-κB from the cytoplasm by phosphorylating a variety of
497 TFs such as RELA and BCL3 [52-54]. Activated NF-κB can induce many
498 proinflammatory cytokines and produce lung inflammation. On the other hand,
499 upon exposome stimulations, bronchial epithelial cells release IL-25, IL-33, and
500 thymic stromal lymphopoietin. These cytokines promote the proliferation of Th17
501 cells and drive them to release IL-17, leading to lung inflammation [55,56]. Therefore,
502 we were able to predict the MOAs of the primary active compounds in Suhuang
503 during the treatment of CVA: quercetin and luteolin could directly inhibit the IL-
504 17 signaling pathway and simultaneously inactivate PI3K, inhibiting Akt and NF-

κB, and thereby preventing the occurrence of lung inflammation and treating CVA. This multi-compound, multi-target, and multi-pathway treatment of CVA provided a scientific explanation for the drug combination of Suhuang.

3.7 Association between TCM properties and the proximity distance of their targets with DEGs

The target-DEG shortest path of different targets varied from 0 to 1.287 (Fig. S4A), indicating that the interactions of these targets with DEGs differed. As the targets were predicted to interact with different ingredients from the nine herbs in Suhuang, we investigated whether the shortest path of target-DEG from different components was related to their respective herbs and properties (Fig. 9A).

Generally speaking, the target-DEG shortest path of targets from Zisuye was significantly lower than that of targets from other herbs (Fig. 9B). Additionally, Zisuye also tended to have more significant target-DEG shortest path in the transformed proximity network than other herbs, according to Fisher's exact test (*p-value* = 3.56×10^{-64}) (Fig. 10A). These results revealed the essential roles of targets from Zisuye, which were consistent with our prior knowledge that Zisuye was the monarch ("JUN") medicine of Suhuang based on TCM theory. In terms of TCM properties, we found that the targets of cold nature, the pungent flavor, and the lung and spleen meridian tended to have significantly lower proximity with DEGs (*p-value* < 10^{-4}) as well as more significant target-DEG pairs (Fisher's exact test, *p*- value< 10^{-4}) when compared to their counterparts in TCM theory (Fig. 10A).

We found that the proximity of target-CVA gene and target-interesting pathway gene interactions was significantly lower than that of target-DEG interactions, suggesting that the significant targets we kept for network perturbation interacted strongly with CVA genes and were more likely to be the therapeutic targets (Fig. 10B). In particular, we identified six exciting targets which were not only DEGs of Suhuang but also related to CVA, including *Il17a*, *Tnf*, *Ifng*, *Il10*, *Ccr4*, and *Ido1*. Thus, we believe that Suhuang's ingredients predicted to interact with these six targets

533 deserve further verification (Fig. 10C).

534 More importantly, we found that targets from the lung and large intestine meridian
535 exhibited significantly lower proximity in the transformed proximity network
536 simultaneously (Fig. 10C). This reminded us that the lung and large intestine
537 meridian played essential roles in Suhuang. This was consistent with the TCM
538 theory of the "lung large intestine connection" [57]. Specifically, based on our
539 previous work, Suhuang could alleviate airway inflammation-associated lung
540 injury by promoting Hepatocyte growth factor (HGF) secretion in an intestinal
541 endocrine-dependent pathway. Moreover, the critical role of the lung meridian was
542 in line with the tissue distribution of DEGs in the lung (Figure 4B).

543 In summary, by comparing the proximity of target-DEG interactions from the
544 different herbs and properties of Suhuang, we found that the properties of
545 Suhuang based on TCM theory could explain its mechanisms of treating CVA to
546 some extent. We also reflected on the multi-component, multi-target, and multi-
547 pathway characteristics of herbal medicine.

548 **4. Discussion**

549 Identifying the bioactive ingredients and critical targets of herbal medicine is
550 difficult due to the lack of target information. Thanks to the development of
551 computational biology and multi-omic methods, these large-scale data at different
552 levels allow us to explore the scientific connotation of herbal medicine for treating
553 diseases. To understand the molecular MOAs of herbal medicine treatments at the
554 phenotype level, we proposed a novel method, Herb-CMap, to infer the therapeutic
555 targets of herbal medicine using a gene perturbation network and the random
556 walk algorithm. After finding potential targets, we worked backward to identify
557 the bioactive ingredients according to the predicted ingredient-target pairs.

558 Herb-CMap completed the multiple data fusions required to uncover the MOAs of
559 herbal medicine, such as PPI, gene profiling, disease genes associations, and herb-
560 ingredient-targets. More importantly, advanced network proximity and network

561 perturbation were used to integrate all the data. Our results showed that our
562 network-based inference method could help to capture some bioactive targets that
563 were not observed as significantly changed genes at the transcript level. For
564 example, *Pik3cb*, a non-differentially expressed gene with a high rank by Diffusion,
565 was closely interactive with adjacent target genes and showed a strong affinity
566 with quercetin in Suhuang. Moreover, network proximity methods could help
567 uncover the underlying association by considering the global topological
568 characteristics. By calculating the network proximity of CVA-DEG, we quantified
569 the importance of CVA-related genes in the gene perturbation of Suhuang. By
570 calculating the target-DEG distance, we quantified the importance of targets of
571 Suhuang for gene perturbation. However, although the "shortest step" network
572 distance was applied in the previous model, the nodes and link attributions from
573 the PPI network were not considered. To further optimize the algorithm, it is
574 necessary to integrate the diversity of biological networks.

575 Herb-CMap can accurately predict targets and active ingredients, further
576 accelerating the MOA exploration of herbal medicine and drug discovery from
577 herbal medicine. In this case study, our method accurately predicted the known
578 vital targets and the active ingredients of Suhuang and uncovered one novel
579 ingredient. We successfully identified 17 active ingredients acting on the IL-17
580 signaling pathway and 21 working on the PI3K-Akt-NFκB signaling pathway, such
581 as decursin and rosmarinic acid. It was reported that quercetin and luteolin could
582 treat allergic asthma by acting as critical targets related to PI3K-Akt and IL-17
583 signaling pathways, which was also verified in vivo [58,59]. Decursin could improve
584 inflammatory diseases by regulating multiple pathways, including the PI3K-Akt-
585 NFκB signaling pathway [60]. Rosmarinic acid could also alleviate airway
586 inflammation by inhibiting NFκB activation [61]. Currently, few studies are available
587 about the therapeutic effects of these compounds on asthma by regulating the IL-
588 17 signaling pathway. In addition, Herb-CMap also helped to discover a novel MOA:
589 quercetin and luteolin could directly inhibit the IL-17 signaling pathway and
590 simultaneously inactivate PI3K, inhibiting Akt and NF-κB, thereby preventing the
591 occurrence of lung inflammation and treating CVA. In summary, Herb-CMap

provides a novel insight into the deep understanding of how the complex systems of TCM prescriptions treat diseases by the MOAs of multi-compound, multi-target, and multi-pathway.

We also proposed a novel paradigm for deciphering TCM theories by investigating the relationships between our prioritized targets and the distributions of these properties. TCM theory was essential in using TCM formulae, such as nature, flavors, and meridians in TCM properties. First, for meridian theory, we found that the tissue expression attendance of DEGs followed the lung meridian of Suhuang. Second, Zisuye had closer interactions with DEGs than other herbs that are considered to play significant therapeutic effects, which is in line with its monarch role in TCM theory. Studies have shown that the extraction of Zisuye could effectively alleviate airway inflammation and high airway responsiveness in the asthma mouse model induced by ovalbumin [62,63]. Specifically, Zisuye contains various active ingredients, including flavonoids and phenolic acids such as luteolin and rosmarinic acid, with anti-inflammatory activity [64,65]. These ingredients in Zisuye may be the therapeutic material basis in Suhuang for treating CVA. Third, in terms of the "JUN-CHEN-ZUO-SHI" theory of combination, we found that NIUBANGZI and WUWEIZI had significant target-DEG shortest path according to Fisher's exact test, with p values 1.63×10^{-53} and 2.08×10^{-18}, respectively. Accordingly, this also suggested that the active compounds corresponding to the crucial targets from NIUBANGZI and WUWEIZI had therapeutic roles in Suhuang for treating CVA, which was also consistent with our prior knowledge that NIUBANGZI and WUWEIZI were the minister ("CHEN") medicines based on TCM theory. Additionally, the main active components of NIUBANGZI and WUWEIZI, such as arctigenin and schizandrin, were found to relieve airway inflammation and treat asthma [66,67]. Fourth, the results of the property distributions of Suhuang's targets based on TCM theory were also consistent with modern scientific research. If the transformed proximity network was noticed, the proximity of targets of the pungent flavor was also lower than that of other flavors (Fig. 9B). It had previously been reported that transient receptor potential (TRP) channels were responsible for various sensory reactions, including heat, cold, pain, stress, vision, and taste [68].

For example, TRP channel subfamily V-1 (TRPV1) was sensitive to capsaicin interrelated with the pungent flavor, and TRP channel subfamily M-8 (TRPM8) was related to the cold sensory [69-71]. They also played a role in respiratory diseases, nervous and mental diseases, diabetes, and cancer. Therefore, the property classifications of Suhuang might be associated with regulating these proteins. To sum up, our Herb-CMap could help us better understand TCM theory in formulae at the molecular level[72].

There are some limitations in this study. For example, our network perturbation method considered first-level neighborhood genes in PPI, namely LR. Many other advanced network perturbation methods are worth trying.

In summary, this study addresses the challenge of identifying bioactive ingredients and key targets in herbal medicine due to limited target information. Leveraging computational biology and multi-omic methods, we introduce a novel approach called Herb-CMap to understand the molecular mechanisms of herbal medicine treatments. Herb-CMap integrates various data sources, including protein-protein interactions, gene profiling, disease-gene associations, and herb-ingredient-targets, employing network proximity and perturbation methods to uncover therapeutic targets not detected at the transcript level. This approach accurately predicts targets and active ingredients, facilitating the exploration of herbal medicine mechanisms and drug discovery, exemplified by identifying active ingredients in Suhuang for treating Cough Variant Asthma and offering insights into TCM theory.

Acknowledgments

This work was supported by the "Double First-Class" University Project (CPU2018GF05) from China Pharmaceutical University and the Program for Innovative Research Team of Jiangsu Province.

Reference

1 Kim, H. U., Ryu, J. Y., Lee, J. O. & Lee, S. Y. A systems approach to traditional oriental medicine. *Nature biotechnology* **33**, 264-268, doi:10.1038/nbt.3167 (2015).

2 Zhang, R., Li, X., Zhang, X., Qin, H. & Xiao, W. Machine learning approaches for elucidating the biological effects of natural products. *Nat Prod Rep* **38**, 346-361, doi:10.1039/d0np00043d (2021).

3 Wang, Y., Jafari, M., Tang, Y. & Tang, J. Predicting Meridian in Chinese traditional medicine using machine learning approaches. *PLoS computational biology* **15**, e1007249, doi:10.1371/journal.pcbi.1007249 (2019).

4 Zhang, R., Zhu, X., Bai, H. & Ning, K. Network Pharmacology Databases for Traditional Chinese Medicine: Review and Assessment. *Frontiers in pharmacology* **10**, 123, doi:10.3389/fphar.2019.00123 (2019).

5 Jafari, M., Wang, Y., Amiryousefi, A. & Tang, J. Unsupervised Learning and Multipartite Network Models: A Promising Approach for Understanding Traditional Medicine. *Frontiers in pharmacology* **11**, 1319, doi:10.3389/fphar.2020.01319 (2020).

6 Menche, J. *et al.* Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science (New York, N.Y.)* **347**, 1257601, doi:10.1126/science.1257601 (2015).

7 Tian, S. *et al.* Exploring pharmacological active ingredients of traditional Chinese medicine by pharmacotranscriptomic map in ITCM. *Brief Bioinform* **24**, doi:10.1093/bib/bbad027 (2023).

8 Cheng, F., Kovács, I. A. & Barabási, A. L. Network-based prediction of drug combinations. *Nature communications* **10**, 1197, doi:10.1038/s41467-019-09186-x (2019).

9 Wang, Y., Yang, H., Chen, L., Jafari, M. & Tang, J. Network-based modeling of herb combinations in traditional Chinese medicine. *Briefings in bioinformatics* **22**, doi:10.1093/bib/bbab106 (2021).

10 Gu, S. & Lai, L. H. Associating 197 Chinese herbal medicine with drug targets and diseases using the similarity ensemble approach. *Acta Pharmacol Sin* **41**, 432-438, doi:10.1038/s41401-019-0306-9 (2020).

11 Luo, Z. *et al.* Integrated Systems Pharmacology and Surface Plasmon Resonance Approaches to Reveal the Synergistic Effect of Multiple Components of Gu-Ben-Ke-Chuan Decoction on Chronic Bronchitis. *J Inflamm Res* **14**, 1455-1471, doi:10.2147/jir.S303530 (2021).

12 Theodoris, C. V. *et al.* Transfer learning enables predictions in network biology. *Nature* **618**, 616-624, doi:10.1038/s41586-023-06139-9 (2023).

13 Selvaraj, N. *et al.* Network-based Drug Repurposing: A Critical Review. *Curr Drug Res Rev* **14**, 116-131, doi:10.2174/2589977514666220214120403 (2022).

14 Ma, X. *et al.* Bioinformatics-assisted, integrated omics studies on medicinal plants. *Brief Bioinform* **21**, 1857-1874, doi:10.1093/bib/bbz132 (2020).

15 Musa, A. *et al.* A review of connectivity map and computational approaches in pharmacogenomics. *Briefings in bioinformatics* **19**, 506-523, doi:10.1093/bib/bbw112 (2018).
16 Xiao, Y. *et al.* Novel Drug Candidate Prediction for Intrahepatic Cholangiocarcinoma via Hub Gene Network Analysis and Connectivity Mapping. *Cancers* **14**, doi:10.3390/cancers14133284 (2022).
17 Isik, Z., Baldow, C., Cannistraci, C. V. & Schroeder, M. Drug target prioritization by perturbed gene expression and network information. *Scientific reports* **5**, 17417, doi:10.1038/srep17417 (2015).
18 Zhu, X. *et al.* Multi-omics approaches for in-depth understanding of therapeutic mechanism for Traditional Chinese Medicine. *Frontiers in pharmacology* **13**, 1031051, doi:10.3389/fphar.2022.1031051 (2022).
19 Qin, W. *et al.* Preservation of mitochondrial homeostasis is responsible for the ameliorative effects of Suhuang antitussive capsule on non-resolving inflammation via inhibition of NF-κB signaling and NLRP3 inflammasome activation. *J Ethnopharmacol* **271**, 113827, doi:10.1016/j.jep.2021.113827 (2021).
20 Gocho, Y. *et al.* Network-based systems pharmacology reveals heterogeneity in LCK and BCL2 signaling and therapeutic sensitivity of T-cell acute lymphoblastic leukemia. *Nature cancer* **2**, 284-299, doi:10.1038/s43018-020-00167-4 (2021).
21 Zhang, W. *et al.* Integrated strategy for accurately screening biomarkers based on metabolomics coupled with network pharmacology. *Talanta* **211**, 120710, doi:10.1016/j.talanta.2020.120710 (2020).
22 Padem, N. & Saltoun, C. Classification of asthma. *Allergy and asthma proceedings* **40**, 385-388, doi:10.2500/aap.2019.40.4253 (2019).
23 Diab, N., Patel, M., O'Byrne, P. & Satia, I. Narrative Review of the Mechanisms and Treatment of Cough in Asthma, Cough Variant Asthma, and Non-asthmatic Eosinophilic Bronchitis. *Lung* **200**, 707-716, doi:10.1007/s00408-022-00575-6 (2022).
24 Qin, W. *et al.* Suhuang antitussive capsule inhibits NLRP3 inflammasome activation and ameliorates pulmonary dysfunction via suppression of endoplasmic reticulum stress in cough variant asthma. *Biomed Pharmacother* **118**, 109188, doi:10.1016/j.biopha.2019.109188 (2019).
25 Tong, X. *et al.* Suhuang Antitussive Capsules-Ameliorative Effects on LPS-Induced Sputum Obstruction in Mice Through Promoting HGF Secretion. *Front Pharmacol* **10**, 1422, doi:10.3389/fphar.2019.01422 (2019).
26 Jiang, H. *et al.* β-Hydroxybutyric acid upregulated by Suhuang antitussive capsule ameliorates cough variant asthma through GSK3β/AMPK-Nrf2 signal axis. *J Ethnopharmacol* **307**, 116013, doi:10.1016/j.jep.2022.116013 (2023).
27 Xu, H. Y. *et al.* ETCM: an encyclopaedia of traditional Chinese medicine. *Nucleic Acids Res* **47**, D976-d982, doi:10.1093/nar/gky987 (2019).
28 Ru, J. *et al.* TCMSP: a database of systems pharmacology for drug discovery

from herbal medicines. *J Cheminform* **6**, 13, doi:10.1186/1758-2946-6-13 (2014).

29 Zhang, L. X. *et al.* TCMSID: a simplified integrated database for drug discovery from traditional chinese medicine. *Journal of cheminformatics* **14**, 89, doi:10.1186/s13321-022-00670-z (2022).

30 Liu, Z. *et al.* TCMIO: A Comprehensive Database of Traditional Chinese Medicine on Immuno-Oncology. *Front Pharmacol* **11**, 439, doi:10.3389/fphar.2020.00439 (2020).

31 Huang, L. *et al.* TCMID 2.0: a comprehensive resource for TCM. *Nucleic Acids Res* **46**, D1117-d1120, doi:10.1093/nar/gkx1028 (2018).

32 Wang, Y. *et al.* PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* **37**, W623-633, doi:10.1093/nar/gkp456 (2009).

33 Szklarczyk, D. *et al.* STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* **44**, D380-384, doi:10.1093/nar/gkv1277 (2016).

34 Dong, J. *et al.* ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *J Cheminform* **10**, 29, doi:10.1186/s13321-018-0283-x (2018).

35 Safran, M. *et al.* GeneCards Version 3: the human gene integrator. *Database (Oxford)* **2010**, baq020, doi:10.1093/database/baq020 (2010).

36 Hamosh, A. *et al.* Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* **30**, 52-55, doi:10.1093/nar/30.1.52 (2002).

37 Piñero, J. *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* **48**, D845-d855, doi:10.1093/nar/gkz1021 (2020).

38 Szklarczyk, D. *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**, D638-d646, doi:10.1093/nar/gkac1000 (2023).

39 Jaaks, P. *et al.* Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature* **603**, 166-173, doi:10.1038/s41586-022-04437-2 (2022).

40 Carlin, D. E., Demchak, B., Pratt, D., Sage, E. & Ideker, T. Network propagation in the cytoscape cyberinfrastructure. *PLoS Comput Biol* **13**, e1005598, doi:10.1371/journal.pcbi.1005598 (2017).

41 Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e1417, doi:10.1016/j.cell.2017.10.049 (2017).

42 Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-461, doi:10.1002/jcc.21334 (2010).

43 Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**, 2, doi:10.1186/1471-2105-4-2 (2003).

44 Burley, S. K. *et al.* RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* **49**, D437-d451, doi:10.1093/nar/gkaa1038 (2021).

45 Kuruvilla, M. E., Lee, F. E. & Lee, G. B. Understanding Asthma Phenotypes, Endotypes, and Mechanisms of Disease. *Clin Rev Allergy Immunol* **56**, 219-233, doi:10.1007/s12016-018-8712-1 (2019).

46 Agache, I. *et al.* Advances and highlights in asthma in 2021. *Allergy* **76**, 3390-3407, doi:10.1111/all.15054 (2021).

47 Niimi, A. Narrative Review: how long should patients with cough variant asthma or non-asthmatic eosinophilic bronchitis be treated? *J Thorac Dis* **13**, 3197-3214, doi:10.21037/jtd-20-2026 (2021).

48 Bu, Y. *et al.* Targeting NF-κB RelA/p65 phosphorylation overcomes RITA resistance. *Cancer Lett* **383**, 261-271, doi:10.1016/j.canlet.2016.10.006 (2016).

49 Poveda, J. *et al.* Bcl3: a regulator of NF-κB inducible by TWEAK in acute kidney injury with anti-inflammatory and antiapoptotic properties in tubular cells. *Exp Mol Med* **49**, e352, doi:10.1038/emm.2017.89 (2017).

50 Winter, N. A. *et al.* Sputum mast cell/basophil gene expression relates to inflammatory and clinical features of severe asthma. *J Allergy Clin Immunol* **148**, 428-438, doi:10.1016/j.jaci.2021.01.033 (2021).

51 Hunninghake, G. M. *et al.* MMP12, lung function, and COPD in high-risk populations. *N Engl J Med* **361**, 2599-2608, doi:10.1056/NEJMoa0904006 (2009).

52 Ozes, O. N. *et al.* NF-kappaB activation by tumour necrosis factor requires the Akt serine-threonine kinase. *Nature* **401**, 82-85, doi:10.1038/43466 (1999).

53 Madrid, L. V., Mayo, M. W., Reuther, J. Y. & Baldwin, A. S., Jr. Akt stimulates the transactivation potential of the RelA/p65 Subunit of NF-kappa B through utilization of the Ikappa B kinase and activation of the mitogen-activated protein kinase p38. *J Biol Chem* **276**, 18934-18940, doi:10.1074/jbc.M101103200 (2001).

54 Wang, V. Y. *et al.* Bcl3 Phosphorylation by Akt, Erk2, and IKK Is Required for Its Transcriptional Activity. *Mol Cell* **67**, 484-497.e485, doi:10.1016/j.molcel.2017.06.011 (2017).

55 Choy, D. F. *et al.* TH2 and TH17 inflammatory pathways are reciprocally regulated in asthma. *Sci Transl Med* **7**, 301ra129, doi:10.1126/scitranslmed.aab3142 (2015).

56 Brandt, E. B., Bolcas, P. E., Ruff, B. P. & Khurana Hershey, G. K. IL33 contributes to diesel pollution-mediated increase in experimental asthma

825 severity. *Allergy* **75**, 2254-2266, doi:10.1111/all.14181 (2020).
826 57 Liu, P. *et al.* Study on traditional Chinese medicine theory of lung being
827 connected with large intestine. *J Tradit Chin Med* **32**, 482-487,
828 doi:10.1016/s0254-6272(13)60059-x (2012).
829 58 Zhou, K. *et al.* Integrated plasma pharmacochemistry and network
830 pharmacology to explore the mechanism of Gerberae Piloselloidis Herba in
831 treatment of allergic asthma. *Journal of ethnopharmacology* **298**, 115624,
832 doi:10.1016/j.jep.2022.115624 (2022).
833 59 Liu, X. *et al.* A systematic pharmacology-based in vivo study to reveal the
834 effective mechanism of Yupingfeng in asthma treatment. *Phytomedicine*
835 **114**, 154783, doi:10.1016/j.phymed.2023.154783 (2023).
836 60 Shehzad, A., Parveen, S., Qureshi, M., Subhan, F. & Lee, Y. S. Decursin and
837 decursinol angelate: molecular mechanism and therapeutic potential in
838 inflammatory diseases. *Inflamm Res* **67**, 209-218, doi:10.1007/s00011-
839 017-1114-7 (2018).
840 61 Liang, Z. *et al.* Therapeutic effects of rosmarinic acid on airway responses
841 in a murine model of asthma. *Int Immunopharmacol* **41**, 90-97,
842 doi:10.1016/j.intimp.2016.10.010 (2016).
843 62 Chen, M. L., Wu, C. H., Hung, L. S. & Lin, B. F. Ethanol Extract of Perilla
844 frutescens Suppresses Allergen-Specific Th2 Responses and Alleviates
845 Airway Inflammation and Hyperreactivity in Ovalbumin-Sensitized Murine
846 Model of Asthma. *Evid Based Complement Alternat Med* **2015**, 324265,
847 doi:10.1155/2015/324265 (2015).
848 63 Yang, H. *et al.* Perilla Leaf Extract Attenuates Asthma Airway Inflammation
849 by Blocking the Syk Pathway. *Mediators Inflamm* **2021**, 6611219,
850 doi:10.1155/2021/6611219 (2021).
851 64 Luo, C. *et al.* A Review of the Anti-Inflammatory Effects of Rosmarinic Acid
852 on Inflammatory Diseases. *Front Pharmacol* **11**, 153,
853 doi:10.3389/fphar.2020.00153 (2020).
854 65 Conti, P. *et al.* Powerful anti-inflammatory action of luteolin: Potential
855 increase with IL-38. *Biofactors* **47**, 165-169, doi:10.1002/biof.1718 (2021).
856 66 Lee, M. Y. *et al.* Anti-asthmatic effect of schizandrin on OVA-induced airway
857 inflammation in a murine asthma model. *Int Immunopharmacol* **10**, 1374-
858 1379, doi:10.1016/j.intimp.2010.07.014 (2010).
859 67 Fang, R. *et al.* PDK1/Akt/PDE4D axis identified as a target for asthma
860 remedy synergistic with β2 AR agonists by a natural agent arctigenin.
861 *Allergy* **70**, 1622-1632, doi:10.1111/all.12763 (2015).
862 68 Koivisto, A. P., Belvisi, M. G., Gaudet, R. & Szallasi, A. Advances in TRP
863 channel drug discovery: from target validation to clinical studies. *Nat Rev*
864 *Drug Discov* **21**, 41-59, doi:10.1038/s41573-021-00268-4 (2022).
865 69 McKemy, D. D., Neuhausser, W. M. & Julius, D. Identification of a cold
866 receptor reveals a general role for TRP channels in thermosensation.
867 *Nature* **416**, 52-58, doi:10.1038/nature719 (2002).
868 70 Peier, A. M. *et al.* A TRP channel that senses cold stimuli and menthol. *Cell*

108, 705-715, doi:10.1016/s0092-8674(02)00652-9 (2002).

71 Salazar, H. *et al.* A single N-terminal cysteine in TRPV1 determines activation by pungent compounds from onion and garlic. *Nat Neurosci* **11**, 255-261, doi:10.1038/nn2056 (2008).

72 Hoel, E. P., Albantakis, L. & Tononi, G. Quantifying causal emergence shows that macro can beat micro. *Proc Natl Acad Sci U S A* **110**, 19790-19795, doi:10.1073/pnas.1314922110 (2013).

Figures legend

Fig. 1. The schema of the Herb-CMap methodology's sequential steps for connecting gene perturbation to therapeutic targets in the complex system of TCM prescriptions. (A) Data Collection: TCM-ingredient-target relationships, disease-gene associations, and protein-protein interactions (PPIs) were extracted. **(B) Gene Perturbation Data:** Gene perturbation data were obtained through RNA-seq analysis on CVA rats, identifying differentially expressed genes (DEGs). **(C) Functional Enrichment:** Various pathway enrichment methods, including KEGG, GSEA, RcisTarget, and CMap, were employed to enrich DEGs with relevant functions at different levels. **(D) DEG-Weighted PPI:** Gene perturbation data were mapped onto the PPI network to construct a DEG-weighted Protein-Protein Interaction Network. **(E) Target-DEG Association Network and network propagation**: The proximity between targets and DEGs was calculated to identify vital target-DEG interactions. Network random walk algorithms (LR and hot diffusion) were used to propagate gene perturbation and score each target, with higher gene perturbation scores indicating more essential roles of therapeutic targets. **(F) Validation:** Prioritized targets and corresponding ingredients were further validated through molecular docking and wet experiments.

Fig. 2. Data Statistics of the Herb-Ingredient-Target Network. (A) The overlapping situation of the ingredients from Suhuang. **(B)** The overlapping of the ingredients from different databases. **(C)** The percentage of OB and DL for ingredients in Suhuang. **(D)** The distribution of the number of targets for each ingredient in Suhuang.

Fig. 3. Transcriptomics analysis of Suhuang-treated OVA-induced CVA rats. (A) DEGs volcano plot of Suhuang *versus*. CVA. **(B)** DEGs volcano plot of CVA *versus*. Control. **(C)** The Venn diagram of DEGs. **(D)** The hierarchical clustering heat map of DEGs. **(E)** The visual analysis of the pathway enrichment results of GSEA analysis for GO: BP gene sets by the plug-in EnrichmentMap of Cytoscape. **(F)** The GO: BP pathway enrichment of the DEGs of Suhuang *versus*. CVA.

Fig. 4. The associations between the tissue-specific expression of DEGs and the meridian types of herbs in Suhuang. (A) The properties distributions of herbs in Suhuang. **(B)** The average expression of DEGs in different tissues.

Fig. 5. The pathway enrichment analysis of the DEGs in Suhuang *versus*. CVA. (A) KEGG pathway enrichment analysis based on the DEGs. **(B)** The gene set enrichment of the Connectivity Map about the modes of action based on the DEGs. The top 10 mechanisms of action (the normalized connectivity score > 0) were positively correlated with Suhuang for treating CVA. **(C)** GSEA plot for the IL-17 Signaling Pathway. **(D)** The expressional heatmap of the DEGs enriched in IL-17

Signaling Pathway. **(E)** GSEA plot for NF-κB Signaling Pathway. **(F)** The expressional heatmap of the DEGs enriched in NF-κB Signaling Pathway. **(G)** The information about the top motif enrichment and its TF annotation. NES: Normalized enrichment score of the enriched motifs in the geneset. The motifs whose NES > 3.0 are considered essential. **(H)** The DEG regulatory network about the motif-TF annotations.

Fig. 6. The association between DEGs and CVA-related genes. (A) The scatter plot of the network distance between the DEGs and CVA genes and the log2 foldchange absolute values of the DEGs. The Pearson correlation coefficient (*R*) and two-sided t-distribution *p* values (*p*) for the comparison are shown in the plot. The linear regression trendline (black) and its 95% confidence interval (shaded gray area) are shown in the graph. The colors suggest the clusters of the DEGs. **(B)** Comparison of the network distance between the DEGs of three clusters and CVA genes. *P-values* were calculated by t.test. ****$p < 0.0001$, vs. the cluster Ⅱ group. **(C)** Comparison of the log2 foldchange absolute values of the DEGs of three clusters. **(D)** The top 20 significant signaling pathways were enriched by introducing 73 DEGs into the KEGG pathway enrichment analysis. **(E)** The heatmap of the network distance between 73 DEGs of cluster 2 and CVA genes. The row genes are CVA genes, and the column genes are the DEGs of cluster 2. **(F)** The distribution of the DEGs of cluster 2. The genes of the left nest were significantly enriched in the IL-17 signaling pathway, and the genes of the right nest were enriched in the PI3K-AKt signaling pathway. The genes of the deepest color belonged to CVA genes. **(G)** The LR of the top 14 CVA-related genes with the significantly low network distance (*p*-value < 0.025) is based on the network propagation method.

Fig. 7. The main active ingredients and the molecular mechanisms of Suhuang for treating CVA. (A) The heat and ranking values of the top 20 targets in the target-DEG interactions by the Cytoscape implementation of Diffusion. The blue genes represented non-differentially expressed genes, and the red genes represented DEGs. (B-C) Sankey diagram of the active ingredients interacted closely with the enriched genes in PI3K-AKt signaling pathway **(B)** and Il-17 signaling pathway **(C)** based on the KEGG pathway enrichment analysis. Active ingredients (right) were connected to their targets (left) according to the compound-target interactions in Suhuang previously mentioned. **(D)** Mocular docking of quercetin on *Il-17a*, *Pik3cb*, *Pik3cd and Akt1* as well as luteolin on *Tnf* and *Akt1*. The essential residues with five colors represented corresponding target genes. Quercetin was shown in blue cartoon structures and luteolin in yellow cartoon structures. The polar interactions between the compounds and the target genes were shown in yellow dashes.

Fig. 8. The novel mechanism of Suhuang for CVA treatment was identified by our Herb-CMap. (A) Real-time PCR determined the relative mRNA level of *Il17a*

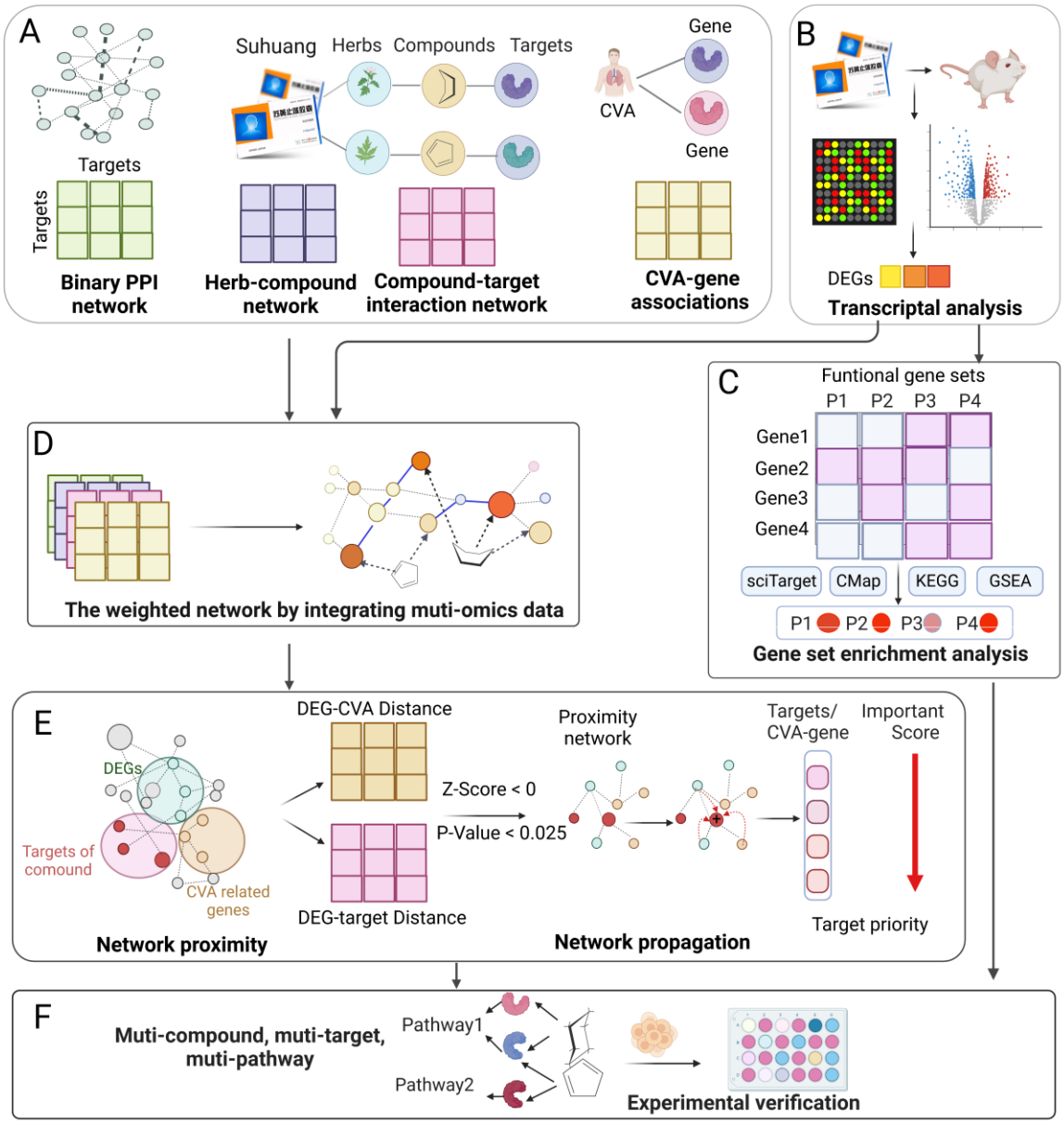
in LPS-induced RAW264.7 cells treated by Suhuang. Data is expressed as mean ± SEM (n = 3, $^{\#}p < 0.05$ vs. the control group; $^{*}p < 0.05$ vs. the LPS group). Quercetin and luteolin could directly inhibit the IL-17 signaling pathway and simultaneously inactivate PI3K, inhibiting Akt and NF-κB, thereby preventing lung inflammation and treating CVA. **(B)** The dose-response curve of quercetin and luteolin on the growth of LPS-induced RAW264.7 cells. **(C)** The dose response curve of quercetin and luteolin on the growth of LPS-induced RAW264.7 cells.

Fig. 9. The calculated network distance of the DEGs of Suhuang and the targets from Suhuang and the TCM property distribution of targets. (A) The property distribution of targets in the target-DEG interactions. **(B)** The herb, nature, flavor, and meridian distributions of the targets in all the target-DEG interactions. *P-values* were calculated by t.test. $^{**}p < 0.01$, $^{****}p < 0.0001$ vs. the ZISUYE herb; $^{****}p < 0.0001$ vs. the Cold nature; $^{\#\#\#\#}p < 0.0001$ vs. the Pungent flavor; $^{++++}p < 0.0001$ vs. the Lung, spleen meridian.

Fig. 10. The important score of ingredients and target from LR methods. (A) The fisher extract test of the correlation between the target-DEG interactions with the significantly low network distance and the properties of targets. Relative log10 *p*.value > 0 & *p*.value < 0.05 suggested a positive correlation. **(B)** The herb, nature, flavor, and meridian distributions of the targets in the target-DEG, target-CVA DEG, and target-pathway DEG interactions with the significantly low network distance. *P-values* were calculated by t.test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$. **(C)** The chord diagram of the active ingredients interacted with three CVA-related targets of Suhuang. Active ingredients (right) were connected to their targets (left) according to the compound-target interactions in Suhuang.

Fig. S1. The workflow of Herb-CMap.
Fig. S2. The connection map of DEGs for MOAs.
Fig. S3. Based on the network proximity model, The calculated network distance between the DEGs of Suhuang and the CVA-related genes.
Fig. S4. The target-DEGs distance analysis. (A) The network distance distribution of all the target-DEG interactions (n = 928,485). The red part of the density plot was the target-DEG interactions with the significantly low network distance (n = 19,009) based on the one-sided test of normal distribution ($z < 0$, $p < 0.025$). **(B)** The topmost significant signaling pathways were enriched by introducing 499 targets that were not DEGs into the KEGG pathway enrichment analysis.

1007 **Fig. 1.**



1008

1009 **Fig. 1. The schema of the TCM-CMap methodology's sequential steps for**
1010 **connecting gene perturbation to therapeutic targets in the complex system**
1011 **of TCM prescriptions. (A) Data Collection:** TCM-ingredient-target
1012 relationships, disease-gene associations, and protein-protein interactions (PPIs)
1013 were extracted. **(B) Gene Perturbation Data:** Gene perturbation data were
1014 obtained through RNA-seq analysis on CVA rats, identifying differentially
1015 expressed genes (DEGs). **(C) Functional Enrichment:** Various pathway
1016 enrichment methods, including KEGG, GSEA, RcisTarget, and CMap, were
1017 employed to enrich DEGs with relevant functions at different levels. **(D) DEG-**
1018 **Weighted PPI:** Gene perturbation data were mapped onto the PPI network to
1019 construct a DEG-weighted Protein-Protein Interaction Network. **(E) Target-DEG**
1020 **Association Network and network propagation**: The proximity between targets
1021 and DEGs was calculated to identify vital target-DEG interactions. Network

random walk algorithms (LR and hot diffusion) were used to propagate gene perturbation and score each target, with higher gene perturbation scores indicating more essential roles of therapeutic targets. **(F) Validation:** Prioritized targets and corresponding ingredients were further validated through molecular docking and wet experiments.

Fig. 2.

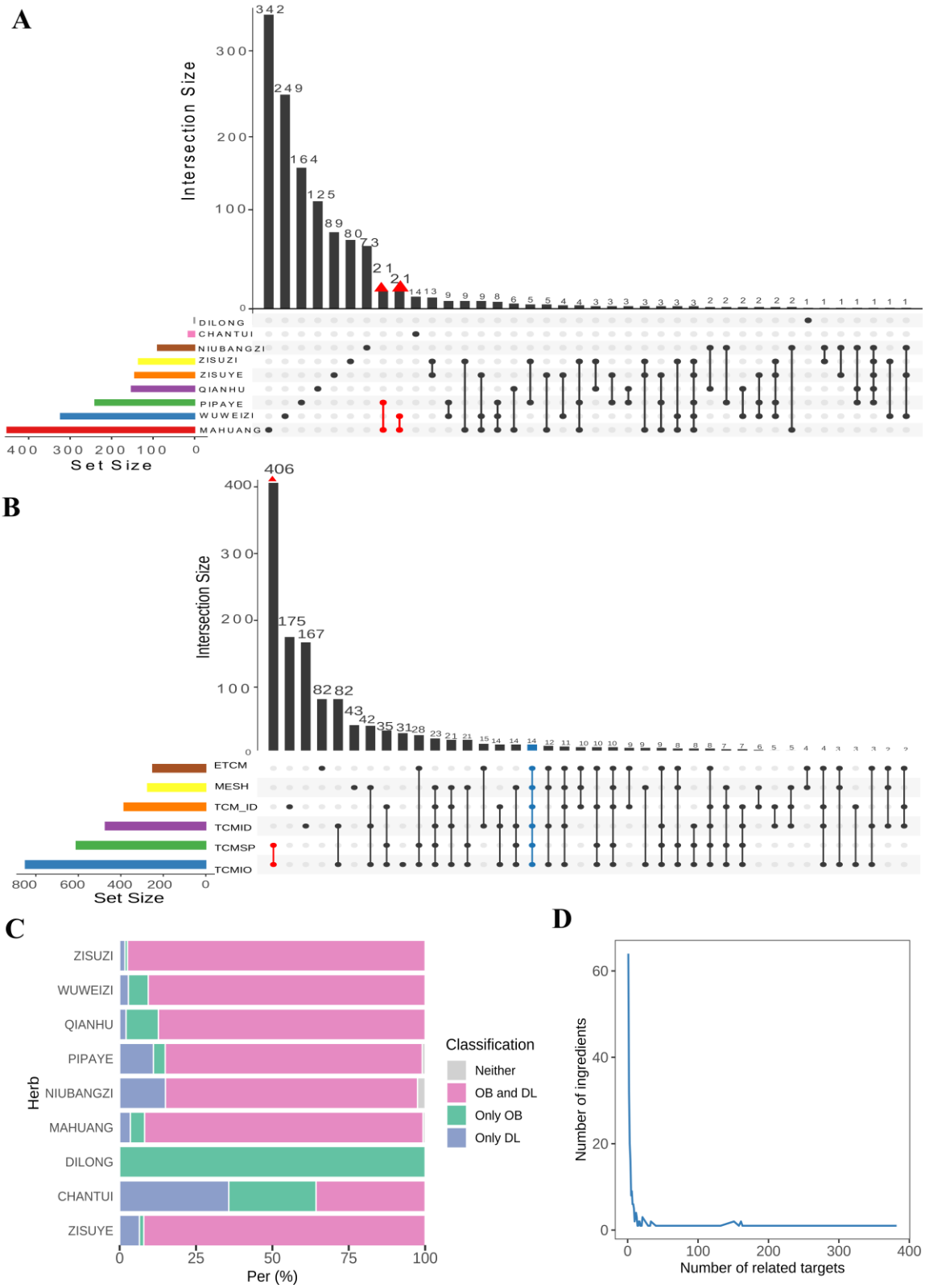
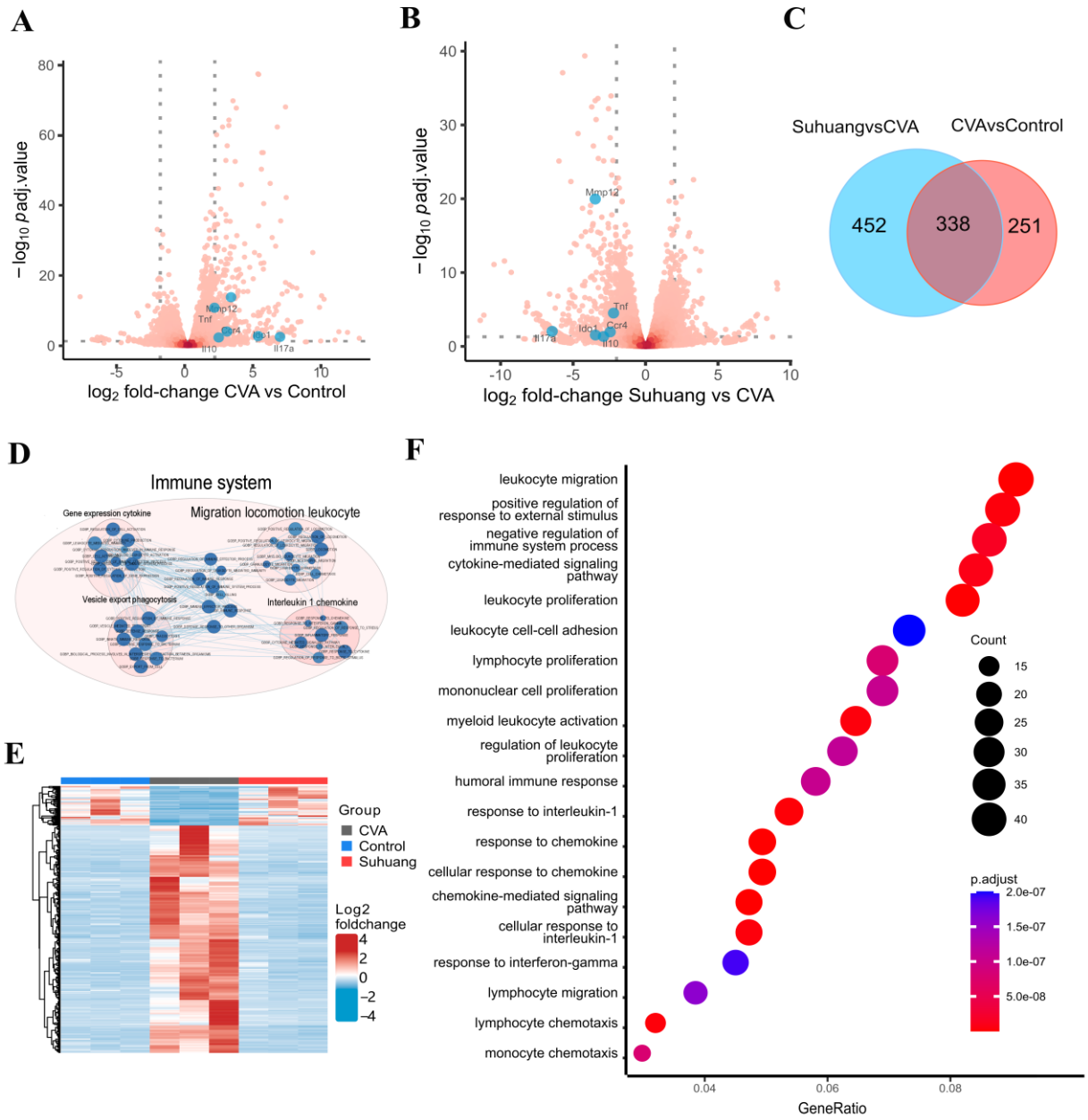


Fig. 2. Data Statistics of the Herb-Ingredient-Target Network. (A) The overlapping situation of the ingredients from Suhuang. **(B)** The overlapping of the ingredients from different databases. **(C)** The percentage of OB and DL for ingredients in Suhuang. **(D)** The distribution of the number of targets for each

ingredient in Suhuang.

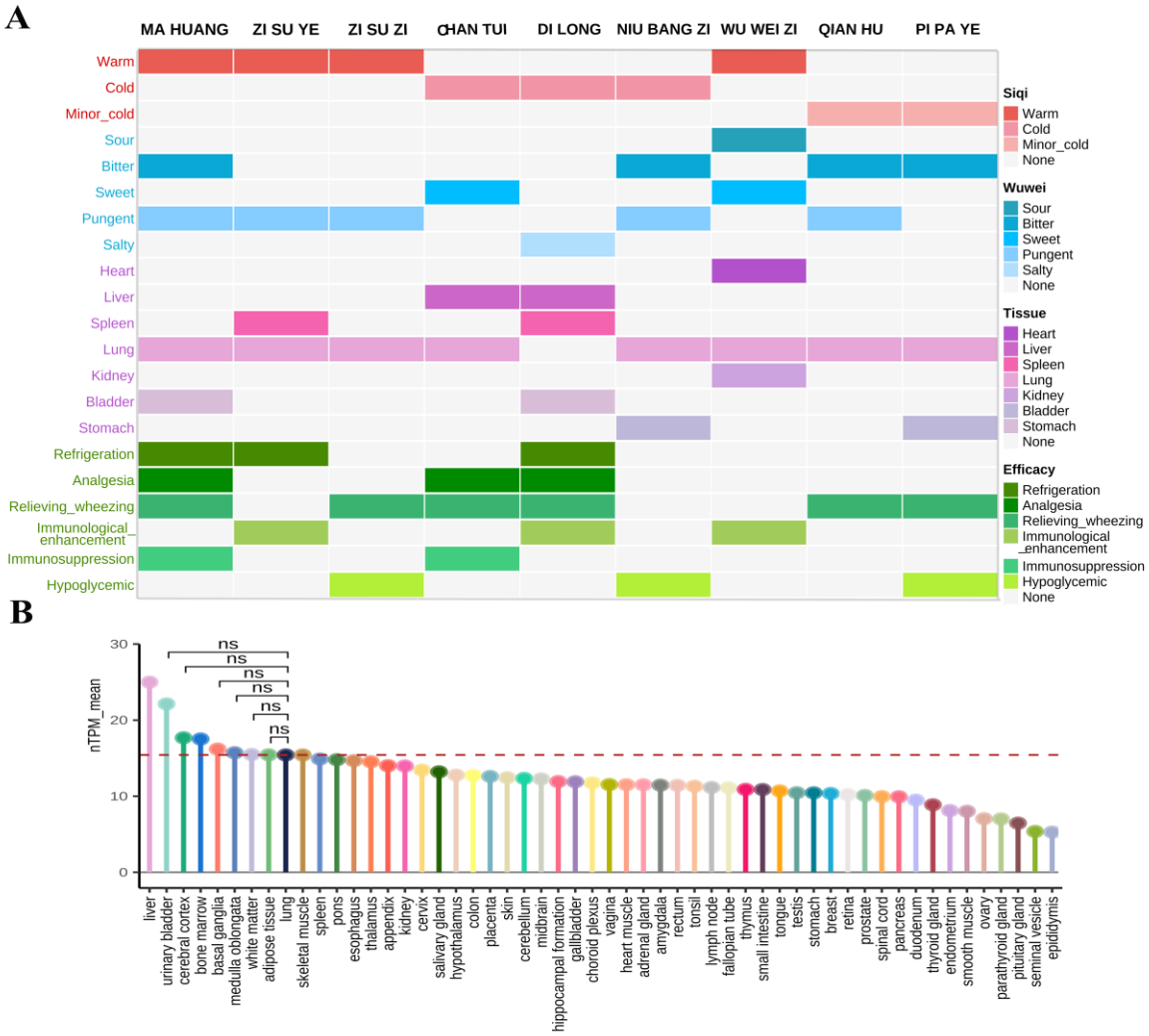
1036 **Fig. 3.**



1037

1038 **Fig. 3. Transcriptomics analysis of Suhuang-treated OVA-induced CVA rats.**
1039 **(A)** DEGs volcano plot of Suhuang *versus*. CVA. **(B)** DEGs volcano plot of CVA *versus*.
1040 Control. **(C)** The Venn diagram of DEGs. **(D)** The hierarchical clustering heat map
1041 of DEGs. **(E)** The visual analysis of the pathway enrichment results of GSEA
1042 analysis for GO: BP gene sets by the plug-in EnrichmentMap of Cytoscape. **(F)** The
1043 GO: BP pathway enrichment of the DEGs of Suhuang *versus*. CVA.

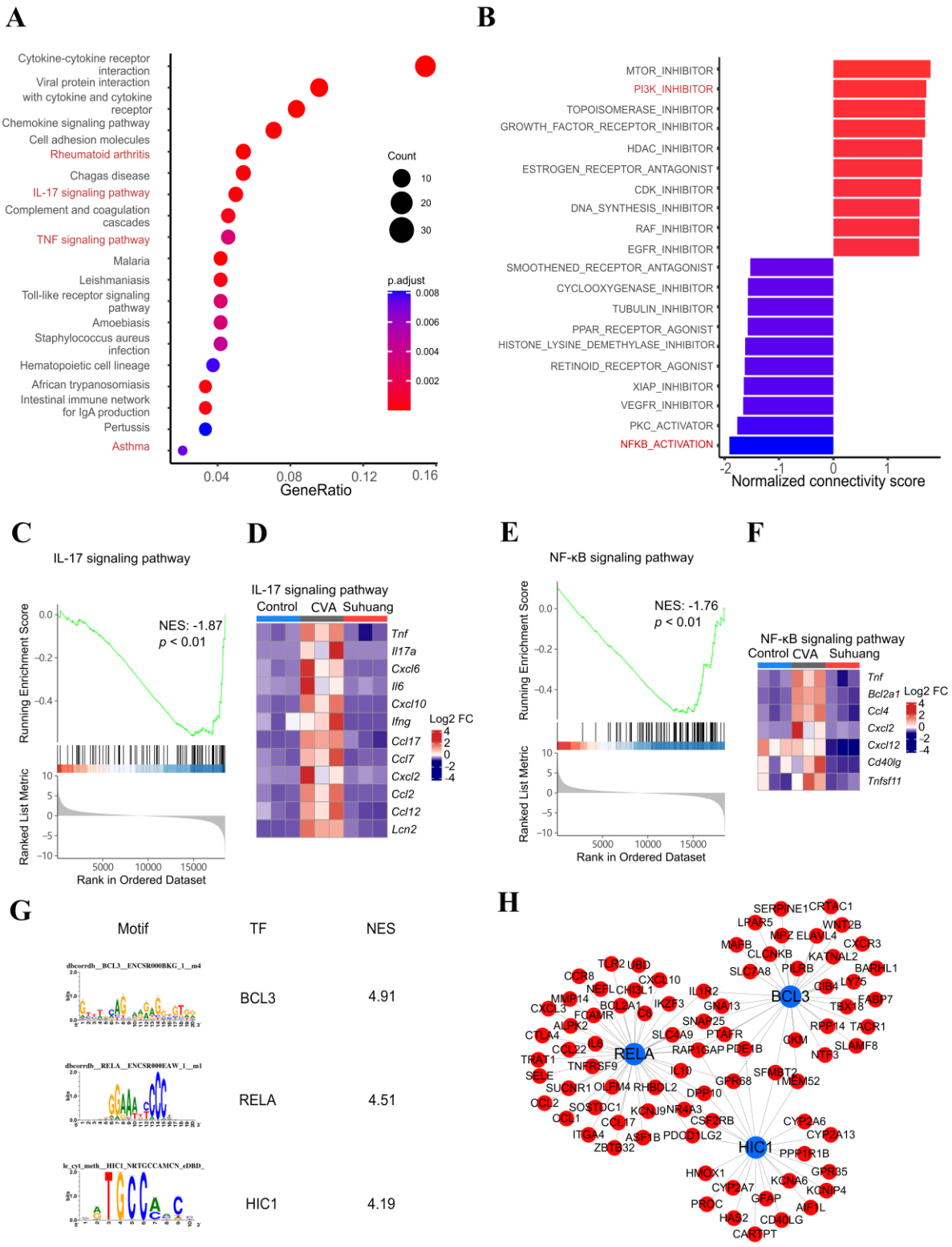
1044 **Fig. 4.**



1045

1046 **Fig. 4. The associations between the tissue-specific expression of DEGs and**
1047 **the meridian types of herbs in Suhuang. (A)** The properties distributions of
1048 herbs in Suhuang. **(B)** The average expression of DEGs in different tissues.

1049 **Fig. 5.**

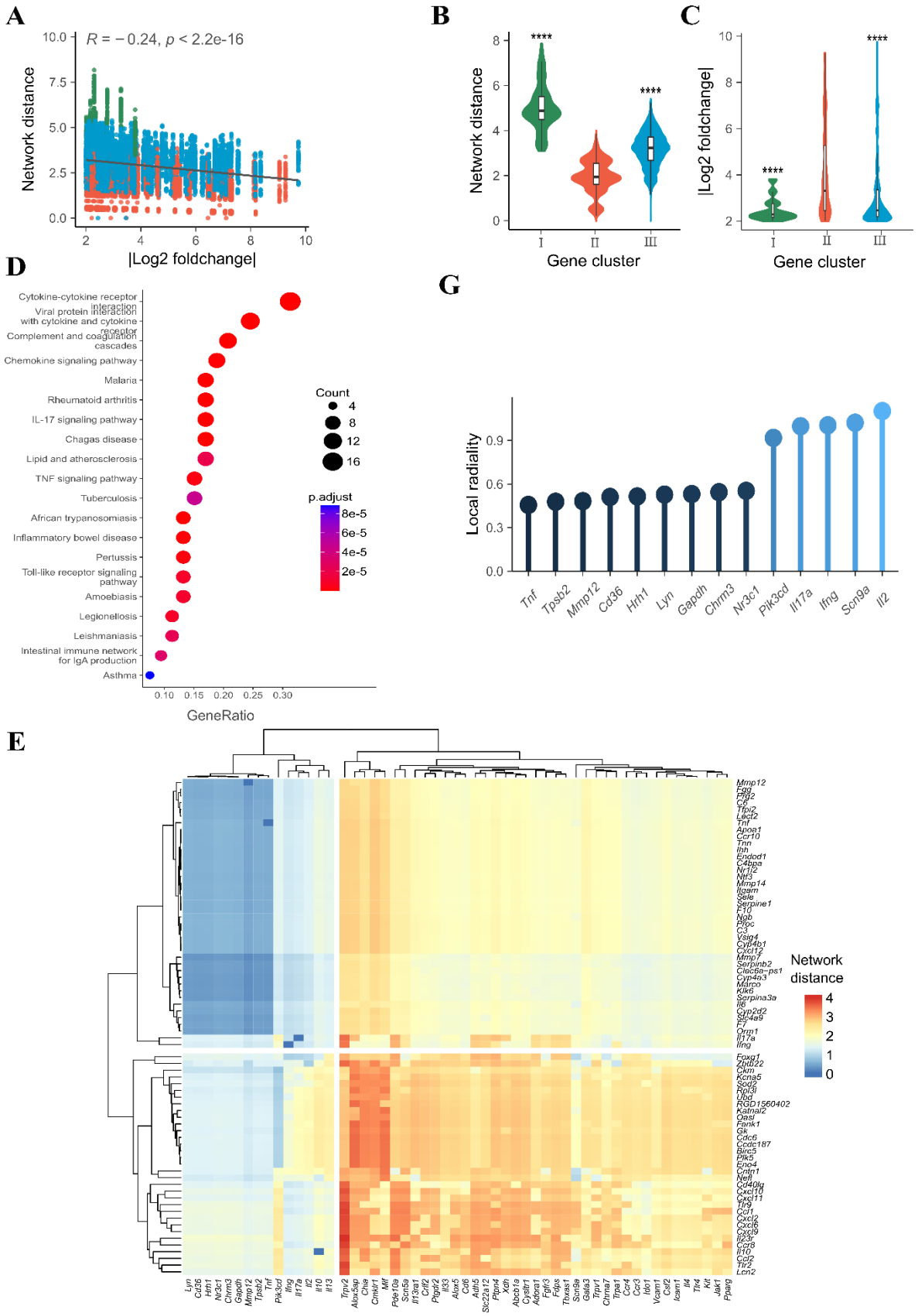


1050

1051 **Fig. 5. The pathway enrichment analysis of the DEGs in Suhuang *versus*. CVA.**
1052 **(A)** KEGG pathway enrichment analysis based on the DEGs. **(B)** The gene set
1053 enrichment of the Connectivity Map about the modes of action based on the DEGs.
1054 The top 10 mechanisms of action (the normalized connectivity score > 0) were

positively correlated with Suhuang for treating CVA. **(C)** GSEA plot for the IL-17 Signaling Pathway. **(D)** The expressional heatmap of the DEGs enriched in IL-17 Signaling Pathway. **(E)** GSEA plot for NF-κB Signaling Pathway. **(F)** The expressional heatmap of the DEGs enriched in NF-κB Signaling Pathway. **(G)** The information about the top motif enrichment and its TF annotation. NES: Normalized enrichment score of the enriched motifs in the geneset. The motifs whose NES > 3.0 are considered essential. **(H)** The DEG regulatory network about the motif-TF annotations.

1063 **Fig. 6.**

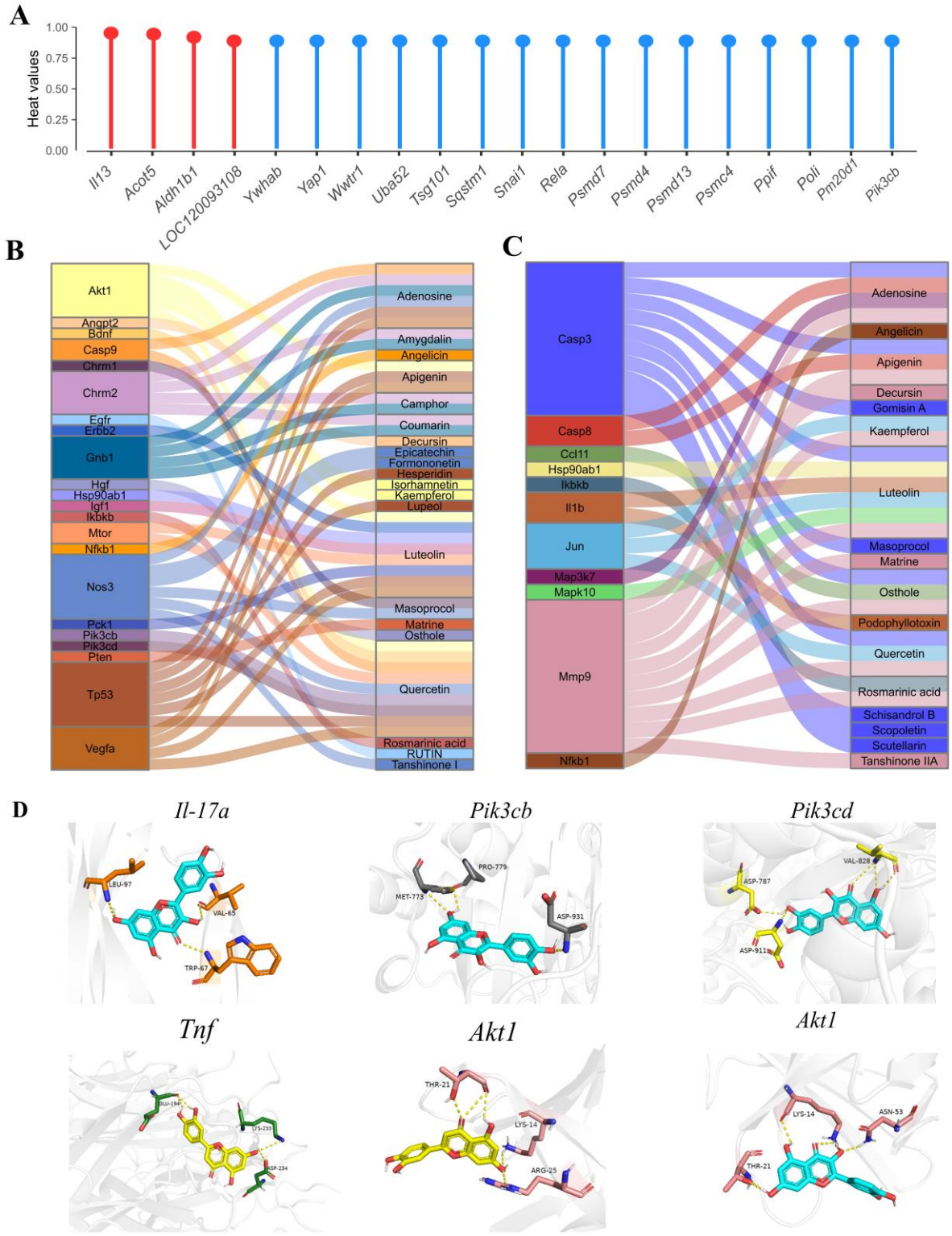


1064

1065 **Fig. 6. The association between DEGs and CVA-related genes. (A)** The scatter

plot of the network distance between the DEGs and CVA genes and the log2 foldchange absolute values of the DEGs. The Pearson correlation coefficient (R) and two-sided t-distribution p values (p) for the comparison are shown in the plot. The linear regression trendline (black) and its 95% confidence interval (shaded gray area) are shown in the graph. The colors suggest the clusters of the DEGs. **(B)** Comparison of the network distance between the DEGs of three clusters and CVA genes. *P-values* were calculated by t.test. ****$p < 0.0001$, vs. the cluster Ⅱ group. **(C)** Comparison of the log2 foldchange absolute values of the DEGs of three clusters. **(D)** The top 20 significant signaling pathways were enriched by introducing 73 DEGs into the KEGG pathway enrichment analysis. **(E)** The heatmap of the network distance between 73 DEGs of cluster 2 and CVA genes. The row genes are CVA genes, and the column genes are the DEGs of cluster 2. **(F)** The distribution of the DEGs of cluster 2. The genes of the left nest were significantly enriched in the IL-17 signaling pathway, and the genes of the right nest were enriched in the PI3K-AKt signaling pathway. The genes of the deepest color belonged to CVA genes. **(G)** The LR of the top 14 CVA-related genes with the significantly low network distance (p-value < 0.025) is based on the network propagation method.

1085 **Fig. 7.**



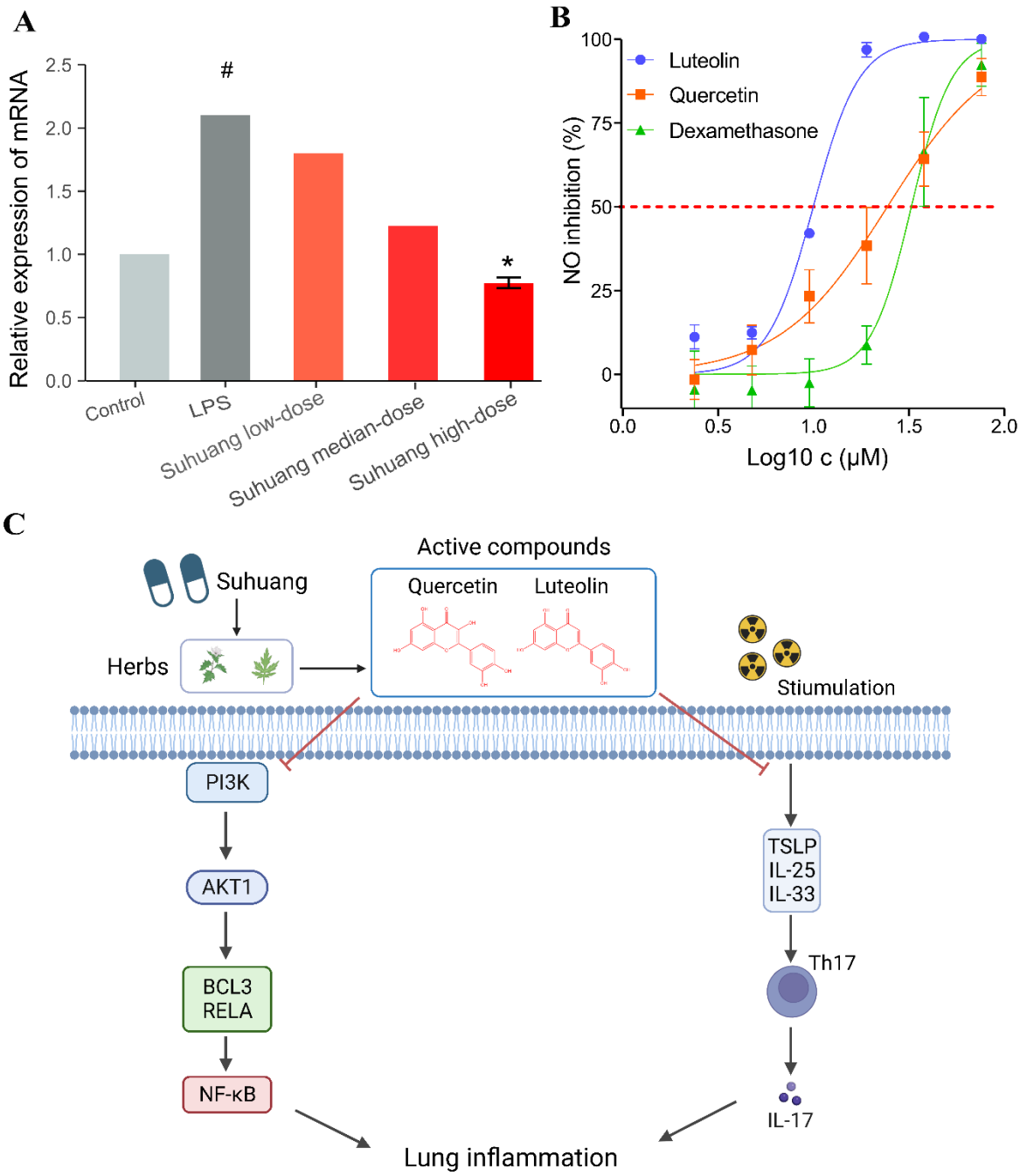
1086

1087 **Fig. 7. The main active ingredients and the molecular mechanisms of**
1088 **Suhuang for treating CVA. (A)** The heat and ranking values of the top 20 targets
1089 in the target-DEG interactions by the Cytoscape implementation of Diffusion. The
1090 blue genes represented non-differentially expressed genes, and the red genes
1091 represented DEGs. (B-C) Sankey diagram of the active ingredients interacted
1092 closely with the enriched genes in PI3K-AKt signaling pathway **(B)** and Il-17

signaling pathway **(C)** based on the KEGG pathway enrichment analysis. Active ingredients (right) were connected to their targets (left) according to the compound-target interactions in Suhuang previously mentioned. **(D)** Mocular docking of quercetin on *Il-17a*, *Pik3cb*, *Pik3cd and Akt1* as well as luteolin on *Tnf* and *Akt1*. The essential residues with five colors represented corresponding target genes. Quercetin was shown in blue cartoon structures, and luteolin in yellow cartoon structures. The polar interactions between the compounds and the target genes were shown in yellow dashes.

1102 **Fig. 8.**

1103



1104

1105 **Fig. 8. The novel mechanism of Suhuang for CVA treatment was identified by**
1106 **our TCM-CMap. (A)** Real-time PCR determined the relative mRNA level of *Il17a* in
1107 LPS-induced RAW264.7 cells treated by Suhuang. Data is expressed as mean ± SEM
1108 (n = 3, $^{\#}p < 0.05$ vs. the control group; $^{*}p < 0.05$ vs. the LPS group). Quercetin and
1109 luteolin could directly inhibit the IL-17 signaling pathway and simultaneously
1110 inactivate PI3K, inhibiting Akt and NF-κB, thereby preventing lung inflammation
1111 and treating CVA. **(B)** The dose-response curve of quercetin and luteolin on the
1112 growth of LPS-induced RAW264.7 cells. **(C)** The dose response curve of quercetin

1113 and luteolin on the growth of LPS-induced RAW264.7 cells.
1114

Fig. 9.

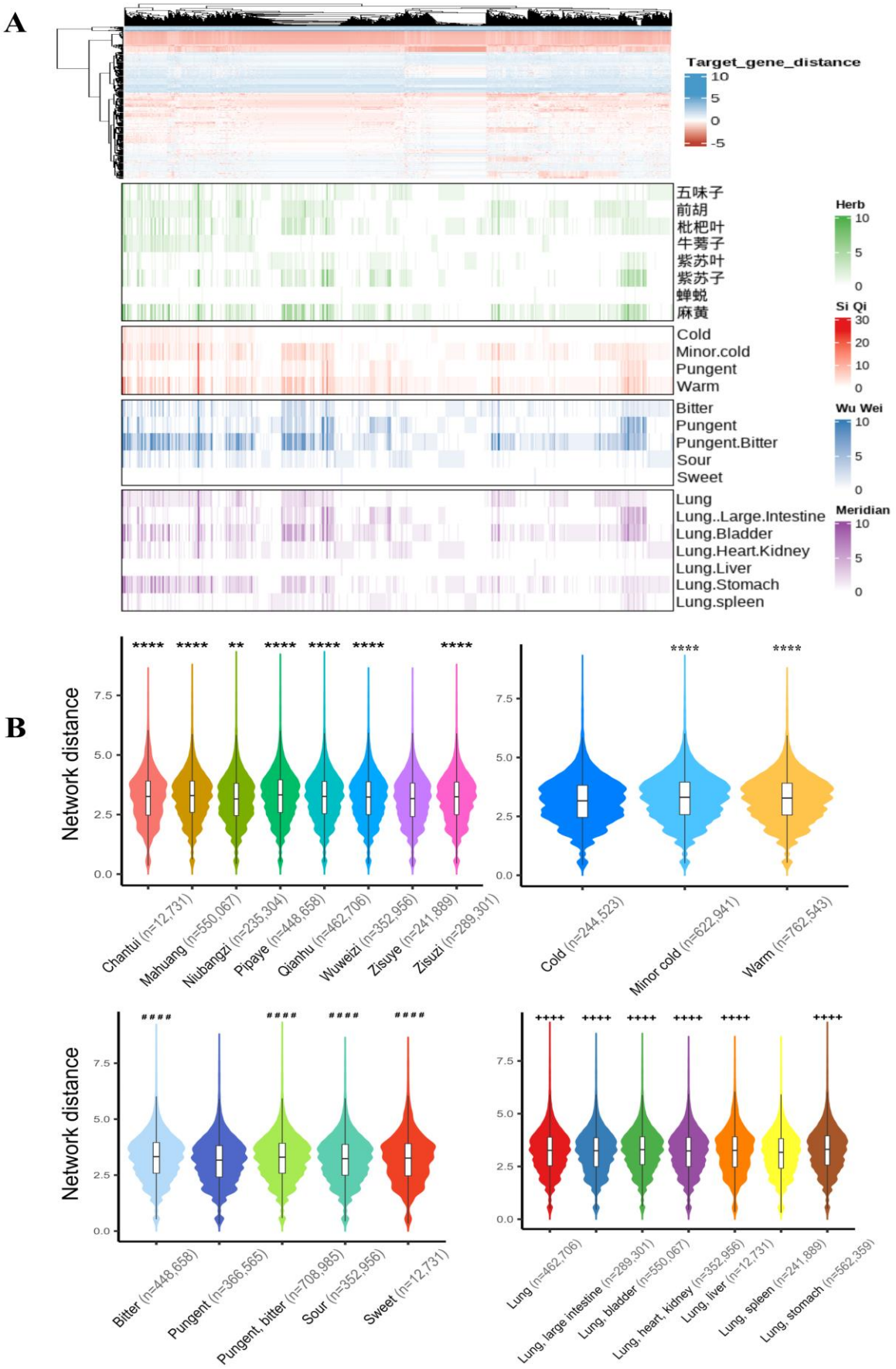

A
Target_gene_distance
10
5
0
-5
五味子
前胡
枇杷叶
牛蒡子
紫苏叶
紫苏子
蝉蜕
麻黄
Herb
Cold
Minor.cold
Pungent
Warm
Si Qi
Bitter
Pungent
Pungent.Bitter
Sour
Sweet
Wu Wei
Lung
Lung..Large.Intestine
Lung.Bladder
Lung.Heart.Kidney
Lung.Liver
Lung.Stomach
Lung.spleen
Meridian
B
Network distance
Chantui (n=12,731)
Mahuang (n=550,067)
Niubangzi (n=235,304)
Pipaye (n=448,658)
Qianhu (n=462,706)
Wuweizi (n=352,956)
Zisuye (n=241,889)
Zisuzi (n=289,301)
Cold (n=244,523)
Minor cold (n=622,941)
Warm (n=762,543)
Bitter (n=448,658)
Pungent (n=366,565)
Pungent, bitter (n=708,985)
Sour (n=352,956)
Sweet (n=12,731)
Lung (n=462,706)
Lung, large intestine (n=289,301)
Lung, bladder (n=550,067)
Lung, heart, kidney (n=352,956)
Lung, liver (n=12,731)
Lung, spleen (n=241,889)
Lung, stomach (n=562,359)

Fig. 9. The calculated network distance of the DEGs of Suhuang and the targets from Suhuang and the TCM property distribution of targets. (A) The property distribution of targets in the target-DEG interactions. **(B)** The herb, nature, flavor, and meridian distributions of the targets in all the target-DEG interactions. *P-values* were calculated by t.test. $^{**}p < 0.01$, $^{****}p < 0.0001$ vs. the ZISUYE herb; $^{****}p < 0.0001$ vs. the Cold nature; $^{\#\#\#\#}p < 0.0001$ vs. the Pungent flavor; $^{++++}p < 0.0001$ vs. the Lung, spleen meridian.

Fig. 10.

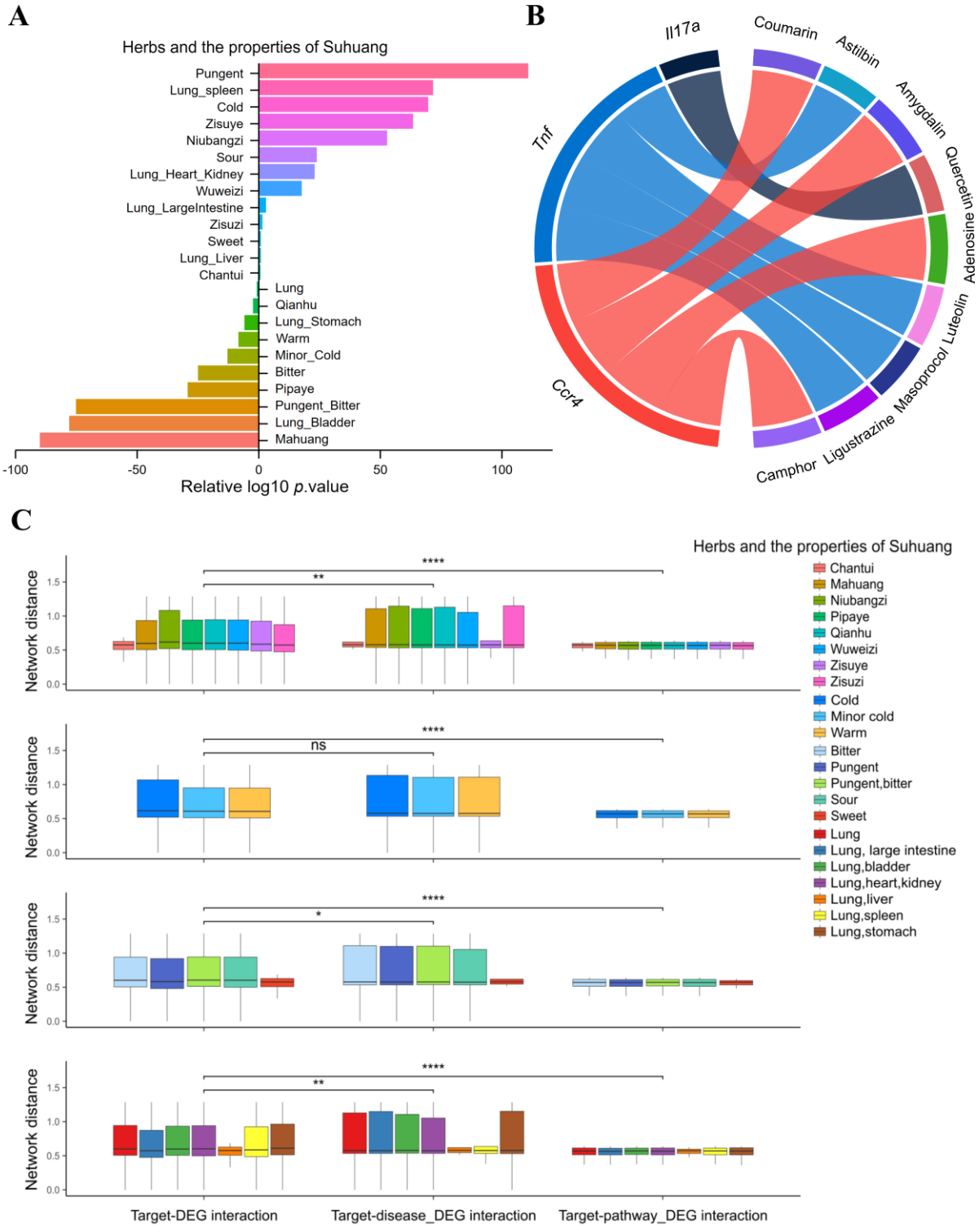
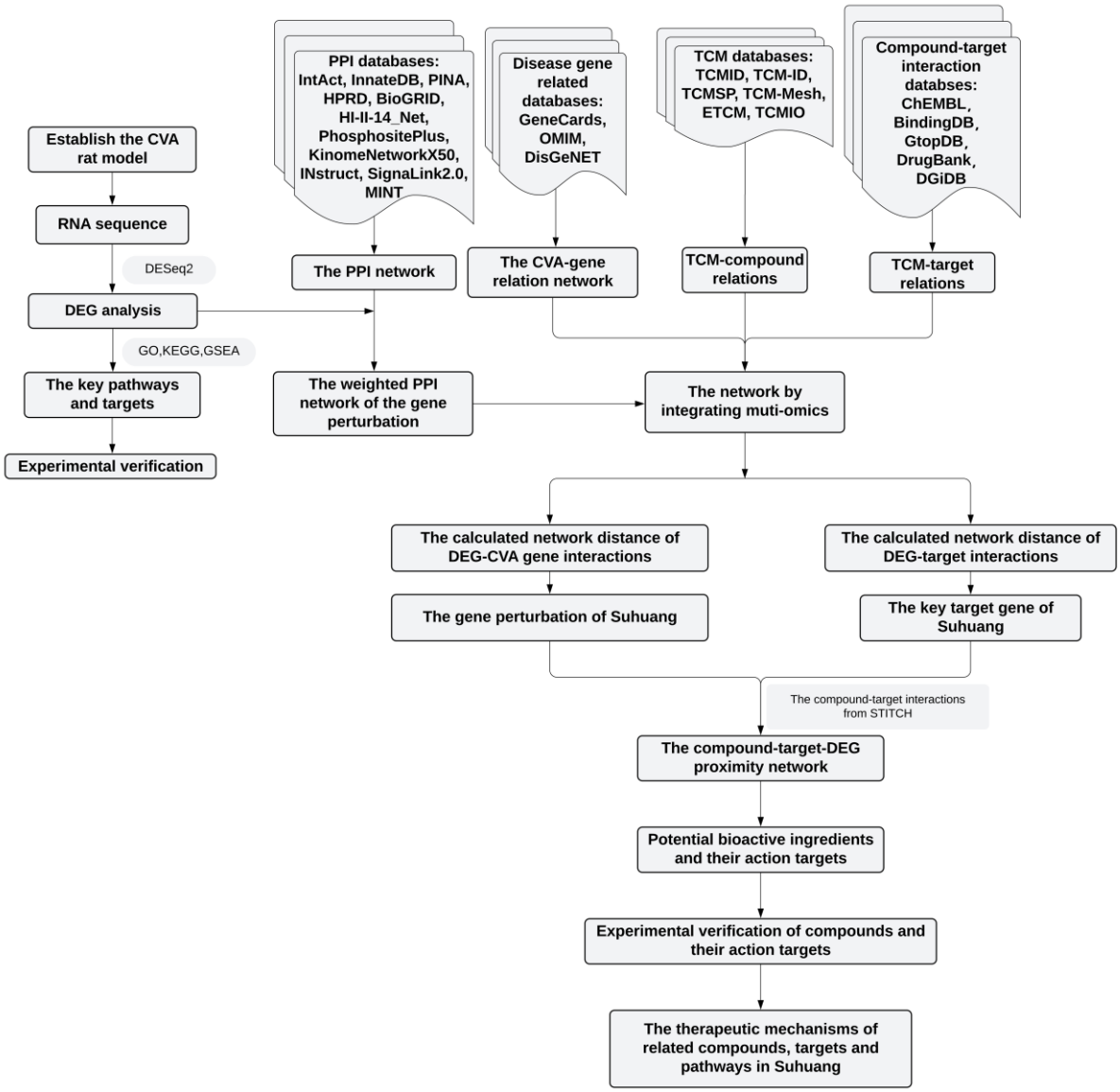


Fig. 10. The important score of ingredients and target from LR methods. (A) The fisher extract test of the correlation between the target-DEG interactions with the significantly low network distance and the properties of targets. Relative log10 *p*.value > 0 & *p*.value < 0.05 suggested a positive correlation. **(B)** The herb, nature, flavor, and meridian distributions of the targets in the target-DEG, target-CVA DEG, and target-pathway DEG interactions with the significantly low network distance.

P-values were calculated by t.test. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$. **(C)** The chord diagram of the active ingredients interacted with three CVA-related targets of Suhuang. Active ingredients (right) were connected to their targets (left) according to the compound-target interactions in Suhuang.

1140 **Supplementary Fig. 1.**



1141

1142 **Supplementary Fig. 1. The workflow of TCM-CMap.**

1143

1144 **Supplementary Fig. 2.**

Perturbagen class

score	type	name	description
99.90		Tumor necrosis factor superfamily LOF	
99.53		Rho GTPase activating proteins LOF	
99.09		Lipocalins GOF	
98.48		Kinase anchoring proteins LOF	
98.37		AHSP Pathway LOF	
2.56		PI3K Signaling LOF	
2.42		GPCR Subset GOF	
1.59		General transcription factors group 2 LOF	
0.94		X linked mental retardation group 2 LOF	
0.76		Trace amine receptors LOF	

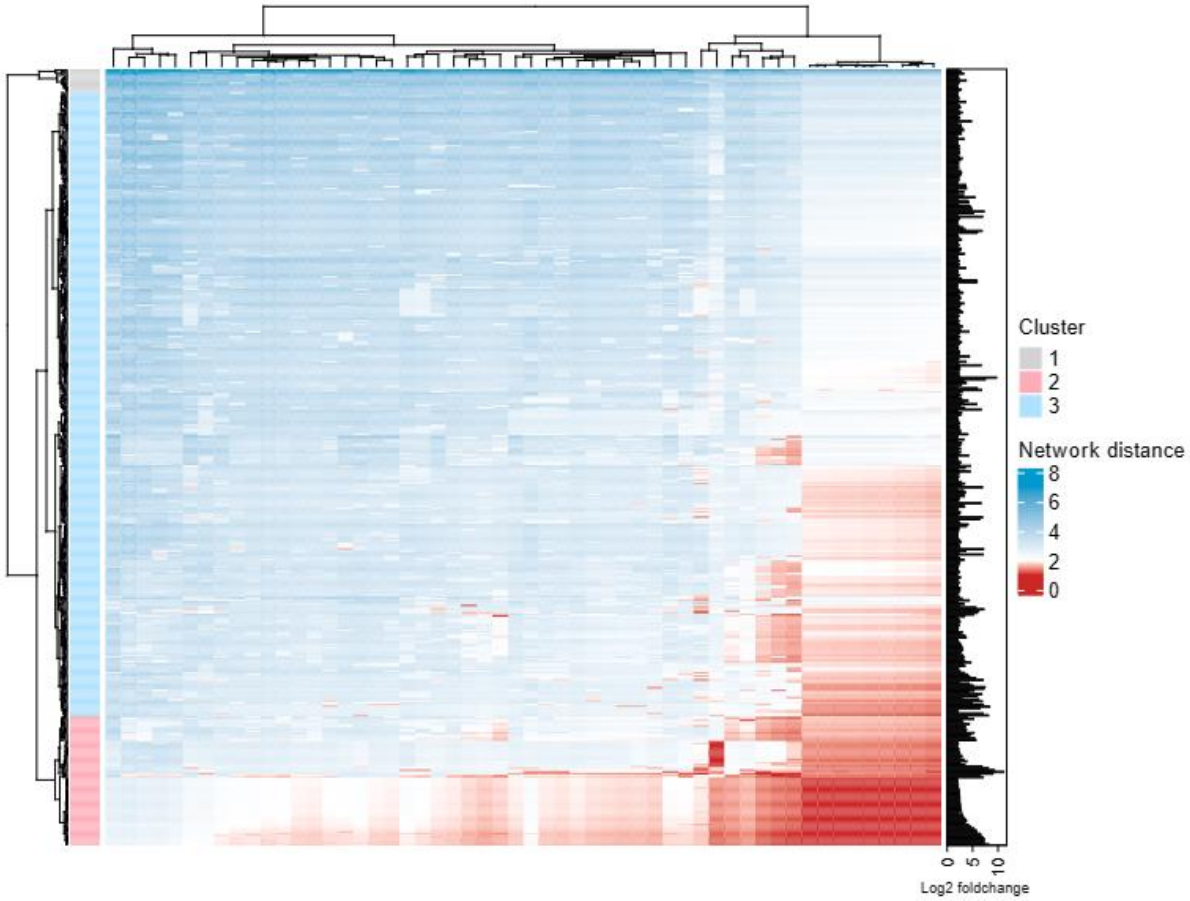
1145

1146

1147 **Supplementary Fig 2. The connection map of DEGs for MOAs.**

1148

1149 **Supplementary Fig 3.**



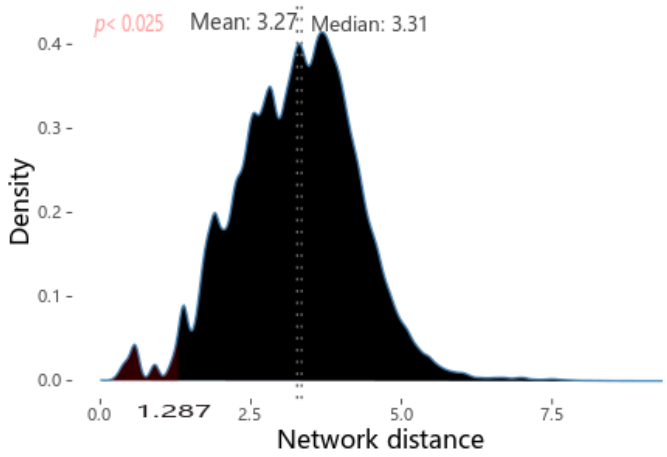
1150

1151 **Supplementary Fig. 3. Based on the network proximity model, The**
1152 **calculated network distance between the DEGs of Suhuang and the CVA-**
1153 **related genes.**

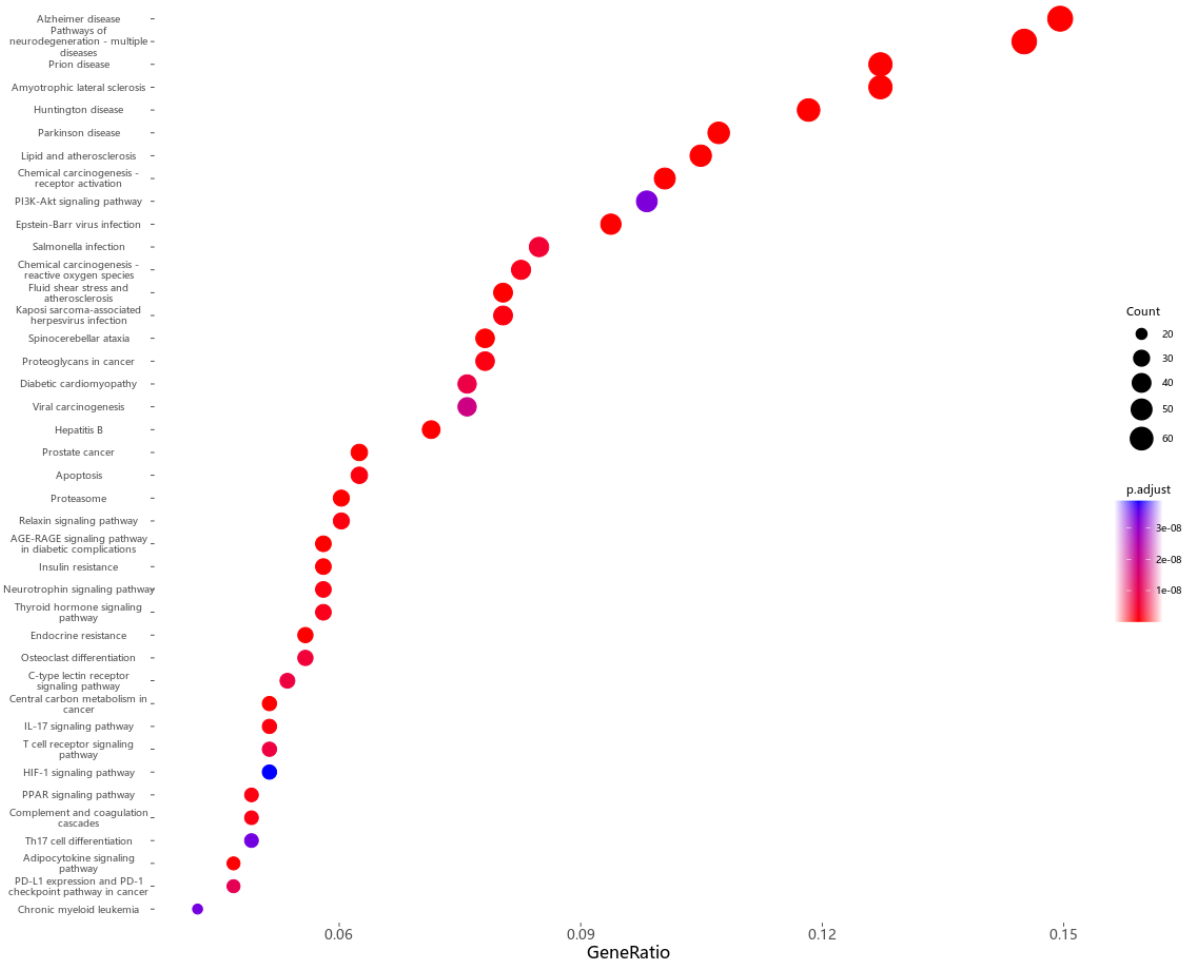
1154

Supplementary Fig 4.

A



B



Supplementary Fig 4. The target-DEGs distance analysis. (A) The network distance distribution of all the target-DEG interactions (n = 928,485). The red part of the density plot was the target-DEG interactions with the significantly low network distance (n = 19,009) based on the one-sided test of normal distribution ($z < 0$, $p < 0.025$). **(B)** The top most significant signaling pathways were enriched

1165 by introducing 499 targets that were not DEGs into the KEGG pathway enrichment
1166 analysis.
1167