

# SCIENTIFIC REPORTS

OPEN

## Combining Chemical Profiling and Network Analysis to Investigate the Pharmacology of Complex Prescriptions in Traditional Chinese Medicine

Received: 14 July 2016  
Accepted: 07 December 2016  
Published: 13 January 2017

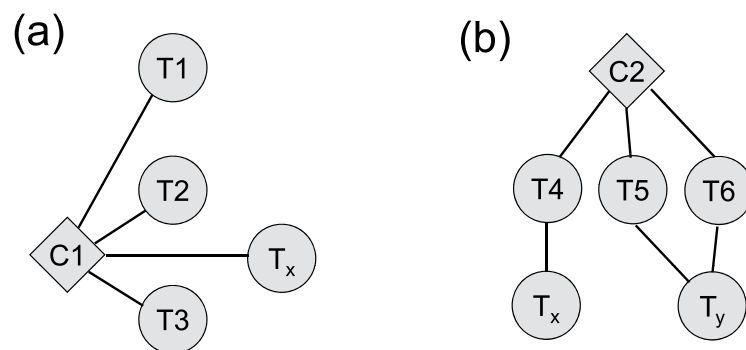
Tongchuan Suo<sup>1,\*</sup>, Jinping Liu<sup>2,\*</sup>, Xi Chen<sup>1</sup>, Hua Yu<sup>2</sup>, Tenglong Wang<sup>2</sup>, Congcong Li<sup>2</sup>, Yuefei Wang<sup>2</sup>, Chunhua Wang<sup>2</sup> & Zheng Li<sup>1,2</sup>

We present a paradigm, combining chemical profiling, absorbed components detection in plasma and network analysis, for investigating the pharmacology of combination drugs and complex formulae. On the one hand, the composition of the formula is investigated comprehensively via mass spectrometry analysis, followed by pharmacological studies of the fractions as well as the plasma concentration testing for the ingredients. On the other hand, both the candidate target proteins and the effective ingredients of the formula are predicted via analyzing the corresponding networks. The most probable active compounds can then be identified by combining the experimental results with the network analysis. In order to illustrate the performance of the paradigm, we apply it to the Danggui-Jianzhong formula (DJF) from traditional Chinese medicine (TCM) and predict 4 probably active ingredients, 3 of which are verified experimentally to display anti-platelet activity, i.e., (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose. Moreover, the 3-compound formulae composed of these 3 chemicals show better anti-platelet activity than DJF. In addition, the paradigm predicts the association between these 3 compounds and COX-1, and our experimental validation further shows that such association comes from the inhibitory effects of the compounds on the activity of COX-1.

Prescriptions in traditional Chinese medicine (TCM) are well known by their adoption of “multi-chemical components” to take “multi-pharmacological effects” on “multi-action targets”<sup>1</sup>. However, the complicated chemical composition also brings great difficulties to the pharmacological investigations of TCM prescriptions. Network pharmacology, which was first proposed by Hopkins<sup>2</sup>, offers an ideal paradigm to deal with multi-target combination drugs and has recently been successfully adopted to investigate the formulae in TCM<sup>3–7</sup>. The core of the scheme is the construction and analysis of the pharmacological network, which is normally composed of the nodes of active ingredients, the nodes of candidate protein targets, the nodes of intermediate proteins transferring protein-protein interactions (PPI) and the connections (i.e., edges) between them. While the PPI can always be collected from online databases, it is essential to have the chemical composition of the prescription and the candidate protein targets *a priori* in order to build the network.

In practice, the chemical ingredients of herbs and other TCM medicinal materials may be found from several databases. However, it is not uncommon that the composition of a TCM prescription differs dramatically from the simple summation of the ingredients of each medicinal component. For example, in the work of Yang *et al.* on Xiao-Ke-An, 40 chemicals were found from relevant databases while only 20 compounds were identified experimentally<sup>5</sup>. Moreover, the compounds identified *in vitro* may not be able to enter the plasma, and thus could not really explain the mechanisms of these TCM formulae. Hence, it is necessary to use chemical profiling to obtain

<sup>1</sup>College of Pharmaceutical Engineering of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, PR China. <sup>2</sup>Tianjin Key Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, PR China. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.W. (email: pharmwch@126.com) or Z.L. (email: lizheng@tjutcm.edu.cn)

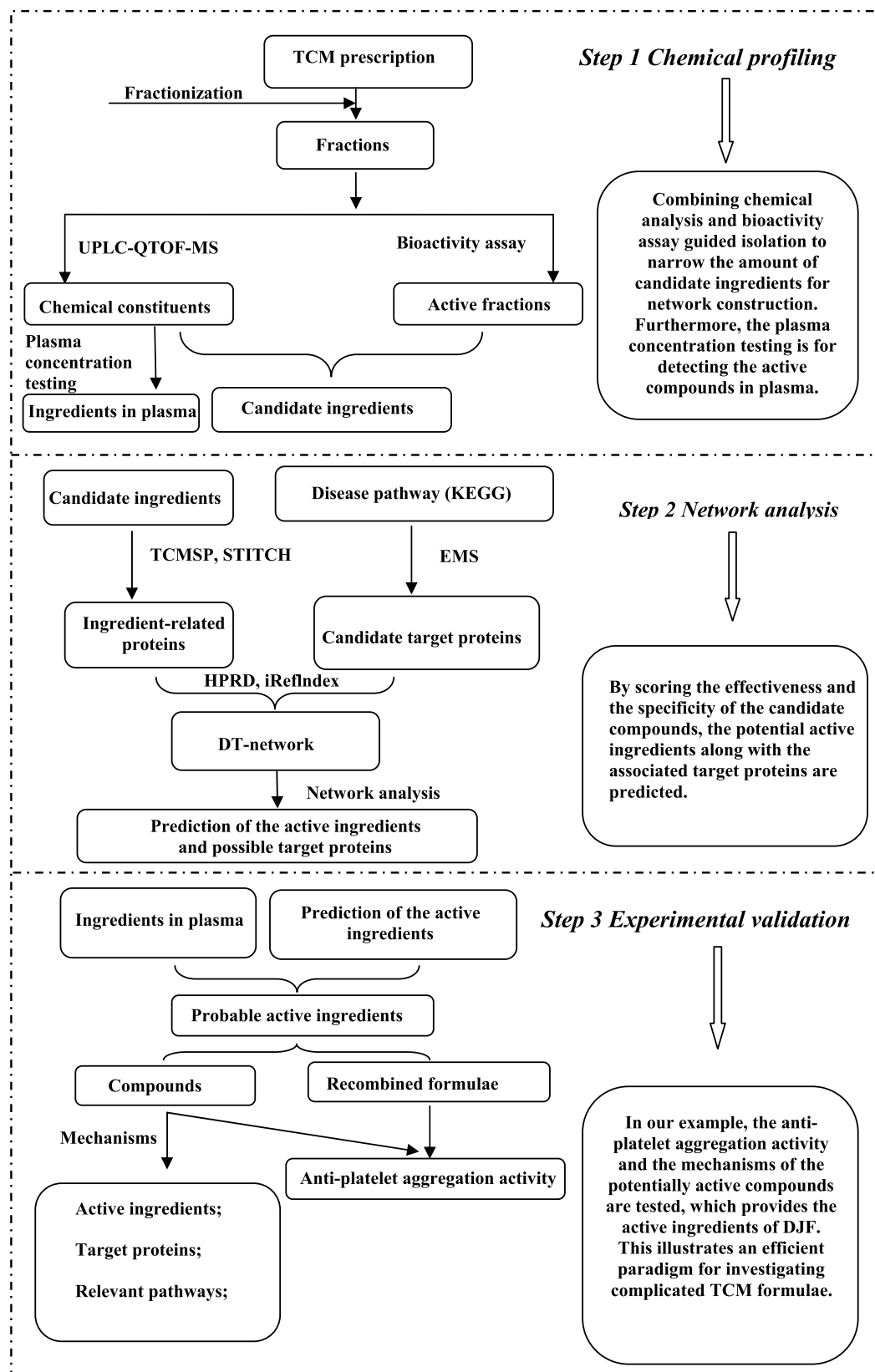


**Figure 1. Illustration of two modes of interaction between ingredients and targets.** C1 and C2 denote two chemical compounds under investigation. T1, T2, T3, T4, T5, T6,  $T_x$  and  $T_y$  denote different proteins, among which  $T_x$  and  $T_y$  are the targets that are directly related to the disease. (a) C1 has direct interaction with four proteins, one of which is the disease-related target  $T_x$ ; (b) C2 interacts with T4, T5 and T6, which further manipulates the disease-related targets  $T_x$  or  $T_y$ .

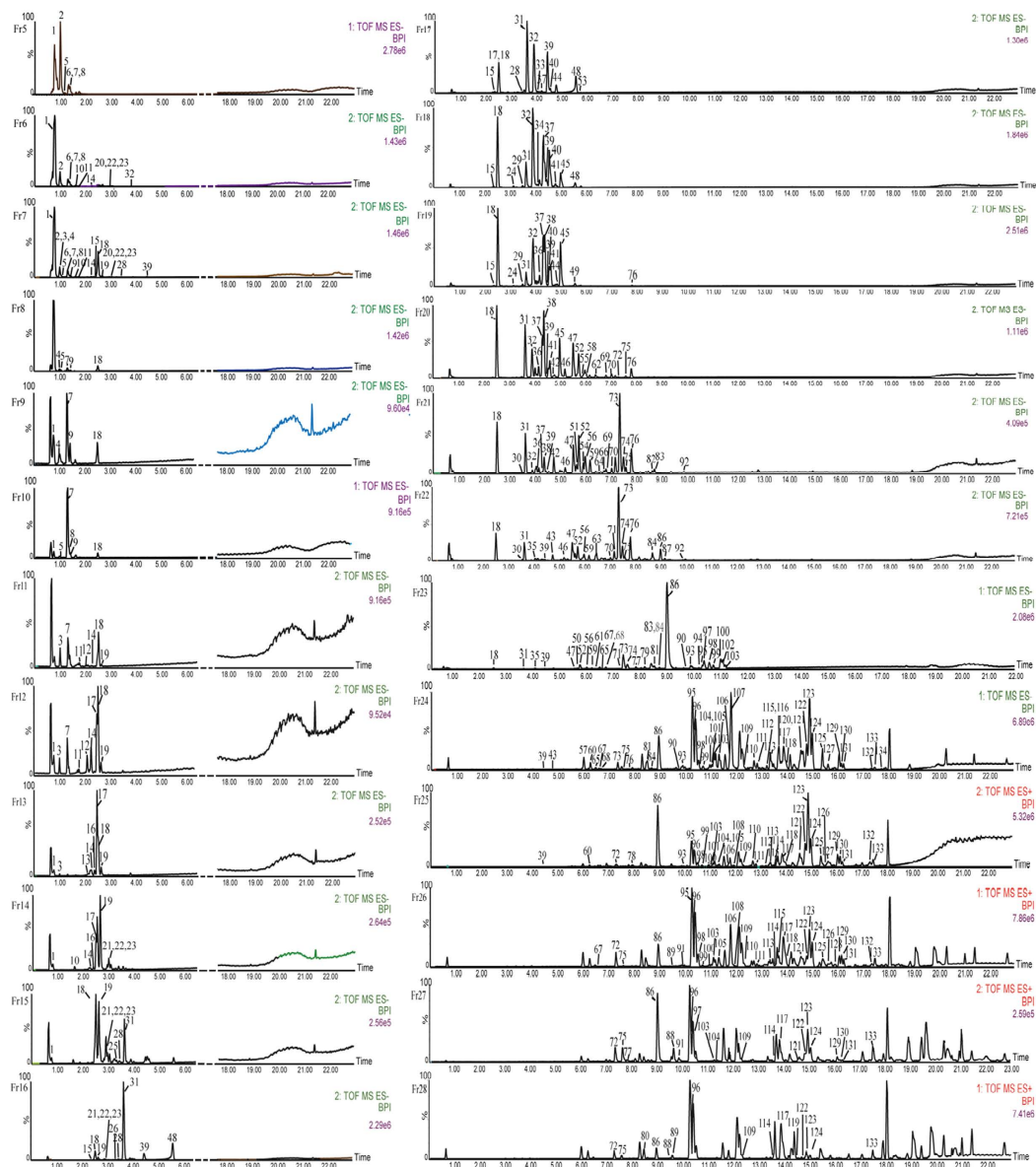
the reliable chemical constitution to construct the network and use absorbed components analysis to validate the bio-active constituents. On the other hand, the candidate targets and other relevant proteins are always collected from databases or by text mining of literature. Although this ensures the relevance of the candidate protein targets, the essentiality of each protein is often poorly assessed, especially when the physiological disorder under investigation is dominant by cascade reactions in which the topological attributes (e.g., degree, closeness, etc.) of each node are not quite related to its importance<sup>8</sup>. Another concern lies in the estimation of the pharmaceutical effectiveness of each chemical ingredient through the analysis of the pharmacological network. In this respect, it is common to require an effective ingredient have direct interaction with the disease-related protein targets, which could lead to neglect of compounds with indirect but significant effectiveness. In order to see it, we illustrate 2 possible interaction modes between ingredients and targets in Fig. 1. Now suppose we are trying to assess the effectiveness of compounds C1 and C2 on the disease-related target  $T_x$ . As can be seen from Fig. 1(a), C1 has interaction with four proteins, only one of which is disease-related. Hence for a given dosage, the effectiveness of C1 on the disease is roughly reduced to 1/4. On the other hand, C2 does not interact with  $T_x$  but has specific interaction with T4, the manipulator of  $T_x$ . Because C2 interacts with only three proteins, its effectiveness on T4 is about 1/3 for a given dosage, which can be completely transferred to  $T_x$ . As a result, it can be expected that C2 is more effective than C1, although it does not affect  $T_x$  directly. Such kind of indirect but essential effectiveness has been receiving attention in the community. For example, in the method proposed by Wang *et al.*, proteins indirectly related to the disease were included in the group of effective targets, although empirical parameters are needed for further quantifying the effectiveness<sup>7</sup>. Indeed, if we regard the chemical ingredients as information sources and the targets as sinks, assessing the effectiveness can be mapped onto the problem of information diffusion through an interaction network<sup>9,10</sup>.

With these considerations, we develop a joint paradigm for the pharmacological studies of TCM prescriptions, integrating chemicals identification and network analysis. The overall procedure is illustrated in Fig. 2. On the one hand, experiments including fractionization, mass spectrometry, pharmacological assay and absorbed components detection are performed to the TCM prescription in order to seek for the probably active chemical ingredients. On the other hand, the candidate disease targets are collected by analyzing existing disease networks (e.g., biological pathways), which are further integrated with the prescription components and the PPI to construct the drug-target (DT) network. With this network in hand, the effectiveness of each candidate ingredient can be assessed by analyzing how much “information” that emits from the ingredient can be received by the key targets, and then the active ingredients can be predicted and finally validated experimentally. The whole procedure dramatically ease the workload to extract the active chemicals from a complicated TCM prescription (always containing hundreds of chemical compounds) and hence provide an effective paradigm to study the pharmacology of TCM.

As an implementation of the paradigm, we apply it to the Danggui-Jianzhong formula (DJF), which is primarily composed of *Angelicae Sinensis Radix* (Danggui), *Cinnamomum Cortex* (Guixin), *Licorice* (Gancao), *Paeonia Radix Alba* (Baishao), *Zingiber Officinale Roscoe* (Shengjiang) and *Jujubae Fructus* (Dazao). In practice, DJF works as a mixture of chemical ingredients. This prescription is extensively adopted in China for gynecological disorders related to blood issues, such as primary dysmenorrhea, with its effectiveness in blood quality promoting and pain releasing. Our focus in this article is on the anti-platelet effect of the formula, especially its effectiveness on platelet aggregation. After experimentally identifying the ingredients of DJF and finding their related proteins from online databases, an elementary-signaling-mode (ESM) analysis is used on the pathway of platelet activation from the Kyoto Encyclopedia of Genes and Genomes (KEGG)<sup>11,12</sup> to extract the disease-relevant candidate targets<sup>8</sup>. Furthermore, an algorithm based on the information flow through the shortest simple paths is developed to study the activity of each candidate ingredient. The active compounds from DJF are finally obtained by



**Figure 2.** The workflow to study TCM prescriptions combining chemicals identification and network analysis.



**Figure 3.** The ion chromatograms of the fractions from F5 to F28, while the signals of F2, F3 and F4 are too weak to be detected.

combining the results of network analysis and absorbed components detection. We finally perform experiments to validate the prediction and discuss the mechanisms of the action of the ingredients.

## Results and Discussion

### Identification of the chemical composition of DJF.

To investigate the chemical ingredients of DJF, we firstly fractionate the formula into 27 fractions (F2–F28) via our preparative chromatography system (Pr-HPLC). Since the chemicals within the same fraction have similar polarity, the subsequent analysis can be performed more efficiently and comprehensively. We then analyze the ingredients of each fraction via ultra performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS). The ion chromatograms are displayed in Fig. 3. By comparing the retention time, fragmentation pathways and MS/MS spectra with the data of reference compounds or literature, a total of 134 chemicals are finally identified, which can be divided into 6 groups, i.e., flavonoid, monoterpene, lignan, terpenoid saponins, coumarin and “others”.

While the details of analyzing the MS data are beyond the scope of this article, we would only take the compounds of flavonoid as a typical case to briefly illustrate the identification procedure. Among those flavonoids, compounds of flavones, dihydroflavones, chalcones and flavanones mainly fragment at their C rings in the positive/negative ion mode, generating CO (28 Da) and CO<sub>2</sub> (44 Da) fragments via the retro-Diels-Alder (RDA) reaction. For example, in the positive ion mode, the [M + H]<sup>+</sup> (*m/z* 257.0809) ion from Liquiritigenin (No. 59 in Fig. 3) further fragments via RDA, losing one C<sub>8</sub>H<sub>8</sub>O<sub>8</sub> (4-vinylphenol) and leading to the generation of abundant ions

Ingredient name	Max. of $I(m \rightarrow n)$	$I(m)$ value
Pentagalloylglucose	0.167 (P00734) <sup>b</sup>	0.207
Glabrone	0.048 (P00734 and P23219)	0.121
(Z)-Ligustilide	0.053 (P23219)	0.099
Glyasperin C	0.038 (P00734)	0.082
Glyasperin F	0.050 (P23219)	0.075
Formononetin	0.023 (P00734 and P23219)	0.069
cAMP <sup>a</sup>	0.004 (P60709)	0.067
Licochalcone A	0.030 (P23219)	0.058
Z,Z'-6,8',7,3'-Diligustilide	0.011 (P17612)	0.052
Sucrose	0.021 (P17612)	0.048
Isoangustone A	0.015 (P17612)	0.047
Levistolide A	0.018 (P17612)	0.036
Gancaonin H	0.005 (P16885)	0.036
(E)-Ligustilide	0.024 (P23219)	0.035
Senkyunolide O	0.013 (P17612)	0.027
Catechin-5-O-glucoside	0.016 (P23219)	0.023
Glycyrrhetic acid	0.003 (P17612)	0.019
Paeoniflorin	—	0
18 $\alpha$ -glycyrrhizin	—	0

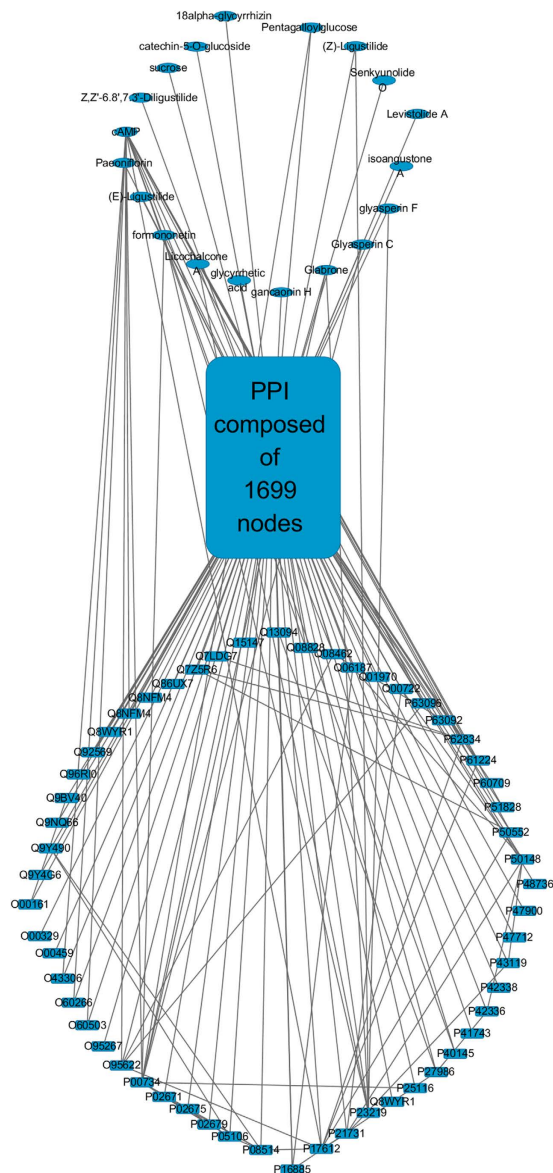
**Table 1. The 19 candidate ingredients and their scores of effectiveness.** <sup>a</sup>Cyclic adenosine monophosphate. <sup>b</sup>The UniProt ID(s) of the target(s) corresponding to the maximum of  $I(m \rightarrow n)$ .

at  $m/z$  137.0231. It can also lose one  $C_6H_6O_2$  (resorcinol) with the generation of ions at  $m/z$  147.0441, or lose one CO with the generation of ions at  $m/z$  119.0499. Meanwhile, in the negative ion mode, the  $[M - H]^-$  ( $m/z$  255.0650) ion from Liquiritigenin generates ions at  $m/z$  119.0493 and 135.0079 via RDA. On the other hand, flavone glycosides, which can be further divided into flavone O-glycosides and flavone C-glycosides, perform relatively simple fragmentation. Specifically, flavone O-glycosides usually lose glucose (162 Da), apiose (132 Da) and rhamnose (146 Da) simultaneously or consecutively, while flavone C-glycosides dehydrates at the hydroxyls of the saccharide groups. The 134 compounds are all obtained through such analysis of the MS data and summarized in Supplementary Table S1.

**Prediction of the active ingredients.** Accompanying the chemical analysis of the 27 fractions, we also explore their anti-platelet activities through the platelet aggregation experiments and find 10 effective fractions (i.e., F11, F12, F14–F17, F24–F26 and F28; see Supplementary Fig. S1). For efficiency, only the 86 compounds from these 10 effective fractions are considered when constructing the DT network, 19 of which have associated proteins reported and hence constitute the set of candidate ingredients (Table 1), along with their 450 associated proteins from the relevant databases (see the section “Methods” for more details). On the other hand, the ESM analysis is performed to the pathway of platelet activation (KEGG entry ID: hsa04611), which results in 56 potential target proteins, as listed in Supplementary Table S2. Further combining these results with the PPI extracted from the Human Protein Reference Database (HPRD)<sup>13</sup> and iRefIndex<sup>14</sup>, we obtain the final DT network (Fig. 4) for DJF. This network is composed of 1774 nodes and 13475 edges, in which the 19 candidate ingredients affect the 56 potential targets either directly or through the intermediate PPI network.

We subsequently calculate the effectiveness scores (i.e.,  $I(m)$  values) and the specificity scores (i.e., the maximal value of  $I(m \rightarrow n)$ ) for these 19 ingredients via a simple algorithm, which maps the problem onto the information flow through an interaction network, as depicted in the section “Methods”. The results are summarized in Table 1. The list is ranked in the order of  $I(m)$  values, and the ingredients have relatively high  $I(m)$  value are expected to be more effective to regulate the disorder under investigation (i.e., platelet aggregation in our case). In the mean time, the specificity scores are also important when assessing the effectiveness of a given compound, since lower specificity may lead to more side effects or weaker pharmaceutical action (since less dosage can reach the key target). In practice, the thresholds for the scores of effectiveness and specificity may be case-dependent. We regard an ingredient as an effective one if its  $I(m)$  is greater than 0.05 and its total specificity score is larger than 0.03. Please note when there are more than one maxima in the  $I(m \rightarrow n)$  of the ingredient, e.g., formononetin, the maximal values are added up as the specificity score. As a result, 7 compounds from DJF are predicted as potentially effective for platelet aggregation, as denoted in italics in Table 1.

Among these 7 ingredients, anti-platelet effects have been reported for Pentagalloylglucose<sup>15</sup>, (Z)-Ligustilide<sup>16</sup> and Licochalcone A<sup>17</sup>. On the other hand, Formononetin is recognized to accelerate wound repair<sup>18</sup>, and interact with thrombin (UniProt ID: P00734)<sup>19</sup>, an essential target for regulating platelet aggregation. There are also reports for the pharmaceutical effects of Glabrone and Glyasperin C. Specifically, Glabrone is known to have anti-influenza activity (probably by inhibiting neuraminidase)<sup>20</sup> and PPAR- $\gamma$  ligand-binding ability<sup>21</sup>. Glyasperin C is reported as a tyrosinase inhibitor<sup>22</sup> and an estrogen antagonist<sup>23</sup>. However, there is little information for the anti-platelet effects of these 2 chemicals. We find no literature about Glyasperin F, although TCMSF predicts its

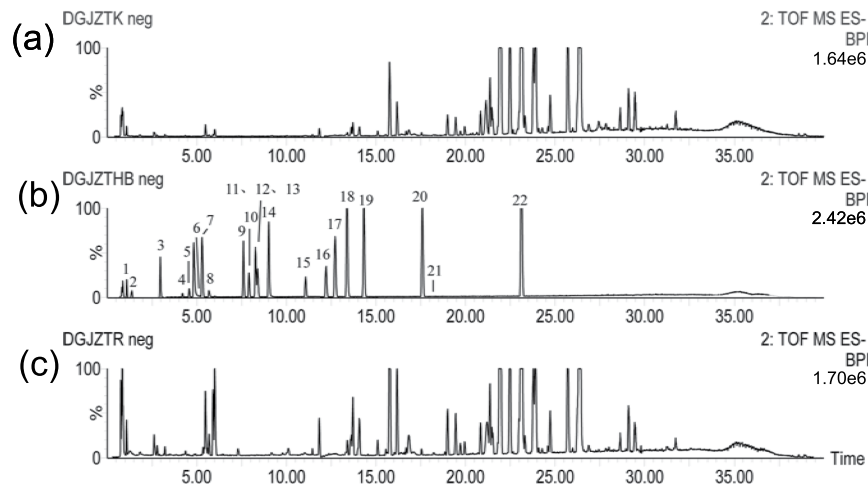


**Figure 4.** The final DT network composed of the nodes of the 19 candidate ingredients (ellipses), the 56 potential targets and the intermediate PPI.

interaction with cyclooxygenase-1 (COX-1, UniProt ID: P23219). With these pieces of information in hand, we regard 4 compounds, i.e., Pentagalloylglucose, (Z)-Ligustilide, Licochalcone A and Formononetin, as the most probable candidates contributing to the anti-platelet effect of DJF.

Accompany with the network analysis, we also investigate the plasma concentration of the ingredients in DJF with a rat model, via the UPLC-QTOF-MS experiments. Since the complicated composition of the plasma brings inherent noises to the MS data, reference chemicals are used to help the identification process. Specifically, a total of 26 reference compounds are available commercially, which are used to make the reference solution for the UPLC-QTOF-MS testing. The MS data of the blank rat plasma, the reference solution and the rat plasma after oral administration of the decoction of DJF are displayed in Fig. 5. After comparing the MS data (retention time, fragments, etc.) of the rat plasma sample with those of the reference solution, 22 chemicals are identified, as listed in Table 2. It can be seen that the 4 candidate compounds are present.

**Experimental validation.** In order to validate the prediction from last subsection and the whole joint paradigm to study DJF, we perform platelet aggregation experiments to investigate the anti-platelet activities of the ingredients found in the plasma concentration testing. Two types of agonist are utilized, i.e., adenosine diphosphate (ADP) and thrombin, and the corresponding activation mechanism is briefly illustrated in Fig. 6(a). It is known that ADP induces platelet activation via P2Y<sub>1</sub> and P2Y<sub>12</sub>, which further causes a series of events including thromboxane A<sub>2</sub> (TXA<sub>2</sub>) synthesis that further promotes the aggregation of platelets. Thrombin, on the other hand, activates platelet through PAR1 and PAR4, which directly couples to G13- and Gq-mediated signaling



**Figure 5.** The ion chromatograms of (a) the blank rat plasma, (b) the reference solution and (c) the plasma sample from the rat after oral administration of DJF decoction. Negative ion mode was used during the testing.

(for shape change and aggregation, respectively)<sup>24–26</sup>. In the platelet aggregation experiments, all of the 22 compounds that are verified by the plasma concentration testing have been tested, including the 4 predicted active ingredients (Table 2). It is worth pointing out that 10 of these 22 compounds are from the less active fractions and hence excluded before constructing the DT network. The inclusion of them in the experiments helps to check the reliability of our pretreatment for screening the fractions *a priori*. The final results are shown in the parts b and c of Fig. 6, which displays the maximum aggregation rate (MAR) of different chemicals. Figure 6(b) presents the results with ADP as the platelet agonist and Fig. 6(c) gives the results for thrombin cases. In both figures, the reagents (i.e., Brilinta and Hirudin) display expected inhibitory effects on platelet aggregation, which verifies the reliability of our experiments. It can be seen from the figure that 3 ingredients dose-dependently suppress platelet aggregation, while other inactive chemicals are not shown for brevity. Specifically, (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose have inhibitory activity on ADP-activated platelet aggregation, and Pentagalloylglucose can also suppress the aggregation induced by thrombin. Namely, the chemicals that are not predicted by our network analysis display no activity in the experiments either, while 3 of the 4 predicted and under-testing compounds are verified to be active. Formononetin is the one that is predicted in Table 1 but does not have inhibitory performance experimentally. Combining with the fact that its interaction with thrombin has been validated elsewhere<sup>19</sup>, our experimental result indicates that interaction (binding, etc.) with a key protein does not always lead to pharmaceutical effects.

With the 3 active ingredients in hand, we would further ask whether the combination of these 3 compounds could have similar anti-platelet activity to DJF. In order to answer this question, we firstly use UPLC with diode array detector (UPLC-DAD), combined with the external standard method, to determine the content of these chemicals in DJF, obtaining a ratio of content as 40.4:1.0:117.3 (Pentagalloylglucose:Licochalcone A:(Z)-Ligustilide, Supplementary Fig. S3). Then, a 3-compound formula with this content ratio is constructed as RF<sub>2</sub>, and another formula with content ratio 1:1:1 (RF<sub>1</sub>) is also constructed for comparison. We further perform the platelet aggregation experiments (ADP-activated) with these two new recombinant formulae and DJF, the results of which are shown in Fig. 6(d). From Supplementary Table S9, we know that the mass concentrations of RF<sub>1</sub> - C<sub>1</sub> and RF<sub>2</sub> - C<sub>1</sub> are 140 µg/mL and 201.25 µg/mL, respectively, which is lower than DJF 250 (250 µg/mL). However, it can be seen from Fig. 6(d) that, both RF<sub>1</sub> - C<sub>1</sub> and RF<sub>2</sub> - C<sub>1</sub> display better anti-platelet activity than DJF 250. Moreover, RF<sub>1</sub> - C<sub>2</sub> and RF<sub>2</sub> - C<sub>2</sub> also have better activity than DJF 250, although their mass concentrations are as low as 70 µg/mL and 100.63 µg/mL, respectively. These results show that the 3-compound formula can have obvious better performance than the original DJF.

In addition to predicting the active ingredients, our network analysis can also hint the key target proteins associated with each compound, as listed in the parentheses in Table 1. It can be seen that, the key target for (Z)-Ligustilide and Licochalcone A is predicted to be cyclooxygenase-1 (COX-1), which modulates the generation of prostaglandin from arachidonic acid (AA) and finally affects the generation of TXA<sub>2</sub>. For Licochalcone A, such prediction is consistent with the observation of its inhibitory effect on the formation of thromboxane B<sub>2</sub>, the metabolite of TXA<sub>2</sub><sup>17</sup>. Nevertheless, we do not find any reference that directly indicates the association between (Z)-Ligustilide and COX-1. There are relevant studies about the modulation of (Z)-Ligustilide on cyclooxygenase-2 (COX-2)<sup>27–29</sup>, a protein that has similar function as COX-1. However, in our experiments of normal platelets, COX-2 is expected to be little expressed. On the other hand, Pentagalloylglucose is predicted by our scheme to mainly affect thrombin, which is consistent with the fact that Pentagalloylglucose is the only active ingredient in the thrombin-activated platelet aggregation experiments. Indeed, the mixed noncompetitive inhibitory effect of Pentagalloylglucose on thrombin has been verified experimentally<sup>30</sup>. In the mean time, the performance of this compound in the ADP-activated platelet aggregation experiments is somewhat unexpected according to the network analysis, although we notice that its inhibitory effect on COX-2 has been uncovered experimentally<sup>31</sup>.

	$T_R$ (min)	Formula	$ES^- (m/z)$	$MS^2$	$ES^+ (m/z)$	$MS^2$	Identification
1	1.091	$C_{10}H_{12}N_5O_6P$	328.0446 (-0.3)	134.0471[ADE] <sup>-</sup>			cAMP
2	1.375	$C_7H_6O_5$	169.0135 (5.3)	125.0262[M - H - CO <sub>2</sub> ] <sup>-</sup>			Gallic acid
3	2.969	$C_{23}H_{38}O_{12}$	495.1472 (-6.3)	137.0259[4 - OH - BZ - H] <sup>-</sup>			Oxypaeoniflorin
4	4.213	$C_{23}H_{38}O_{11}$	525.1625[M + HCOO <sup>-</sup> ] <sup>-</sup>	121.0299[BZ - H] <sup>-</sup>	481.1685 (-5.2)	319.1195[M + H - Glc] <sup>+</sup>	Albiflorin
5	4.583	$C_{23}H_{38}O_{11}$	525.1614[M + HCOO <sup>-</sup> ] <sup>-</sup>	449.1448[M - H - HCHO] <sup>-</sup>	503.1523[M + Na]	301.1110[M + H - Glc - H <sub>2</sub> O + H] <sup>+</sup>	Paeoniflorin
6	5.185	$C_{21}H_{22}O_9$	417.1177 (-2.1)	255.0667[M - H - Glc] <sup>-</sup>	419.1327 (-3.6)	257.0776[M + H - Glc] <sup>+</sup>	Liquiritin
7	5.327	$C_{26}H_{30}O_{13}$	549.1616 (1.4)	417.1204[M - H - Api] <sup>-</sup>	551.1766 (-5.3)	257.0788[M + H - Api - Glc] <sup>+</sup>	Liquiritin apioside
8	5.693	$C_{41}H_{52}O_{26}$	939.1078 (-2.8)	169.016[GA - H] <sup>-</sup>			Pentagalloylglucose
9	7.617	$C_{26}H_{30}O_{13}$	549.1638 (5.5)	417.1241[M - H - Api] <sup>-</sup>	551.1758 (-1.3)	257.078[M + H - Api - Glc] <sup>+</sup>	Isoliquiritin apioside
10	7.923	$C_{21}H_{22}O_9$	417.1193 (1.7)	255.0663[M - H - Glc] <sup>-</sup>	419.1311 (-7.4)	257.0795[M + H - Glc] <sup>+</sup>	Isoliquiritin
11	8.271	$C_{21}H_{22}O_9$			419.1318 (-5.7)	257.0800[M + H - Glc] <sup>+</sup>	Neoisoliquiritin
12	8.286	$C_{22}H_{22}O_9$	475.1277[M + HCOO <sup>-</sup> ] <sup>-</sup>	267.0660[M - H - Glc] <sup>-</sup>	431.1333 (-2.1)	269.0769[M + H - Glc] <sup>+</sup>	Ononin
13	8.41	$C_{15}H_{12}O_4$	255.0664 (2.7)	135.0101[M - H - VP] <sup>-</sup>		137.0207[M + H - VP] <sup>+</sup>	Liquiritigenin
14	9.022	$C_{30}H_{32}O_{13}$	599.1772 (1.2)	137.0209[4 - OH - BZ - H] <sup>-</sup>			Benzoyloxypaeoniflorin
15	11.07	$C_{30}H_{32}O_{12}$	629.1867[M + HCOO <sup>-</sup> ] <sup>-</sup>	121.0301[BZ - H] <sup>-</sup>			Benzoylpaeniflorin
16	12.229	$C_{15}H_{12}O_4$	255.0674 (6.7)	135.0197[M - H - VP] <sup>-</sup>	257.0824 (3.9)		Isoliquiritigenin
17	12.734	$C_{16}H_{12}O_4$	267.0657 (0.0)	252.0461[M - H - CH <sub>3</sub> ] <sup>-</sup>	269.0813 (6.3)	237.0540[M + H - CH <sub>3</sub> OH] <sup>+</sup>	Isoliquiritigenin
18	13.382	$C_{42}H_{62}O_{17}$	837.3922 (1.5)	351.0577[2GlcA - H] <sup>-</sup>	839.4033 (-3.8)	663.3651[M + H - GlcA] <sup>+</sup>	Licorice saponin G2
						487.3369[M + H - 2GlcA] <sup>+</sup>	
						469.3370[M + H - 2GlcA - H <sub>2</sub> O] <sup>+</sup>	
19	14.321	$C_{42}H_{62}O_{16}$	821.3959 (-0.1)	351.0578[2GlcA - H] <sup>-</sup>	823.4133 (-4.0)	453.3361[M + H - 2GlcA - H <sub>2</sub> O] <sup>+</sup>	Glycyrrhizin
20	17.614	$C_{21}H_{22}O_4$	337.1458 (5.3)	187.0720[M - H - 2CH <sub>3</sub> - OCD] <sup>-</sup>			Licochalcone A
21	18.034	$C_{12}H_{14}O_2$			191.1065 (-3.7)	173.0897[M + H - H <sub>2</sub> O] <sup>+</sup>	(Z)-ligustilide
22	23.141	$C_{30}H_{46}O_4$	469.3343 (5.3)	355.2655[M - H - CO <sub>2</sub> - ME] <sup>-</sup>	471.3472 (-0.4)	407.3469[M + H - HCOOH - H <sub>2</sub> O] <sup>+</sup>	Glycyrrhetic acid

**Table 2. Chemical information of the 22 compounds from the plasma concentration testing.** Glc: Glucose, ADE: Adenine, GA: Gallic acid, BZ: Benzoic Acid, 4-OH-BZ: 4-hydroxyl-benzoic acid, VP: 4-vinylphenol, Api: apiose, GlcA: glucuronic acid, OCD: 4-oxomethylidencyclohexa-2,5-dienone, ME: 3-methylbut-1-ene.

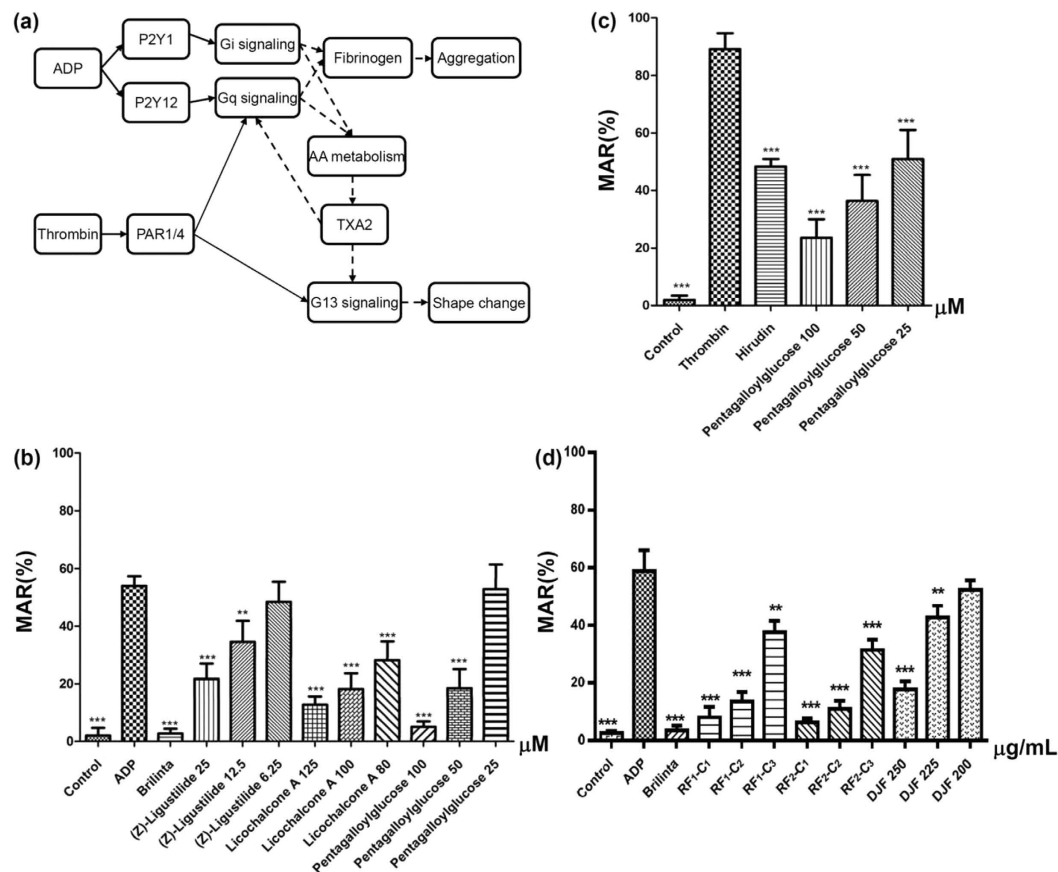
In order to validate the predictions of the key target proteins for the 3 active ingredients, we perform further experiments including real-time quantitative PCR (qPCR) and COX activity assay kit. The former is to test the effects of the 3 compounds on the mRNA expression level of COX-1 and the latter is to analyze whether the compounds affect the activity of COX-1. The results are summarized in Fig. 7. Specifically, Fig. 7(a) shows the qPCR results, which indicates that (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose hardly affect the mRNA expression level of COX-1. However, it can be seen from Fig. 7(b) that these 3 compounds display obvious inhibitory effects on the activity of COX-1, comparable to that of aspirin. These results not only verify the prediction of our network analysis about the target proteins, but also indicate that the 3 active ingredients affect COX-1 by suppressing the activity rather than its mRNA expression level.

## Conclusion

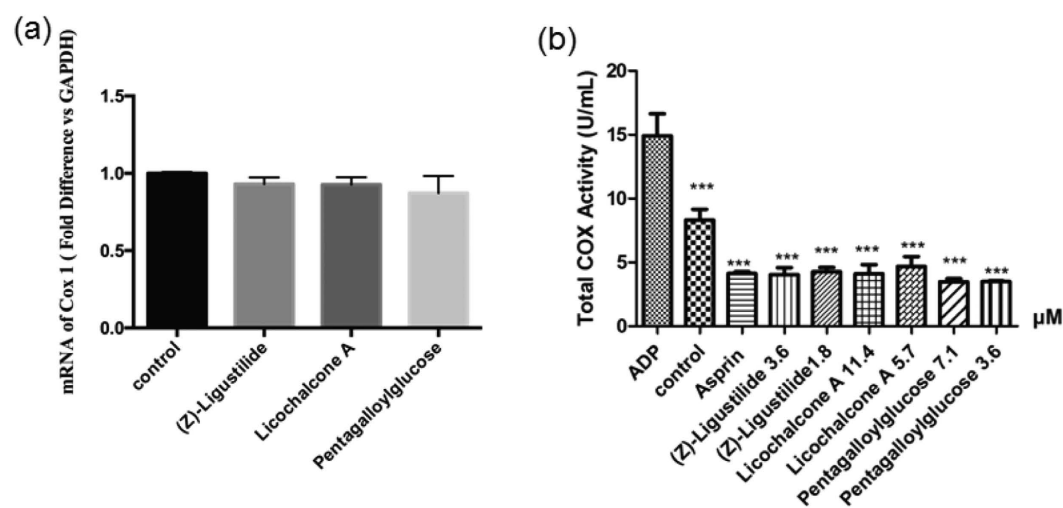
In this article, we combine chemical profiling and network analysis to comprehensively investigate the pharmacology of combination drugs and complex formulae. On the one hand, the chemical ingredients are totally identified experimentally to confirm their presence in the formula. On the other hand, the potential targets are determined through the analysis of the disease pathway such that their relevance is certificated. Moreover, the effectiveness of each candidate chemical ingredient is assessed via the analysis of the DT network, and the specificity of the ingredient to any target is simultaneously scored, which is further combined with the plasma concentration testing in order to efficiently extract the active chemicals from the complicated formula. In addition, the scheme can also provide information on the key targets associated with each active ingredient, which helps to decipher the mechanism of the pharmaceutical action of the compound.

As an application, we use the scheme to study a traditional Chinese medicine DJF. We fractionate the formula into 27 fractions, which makes the chemical identification procedure more efficient and allows us to perform elementary study on the pharmaceutical activities of the fractions. Combining the experimental testing and the network analysis results, we predict 4 chemicals as the most probable active ingredients. Through the subsequent platelet aggregation experiments, 3 of them are verified to have anti-platelet activity, i.e., (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose. Bearing in mind that all other compounds tested in the platelet aggregation experiments display no activity, these results show that the prediction of our scheme is indeed robust. In addition, according to the network analysis, (Z)-Ligustilide and Licochalcone A modulate ADP-activated platelet aggregation primarily through COX-1 and Pentagalloylglucose reduces the thrombin-activated platelet mainly via thrombin, which are consistent with the known experiments. We further experimentally test such predictions by investigating the effects of these 3 compounds on the mRNA expression and the activity of COX-1. The results indicate that they have little modulation on the mRNA expression level of COX-1 but present strong inhibitory effects on the activity of the protein.





**Figure 6.** (a) The illustration of the platelet activation process induced by ADP and/or thrombin. The solid arrow denotes direct association, while the dashed arrow stands for indirect connection mediated by other proteins and/or compounds; (b) The MAR of the chemicals that have anti-platelet activity in the ADP-activated experiments; (c) The MAR of the chemicals that have anti-platelet activity in the thrombin-activated experiments; (d) The MAR of the 3-compound formulae in the ADP-activated experiments with comparison to DJF. Significant difference with respect to the corresponding agonist group: \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .



**Figure 7.** The experimental results of (a) the real-time quantitative PCR for detecting the mRNA level of COX-1 in the platelet treated by the active ingredients, and (b) the total COX activity in the platelet treated by various chemicals, with the chemical concentrations indicating in the sample names. Significant difference with respect to the corresponding agonist group: \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

## Methods

**Network construction.** The chemical ingredients of DJF are identified experimentally. With their names and/or chemical structures (denoted by the simplified molecular-input line-entry system (SMILES) strings) as key words, ingredient-related proteins are collected from the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP)<sup>32</sup> and the Search Tool for Interactions of Chemicals (STITCH)<sup>33</sup>.

On the other hand, disease-related proteins are collected from the pathway of platelet activation. Specifically, we choose the nodes “ADP” and “Aggregation” as starting and ending points, respectively, and perform an ESM analysis to the pathway according to the algorithm in ref. 8. The essentiality of node  $\nu$  is assessed by the reduction in the number of ESMs following the removal of  $\nu$ , i.e.,

$$E^{\text{ESM}}(\nu) = \frac{N_{\text{ESM}}(G) - N_{\text{ESM}}(G_{\Delta\nu})}{N_{\text{ESM}}(G)} \quad (1)$$

where  $N_{\text{ESM}}(G)$  and  $N_{\text{ESM}}(G_{\Delta\nu})$  are the total number of ESMs from the starting point to the ending point in the original network  $G$  and the one after deleting  $\nu$ , respectively. Obviously, node  $\nu$  is essential if  $E^{\text{ESM}}(\nu) = 1$ , because its absence disrupts all ESMs. After the analysis, we identified 56 proteins as potential targets, as listed in Supplementary Table S2.

In addition, relevant PPI network is extracted from HPRD and iRefIndex. This along with the active ingredients, ingredient-related proteins and disease-related proteins are input into Cytoscape<sup>34</sup> to build the complete DT network for DJF.

**Scoring the effectiveness of each ingredient.** In order to quantitatively assess the effectiveness of each active ingredient, we now consider the question that, after emitting one unit information by one of the 19 nodes of the candidate compounds, how much effect each disease-related target can receive. We assume that the information propagates from one node to the other mainly through the simple paths connecting them in the DT network, and each node can merely and equivalently affect its neighbors, which is indeed a simple case of information flow through interaction networks<sup>9,10</sup>. Hence, if one unit of information comes to a node of degree  $k$ , it flows downstream through  $k - 1$  branches, each of which transfers  $1/k - 1$  unit of the original information. Then, the information that target  $n$  receives from ingredient  $m$  through a simple path  $i$  connecting them can be evaluated as,

$$I_i(m \rightarrow n) = \frac{1}{k_m} \prod_{j \in V(i)} \frac{1}{k_j - 1} \quad (2)$$

where  $V(i) = \{\text{protein nodes between } n \text{ and } m \text{ in path } i\}$  and it should be pointed out that we are considering the paths that have no other ingredient nodes than the starting point. We further take the approximation that only the  $I_i(m \rightarrow n)$  from the shortest paths are significant and thus the effectiveness of ingredient  $m$  on target  $n$  can be estimated by,

$$I(m \rightarrow n) = \sum_i I_i(m \rightarrow n) \quad (3)$$

where the summation runs over all the shortest paths between  $m$  and  $n$ . The total effectiveness of ingredient  $m$  is given by,

$$I(m) = \sum_n I(m \rightarrow n) \quad (4)$$

These constitute the scoring scheme for the active chemical ingredients. Specifically,  $I(m)$  shows the overall effectiveness of ingredient  $m$  on the disorder under investigation, and  $I(m \rightarrow n)$  gives the specificity of ingredient  $m$  to target  $n$ .

We can take Fig. 1(b) as an instance to illustrate the above procedure. In this example, we would like to calculate the effectiveness score of C2, i.e.,  $I(C2)$ . Because we have 2 disease-related targets  $T_x$  and  $T_y$ , we need  $I(C2 \rightarrow T_x)$  and  $I(C2 \rightarrow T_y)$  to obtain  $I(C2)$  via Eq. 4. It can be seen that node C2 has the degree of 3 and it indirectly affects  $T_x$  and  $T_y$  through 3 mediate nodes T4, T5 and T6, the degrees of which are all 2. Since there is one shortest path connecting C2 and  $T_x$ , i.e., C2 – T4 –  $T_x$ , we have

$$I(C2 \rightarrow T_x) = \frac{1}{3} \times \frac{1}{2 - 1} = \frac{1}{3} \quad (5)$$

On the other hand, there are 2 shortest paths connecting C2 and  $T_y$ , i.e., C2 – T5 –  $T_y$  and C2 – T6 –  $T_y$ . According to Eqs 2 and 3, we have

$$I(C2 \rightarrow T_y) = \frac{1}{3} \times \frac{1}{2 - 1} + \frac{1}{3} \times \frac{1}{2 - 1} = \frac{2}{3} \quad (6)$$

Hence,  $I(C2)$  is given by

$$I(C2) = I(C2 \rightarrow T_x) + I(C2 \rightarrow T_y) = 1 \quad (7)$$

In practice, the DT network is much more complicated than Fig. 1(b), but the procedure of calculating the effectiveness score is quite similar.

**Experiments. Materials and chemicals.** *Angelicae Sinensis Radix*, *Cinnamomum Cortex*, *Licorice*, *Paeoniae Radix Alba* and *Jujubae Fructus* were purchased from Jinshaotang Herbal Medicine Co. Ltd. (Bozhou, China). *Zingiber Officinale Roscoe* was purchased from Tesco China (Tianjin, China). Pentagalloylglucose, Formononetin, Glycyrrhizin, Benzoylpaeoniflorin, Albiflorin, Isoliquiritigenin and Ononin were purchased from Yifang Technology Co. Ltd. (Tianjin, China). (Z)-Ligustilide was purchased from Jianglai-bio Co. Ltd. (Shanghai, China). Licochalcone A and Glycyrrhetic acid were purchased from Zhongxin Innova (Tianjin, China). Licorice saponin G2 was purchased from Tokiwa Phytochemical Co. Ltd. (Sakura-shi, Japan). Cyclic adenosine monophosphate (cAMP) was purchased from Harvey-bio Co. Ltd. (Beijing, China). Oxypaeoniflorin, Levistolide A, Liquiritin apioside, Glycyrrhizic acid, Ellagic acid and Neoisoliquiritin were purchased from Yuanye-bio Co. Ltd. (Shanghai, China). Sucrose was purchased from Sigma-Aldrich Co. LLC. (St. Louis, United States). Paeoniflorin, Liquiritin, Gallic acid and Citric acid were purchased from National Institutes for Food and Drug Control (Beijing, China). Isoliquiritin was purchased from Tauto Biotech Co. Ltd. (Shanghai, China). Isoliquiritin apioside was purchased from Sichuan Victory Biological Technology Co. Ltd. (Chengdu, China). Benzoyloxypaeoniflorin was purchased from Spring & Autumn Biological Engineering Co. Ltd. (Nanjing, China). Brilinta and Aspirin were purchased from Sigma-Aldrich Co. LLC. (St. Louis, MO, USA). Hirudin was purchased from GL Biochem Ltd. (Shanghai, China). TRIZOL reagent was purchased from Thermo Fisher Scientific Inc. (Carlsbad, CA, USA).

Acetonitrile and methanol, both of HPLC grade, were purchased from Sigma-Aldrich Co. LLC. (St. Louis, MO, USA). Ethanol was purchased from Real&Lead Chemical Co. Ltd. (Tianjin, China).

Sprague-Dawley rats (200–200 g) were purchased from the Academy of Military Science of People's Liberation Army (Beijing, China) with production license No: SCXK-(Jun)-2012-0004. All the rats are maintained in a room equipped with an air-filtering system, and the cages and water were sterilized. The study is approved by the Research Ethics Committee of Tianjin International Joint Academy of Biomedicine. All animals are treated in accordance with the guidelines and regulations for the use and care of animals of the Center for Laboratory Animal Care, Tianjin International Joint Academy of Biomedicine.

The Quantscript RT Kit and SYBR Premix HotmasterTaq were purchased from TIANGEN Biotech (Beijing) Co., Ltd. (Beijing, China). The COX Activity Assay Kit was purchased from Cayman Chemical (Ann Arbor, MI, USA).

**Identification of the chemical ingredients of DJF.** We use UPLC-QTOF-MS system to identify the chemical ingredients of DJF and the analysis is performed through 3 main steps as follows.

Firstly, *Angelicae Sinensis Radix*, *Cinnamomum Cortex*, *Licorice*, *Paeoniae Radix Alba*, *Zingiber Officinale Roscoe* and *Jujubae Fructus* are blended with mass ratio 4:3:2:6:3:2 and then immersed in 50% (V/V) ethanol solution with mass ratio 1:10 for 30 min. The mixture is heated under reflux for 2 hours followed by filtration. The residue is immersed in 50% (V/V) ethanol solution with mass ratio 1:8 and heated under reflux for 2 hours. After another filtration, the two filtrates are combined, distilled and dried. The obtained solid sample is mixed with 80% (V/V) methanol up to a concentration of 53 mg/mL, followed by centrifugation at 14,000 rpm for 10 min. The supernatant is further filtered with pore size of 0.24  $\mu\text{m}$  and the filtrate (F1) is collected for the subsequent analysis.

Secondly, F1 is fractionated via the GX-281 preparative high-performance liquid chromatography (Pr-HPLC, Gilson Inc., Middleton, WI, USA), equipped with a Agilent Prep- C<sub>18</sub> column (21.2  $\times$  250 mm, 10  $\mu\text{m}$ ). The elution rate and injection volume are 10 mL/min and 2.0 mL, respectively. The mobile phase contains ultrapure water (A) and methanol (B), and the detailed elution condition can be found in Supplementary Table S3. A total of 27 fractions are collected with a sampling rate of 10 mL/min, which are further solidified via distillation and drying process. The sample solution of each fraction is prepared with 1 mg fraction sample well dissolved in 2 mL 50% (V/V) methanol. Each solution is then centrifuged twice at 14,000 rpm for 10 min and the supernatant is collected.

Thirdly, the sample solutions are analyzed via our UPLC-QTOF-MS platform to identify their ingredients. We use a Waters ACQUITY UPLC system (Waters Co., Milford, MA, USA) equipped with ACQUITY UPLC BEH Shield RP 18 column (100  $\times$  2.1 mm, 1.7  $\mu\text{m}$ ). The injection volume is 2.0  $\mu\text{L}$ , and the mobile phase contains 0.1% (V/V) formic acid (A) and acetonitrile (B). The detailed elution condition can be found in Supplementary Table S4. On the other hand, the operation parameters for the mass detection are: capillary voltage, -2500 V; cone voltage, 40 V; ion source temperature, 100 °C; desolvation gas (nitrogen) temperature, 400 °C; desolvation gas flow rate, 600 L/h; lower collision energy, 0 eV; upper collision energy, 15–60 eV; analysis mode, elevated-energy mass spectrometry (MS<sup>E</sup>).

**Absorbed component in rat plasma assay.** We immerse 80 g raw materials of DJF (16 g *Angelicae Sinensis Radix*, 12 g *Cinnamomum Cortex*, 8 g *Licorice*, 24 g *Paeoniae Radix Alba*, 12 g *Zingiber Officinale Roscoe* and 8 g *Jujubae Fructus*) in 800 ml water for 30 min, and then boil the mixture for 2 hours. After filtration, the residue is immersed in enough water (mass ratio 1:8) and boiled for 2 hours followed by another filtration. The two filtrates are combined and distilled to a DJF solution with drug ratio 1.06 g/ml.

On the other hand, after weighing the rats, they are fasted for 12 hours with free access to water. Subsequently, we administer the DJF decoction to 6 rats, at a dose of 10 ml/kg body weight, and take blood samples from the orbital venous plexus before and after the administration (0, 5, 10, 20, 30 (min), 1, 2, 4, 6 (h)) for a total of 10 times, 0.5 ml for each time. We then put these blood samples in heparinized tubes, centrifuge them at 4 °C, 8,000 rpm for 10 min, and collect the supernatant plasma.

To analyze the DJF ingredients in the plasma, the blood samples are pretreated by mixing 100  $\mu\text{L}$  of each of the 10 blood samples with 300  $\mu\text{L}$  methanol, followed by centrifugation at 14,000 rpm for 10 min. The supernatant of each sample is dried by nitrogen at room temperature, and dissolved in 50 mL 50% (V/V) methanol. After another centrifugation at 14,000 rpm for 10 min, we collect 40  $\mu\text{L}$  supernatant of each sample and combined them

together as the mixed plasma for the MS analysis (with the same parameters as the ingredients identification experiment). The detailed elution condition can be found in Supplementary Table S5.

**Platelet aggregation experiment.** Twenty-six of the 134 chemical ingredients of DJF are commercially available, as noted in the subsection of *Materials and chemicals*, and the corresponding solutions are prepared by separately diluting these compounds with Tris buffer solution (TBS, Supplementary Table S6) to 0.1 mM. On the other hand, platelets are taken from a male Sprague-Dawley rat (weighing 200~220 g). Whole blood obtained by arteri-puncture at the abdominal aorta of the rat is collected in a tube (about 10 ml) with 10% volume of ACD solution (Supplementary Table S7), from which the platelet-rich plasma (PRP) is isolated by 10 min centrifugation at  $200 \times g$  and room temperature. The PRP fraction is further processed by 10 min centrifugation at  $800 \times g$  and the final isolated platelets are suspended in TBS.

The grouping of the experiment samples are summarized in the Supplementary Table S8. Specifically, platelets of  $100 \mu\text{L}$  (about  $10^7$  platelets) are added in each well of a 96-well plate, followed by adding  $50 \mu\text{L}$  solution of the investigated chemicals (ingredients of DJF or the reagents) or  $50 \mu\text{L}$  Tris buffer solution (for the negative sample), such that each well contains  $150 \mu\text{L}$  liquid. The plate is then incubated in a multi-mode microplate reader (FlexStation 3, Molecular Devices LLC.) for 10 min at  $37^\circ\text{C}$  in order to get a well-blended platelet-chemical mixture. Subsequently,  $50 \mu\text{L}$  agonist solution is added to each well and the optical density (OD) value under 405 nm is measured continuously for 30 min at  $37^\circ\text{C}$ . Both types of agonist solution used in our experiments, i.e., a solution of  $25 \mu\text{M}$  ADP and a solution of 0.1 U thrombin, contain 2.0 mM  $\text{CaCl}_2$ .

The aggregation rate (AR) at time  $t$  is derived from the OD values, i.e.,

$$\text{AR}(t) = \frac{\text{OD}_{t=0} - \text{OD}_t}{\text{OD}_{t=0}} \times 100\% \quad (8)$$

And the maximum aggregation rate (MAR) is obtained by extracting the maximum of  $\text{AR}(t)$  and taking the average over the wells.

In addition, we also investigate the anti-platelet activities of the 27 fractions through the above procedure, with the following distinctions:

- The concentration of each fraction solution is  $225 \mu\text{g}/\text{mL}$ .
- Aspirin is used as the reagent for the fractions.

**Real-time quantitative PCR.** Total RNA is extracted by using TRIZOL reagent from the platelets cultured with the active compounds (i.e., (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose), and then reverse-transcribed using the Quantscript RT Kit. mRNA expression is quantified using SYBR Premix HotmasterTaq, and the gene expression of ribosomal protein large P0 is used as an internal control. Primer concentrations are  $10 \mu\text{mol}/\text{L}$ . The primer sequences used are as follows: GAPDH forward: 5'-GATTTGGCCGTATCGGAC-3', reverse: 5'-GAAGA CGCCAGTAGACTC-3'; Cox-1 forward: 5'-TGCATGTGGCTGTGGATGTCATCAA-3', reverse: 5'-CACTA AGACAGACCCGTCATCTCCA-3'.

**COX Activity Assay.** We firstly activate the platelets by ADP *in vitro*, and then treat them with the active ingredients (i.e., (Z)-Ligustilide, Licochalcone A and Pentagalloylglucose) and the reagent aspirin. The total activity of COX-1 is analyzed with the COX Activity Assay Kit.

## References

1. Leung, E. L.-H., Wong, V. K.-W., Jiang, Z.-H., Li, T. & Liu, L. Integrated network-based medicine: The role of traditional Chinese medicine in developing a new generation of medicine. *Science* **346** (6216 Suppl), S16–S18 (2014).
2. Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682–690 (2008).
3. Wang, L. *et al.* A network study of Chinese medicine Xuesaitong injection to elucidate a complex mode of action with multicomponent, multitarget, and multipathway. *Evid.-based Complement Altern. Med.* **2013**, 652373 (2013).
4. Wang, L. *et al.* Dissecting active ingredients of Chinese medicine by content-weighted ingredient-target network. *Mol. BioSyst.* **10**, 1905–1911 (2014).
5. Yang, Z.-Z., Liu, W., Zhang, F., Li, Z. & Cheng, Y.-Y. Deciphering the therapeutic mechanisms of Xiao-Ke-An in treatment of type 2 diabetes in mice by a fangjiomics approach. *Acta Pharmacol. Sin.* **36**, 699 (2015).
6. Luo, F. *et al.* Multiscale modeling of drug-induced effects of ReDuNing injection on human disease: From drug molecules to clinical symptoms of disease. *Sci. Rep.* **5**, 10064 (2015).
7. Wang, C.-H. *et al.* A network analysis of the Chinese medicine Lianhua-Qingwen formula to identify its main effective components. *Mol. BioSyst.* **12**, 606–613 (2016).
8. Wang, R.-S. & Albert, R. Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst. Biol.* **5**, 44 (2011).
9. Ahlswede, R., Cai, N., Li, S.-Y. R. & Yeung, R. W. Network information flow. *IEEE Trans. Inf. Theory* **46**, 1204–1216 (2000).
10. Stojmirović, A. & Yu, Y.-K. Information flow in interaction networks. *J. Comput. Biol.* **14**, 1115–1143 (2007).
11. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* **28**, 27–30 (2000).
12. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucl. Acids Res.* **44**, D457–D462 (2016).
13. Prasad, T. S. K. *et al.* Human Protein Reference Database–2009 update. *Nucl. Acids Res.* **37**, D767–D772 (2009).
14. Razick, S., Magklaras, G. & Donaldson, I. M. iRefIndex: A consolidated protein interaction database with provenance. *BMC Bioinformatics* **9**, 405 (2008).
15. Jeon, W. K. *et al.* Anti-platelet effects of bioactive compounds isolated from the bark of rhus verniciflua stokes. *J. Ethnopharmacol.* **106**, 62–69 (2006).
16. Zhang, L. *et al.* Z-ligustilide extracted from radix angelica sinensis decreased platelet aggregation induced by ADP *ex vivo* and arterio-venous shunt thrombosis *in vivo* in rats. *Yakugaku Zasshi-J. Pharm. Soc. Jpn.* **129**, 855–859 (2009).

17. Kimura, Y., Okuda, T. & Okuda, H. Effects of flavonoids from licorice roots (*glycyrrhiza inflata* bat.) on arachidonic acid metabolism and aggregation in human platelets. *Phytother. Res.* **7**, 341–347 (1993).
18. Huh, J.-E. *et al.* Formononetin accelerates wound repair by the regulation of early growth response factor-1 transcription factor through the phosphorylation of the ERK and p38 MAPK pathways. *Int Immunopharmacol.* **11**, 46–54 (2011).
19. Zhang, S. *et al.* Interactions between thrombin and natural products of *milletia nitita* var. *hirsutissima* using capillary zone electrophoresis. *J. Chromatogr. B* **877**, 4107–4114 (2009).
20. Grienke, U. *et al.* Computer-guided approach to access the anti-influenza activity of licorice constituents. *J. Nat. Prod.* **77**, 563–570 (2014).
21. Kuroda, M. *et al.* Phenolics from *glycyrrhiza glabra* roots and their ppar- $\gamma$  ligand-binding activity. *Bioorg. Med. Chem.* **18**, 962–970 (2010).
22. Kim, H. J., Seo, S. H., Lee, B.-G. & Lee, Y. S. Identification of tyrosinase inhibitors from *glycyrrhiza uralensis*. *Planta Med.* **71**, 785–787 (2005).
23. Boonmuen, N. *et al.* Licorice root components in dietary supplements are selective estrogen receptor modulators with a spectrum of estrogenic and anti-estrogenic activities. *Steroids* **105**, 42–49 (2016).
24. Varga-Szabo, D., Pleines, I. & Nieswandt, B. Cell adhesion mechanisms in platelets. *Arterioscler. Thromb. Vasc. Biol.* **28**, 403–412 (2008).
25. Rivera, J., Lozano, M. L., Navarro-Núñez, L. & Vicente, V. Platelet receptors and signaling in the dynamics of thrombus formation. *Haematologica* **94**, 700–711 (2009).
26. Li, Z., Delaney, M. K., O'Brien, K. A. & Du, X. Signaling during platelet adhesion and activation. *Arterioscler. Thromb. Vasc. Biol.* **30**, 2341–2349 (2010).
27. Wang, J., Du, J.-R., Wang, Y., Kuang, X. & Wang, C.-Y. Z-ligustilide attenuates lipopolysaccharide-induced proinflammatory response via inhibiting NF- $\kappa$ B pathway in primary rat microglia. *Acta Pharmacol. Sin.* **31**, 791–797 (2010).
28. Ma, Z. & Bai, L. The anti-inflammatory effect of Z-ligustilide in experimental ovariectomized osteopenic rats. *Inflammation* **35**, 1793–1797 (2012).
29. Chung, J. W. *et al.* Anti-inflammatory effects of (Z)-ligustilide through suppression of mitogen-activated protein kinases and nuclear factor- $\kappa$ B activation pathways. *Arch. Pharm. Res.* **35**, 723–732 (2012).
30. Dong, H., Chen, S.-X., Kini, R. M. & Xu, H.-X. Effects of tannins from *geum japonicum* on the catalytic activity of thrombin and factor Xa of blood coagulation cascade. *J. Nat. Prod.* **61**, 1356–1360 (1998).
31. Lee, S.-J., Lee, I.-S. & Mar, W. Inhibition of inducible nitric oxide synthase and cyclooxygenase-2 activity by 1,2,3,4,6-penta-*o*-galloyl- $\beta$ -*d*-glucose in murine macrophage cells. *Arch. Pharm. Res.* **26**, 832–839 (2003).
32. Ru, J. *et al.* Tcmisp: a database of systems pharmacology for drug discovery from herbal medicines. *J. Cheminformatics* **6**, 13 (2014).
33. Kuhn, M. *et al.* STITCH 4: integration of protein - chemical interactions with user data. *Nucl. Acids Res.* **42**, D401–D407 (2014).
34. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

## Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFC) (Grant Nos 81403059 and 81573826) and Tianjin Technology Innovation System and Platform Development Plan (15PTCYSY00030 and 14TXZYJC00440).

## Author Contributions

C.W. and Z.L. conceived the study; T.S. and Z.L. designed the algorithm for network analysis; J.L. and C.W. designed the experiments; T.S. performed the network analysis with the help from Z.L.; J.L. performed the experiments with the help from C.W., Y.W., H.Y., T.W., C.L. and X.C.; T.S. and J.L. analyzed the data and wrote up the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Suo, T. *et al.* Combining Chemical Profiling and Network Analysis to Investigate the Pharmacology of Complex Prescriptions in Traditional Chinese Medicine. *Sci. Rep.* **7**, 40529; doi: 10.1038/srep40529 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017