Citation: CPT Pharmacometrics Syst. Pharmacol. (2017) 00, 00; doi:10.1002/psp4.12253

ORIGINAL ARTICLE

Network-Wide Screen Identifies Variation of Novel Precise On-Module Targets Using Conformational Modudaoism

Bing Li[1,2,3], Jun Liu[1], Yanan Yu[1], Pengqian Wang[4], Yingying Zhang[5], Xumin Ni[6], Qiong Liu[1], Xiaoxu Zhang[1], Zhong Wang[1] and Yongyan Wang[1]*

Modular targeting is promising in drug research at the network level, but it is challenging to quantificationally identify the precise on-modules. Based on a proposed Modudaoism (MD), we defined conserved MD (MDc) and varied MD (MDv) to quantitatively evaluate the conformational and energy variations of modules, and thereby identify the conserved and discrepant allosteric modules (AMs). Compared to the $Z_{summary}$, MDc/MDv got an optimized result of module preserved ratio and modular structure. In the mice anti-ischemic networks, 3, 5, and 1 conserved AMs as well as 4, 1, and 3 on-modules of baicalin (BA), jasminoidin (JA), and ursodeoxycholic acid (UA) were identified by MDc and MDv, 5 unique AMs and their characteristic actions were revealed. Besides, co-immunoprecipitation (Co-IP) experiments validated the representative modular structure. MDc/MDv method can quantitatively define the conformational variations of modules and screen the precise on-modules network-wide, which may provide a promising strategy for drug discovery.
CPT Pharmacometrics Syst. Pharmacol. (2017) **00**, 00; doi:10.1002/psp4.12253; published online on 0 Month 2017.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Modular targeting is promising in drug research at the network level, but it is challenging to quantificationally identify the precise on-modules.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ Based on a proposed MD, we defined MDc and MDv to quantitatively evaluate the conformational and energy variations of modules, and thereby identify the conserved and discrepant AMs.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

☑ The MDc/MDv methods can quantitatively screen the precise on-modules, which outperform the existing method. On-module screening revealed unique AMs of BA, JA, and UA in ischemia treating.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

☑ The MD-based method may help to explore the therapeutic target modules rather than independent gene or protein in drug discovery. In addition, this module-centric analysis may provide a promising strategy for pharmacological research and disease therapy.

Modular targeting that goes beyond individual genes is critical to clarify the flexible mechanisms of drugs from a systematic point of view.[1] Several studies have integrated biological network and gene expression data to identify the module biomarkers or targets in cancers and other complex diseases,[2–6] but it is still challenging to quantificationally measure the topological structural and energy variations of modules, so as to identify the precise on-modules. In response to diverse drug therapies, the network structure may alter, and modular allosterism and rewiring may be triggered.[7,8] The intramodular structure and energy transformation in a holistic network system should have its inherent balancing and coordination rule, which are analogous to the Yin and Yang activity in Chinese philosophical law of "Daoism," and quantitative indicators for this modular level variation are needed. We defined the degree of modular structural and energy variations under different conditions as Modudaoism (MD), which may quantitatively reflect the dynamic drug sensitivity and mechanisms at the network module level.

Compared with various types of module detection algorithms,[9] there are few methods for module evaluation,[10] such as identification of target modules. Biological experiment-based module evaluation methods are merely applicable for small modules that consist of only a few nodes. Previous methods viewed gene expression level as their features for identifying drug-related modules, but neglected the topological structural rewiring of modules.[11,12] A Jaccard's similarity coefficient-based method of SimiNEF has been proposed to determine the topological variations in different compound-dependent protein-protein interaction (PPI) networks.[13] The

[1]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China; [2]Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing, China; [3]ShanXi Buchang Pharmaceutical Co., Ltd., Heze, China; [4]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China; [5]Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China; [6]Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing, China. *Correspondence: Z Wang (zhonw@vip.sina.com)
Received 13 June 2017; accepted 13 September 2017; published online on 0 Month 2017. doi:10.1002/psp4.12253

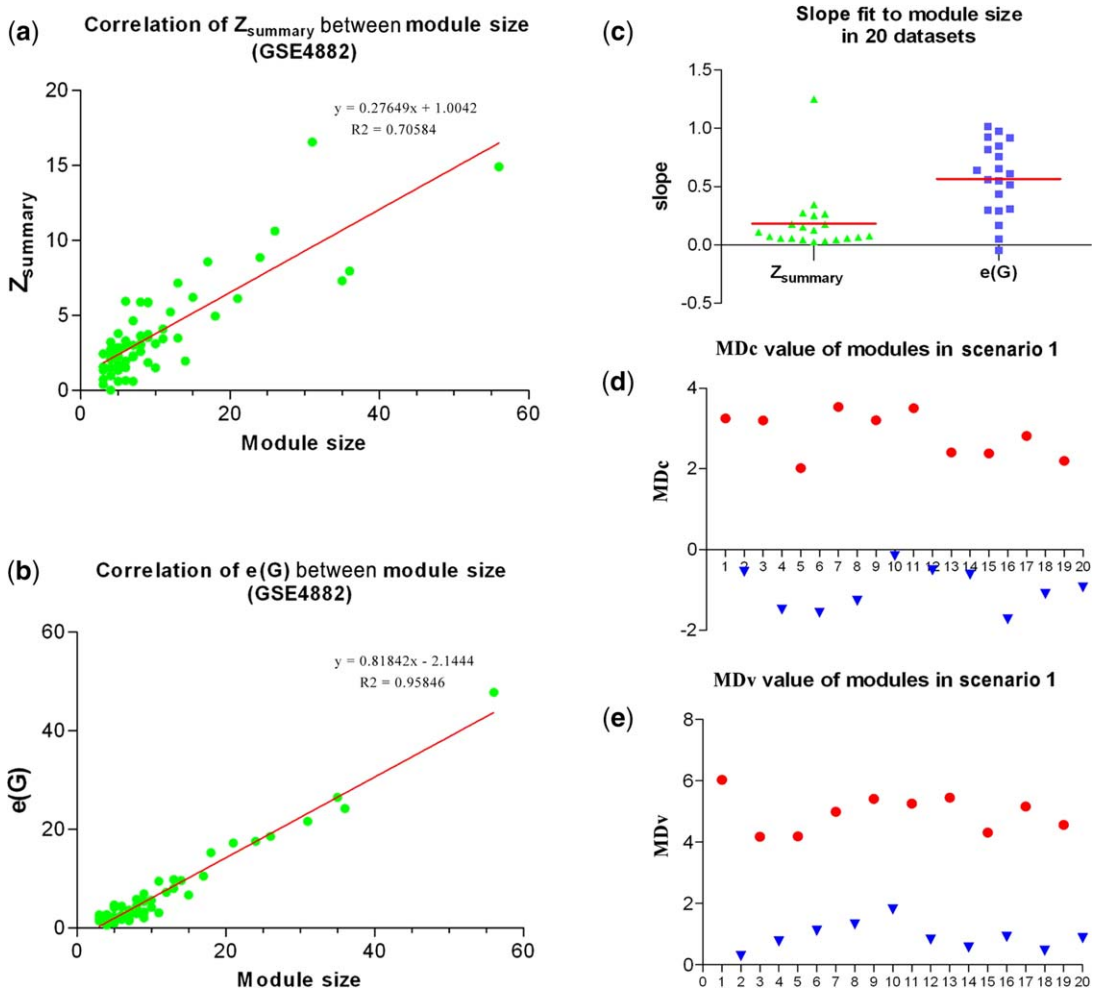

Figure 1 (**a,b**) Correlation between $Z_{summary}$/e(G) (y-axis) and module size (x-axis) using GSE4882 as an example. (**c**) The slope fit value (y-axis) between $Z_{summary}$/e(G) and module size in 20 datasets. The red line represents their mean value. (**d,e**) The conserved Modudaoism (MDc)/varied Modudaoism (MDv) values of the predefined preserved modules (in red circle) or discrepant modules (in blue triangle) in simulation scenario 1.

integrated $Z_{summary}$ index can assess whether modules are preserved under different conditions, but it is dependent on module size and there is a lack of internal module quality assessment.[14] Methods for identifying conserved or characteristic allosteric modules (AMs) based on topological variation are not yet available.

Our previous studies showed that baicalin (BA), jasminoidin (JA), and ursodeoxycholic acid (UA) had both similar and differential pharmacological mechanisms in treating cerebral ischemia at the pathway and network levels.[15–17] A $Z_{summary}$-based method was selected to identify the convergent or divergent modules, but the internal quality of modules was not considered.[18] In this paper, based on the novel concept of MD, we proposed conserved MD (MDc) and varied MD (MDv) values to quantitatively discriminate the conserved and discrepant AMs of BA, JA, and UA in mice anti-ischemic gene co-expression networks, so as to further elucidate and compare their pharmacological mechanisms.

MATERIALS AND METHODS

MDc and MDv module discrimination methods

In order to quantitatively define the conformational and energy variations of modules in networks, based on the MD concept, we proposed the optimized MDc value and MDv values by integrating the statistic $Z_{summary}$[14] and topological entropy e(G)[19] (**Supplementary Text S1**). The $Z_{summary}$ is an external module screening index composed of four statistics related to density and three statistics related to connectivity, which can quantitatively assess whether modules defined in a reference data set are preserved in a test dataset. However, $Z_{summary}$ is more or less dependent on module size (**Figure 1a**), which means that larger modules are more likely to be supposed as preserved. In addition, e(G) is an information-theoretic index to assess the modularity of a graph, and lower entropy indicates that the module has more inner links and less outer links, so it can measure the cluster quality effectively. As no test data are needed, e(G) is an internal module

evaluation index and it also has positive correlation with module size (**Figure 1b**).[14]

$$Z_{summary} = \frac{\text{median}(Z_{meanCor}, Z_{meanAdj}, Z_{propvarExpl}, Z_{meamKME}) + \text{median}(Z_{cor.KIM}, Z_{cor.KME}, Z_{cor.cor})}{2} \quad (1)$$

$$e(G) = \sum_{v \in V} e(v) \quad (2)$$

In view of the characteristics of $Z_{summary}$ and e(G) (Eqs. 1 and 2), a combination of these two complementary methods would be optimal. Therefore, we proposed the MDc value to identify preserved modules and the MDv value to define discrepant AMs. By computing the correlation slopes of $Z_{summary}$ and e(G) to module size in multiple datasets to estimate their respective weights, the extent of module size contributing to $Z_{summary}$ and e(G) can be determined.

In 20 gene expression datasets (**Supplementary Table S1**), the average correlation slopes of $Z_{summary}$ and e(G) to module size were 0.18 and 0.56, respectively, at a ratio of $\sim$1:3 (**Figure 1c**). Based on this contribution ratio, the weights of $Z_{summary}$ and e(G) were set at 3 and 1, respectively. In order to screen the significantly preserved modules, a larger positive $Z_{summary}$ and a smaller e(G) were expected, so the MDc value was defined as Eq. 3:

$$MDc = \frac{3}{4} Z_{summary} - \frac{1}{4} e(G) \quad (3)$$

On the other hand, in order to identify the significantly varied modules (AMs) under different conditions, such as pretreatment and post-treatment modules, a smaller negative $Z_{summary}$ and a smaller e(G) were expected, so the MDv value was defined as Eq. 4:

$$MDv = \frac{3}{4} Z_{summary} + \frac{1}{4} e(G) \quad (4)$$

Cutoff value of MDc and MDv

To quantitatively discriminate between preserved and discrepant AMs, we designed different simulation scenarios to determine the thresholds of MDc and MDv values. Three scenarios with different numbers of genes, modules, and sizes were designed: (1) 100 samples, 2,000 genes with 20 homogeneous size modules in the reference network; (2) 100 samples, 10 modules with varied sizes in the reference network, randomly generating 1,680 genes; and (3) 100 samples, 20 modules with varied sizes in the reference network, randomly generating 2,964 genes. For each scenario, one-half of the modules were simulated to be preserved in the test network, whereas the other half was not preserved. The simulated datasets were generated by the functions in the weighted gene co-expression network analysis (WGCNA) R software package[20] (**Supplementary Text S1**). Each module is simulated around a randomly chosen "seed eigenmode," in-module nodes are set to varied intramodular correlation levels. An empirical higher intramodular correlations (tightly coexpressed) are assigned to the preserved modules, as for the nonpreserved modules, they are expected to be zero.

In all of the three simulation scenarios, MDc/MDv performed well in distinguishing preserved from nonpreserved modules. The level of MDc/MDv that can differentiate between the predefined preserved and discrepant AMs was selected as the cutoff value. In scenario 1, the lowest MDc value of preserved module was 2.01, the highest MDc value of unpreserved module was -0.15 (**Figure 1d**), and the lowest MDv value of unpreserved module was 0.29 (**Figure 1e**). The results of scenarios two and three are shown in **Supplementary Figure S1**. Considering all these values, the following thresholds were defined: modules with an MDc value >0 were viewed as preserved modules, with an MDc >2 indicating strong preservation; modules with an MDv value <0 were considered as discrepant AMs.

Datasets. Twenty spotted DNA/complementary DNA expression datasets used to calculate the weights of $Z_{summary}$ and e(G) as well as the preservation ratio of MDc were obtained from GEO (http://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress databases (http://www.ebi.ac.uk/arrayexpress/). The raw gene expression datasets from different organisms and information about experiment platforms are shown in **Supplementary Table S1**. In these datasets, the sample size ranged from 12–597, and the number of genes ranged from 374–3,040. As a test dataset was needed for $Z_{summary}$ calculation, one half of the samples in each dataset were selected as reference and the other half as test.

The gene expression data of multicompounds in anti-ischemia models were derived from our previous studies,[21] which was obtained from the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/, E-TABM-612). The datasets of five groups were included in this study: (1) the sham group; (2) the vehicle group (0.9% NaCl); (3) the BA-treated group (5 mg/mL); (4) the JA-treated group (25 mg/mL); and (5) the UA-treated group (7 mg/mL). The procedures of MCAO model preparation, RNA isolation, microarray preparation, and gene collections were described previously.[21] Each dataset consisted of 374 cerebral ischemia-related complementary DNA expression profile data.

Co-expression module detection and discrimination

The gene co-expression network construction and module detection were implemented by WGCNA R package.[20] The topological overlap measure and Dynamic Hybrid Tree Cut algorithm were used to perform average linkage hierarchical clustering and partition the branches of dendrogram as modules.[22] Proper soft-thresholds were selected for each dataset when the network met the best scale-free topology criterion, and the minimum module size was set at three.

For each detected module in the reference datasets, the MDc and MDv values were calculated. Compared with the test datasets, modules with an MDc value >0 were defined as preserved or conserved AMs, with an MDc ≥ 2 indicating strong preservation; and modules with an MDv <0 were defined as discrepant AMs.

Conserved and on-module discrimination

For the gene expression datasets of multicompounds in anti-ischemia models, modules with an MDc value >0 were considered to be conserved allosteric modules (CAMs) compared with other groups. Compared with the vehicle group, modules in the drug groups with an MDv <0 were considered to be on-modules, representing structural disruption activated by the drug. If an on-module also had a negative MDv value compared with any other groups, this module was defined as a unique allosteric module (UAM) of this drug, which might reflect its specific mechanisms.

Functional analysis of modules

To characterize the functions of modules, GO and KEGG pathway enrichment analysis were performed by the Database for Annotation, Visualization, and Integrated Discovery (DAVID) platform.[23] For each module, an over-representation of a functionally relevant annotation was defined by a modified Fisher's exact *P* value with an adjustment for multiple tests using the Benjamini method, and GO terms and pathways with a $P < 0.05$ were considered as significant functions.

Co-immunoprecipitation experimental validation

A representative preserved module in the BA group (MDc = 0.97) consisting of JunD, FMO2, and FRAT1 was selected for experimental validation by co-immunoprecipitation (Co-IP). The edge between JunD and FMO2 was also observed in BA_9 UAM (MDv = −0.05). The Co-IP method can directly test whether two target proteins are combined or not, so as to prove the interaction relationship of proteins within a module. Standard Co-IP analyses were performed as described previously.[24] In brief, protein extracts were incubated with the agarose-conjugated antibodies in Lysis buffer at 4°C for 3 hours and precipitated by centrifugation. The precipitant was washed 4 times and then boiled for 5 minutes in 1 × SD Sample buffer. After centrifugation, the supernatant was run on an sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel, followed by Western blotting. The antibodies, including anti-JunD (rabbit, sc-74; SantaCruz), anti-FMO2 (goat, sc-83827; SantaCruz), and anti-FRAT1 (rabbit, ab108405; Abcam) were used in the Co-IP experiment.

RESULTS

Preserved module screening by MDc and $Z_{summary}$ in 20 datasets

To demonstrate the effectiveness of MDc for module screening, we used MDc to define the preserved modules in 20 gene expression datasets. All modules were identified by WGCNA with the same parameter settings as described in the Methods section. Based on the MDc cutoff value, the preserved modules were identified from all the datasets, including that of the BA group in E-TABM-612, which was originated from our previous studies.[20] A preserved module of BA was selected for Co-IP experimental validation. The proportions of strong preserved modules (MDc ≥ 2; $Z_{summary} \geq 10$) and weak preserved modules (MDc >0; $Z_{summary} \geq 2$) are shown in **Figure 2a,b**. Compared with $Z_{summary}$, the MDc method resulted in a higher proportion of strong preserved modules (MDc = 25.85%; $Z_{summary}$ = 17.95%) and a similar proportion of weak preserved modules (MDc = 62.50%; $Z_{summary}$ = 63.25%; **Figure 2c**). This indicated that the MDc method might identify more preserved modules than $Z_{summary}$.

MDc was less affected by module size than $Z_{summary}$

As module size is positively correlated with $Z_{summary}$, larger modules are more likely to be judged as preserved by $Z_{summary}$, but e(G) may adjust this bias, using GSE4882 as an example (**Figure 1a,b**). In the 20 datasets, the average linear correlation coefficient R^2 between MDc and module size was 0.51, whereas that between $Z_{summary}$ and module size was 0.55 (**Figure 2d**). This illustrated that MDc was less affected by module size than $Z_{summary}$.

Co-expression modules of BA, JA, and UA

Detection of gene co-expression modules in the BA, JA, and UA groups were performed by WGCNA, as described in the Methods section. The network hotmaps of all genes in the three groups are shown in **Supplementary Figure S2**. With the appropriate soft-thresholding for each group ($\beta = 4$ for BA, 12 for JA, and 8 for UA), hierarchical clustering procedure obtained 23, 42, and 15 modules in the BA, JA, and UA groups, respectively. Detailed information about the gene members of modules in each group labeled by colors and numbers can be found in **Supplementary Table S2**.

MDc-based CAMs of BA, JA, and UA

In order to test the MDc method in common module discrimination for multi-drugs, we used MDc to identity the CAMs of BA, JA, and UA in anti-ischemia models, which were originated from our previous study.[20] Modules with an MDc >0 were judged as conserved. Compared with the vehicle group, the BA group had 3 CAMs (i.e., BA_7, BA_10, and BA_14), the JA group had 5 CAMs (i.e., JA_7, JA_16, JA_17, JA_25, and JA_33), and the UA group had 1 CAM (i.e., UA_6). When pairwise comparisons were performed among the 3 drug groups, BA had 6 CAMs as compared with JA and UA, of which BA_6, BA_8, BA_10, and BA_12 were conserved in all the 3 groups; JA had 3 and 5 CAMs when compared with BA and UA, respectively, of which JA_16 was conserved in all the 3 groups; UA had 3 CAMs compared with BA, and no CAM was found when compared with JA. Detailed MDc values of modules in one drug group when compared with the vehicle and other drug groups are listed in **Supplementary Table S3**.

Next, we compared the proportion of CAMs in each group identified by MDc (>0) and $Z_{summary}$ (≥ 2). Details on the proportion of CAMs in each group are shown in **Figure 3a**. Except the UA-JA module comparison, MDc found more CAMs than $Z_{summary}$. Among the vehicle and 3 drug groups, the average proportions of CAMs identified by MDc and $Z_{summary}$ were 13.6% and 3.5%, respectively (**Figure 3b**). Four modules were defined as conserved by both MDc and $Z_{summary}$, and their detailed values and compositions are shown in **Figure 3c**.

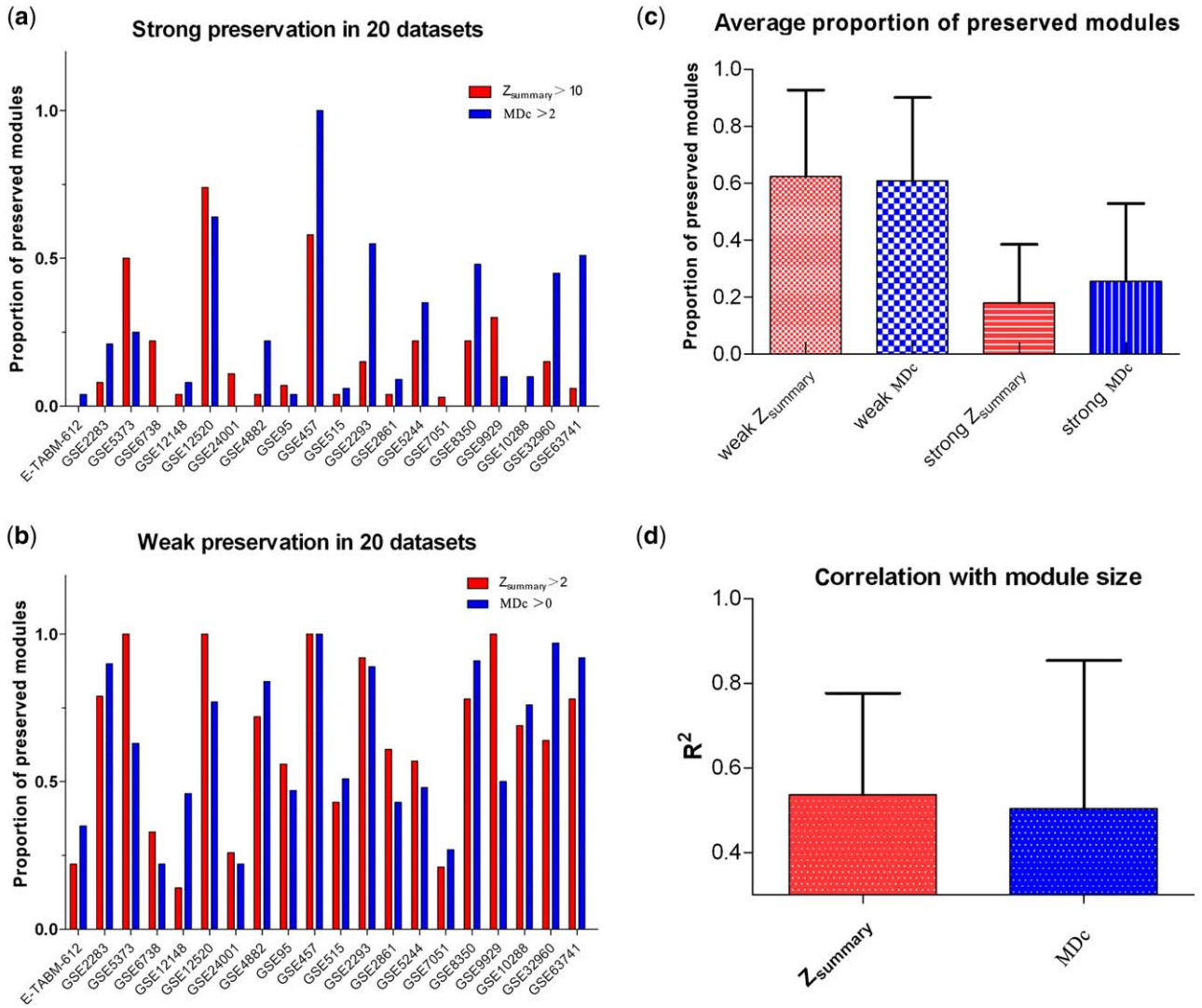

Figure 2 (**a**,**b**) Comparison of the proportions of strong preserved modules ($Z_{summary} \geq 10$, conserved Modudaoism (MDc) ≥ 2) and weak preserved modules ($Z_{summary} \geq 2$, MDc > 0) in 20 datasets. (**c**) The average proportions of strong and weak preserved modules identified by $Z_{summary}$ and MDc in 20 datasets. (**d**) The correlation coefficient R^2 between $Z_{summary}$/MDc and module size in 20 datasets.

Biological functions of CAMs of each group

GO term and KEGG pathway enrichment analysis were performed to characterize the biological functions of the identified CAMs. The number of CAMs and their functions among in each group are shown in **Figure 4**. Two modules (BA_10 and JA_16) were conserved in all the four groups, but no annotation of function was enriched in both modules. We listed all of the significantly enriched GO terms and pathways ($P < 0.05$) in **Supplementary Table S4**. In the vehicle group, the enriched GO terms of CAMs included cytoskeleton organization, protein amino acid phosphorylation, and regulation of mitochondrial membrane permeability; and the enriched KEGG pathways included regulation of actin cytoskeleton, mitogen-activated protein kinase (MAPK) signaling pathway, neurotrophin signaling pathway, amyotrophic lateral sclerosis, and RIG-I-like receptor signaling pathway.

Among the 3 drug groups, 3 modules (BA_6, BA_8, and BA_12) were commonly conserved, and their enriched GO functions included phosphorylation, phosphate metabolic process, and phosphorus metabolic process. The CAMs between BA and JA (BA_4, BA_11, JA_5, and JA_40) were most significantly enriched in the GO function of intracellular signaling cascade and the pancreatic cancer pathway. The CAMs between BA and UA (BA_13, UA_8, and UA_13) were most significantly enriched in the GO function of protein kinase cascade and the KEGG pathways of colorectal cancer and viral myocarditis. The CAMs between JA and UA (JA_6, JA_8, JA_31, and JA_42) were most significantly enriched in the GO function of ATP binding and the long-term potentiation pathway.

MDv-based discrepant modules of BA, JA, and UA

The co-expression pattern variation of modules might be reflected by structural rewiring. A negative $Z_{summary}$ value

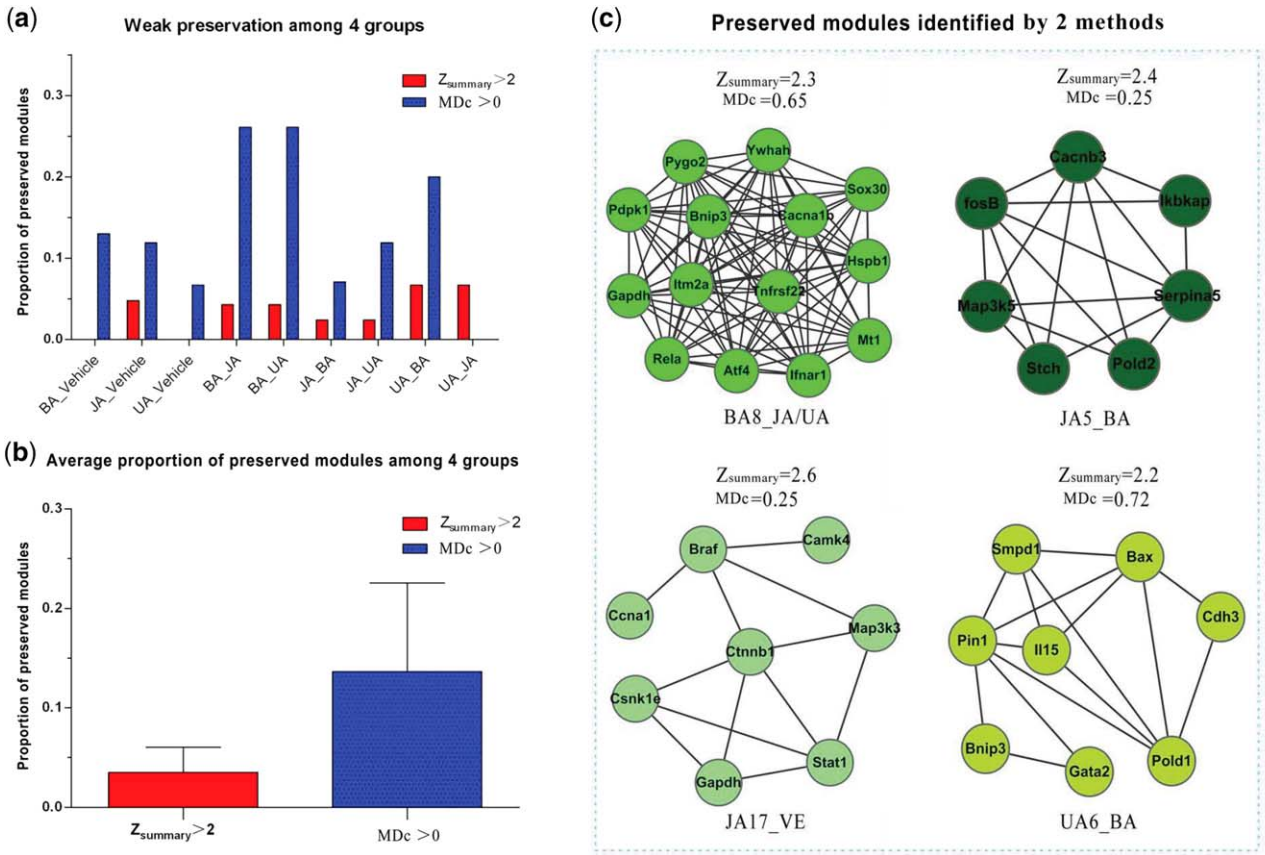

Figure 3 (**a**) Comparison of the proportions of conserved allosteric modules (CAMs; $Z_{summary} \geq 2$ or conserved Modudaoism (MDc) >0) among baicalin (BA), jasminoidin (JA), ursodeoxycholic acid (UA), and vehicle (VE) groups. (**b**) The average proportions of CAMs identified by $Z_{summary}$ and MDc in BA, JA, UA, and vehicle groups. (**c**) The CAMs identified by both $Z_{summary}$ and MDc.

may indicate module disruption,[25] but the topological quality of modules in the network has not been taken into consideration. Thus, we used the MDv value to quantitatively define modules whose co-expression pattern was significantly changed as discrepant modules (significantly changed modules). With the addition of e(G), the MDv value was a more stringent criterion than $Z_{summary}$.

We compared the proportion of discrepant modules in each group identified by MDv (<0) and $Z_{summary}$ (<0). Details on the number and proportion of discrepant modules in each group are shown in **Figure 5a**. Among the vehicle and 3 drug groups, the average proportion of discrepant modules identified by MDv and $Z_{summary}$ were 47.9% and 9.9%, respectively (**Figure 5b**). The discrepant modules defined by MDv had better topological structural quality than those defined by $Z_{summary}$, and their average e(G)s were 8.18 and 1.15, respectively (**Figure 5c**).

MDv-based on-modules and UAMs of BA, JA, and UA

The co-expression pattern variation of modules independent of the vehicle group might reflect the pharmacological actions of the three drugs, so the discrepant modules between drug and vehicle were responsive. Compared with the vehicle group, we found 4 (BA_4, BA_6, BA_9, and BA_19), 1 (JA_35), and 3 (UA_4, UA_8, and UA_12) on-modules (MDv <0) in the BA, JA, and UA groups, respectively. Using BA_6 and UA_12 as examples, their MDv values were -0.46 and -0.39, and their e(G)s were also small (i.e., 0.37 and 1.51, respectively; **Figure 5d**). Modules with a negative $Z_{summary}$ but a large e(G) were not judged as on-modules, such as JA_28 and UA_5 (**Figure 5e**).

Next, we pairwise compared the on-modules among the three drug groups to screen out the characteristic modules (UAMs) of each drug, and these unique target modules may reflect the distinct actions of different drugs in treating a same disease. The detailed MDv values of modules in each group are listed in **Supplementary Table S3**. Among these on-modules, 2 modules of BA (BA_9 and BA_19), 1 module of JA (JA_35), and 2 modules of UA (UA_4 and UA_12) were discrepant when compared with the other 2 groups, which were considered the UAMs (**Figure 6a**). Among the member genes in the UAMs, Gpx2 and JunD in BA_9, Htr2c in BA_19, B230120H23Rik in JA_35, and Dusp10 in UA_12 were significantly differentially expressed compared with vehicle based on one-way analysis of variance.

Characteristic functions of BA, JA, and UA in treating cerebral ischemia

To characterize the divergent biological functions of different drugs, we compared the GO functions and pathways of

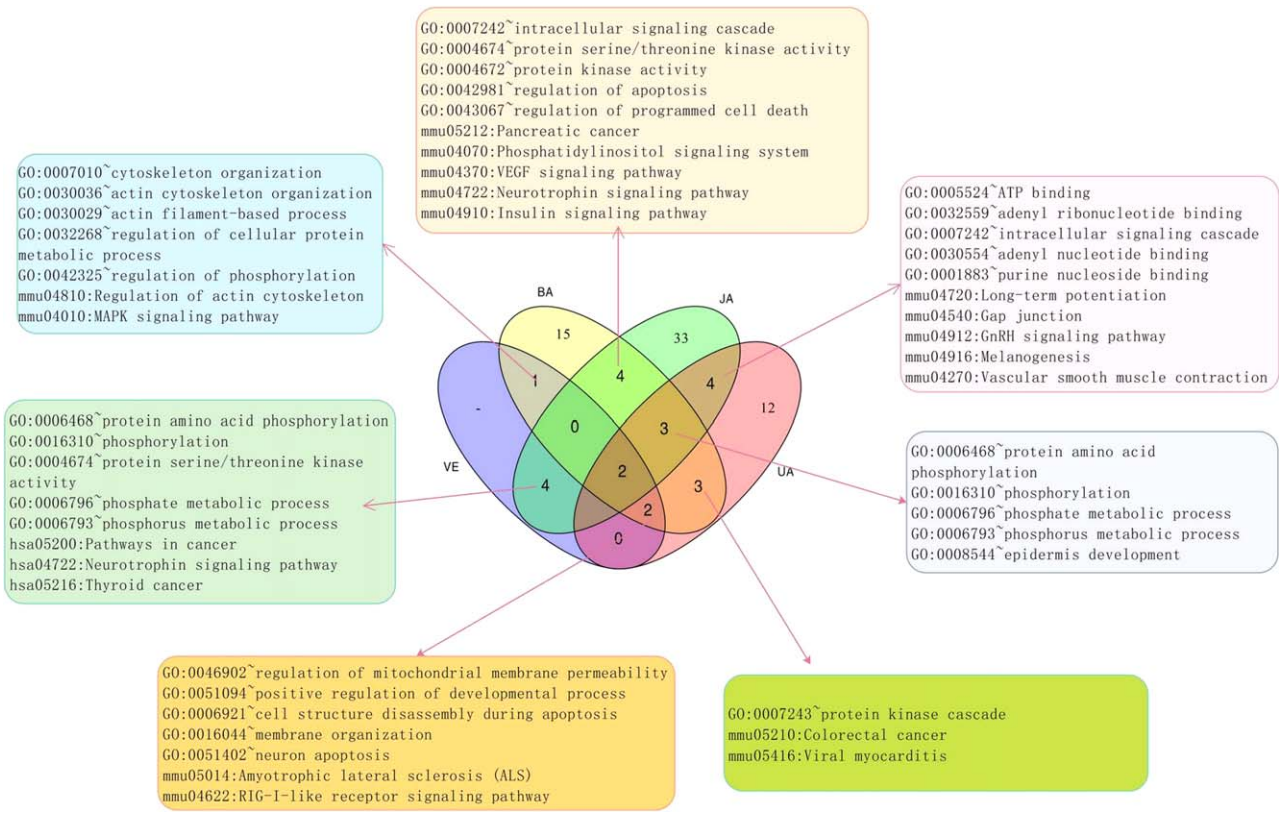

Figure 4 The number of conserved allosteric modules in baicalin (BA), jasminoidin (JA), ursodeoxycholic acid (UA), and vehicle (VE) groups and their significant biological functions. The top five significantly enriched GO terms and KEGG pathways of the unique allosteric modules (UAMs) are listed. The venn diagram in the middle indicates the group that the UAM belongs to. MAPK, mitogen-activated protein kinase; VEGF, vascular endothelial growth factor.

the on-modules among the three drug groups. The significantly enriched GO terms and pathways of on-modules are listed in **Supplementary Table S4**. The BA_4 module was preserved in the JA group, in which several GO terms, such as intracellular signaling cascade, regulation of apoptosis, and regulation of cell death as well as KEGG pathways, such as phosphatidylinositol signaling system, neurotrophin signaling pathway, and vascular endothelial growth factor signaling pathway were significantly enriched. The UA_8 module was conserved in the BA group, in which protein kinase cascade was significantly enriched.

As for the UAMs (**Figure 6a**), Wnt receptor signaling pathway, negative regulation of nitrogen compound metabolic process, and anterior/posterior pattern formation were significantly enriched in the BA_9 and BA_19 modules. The GO term of magnesium ion binding and the pathway of progesterone-mediated oocyte maturation were significantly enriched in the JA_35 module. Negative regulation of cellular component organization, regulation of organelle organization and regulation of cell cycle were significantly enriched in the UA_4 and UA_12 modules. These characteristic functions may reveal the specific mechanisms of BA, JA, and UA in the treatment of cerebral ischemia.

Co-IP experimental validation

First, goat anti-FMO2 was used to assess whether the FMO2 protein had any interactions with JunD and FRAT1 by Co-IP. Results showed that a heavy chain of antibodies was found in the negative group and experimental group (Co-IP), and the JunD and FRAT1 proteins were also found in the two positive controls (**Figure 6b**). We then used rabbit anti-FRAT1 as the input protein to observe its interaction with JunD, and the JunD expression was determined by Western blotting (**Figure 6c**). Our findings revealed that FMO2, JunD, and FRAT1 interacted with one another, and, thus, the modular structural relationship was demonstrated.

DISCUSSION

Considering the modular basis of diseases or drug intervention networks, changes in modular topology or network rewiring could better reflect the actions of drugs.[26–28] In the modular pharmacological paradigm, module evaluation or target module identification according to modular topological variation is regarded as a critical step. Structural variation necessarily causes energy fluctuation in a biomolecular system, such as human chromosomes, proteins, or PPIs,[29–31] and, thus, energy variation of a module is another factor to be considered. This study proposed an optimized MD-based method to quantitatively evaluate the topological structural and energy variations under different conditions. Compared with $Z_{summary}$, the MDc/MDv value was less affected by module size with a better topological quality of modules. When

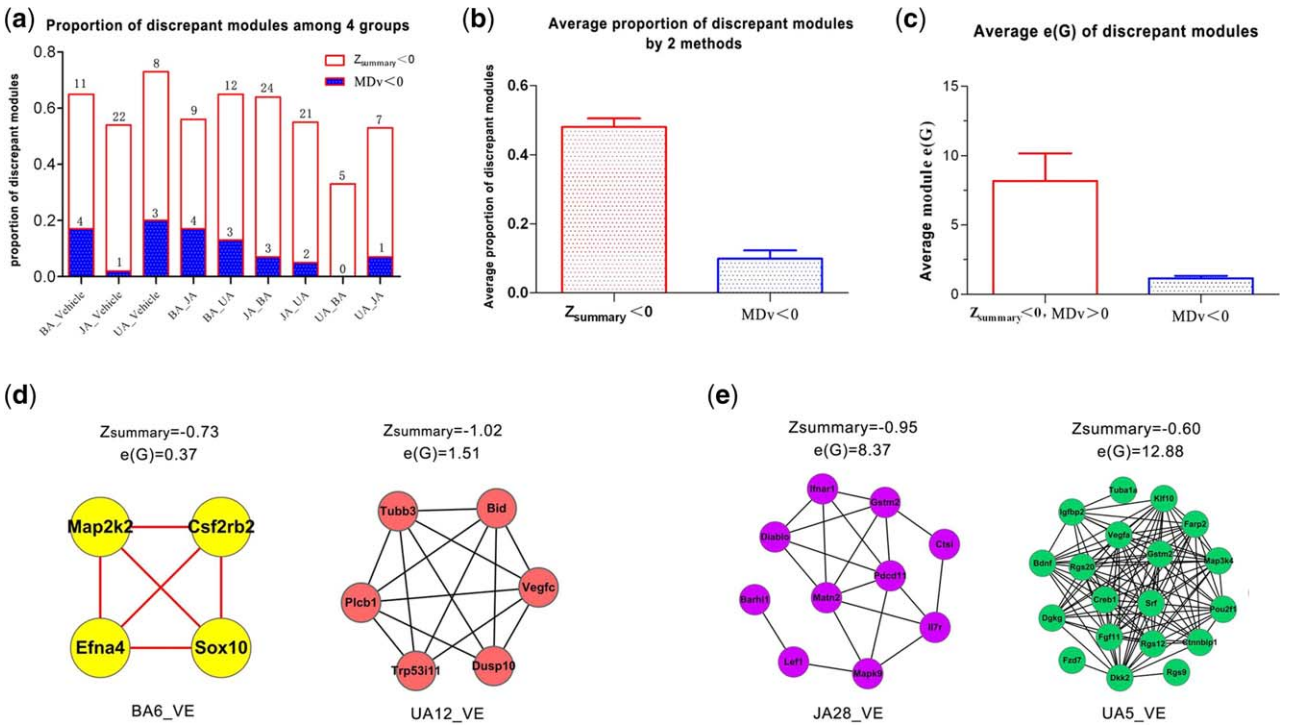

Figure 5 (**a**) Comparison of the proportions of differential modules ($Z_{summary}$ <0 or varied Modudaoism (MDv) <0) among baicalin (BA), jasminoidin (JA), ursodeoxycholic acid (UA) and vehicle (VE) groups. (**b**) The average proportions of differential modules identified by $Z_{summary}$ and MDv in the four groups. (**c**) The average e(G) of differential modules in the four groups. (**d**) The sample on-modules with negative MDv and $Z_{summary}$ values as well as a small e(G). (**e**) The sample modules with a negative $Z_{summary}$ value but a large e(G), which were not identified as on-modules.

used for comparison of multiple compounds, MDc/MDv can effectively discriminate between CAMs and UAMs of BA, JA, and UA, so as to reveal the diverse drug-induced co-expression patterns and elucidate the pharmacological effects of drugs from the modular perspective.

As a module is generally considered as a closely linked unit to perform biological functions at the network level, drug-induced network rewiring may lead to changes in modular structure.[7] The equilibrium in a biomolecular system results from a balance between energy and entropy,[32–34] and entropy-driven structural and energy variations may be a critical factor to identify the active modules in biological networks. Many previous studies have discussed the conserved gene expressions[35] or modules,[36–39] but few methods have been exploited to measure the structural and energy variations of modules. Methods based on overlapping nodes or edges may not reflect the modular topological properties globally,[40,41] and optimized models are required to quantitatively evaluate module variation from a topological structure and energy perspective. In this case, the MDc/MDv method may provide a novel framework to explore drug's modular targets and compare the actions of multiple drugs in disease treatment.

Based on the MDc value, more CAMs in BA, JA, and UA were identified than the $Z_{summary}$-based method.[18] Some modules were consistently defined as CAMs by both methods, such as the BA_8, JA_5, and UA_6 modules. Among them, the BA_8 module was conserved in all the 3 compound groups, and its functions involved regulation of apoptosis and cell death, transcription activator activity, MAPK, neurotrophin signaling pathway, etc. These functions were shown to be closely related to cerebral ischemia and also consistent with the findings from our previous studies.[18,42] As for other CAMs in the three compound groups, the phosphorylation process was significantly enriched, which was a basically pathologic process in cerebral ischemia.[42] These CAMs between different drugs reflect their common or synergetic pharmacological actions in treating a disease.

With the integration of network entropy index, the MDv value used for on-module identification was considered a more stringent criterion, so fewer on-modules were identified by MDv, but the reduced entropy represented better modular structure.[19] According to the MDv value, the on-modules and characteristic UAMs of different drugs may help to elucidate their specific or unique actions. In the on-module of BA_4, anti-apoptosis and regulation of apoptosis functions were enriched, consistent with our previous findings that BA had anti-oxidative and anti-apoptotic actions to exert neuroprotective effects.[43] Among the UAMs of BA, JA, and UA, no overlap of enriched GO terms and pathways was found, indicating the characteristic role of UAMs. In the UAMs of all three drugs, cerebral ischemia-related functions were consistently enriched, such as Wnt signaling pathway in BA,[44,45] magnesium ion binding in JA,[46] and negative regulation of apoptosis in UA.[47,48] Besides, other

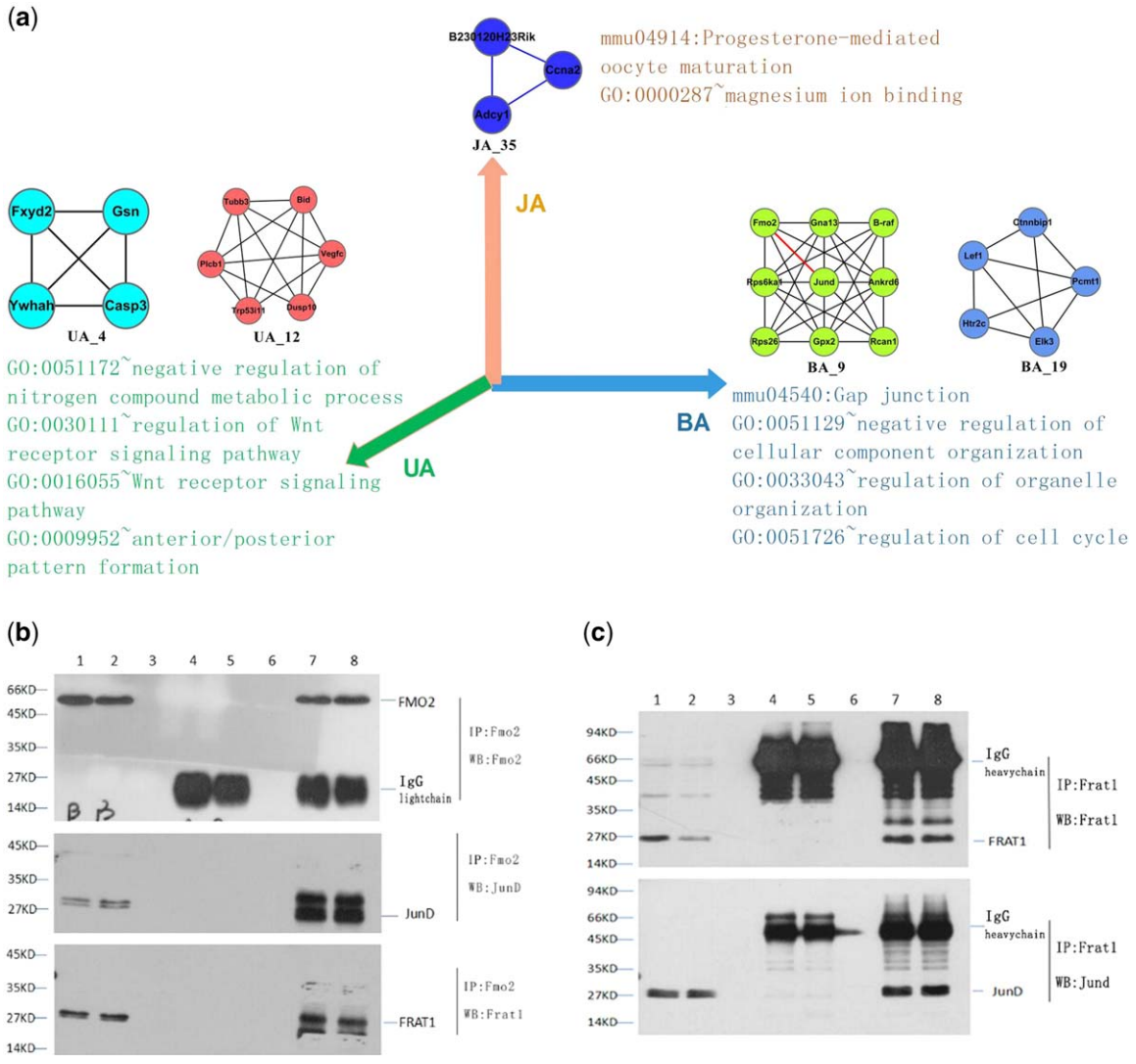

Figure 6 (**a**) The unique allosteric modules of baicalin (BA), jasminoidin (JA), ursodeoxycholic acid (UA), and their significant biological functions. The top five significantly enriched GO terms and KEGG pathways are listed. (**b,c**) The interactions among JunD, FMO2, and FRAT1 as determined by co-immunoprecipitation. 1 and 2 = input the samples; 3 = loading buffer; 4 and 5 = input the negative immunoglobulin G (IgG); 6 = loading buffer; 7 and 8 = input the positive antibody.

functions that were also found related to the three compounds need to be confirmed by further studies, such as negative regulation of nitrogen compound metabolic process in BA, progesterone-mediated oocyte maturation in JA, and regulation of cellular component organization in UA.

Although several attempts have been made to optimize the identification of target on-modules, this study still has some limitations. As the samples and gene numbers in microarray were relatively small, the modular connectivity might be inadequate. In addition, regulation of biological networks should be a dynamic process,[49,50] and, thus, the dynamic factor should also be considered when characterizing modular topological transformation, such as environmental changes, different time series, or evolutionary processes.

In summary, the MD-based method is effective in quantitatively discriminating between the CAMs and UAMs in multiple network comparison. The UAMs of BA, JA, and UA identified by MDc/MDv values revealed their divergent and characteristic actions in treating cerebral ischemia, such as negative regulation of nitrogen metabolism and Wnt signaling pathways of BA, magnesium ions binding and the progesterone-mediated oocyte maturation pathway of JA, and regulation of cell organization, organ organization, and cell cycle of UA. This MD-based modular analysis may provide a unique and promising strategy for pharmacological research and development.

Author Contributions. B.L. wrote the manuscript. Z.W., J.L., and Y.W. designed the research. B.L., P.W., and Y.Y. performed the research. B.L., P.W., Y.Y., X.N., Q.L., and X.Z. analyzed the data. J.L., Y.Z., and X.N. contributed new reagents/analytical tools.

Source of Funding. The authors' work is funded by the National Natural Science Foundation of China (81673833) and the Fundamental

Research Funds for the Central Public Welfare Research Institutes (ZZ0908029).

Conflict of Interest. The authors declare no competing financial interests.

1. Wang, Z. & Wang, Y.Y. Modular pharmacology: deciphering the interacting structural organization of the targeted networks. *Drug Discov. Today* **18**, 560–566 (2013).
2. Mousavian, Z., Nowzari-Dalini, A., Stam, R.W., Rahmatallah, Y. & Masoudi-Nejad, A. Network-based expression analysis reveals key genes related to glucocorticoid resistance in infant acute lymphoblastic leukemia. *Cell. Oncol. (Dordr).* **40**, 33–45 (2017).
3. Iskar, M. *et al.* Characterization of drug-induced transcriptional modules: towards drug repositioning and functional understanding. *Mol. Syst. Biol.* **9**, 662 (2013).
4. Feng, L. *et al.* A network-based method for identifying prognostic gene modules in lung squamous carcinoma. *Oncotarget* **7**, 18006–18020 (2016).
5. Pan, J., Cong, Z., Zhong, M. & Sun, Y. Analysis of hepatocellular carcinoma and metastatic hepatic carcinoma via functional modules in a protein-protein interaction network. *J. Cancer Res. Ther.* **10**(suppl.), C186–C194 (2014).
6. Wen, Z., Liu, Z.P., Liu, Z., Zhang, Y. & Chen, L. An integrated approach to identify causal network modules of complex diseases with application to colorectal cancer. *J. Am. Med. Inform. Assoc.* **20**, 659–667 (2013).
7. Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A. & Gerstein, M. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308–312 (2004).
8. Zeng, T., Wang, D.C., Wang, X., Xu, F. & Chen, L. Prediction of dynamical drug sensitivity and resistance by module network rewiring-analysis based on transcriptional profiling. *Drug Resist. Updat.* **17**, 64–76 (2014).
9. Chen, Y., Wang, Z. & Wang, Y. Spatiotemporal positioning of multipotent modules in diverse biological networks. *Cell. Mol. Life Sci.* **71**, 2605–2624 (2014).
10. Li, B. *et al.* Quantitative assessment of gene expression network module-validation methods. *Sci. Rep.* **5**, 15258 (2015).
11. Yu, S. *et al.* Causal co-expression method with module analysis to screen drugs with specific target. *Gene* **518**, 145–151 (2013).
12. Liu, R., Cheng, Y., Yu, J., Lv, Q.L. & Zhou, H.H. Identification and validation of gene module associated with lung cancer through coexpression network analysis. *Gene* **563**, 56–62 (2015).
13. Chen, Y.Y. *et al.* Quantitative determination of flexible pharmacological mechanisms based on topological variation in mice anti-ischemic modular networks. *PLoS One* **11**, e0158379 (2016).
14. Langfelder, P., Luo, R., Oldham, M.C. & Horvath, S. Is my network module preserved and reproducible? *PLoS Comput. Biol.* **7**, e1001057 (2011).
15. Wang, P.Q. *et al.* Phenotype-dependent alteration of pathways and networks reveals a pure synergistic mechanism for compounds treating mouse cerebral ischemia. *Acta Pharmacol. Sin.* **36**, 734–747 (2015).
16. Chen, Y. *et al.* Hierarchical profiles of signaling pathways and networks reveal two complementary pharmacological mechanisms. *CNS Neurol Disord Drug Targets* **12**, 882–893 (2013).
17. Li, B. *et al.* Vertical and horizontal convergences of targeting pathways in combination therapy with baicalin and jasminoidin for cerebral ischemia. *CNS Neurol. Disord. Drug Targets* **15**, 740–750 (2016).
18. Li, B. *et al.* Quantitative identification of compound-dependent on-modules and differential allosteric modules from homologous ischemic networks. *CPT Pharmacometrics Syst. Pharmacol.* **5**, 575–584 (2016).
19. Kenley, E.C. & Cho, Y.R. Detecting protein complexes and functional modules from protein interaction networks: a graph entropy approach. *Proteomics* **11**, 3835–3844 (2011).
20. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
21. Wang, Z. *et al.* Fusion of core pathways reveals a horizontal synergistic mechanism underlying combination therapy. *Eur. J. Pharmacol.* **667**, 278–286 (2011).
22. Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2008).
23. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
24. Wong, E.W., Mruk, D.D., Lee, W.M. & Cheng, C.Y. Par3/Par6 polarity complex coordinates apical ectoplasmic specialization and blood-testis barrier restructuring during spermatogenesis. *Proc. Natl. Acad. Sci. USA* **105**, 9657–9662 (2008).
25. Iancu, O.D. *et al.* Differential network analysis reveals genetic effects on catalepsy modules. *PLoS One* **8**, e58951 (2013).
26. Wang, Z., Liu, J., Yu, Y., Chen, Y. & Wang, Y. Modular pharmacology: the next paradigm in drug discovery. *Expert Opin. Drug Discov.* **7**, 667–677 (2012).
27. Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N. & Barabási, A.L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002).
28. Lorenz, D.M., Jeng, A. & Deem, M.W. The emergence of modularity in biological systems. *Phys. Life Rev.* **8**, 129–160 (2011).
29. Zhang, B. & Wolynes, P.G. Topology, structures, and energy landscapes of human chromosomes. *Proc. Natl. Acad. Sci. USA* **112**, 6062–6067 (2015).
30. Yoshidome, T. *et al.* Free-energy function based on an all-atom model for proteins. *Proteins* **77**, 950–961 (2009).
31. Kamisetty, H., Ramanathan, A., Bailey-Kellogg, C. & Langmead, C.J. Accounting for conformational entropy in predicting binding free energies of protein-protein interactions. *Proteins* **79**, 444–462 (2011).
32. Henderson, D. Attractive energy and entropy or particle size: the yin and yang of physical and biological science. *Interdiscip. Sci.* **1**, 1–11 (2009).
33. Choi, H., Kang, H. & Park, H. Computational prediction of molecular hydration entropy with hybrid scaled particle theory and free-energy perturbation method. *J. Chem. Theory Comput.* **11**, 4933–4942 (2015).
34. Ovchinnikov, V., Cecchini, M. & Karplus, M. A simplified confinement method for calculating absolute free energies and free energy and entropy differences. *J. Phys. Chem. B* **117**, 750–762 (2013).
35. Herschkowitz, J.I. *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.* **8**, R76 (2007).
36. Hou, L. *et al.* Modular analysis of the probabilistic genetic interaction network. *Bioinformatics* **27**, 853–859 (2011).
37. Azuaje, F., Devaux, Y. & Wagner, D.R. Coordinated modular functionality and prognostic potential of a heart failure biomarker-driven interaction network. *BMC Syst. Biol.* **4**, 60 (2010).
38. Beltrao, P., Cagney, G. & Krogan, N.J. Quantitative genetic interactions reveal biological modularity. *Cell* **141**, 739–745 (2010).
39. Huang, H. *et al.* Identification of polycystic ovary syndrome potential drug targets based on pathobiological similarity in the protein-protein interaction network. *Oncotarget* **7**, 37906–37919 (2016).
40. Kovács, I.A., Palotai, R., Szalay, M.S. & Csermely, P. Community landscapes: an integrative approach to determine overlapping network module hierarchy, identify key nodes and predict network dynamics. *PLoS One* **5**, e12528 (2010).
41. Liu, C., Li, J. & Zhao, Y. Exploring hierarchical and overlapping modular structure in the yeast protein interaction network. *BMC Genomics* **11** Suppl 4, S17 (2010).
42. Kalogeris, T., Baines, C.P., Krenz, M. & Korthuis, R.J. Cell biology of ischemia/reperfusion injury. *Int. Rev. Cell Mol. Biol.* **298**, 229–317 (2012).
43. Cao, Y. *et al.* Baicalin attenuates global cerebral ischemia/reperfusion injury in gerbils via anti-oxidative and anti-apoptotic pathways. *Brain Res. Bull.* **85**, 396–402 (2011).
44. Liu, B. *et al.* Involvement of the Wnt signaling pathway and cell apoptosis in the rat hippocampus following cerebral ischemia/reperfusion injury. *Neural Regen. Res.* **8**, 70–75 (2013).
45. Sun, F.L. *et al.* Promoting neurogenesis via Wnt/β-catenin signaling pathway accounts for the neurorestorative effects of morroniside against cerebral ischemia injury. *Eur. J. Pharmacol.* **738**, 214–221 (2014).
46. Kuzenkov, V.S. & Krushinskii, A.L. Protective effect of magnesium nitrate against neurological disorders provoked by cerebral ischemia in rats. *Bull. Exp. Biol. Med.* **157**, 721–723 (2014).
47. Zhu, C. *et al.* The influence of age on apoptotic and other mechanisms of cell death after cerebral hypoxia-ischemia. *Cell Death Differ.* **12**, 162–176 (2005).
48. Nakka, V.P., Gusain, A., Mehta, S.L. & Raghubir, R. Molecular mechanisms of apoptosis in cerebral ischemia: multiple neuroprotective opportunities. *Mol. Neurobiol.* **37**, 7–38 (2008).
49. Jin, R., McCallen, S., Liu, C.C., Xiang, Y., Almaas, E. & Zhou, X.J. Identifying dynamic network modules with temporal and spatial constraints. *Pac. Symp. Biocomput.* 203–214 (2009).
50. Liu, J. & Wang, Z. Diverse array-designed modes of combination therapies in Fangjiomics. *Acta Pharmacol. Sin.* **36**, 680–688 (2015).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://psp-journal.com)