

SCIENTIFIC REPORTS

OPEN

A Systems Biology-Based Investigation into the Pharmacological Mechanisms of Sheng-ma-bie-jia-tang Acting on Systemic Lupus Erythematosus by Multi-Level Data Integration

Lin Huang^{1,*}, Qi Lv^{2,3,*}, Fenfen Liu^{1,*}, Tieliu Shi^{2,4} & Chengping Wen¹

Sheng-ma-bie-jia-tang (SMBJT) is a Traditional Chinese Medicine (TCM) formula that is widely used for the treatment of Systemic Lupus Erythematosus (SLE) in China. However, molecular mechanism behind this formula remains unknown. Here, we systematically analyzed targets of the ingredients in SMBJT to evaluate its potential molecular mechanism. First, we collected 1,267 targets from our previously published database, the Traditional Chinese Medicine Integrated Database (TCMID). Next, we conducted gene ontology and pathway enrichment analyses for these targets and determined that they were enriched in metabolism (amino acids, fatty acids, etc.) and signaling pathways (chemokines, Toll-like receptors, adipocytokines, etc.). 96 targets, which are known SLE disease proteins, were identified as essential targets and the rest 1,171 targets were defined as common targets of this formula. The essential targets directly interacted with SLE disease proteins. Besides, some common targets also had essential connections to both key targets and SLE disease proteins in enriched signaling pathway, e.g. toll-like receptor signaling pathway. We also found distinct function of essential and common targets in immune system processes. This multi-level approach to deciphering the underlying mechanism of SMBJT treatment of SLE details a new perspective that will further our understanding of TCM formulas.

Traditional Chinese Medicine (TCM) is an ancient practice based on extensive knowledge and experiences accumulated over several thousand years and is both efficient and safe for the treatment of chronic diseases¹. More than 100,000 TCM formulas take effect by concocting different natural products which are essential for therapy². A formula is composed of herbs; thus, some studies have focused on the herbal extracts³ and their biological functions⁴. TCM formulas typically utilize multi-component therapeutics. These combinatorial treatments tend to increase therapeutic effects through the synergism of two or more herbs and decrease side effects through antagonism. The application of TCM in clinical

¹TCM Clinical Basis Institute, Zhejiang Chinese Medicine University, 548 Binwen Road, Hangzhou, Zhejiang, 310000, China. ²Center for Bioinformatics and Computational Biology, and the Institute of Biomedical Sciences, School of Life Science, East China Normal University, 500 Dongchuan Road, Shanghai, 200241, China. ³School of Finance and Statistics, East China Normal University, 500 Dongchuan Road, Shanghai, 200241, China. ⁴Biological Targeting Diagnosis and Therapy Research Center, Guangxi Medical University, Nanning, Guangxi, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to T.S. (email: tieliushi01@gmail.com) or C.W. (email: wengcp@163.com)

practice has become increasingly important and has attracted attention as a source for drug development, such as the widely used drug, total glucosides of paeonia extracted from the root of *Paeonia lactiflora* that affects anti-inflammation, analgesia and immuno-regulation⁵. However, further investigation is still required to define the mechanism how herbs comprise any given formula function together⁶. Different herbal combinations have distinct treatment effects on individuals and contribute to the complexity of the mechanisms of TCM. As described in the “multiple components and multiple targets” concept of TCM, a formula exerts its therapeutic effect through interactions among the complex compounds of the medical herbs and the complex system of the diseased organism, which makes it extremely difficult for us to understand how these formulas treat diseases.

Although a great deal of effort has gone into unveiling the mechanisms behind TCM formulas, most remain unknown⁷. Thus, a systematic analysis of the complex mechanism behind TCM formulas is required. The rapid development of computational analyses and systems modeling approaches provide rich and substantial content of “biological networks”, and generate a new view of the life sciences and medical research⁸. Recently, the robustness of multiple systems biology platforms ensures the discovery of underlying molecular mechanisms and connections between drugs and their targets (e.g. proteomics studies on activated blood circulation of Chinese medicinal herbs)⁹. Newly developed algorithms and network-based computational models can integrate multi-level omics data and can optimize combinational regimens of drug development. The development of these tools encourages us to study medicinal herbs in view of network-based multi-target drug development¹⁰. Due to its robustness, sensitivity and adaptability, network-based drug combination discovery has the potential to provide a better understanding of the effects of Chinese herbal formulas, such as the neuroprotective mechanism behind Sheng-yu-tang¹¹. Resources are available for these systematic analyses. For example, we previously published Traditional Chinese Medicine Integrated Database (TCMID, <http://www.megabionet.org/tcmid/>)¹². TCMID is the largest integrated TCM database and provides the information of formulas, herbs, herbal ingredients, disease-related genes or proteins, diseases and drugs from large scale manually text mining and databases¹³. To be specific, we totally collected 46,914 formulas and 8,159 herbs by manual text mining, and data integration from various databases, including TCM-ID database (Traditional Chinese Medicine Information Database, <http://tcm.cz3.nus.edu.sg/group/tcm-id/>)¹⁴ and *Encyclopedia of Traditional Chinese Medicines*¹². The ingredients of herbs were derived from text mining, TCM-ID database, Herb Ingredients’ Targets database (HIT, <http://lifecenter.sgst.cn/hit/>)¹⁵ and TCM@Taiwan database (<http://tcm.cmu.edu.tw/>)¹⁶. The targets of ingredients were identified with literature evidences, HIT, STITCH (<http://stitch.embl.de/>)¹⁷, Online Mendelian Inheritance in Man (OMIM, <http://omim.org/>)¹⁸ and DrugBank (<http://www.drugbank.ca/>)¹⁹. Besides the comprehensive information, the database also contains online tools to present the relationships between herbs, ingredients and target genes in view of visual networks.

Systemic lupus erythematosus (SLE), a serious disease with no effective cure, is a multi-system autoimmune disease characterized by accumulation of anti-nuclear autoantibodies and various immunological abnormalities and is accompanied by excessive inflammatory responses in a wide range of organs²⁰. This disease can be treated with immunosuppressant drugs, including cyclophosphamide, corticosteroids and other immunosuppressants²¹. However, the exact cause of SLE is not completely understood. Omics data has the potential to uncover the mechanism of SLE. For example, analysis of patient transcriptional profiles offered a means to investigate mechanisms relevant to human diseases on a genome-wide scale²². Indeed, by using genome-wide expression profiles, Damien Chaussabel and his colleagues identified transcriptional modules based on genes co-expressed in multiple disease datasets. They found 11 modules that contained differentially expressed genes related to SLE. A total of 9 out of the 11 modules were directly connected with the immune system (e.g., M1.1 included genes encoding immunoglobulin chains and the plasma cell marker)²². In TCM studies, Sheng-ma-bie-jia-tang (SMBJT), a popular TCM formula developed by Chinese medical sage Zhang Zhongjing, is widely used for the effective treatment of SLE^{23,24}. It comprises six herbs with suggestive dose: *Cimicifuga foetida* (SHENG MA, 6g), *Carapax trionycis* (BIE JIA, 3g), *Pericarpium zanthoxyli* (SHU JIAO, 3g), *Realgar* (XIONG HUANG, 0.5g), *Glycyrrhiza uralensis* (GAN CAO, 6g) and *Angelica sinensis* (DANG GUI, 3g)²⁵. The dose should be slightly modified according to the specific physical condition of each patient. To date, no significant side effects of SMBJT have been reported²³. Although several studies focused on the experiments of SMBJT²⁴, the pharmacological mechanism of SMBJT has not yet been fully elucidated.

To explore the potential pharmaceutical mechanism of SMBJT for the treatment of SLE, we analyzed its herbal targets obtained from the TCMID database. First, we discovered that the targets of SMBJT were significantly enriched in SLE related biological processes, e.g. regulation, signal transduction and metabolism. Considering whether targets were previously defined as SLE proteins, we classified targets into essential or common ones. Next, using a protein-protein interaction network we investigated the potential mechanism of essential and common targets. Finally, we adopted a functional module concept and surveyed module enrichment of SMBJT targets. Overall, this study proposed a systemic method to investigate the molecular mechanisms of SMBJT, which improves the understanding of this TCM formula.

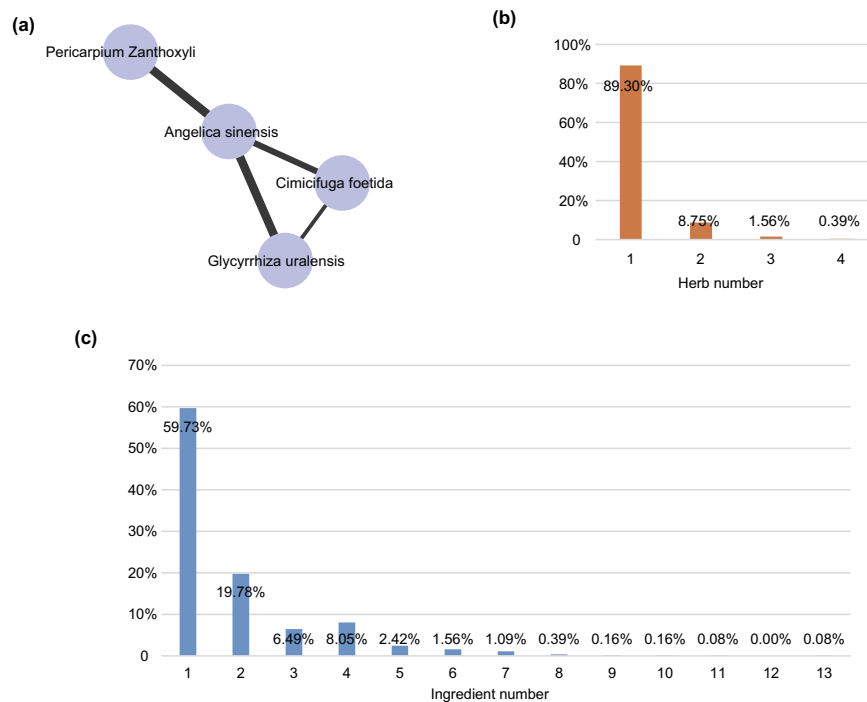


Figure 1. The relationships between herbs, ingredients and targets. (a) The relationships between the 4 herbs sharing the same ingredients. The links between herbs were constructed with the same ingredients in Supplementary Table S2. The width of the edge means the similarity between herbs in ingredients. (b) The component of targets correlated to different number of herbs. The numbers of herbs for targets were from 1 to 4, and each percentage was presented in the pie chart. (c) The component of targets correlated to different number of ingredients. The numbers of ingredients for targets were from 1 to 13, and each percentage was presented in the pie chart.

Results

The herbs, ingredients and targets of the SMBJT formula. Six herbs are included in the SMBJT formula: *Cimicifuga foetida*, *Carapax trionycis*, *Pericarpium zanthoxyli*, *Realgar*, *Glycyrrhiza uralensis* and *Angelica sinensis*. Traditionally, *Cimicifuga foetida* has been used as a cooling and detoxifying remedy herb²⁶ while *Carapax trionycis* can relieve fever²⁷. Due to the significant functions in these effects, *Cimicifuga foetida* and *Carapax trionycis* are thought as the monarch herbs. At the same time, *Pericarpium zanthoxyli*, *Realgar* and *Glycyrrhiza uralensis* all assist in the clearance of heat and toxic materials. In addition to the main effects, *Pericarpium zanthoxyli* also relieves pain, while *Angelica sinensis* promotes blood circulation to remove blood stasis.

We obtained a total of 390 different ingredients of the six herbs from TCMID (Supplementary Table S1). Specifically, *Cimicifuga foetida* contained 42 ingredients, including bicuculline, visnagin, caffeic acid, etc. *Carapax trionycis* contains 3 ingredients, including vitamin D, collagen and keratin. Furthermore, there were 172 ingredients in *Glycyrrhiza uralensis*, 163 in *Angelica sinensis*, 16 in *Pericarpium zanthoxyli* and 3 in *Realgar*. Among the 390 different ingredients, we identified 8 ingredients that were present in more than one herb (Supplementary Table S2). These ingredients were obviously correlated with SLE. An ingredient (glycyrrhizic) inhibited the immunocomplex formation of 60S acidic ribosomal P proteins with their specific antibodies in sera from patients with SLE²⁸, and another ingredient (limonene) had significant anti-inflammatory effects in murine dermal inflammation and wound-healing²⁹. Besides, these ingredients that were shared between herbs might be the key ingredients in the biological function of the herbs. We further identified herb pairs sharing the same ingredients (Fig. 1a): *Glycyrrhiza uralensis* and *Angelica sinensis* (3 ingredients); *Cimicifuga foetida* and *Glycyrrhiza uralensis* (1 ingredient); *Cimicifuga foetida* and *Angelica sinensis* (2 ingredients); and *Angelica sinensis* and *Pericarpium zanthoxyli* (3 ingredients). Compared with single target therapies, these TCM ingredients shared by different herbs systematically promoted different biological responses (cooling, detoxifying, blood circulation and pain relief) in human body, which could be more efficient in complex disease treatment.

In TCMID, half of the 8 ingredients shared between herbs had related targets (Supplementary Table S2), and we totally obtained 1,267 highly confident targets of 93 ingredients. 89.3% of the targets were correlated with one herb, indicating the herb specific property of targets (Fig. 1b). However, 40.27% of the targets (Fig. 1c) were correlated with more than one ingredient, suggesting the “multiple components and multiple targets” in SMBJT formula. For example, interleukin (IL)-6 correlated

KEGG pathway	Counts of genes	p-value	Class
Chemokine signaling pathway	55	5.80E-06	Organismal Systems; Immune system
Toll-like receptor signaling pathway	35	9.70E-06	Organismal Systems; Immune system
Adipocytokine signaling pathway	34	2.30E-10	Organismal Systems; Endocrine system
PPAR signaling pathway	33	3.00E-09	Organismal Systems; Endocrine system
Drug metabolism	23	1.00E-07	Metabolism; Xenobiotics biodegradation and metabolism
Metabolism of xenobiotics by cytochrome P450	28	1.10E-07	Metabolism; Xenobiotics biodegradation and metabolism
Drug metabolism	27	1.10E-06	Metabolism; Xenobiotics biodegradation and metabolism
Fatty acid metabolism	23	1.80E-08	Metabolism; Overview
Retinol metabolism	38	6.60E-18	Metabolism; Metabolism of cofactors and vitamins
Pyruvate metabolism	24	2.60E-09	Metabolism; Carbohydrate metabolism
Citrate cycle (TCA cycle)	20	1.50E-08	Metabolism; Carbohydrate metabolism
Propanoate metabolism	19	2.30E-07	Metabolism; Carbohydrate metabolism
Glycolysis / Gluconeogenesis	26	0.000002	Metabolism; Carbohydrate metabolism
Tryptophan metabolism	19	0.000015	Metabolism; Amino acid metabolism
Tyrosine metabolism	21	4.20E-06	Metabolism; Amino acid metabolism
Arginine and proline metabolism	22	0.000033	Metabolism; Amino acid metabolism
Cytokine-cytokine receptor interaction	69	0.000021	Environmental Information Processing; Signaling molecules and interaction
Neuroactive ligand-receptor interaction	133	7.90E-42	Environmental Information Processing; Signaling molecules and interaction

Table 1. The top 20 KEGG pathways with p-values < 0.05 generated by DAVID.

with 13 different ingredients, e.g. glycyrrhizin and retinol (Supplementary Table S3). IL-6 is cytokine with the anti-inflammation and anti-allergy effects that were correlated with arthritis manifestation of SLE³⁰. Another 4 proteins (IL-8, prostaglandin endoperoxide synthase 2 [PTGS2], tumor necrosis factor [TNF] and vascular endothelial growth factor A [VEGFA]) were correlated by at least 10 ingredients (Supplementary Table S3). The targets could be the most distinct factors for the formula, thus they were comprehensively investigated in the following analyses to uncover the molecular mechanism.

Functional analysis of the 1,267 targets. We performed GO enrichment analysis of the 1,267 targets of SMBJT and found that they were significantly enriched in biological processes of signal transduction, biological regulation and metabolic processes (Supplementary Table S4). Formula targets were significantly enriched in top-ranked biological processes, such as signal transduction (29.1%), cell surface receptor-linked signal transduction (24.4%) and intracellular signaling cascade (16.41%). Additionally, the top enriched biological processes included many regulation-related processes, such as positive regulation, negative regulation, regulation of cell communication, regulation of cell proliferation and regulation of programmed cell death. In addition to signal transduction and biological regulation, the targets also participated in metabolic biological processes (e.g. oxoacid and cellular lipid metabolic processes).

The signaling and metabolic pathways were then investigated in detail. The targets were enriched in 40 pathways with p-values < 0.05. The top 20 pathways (Table 1) belonged to three categories: organismal systems, metabolism, and environmental information processing and cellular processes. The identification of these pathways helped us to understand the functional contributions of the targets to SLE treatment. For example, two immune system pathways (chemokine signaling pathway and Toll-like receptor [TLR] signaling pathway) and two endocrine system pathways (adipocytokine signaling pathway and peroxisome proliferator-activated receptor [PPAR] signaling pathway) were implicated. Chemokines were small chemoattractant peptides about cell trafficking and thus were vital for protective host response in effective inflammatory immune response³¹. According to the KEGG database, the chemokine signaling pathway together with the adipocytokine signaling pathway were associated with a disease called Nuclear Factor kappaB Essential Modulator (NEMO) syndrome (incontinentia pigmenti). Interestingly, NEMO syndrome (an acronym of a mutated, non-functioning NF-kappaB gene) was associated with SLE³², and NF-kappaB was an essential modulator in the pathogenesis of SLE in the context of the complex immune deficiencies increasingly³². On the other hand, TLR family was responsible for sensing microbial pathogens and occupied an important position in innate immune responses³³. TLR signals in B cells amplified anti-dsDNA autoantibody and enhanced one SLE characteristic, autoantibody production³⁴. PPAR

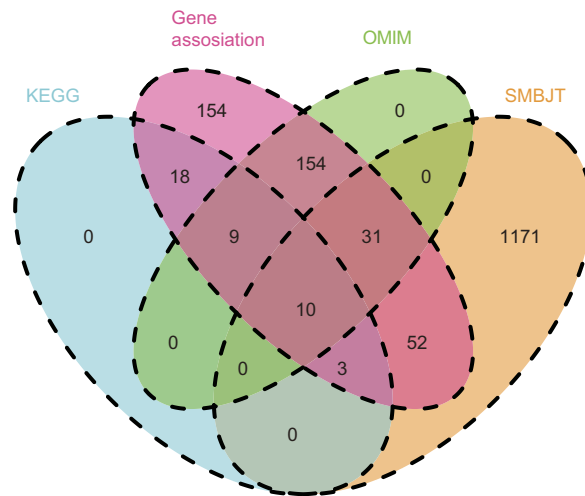


Figure 2. Comparisons between SMBJT targets and SLE disease proteins from different databases. The data in Gene association database are the most complete among the three databases. 96 targets in SMBJT are known SLE disease proteins. The remaining targets could be explained by the PPIs with SLE disease proteins.

gamma expression, one of PPARs subtypes, was increased in SLE patients and regulated the inflammatory signal³⁵.

Metabolism accounted for more than half of the top 20 pathways. These pathways belonged to five different categories (xenobiotics biodegradation and metabolism, metabolism of cofactors and vitamins, fatty acid metabolism, carbohydrate metabolism and amino acid metabolism) and were perturbed in SLE patients with literatures evidences. Metabolic disorders were detected in both human patients and mice with SLE, including disorders in amino acid metabolism in the SLE mouse model and significantly decreased tryptophan in human SLE patients^{36,37}. In SLE patients, two important metabolites (citrate and pyruvate) were also decreased³⁸, while essential fatty acid metabolism altered³⁹.

Overlaps between targets and SLE disease genes. We collected SLE disease genes from three databases (204 from OMIM, 431 from GAD and 40 from KEGG), and obtained a total of 431 SLE disease genes. We found that 96 targets of SMBJT were also defined as SLE disease genes in these databases (Fig. 2) and considered as essential targets. These targets included TLRs, ILs, chemokines, cytochrome P450 family, glutamate receptor, and adiponectin. These essential targets had the potential to significantly affect SLE with different mechanisms. For example, apolipoprotein E (APOE) was involved in the stimulus responses leading to changes in the state or activity of a cell or organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus⁴⁰. The polymorphisms of another essential target ATP-binding cassette, sub-family B (MDR/TAP), member 1 (ABCB1) could interfere with the clinical features of SLE⁴¹. The increased expression of one essential target (calcium/calmodulin-dependent protein kinase IV, CAMK4) was necessary for Th17 cell differentiation that was correlated to tissue inflammation in several autoimmune diseases, including SLE⁴². Three essential targets (vitamin D receptor [VDR], TLRs and APOE) played important roles in SLE pathogenesis⁴³. Moreover, signal transducer and activator of transcription 1 (STAT1) were potentially useful indicators about the initiation, progression and maintenance of inflammation in SLE patients⁴⁴. Beside essential targets, the remaining 1,171 targets (e.g. myeloid differentiation primary response 88[MYD88D], SRC proto-oncogene, non-receptor tyrosine kinase [p60-src], and macrophage migration inhibitory factor-like protein [MIF2]) were defined as common targets.

Direct interactions between essential targets and SLE disease proteins. To uncover the underlying mechanism of the targets, we applied PPIs to explore the functional relationships between targets and SLE disease proteins. A total of 910 targets and 329 disease proteins had PPIs. Among them, 94.7% of the targets either indirectly or directly interacted with 98.5% of the SLE disease proteins. The direct interactions between essential targets and disease proteins were presented in the network (Supplementary Figure S1). The network contained a highly connected PPI cluster, as well as PPI branches originated from the highly connected PPI cluster. The highly connected PPI cluster (Fig. 3a) included 49 nodes (35 SLE disease proteins and 14 essential targets). In this highly connected PPI cluster, 5 proteins (two SLE disease proteins: protein tyrosine phosphatase, non-receptor type 6 [PTPN6] and LYN proto-oncogene, Src family tyrosine kinase [LYN]; and three essential targets: STAT1, Janus kinase 2 [JAK2] and FYN proto-oncogene, Src family tyrosine kinase [FYN]) appeared to act as hub proteins based on their high degrees of association. These 5 proteins were closely correlated with SLE. For example, PTPN6, STAT1

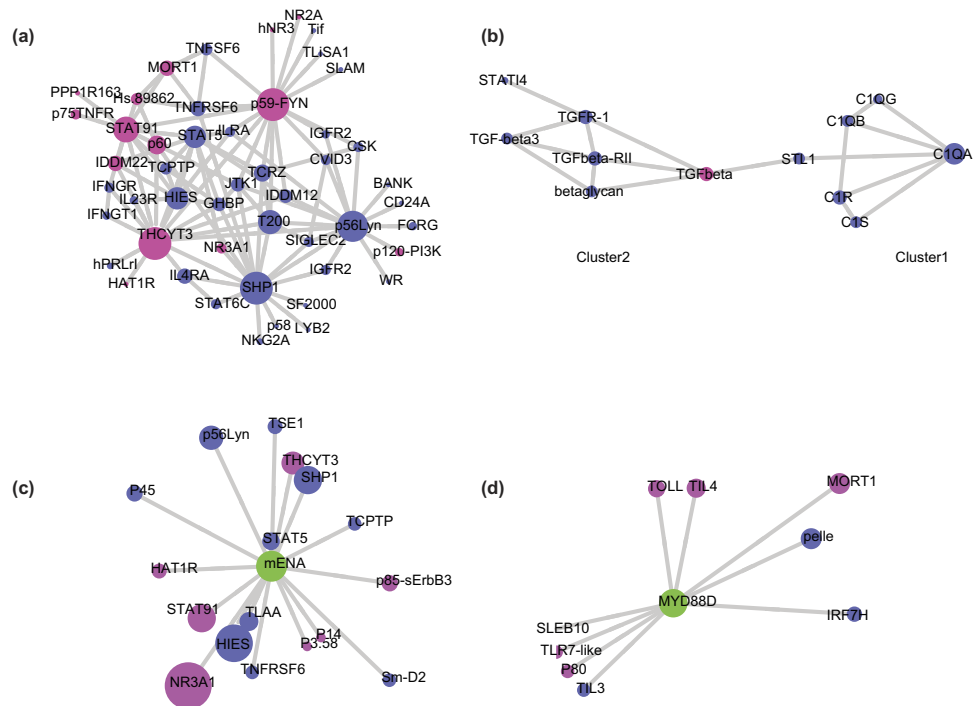


Figure 3. PPI sub-networks of SLE disease genes, essential targets and common targets. (a) A highly connected cluster in the PPI network of SLE disease genes and essential targets. (b) Two disease protein clusters connected by TGFbeta. (c) Sub-network of high-degree common target, mENA. (d) Sub-network of the high-degree, common target MYD88D. Blue, the products of SLE disease genes; red, essential targets; green, common targets.

and JAK2 were involved in JAK-STAT1 signaling pathway. Experiments confirmed the detection of hyperactivation of the JAK-STAT1 signaling pathway might involve in the pathogenesis of SLE⁴⁴. This highly connected PPI cluster could be conceived as the core of the formula.

Additionally, this core PPI cluster was connected to PPI branches that included different disease protein clusters. For example, one PPI branch includes two clusters (Fig. 3b): cluster1: collagen, type II, alpha 1 (STL1), complement component 1, r subcomponent (C1R), complement component 1, s subcomponent (C1S), complement component 1, q subcomponent, B chain (C1QB), complement component 1, q subcomponent, C chain (C1QG) and complement component 1, q subcomponent, A chain (C1QA); and cluster2: transforming growth factor, beta 1 (TGF-beta), beta-glycan, transforming growth factor, beta receptor II (TGFbeta-RII), transforming growth factor, beta receptor 1 (TGFR-1), transforming growth factor, beta 3 (TGFbeta3) and suppressor of cytokine signaling 6 (STAT14). The proteins in cluster1 were involved in the immune response, innate immunity and the complement pathway, which were closely correlated with the immune system. In details, reduced central tolerance to STL1 might induce and aggravate arthritis caused by cross-reactive autoantibody production⁴⁵, and C1Q deficiency was prevalence in SLE patients while Anti-C1q antibodies was identified as a marker in SLE patients^{46,47}. The proteins in cluster2 stimulated invasion and metastasis during carcinogenesis and promoted many pathological fibrotic diseases when overstimulated⁴⁸. To be specific, STAT14 was associated with the development of SLE by inhibiting growth of NK cells^{49,50}. One target (TGFbeta) linking the two disease protein clusters together could be the key protein connecting the two functional modules. Indeed, autoimmunity was triggered as a result of the decreased immunosuppressive effect induced by depressed TGFbeta levels in patients with SLE⁵¹.

Interactions between common targets and SLE disease proteins. For common targets, we discovered that they closely interacted with essential targets and SLE disease proteins (Supplementary Fig. S2), however, their network properties were different. The average degree of common targets (2.7) was smaller than that of essential targets (4.9) and SLE disease proteins (3.3); the average degree of essential targets was the highest of the three. A total of 45% of common targets with low degrees (<3) interacted with specific essential targets or SLE disease proteins. For instance, the essential target chemokine receptor 2 (IL8RB) interacted with 10 common targets (guanine nucleotide binding protein [G protein], alpha inhibiting activity polypeptide 2 [H_LUCA16.1], chemokine ligand 3 [SCYB3], chemokine ligand 1 [SCYB1], guanine nucleotide binding protein alpha 15 [GNA16], chemokine ligand 2 [SCYB2], chemokine ligand 5 [SCYB5], chemokine receptor 1 [IL8RBA], guanine nucleotide binding protein alpha 14 [GNA14],

pro-platelet basic protein [THBGB1] and chemokine ligand 5 [SCYB6]) (Supplementary Figure S3a). IL-8RBA was involved in the augmentation of tissue injury in SLE during inflammatory responses by cooperating with FcγRIIa and enhancing polymorphonuclear leukocyte (PMN) recruitment in the presence of anti-endothelial cell antibodies (AECAs), which were commonly detected in diseases associated with vascular injury, including SLE⁵². Seven of the 10 common targets (guanine nucleotide binding protein [G protein], alpha inhibiting activity polypeptide 2 [H_LUCA16.1], chemokine ligand 3 [SCYB3], chemokine ligand 1 [SCYB1], guanine nucleotide binding protein alpha 15 [GNA16], chemokine ligand 2 [SCYB2], chemokine ligand 5 [SCYB5], chemokine receptor 1 [IL8RBA], guanine nucleotide binding protein alpha 14 [GNA14], pro-platelet basic protein [THBGB1] and chemokine ligand 5 [SCYB6]) participated in the chemokine signaling pathway that was involved in SLE pathogenesis⁵³. The function of most the proteins in this module was about chemokine, indicating this module could be considered as chemokine module. Another essential target (tyrosine 3-monooxygenase [YWHAA]) interacted with 15 common targets (centromere protein C [MIF2], cyclin-dependent kinase inhibitor 1B [P27KIP1], uncoupling protein 2 [UCPH], microtubule-associated protein tau [TAU], uncoupling protein 3 [SLC25A9], tyrosine hydroxylase [TYH], protein tyrosine phosphatase, non-receptor type 3 [PTPH1], insulin receptor substrate 2 [IRS-2], insulin receptor [HHF5], insulin receptor substrate 1 [HIRS-1], Raf-1 proto-oncogene, serine/threonine kinase [C-Raf], protein kinase C, delta [nPKC-delta], ABL proto-oncogene 1, non-receptor tyrosine kinase [v-abl], cell division cycle 25C [PPP1R60] and p60-Src) (Supplementary Figure S3b). Four of the 15 common targets were members of the PI3K-AKT signaling pathway. The PI3K/Akt/mTOR signaling pathway was essential to cellular proliferation and growth signaling and was correlated with autoimmune diseases due to its activation in lymphocytes that developed features of systemic autoimmunity⁵⁴. These two examples demonstrate that the low-degree, common targets mostly interact with disease or essential targets, suggesting they mainly play supporting roles in the formula.

In addition to the low-degree common targets, common targets with higher degrees (≥ 3) could also play essential roles in connecting SLE-related proteins. Common target epidermal growth factor receptor (mENA) interacted with 16 essential targets (estrogen receptor 1 [NR3A1], signal transducer and activator of transcription 3 [HIES], signal transducer and activator of transcription 1, 91kDa [STAT91], Fas cell surface death receptor [TNFRSF6], intercellular adhesion molecule 1 [P3.58], cyclin-dependent kinase inhibitor 2A [P14], small nuclear ribonucleoprotein D2 polypeptide 16.5kDa [SM-D2], erb-b2 receptor tyrosine kinase 3 [p85-sErbB3], protein tyrosine phosphatase, non-receptor type 2 [TCPTP], signal transducer and activator of transcription 5 [STAT5], protein tyrosine phosphatase, non-receptor type 6 [SHP1], Janus kinase 2 [THCYT3], protein kinase, cAMP-dependent, regulatory, type I, alpha [TSE1], S-phase kinase-associated protein 2, E3 ubiquitin protein ligase [P45], angiotensin II receptor, type 1 [HAT1R] and estrogen receptor 1 [NR3A1]) (Fig. 3c). These essential targets were involved in SLE pathogenesis; moreover, genetic variants of NR3A1 (estrogen receptor, ranked as the highest degree of the 16 essential targets) might influence its susceptibility for involvement in pathogenesis⁵⁵. THCYT3, SHP1 and STAT91 influenced the pathogenesis of SLE through the JAK-STAT1 signaling pathway⁵⁶, while mENA (usually known as a direct regulator of microfilament polymerization and bundling) promoted metastasis in various cancers⁵⁷. Another common target, MYD88B, interacted with 9 essential targets (interferon regulatory factor 5 [SLEB10], toll-like receptor 7 [TLR7-like], interleukin 1 receptor, type I [P80], toll-like receptor 5 [TIL3], toll-like receptor 4 [TOLL], toll-like receptor 4 [TIL4], Fas (TNFRSF6)-associated via death domain [MORT1], interleukin-1 receptor-associated kinase 1 [pelle] and interferon regulatory factor 7 [IRF7H]) (Fig. 3d); MYD88B participated in the TLR signaling pathway with 8 of these targets (excluding P80). SLEB10 deficiency inhibited autoantibody production and ameliorates SLE disease likely due to its effects on TLR7 and TLR9⁵⁸. This example presents a potential key role for MYD88B in the TLR signaling pathway during SLE treatment. One final complicated example was about the interactions between 84 common targets and 24 essential targets (Supplementary Figure S4). Common target p60-src interacted with all 24 essential targets, demonstrating its essential role in this cluster. P60-src (Csk) was critical to multiple maturation and activation steps in B cells and physically interacted with the intracellular phosphatase encoded by protein tyrosine phosphatase, non-receptor type 22 (PTPN22) to form the Lyp-Csk complex that was capable of modifying the activation state of downstream SRC proto-oncogene, non-receptor tyrosine kinase (Src kinases) (e.g. Lyn in lymphocytes) and increasing susceptibility to SLE⁵⁸. Overall, these examples demonstrate that the non-key high-degree targets connect different functional modules related to SLE biological processes, suggesting their important functions in SLE treatment.

Evaluation of targets with SLE functional modules. Analysis of patient transcriptional profiles offers a means to investigate mechanisms relevant to human diseases on a genome-wide scale. Using genome-wide expression profiling, Damien Chaussabel and his colleagues identified transcriptional modules based on genes co-expressed in multiple disease data sets. They constructed peripheral blood mononuclear cell (PBMC) transcriptional modules focusing on small sets of coordinately expressed transcripts. Using the “data-driven” and 3-round module-selection process, they finally obtained 28 functional modules, including 3,485 proteins, which were then used to map transcriptional changes between patients and healthy subjects. A total of 235 SBJT targets were included in these 28 functional modules (Supplementary Table S5). Among the 28 modules, 11 modules were related to SLE; a total of 9 out of

the 11 modules were directly correlated to immune system. For example, M1.1 included genes encoding for immunoglobulin (Ig) chains and the plasma cell marker. Other than healthy individuals and patients suffering from other systemic autoimmune disease, light chain Ig κ and Ig λ were expressed simultaneously in B cells from SLE patients, and the allelic inclusion exhibited high frequencies of having lupus nephritis⁵⁹. Nevertheless, inhibition of long-lived plasma cells was involved in several therapeutic strategies of SLE⁶⁰.

Furthermore, we identified targets of SMBJT that were significantly enriched in three modules (M1.5, M2.6 and M3.3). The functions of the proteins in these three modules were related to the myeloid lineage (M1.5 and M2.6) and Inflammation II (M3.3), suggesting the involvement of SMBJT targets in these major biological processes. Genes in M1.5 encode proteins expressed in myeloid-lineage cells (CD86 molecule, CD163 molecule and Fc fragment of IgG, low affinity IIa, receptor [FCGR2A]), proteins involved in pathogen recognition (CD14 molecule, toll-like receptor 2 [TLR2] and MYD88) and TNF-family members (tumor necrosis factor receptor superfamily, member 1B [TNFR2] and tumor necrosis factor superfamily, member 13b [BAFF]). Genes in M2.6 (IGTB2/CD18, Lymphotoxin beta receptor, Myeloid related proteins 8/14, and Formyl peptide receptor 1) were expressed in myeloid-lineage cells such as monocytes and neutrophils. Genes in M3.3 encode proteins (interleukin 18 [IL-18], arachidonate 5-lipoxygenase [ALOX5], alanyl aminopeptidase [ANPEP], acylglycerol hydrolase [AOAH], heme oxygenase 1 [HMOX1] and serpin peptidase inhibitor, clade B, member 1 [SERPINB1]) that were induced by inflammation or lysosomal enzymes (palmitoyl-protein thioesterase 1 [PPT1], cathepsin B/S [CTSB/S], sialidase 1 [NEU1], N-acylsphingosine amidohydrolase 1 [ASAH1], lysosomal-associated membrane protein 2 [LAMP2] and calpastatin [CAST]).

Discussion

SMBJT is a TCM formula that is widely used for SLE treatment. In this study, we evaluated the potential molecular mechanism of SMBJT. Functional analysis of the 1,267 targets of SMBJT highlighted the importance of these targets in regulation, signaling and metabolism. 96 essential and 1,171 common targets of this formula were identified, and their relationships with SLE disease genes in a PPI network were investigated. Hub proteins were identified in network clusters to play important roles in this formula. The small number of targets and SLE disease proteins that do not interact with one another could contribute to SMBJT's side effects and treatment failure in a small number of clinical cases. The biological function of the SMBJT targets identified in this work will help us to understand how this TCM formula treats the complex disease SLE.

Considering the significance of immune system processes in the formula targets, we compared the ratio of essential targets and common targets in the whole immune system to discover the functional relationships between two types of targets (Fig. 4). We found that key targets were involved in the core component of the immune upstream regulation processes significantly correlated with SLE, e.g. toll-like receptor signaling pathway, myeloid leukocyte cytokine, chronic inflammatory response, leukocyte mediated immunity, immunoglobulin, etc. The chronic inflammatory response was involved in the development and prognosis of SLE⁶¹. Fragments of immunoglobulins were significantly correlated with SLE disease activity⁶², and complement receptor of the immunoglobulin superfamily could diminish inflammation and reversed established bone destruction⁶³. On the other hand, common targets normally participated in the peripheral immune processes that were more specific to leukocyte, e.g. leukocyte degranulation, leukocyte chemotaxis, etc. We also discovered that several immune processes dominated by essential targets served as the important regulators in the peripheral immune processes, e.g. negative regulation of leukocyte chemotaxis and regulation of cellular extravasation. The overview of targets in immune system suggests essential and common targets cover the most important processes with distinct immune function, and indicating the SMBJT might provide more systematic therapy for SLE patients in TCM.

In the further investigation of the relationships between essential and common targets in toll-like receptor signaling pathway (Supplementary Figure S5) and chemokine signaling pathway (Supplementary Figure S6), we discovered that the common targets largely filled the empty between essential targets and composed a more complete target path. In toll-like receptor signaling pathway (Supplementary Figure S5), the common target, MYD88 filled the blank step between two essential targets (TLR2 and Fas-associated via death domain [FADD]) that formed a complete signaling path lead to apoptosis. Thus, we believe that in the treatment of complex disease, SMBJT has more drug targets (especially common targets that were not defined as SLE disease gene before) and provides more systematic therapy for patients. This more genome-wide TCM drug mechanism could also avoid the palliatives in single target medical service for the complex autoimmune disease.

Traditional research into the mechanism behind TCM formulas was based on single-ingredient and single-target models. However, the mechanism should be investigated using complex multi-target interactions, which has proven challenging for researchers. The rapid development of bioinformatics offers an alternative approach for the study of the mechanisms of TCM formulas using multi-level data. The method we proposed here provides a new way to comprehensively understand the therapeutic effects of TCM formulas in SLE treatment. This work could serve as a good example for the studies of the mechanisms of TCM formulas on other complex diseases.



Figure 4. The ratio of essential targets and common targets in immune system processes. The network was generated with ClueGO in Cytoscape.

Materials and Methods

Data collection. The TCMID database (TCMID, <http://www.megabionet.org/tcmid/>) contains a large amount of information concerning formulas and their herbal ingredients¹². We retrieved the ingredients of six herbs and their protein targets from TCMID for further analysis. The STITCH defined targets of ingredients with confidence ranges for data scores (low confidence: scores <0.4 ; medium: 0.4 to 0.7 ; high: >0.7)¹⁷. Based on these scores, we chose 1,267 high confident targets with comprehensive scores >0.7 . Genes associated with SLE were collected from three databases: OMIM (<http://www.omim.org/>)¹⁸, GAD (<http://geneticassociationdb.nih.gov/>)⁶⁴ and KEGG (<http://www.kegg.jp/>)⁶⁵.

Gene ontology and pathway enrichment analysis. DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/>)⁶⁶, which represents a comprehensive set of functional annotation tools for understanding the biological meanings behind large gene datasets, were used to perform Gene Ontology (GO) enrichment analysis for the 1,267 genes targeted by SMBJT. DAVID was also applied to conduct pathway enrichment analysis for the same 1,267 gene products to identify associations between SLE and SMBJT. Enriched GO terms and pathways were defined as those with adjusted p-values <0.05 .

Network construction. Based on the protein-protein interaction data in HPRD (<http://www.hprd.org/>)⁶⁷ and STRING (<http://string-db.org/>)⁶⁸, we constructed a protein-protein interaction (PPI) network for the 1,267 target proteins of SMBJT and the SLE susceptibility genes to obtain the hub effective proteins. The networks were visualized with Cytoscape⁶⁹.

Module enrichment analysis. We adopted the 28 functional modules identified by Damien Chaussabel and his colleagues²². A total of 11 modules contained differentially expressed genes related to SLE. Among the 4,742 transcripts contained in the 28 functional modules, 3,485 have been annotated. We compared the 3,485 proteins with the target proteins of SMBJT and identified 235 common genes (Supplementary Table S5). We selected modules containing at least one SMBJT target gene. Next, we performed module enrichment analysis for the 28 functional modules.

The enrichment analysis was designed to filter out differences in expression levels of gene sets between two or more groups (especially enriched gene sets). Common gene enrichment analysis methods can be

summarized into two categories: the bottom-up method and the top-down method. The bottom-up method is the most widely used; this method starts with single-gene analysis and then continues with further analyses based on biological information, such as GO annotation gene sets. Common bottom-up analysis methods include Fisher's exact test, gene enrichment analysis, gene enrichment parameter analysis and gene set analysis⁷⁰. Here, we chose Fisher's exact test. Let G_c denote the set of extracted proteins in component c and G denote the set of proteins in a functional unit (functional module). Let $r = |G_c|$, $k = |G|$, $z = |G_c \cap G|$ and l be the total number of proteins in the whole dataset. We assumed that z followed a hypergeometric distribution. The probability of observing an intersection of size z between G and G_c was computed as equation (1). The enrichment score of component c was defined in equation (2). Enrichment of the target proteins of TCM in the identified functional modules (adjusted p -values < 0.05) indicated that these proteins contributed to the therapeutic mechanism of TCM.

$$P(G, G_c) = \sum_{i=z}^{\min(k,r)} \frac{\binom{k}{i} \binom{l-k}{r-i}}{\binom{l}{r}} \quad (1)$$

$$s(c) = -\log_{10} P^{FDR}(G, G_c) \quad (2)$$

References

- Zhao, J., Jiang, P. & Zhang, W. Molecular networks for the study of TCM pharmacology. *Brief Bioinform* **11**, 417–430 (2010).
- Wen, Z. *et al.* Discovery of molecular mechanisms of traditional Chinese medicinal formula Si-Wu-Tang using gene expression microarray and connectivity map. *PLoS One* **6**, doi: 10.1371/journal.pone.0018278 (2011).
- Wang, L. *et al.* Dissection of mechanisms of Chinese medicinal formula Realgar-Indigo naturalis as an effective treatment for promyelocytic leukemia. *Proc Natl Acad Sci USA* **105**, 4826–4831 (2008).
- Chen, J. Y. *et al.* The herbal compound geniposide rescues formaldehyde-induced apoptosis in N2a neuroblastoma cells. *Sci China Life Sci* **57**, 412–421 (2014).
- Wang, C. *et al.* Total glucosides of paeony inhibit the inflammatory responses of mice with allergic contact dermatitis by restoring the balanced secretion of pro-/anti-inflammatory cytokines. *Int Immunopharmacol* **24**, 325–334 (2015).
- Song, J. *et al.* A Modularity-Based Method Reveals Mixed Modules from Chemical-Gene Heterogeneous Network. *PLoS One* **10**, doi: 10.1371/journal.pone.0125585 (2015).
- Li, J. *et al.* Metabolomics study on model rats of chronic obstructive pulmonary disease treated with BuFei JianPi. *Mol Med Rep* **11**, 1324–1333 (2015).
- Li, S. Framework and practice of network-based studies for Chinese herbal formula. *chin J Integr Med* **5**, 489–492 (2007).
- Liu, Y., Yin, H. & Chen, K. Platelet proteomics and its advanced application for research of blood stasis syndrome and activated blood circulation herbs of Chinese medicine. *Sci China Life Sci* **56**, 1000–1006 (2013).
- Leung, E. L., Cao, Z. W., Jiang, Z. H., Zhou, H. & Liu, L. Network-based drug discovery by integrating systems biology and computational technologies. *Brief Bioinform* **14**, 491–505 (2013).
- Hou, Y. C., Lu, C. K. & Wang, Y. H. Unique Mechanisms of Sheng Yu Decoction on Ischemic Stroke Mice Revealed by an Integrated Neurofunctional and Transcriptome Analysis. *J Tradit Complement Med* **3**, 240–249 (2013).
- Xue, R. *et al.* TCMID: Traditional Chinese Medicine integrative database for herb molecular mechanism analysis. *Nucleic Acids Res* **41**, D1089–1095 (2013).
- Fang, Z. *et al.* Evaluating the pharmacological mechanism of Chinese medicine Si-Wu-Tang through multi-level data integration. *PLoS One* **8**, doi: 10.1371/journal.pone.0072334 (2013).
- Chen, X. *et al.* Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation. *Br J Pharmacol* **149**, 1092–1103 (2006).
- Ye, H. *et al.* HIT: linking herbal active ingredients to targets. *Nucleic Acids Res* **39**, D1055–1059 (2011).
- Chen, C. Y. TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PLoS One* **6**, doi: 10.1371/journal.pone.0015939 (2011).
- Kuhn, M., von Mering, C., Campillos, M., Jensen, L. J. & Bork, P. STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res* **36**, D684–688 (2008).
- Baxevanis, A. D. Searching Online Mendelian Inheritance in Man (OMIM) for information on genetic loci involved in human disease. *Curr Protoc Hum Genet* **9**, doi: 10.1002/0471142905.hg0471140913s0471142973 (2012).
- Knox, C. *et al.* DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* **39**, D1035–1041 (2011).
- Huang, X. *et al.* Dihydroartemisinin inhibits activation of the Toll-like receptor 4 signaling pathway and production of type I interferon in spleen cells from lupus-prone MRL/lpr mice. *Int Immunopharmacol* **22**, 266–272 (2014).
- Loram, L. C., Culp, M. E., Connolly-Strong, E. C. & Sturgill-Koszycki, S. Melanocortin peptides: potential targets in systemic lupus erythematosus. *Inflammation* **38**, 260–271 (2015).
- Chaussabel, D. *et al.* A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* **29**, 150–164 (2008).
- Wang, J. Z. & Zhou, Q. Y. Theory research between Sheng Ma Biejia Tang and SLE. *Acta Chin Med& Pharm* **38**, 3–5 (2010).
- Kang, Q. The influence that Sheng-ma-bie-jia-tang work on the cytokine imbalance of Th1/Th2 in mice. *Chin J Traditi Med Sci&Tech* **20**, 456–458 (2013).
- Wang, X. H., Wang, J. Z., Zhou, Q. Y., Du, Y. Q. & Zhao, Y. J. Theoretical Research of Sheng-ma-bie-jia-tang and Systemic lupus erythematosus. *J Chin Med Pharm* **38**, 3–5 (2010).
- Zhu, D. F. *et al.* Cycloartane Glycosides from the Roots of *Cimicifuga foetida* with Wnt Signaling Pathway Inhibitory Activity. *Nat Prod Bioprospect* **5**, 61–67 (2015).
- Wang, P. *et al.* Protection of a new heptapeptide from Carapax trionycis against carbon tetrachloride-induced acute liver injury in mice. *Chem Pharm Bull (Tokyo)* **61**, 1130–1135 (2013).
- Maekawa, T. *et al.* Biochemical characterization of 60S acidic ribosomal P proteins from porcine liver and the inhibition of their immunocomplex formation with sera from systemic lupus erythematosus (SLE) patients by glycyrrhizin *in vitro*. *Biol Pharm Bull* **23**, 27–32 (2000).

29. d'Alessio, P. A., Mirshahi, M., Bisson, J. F. & Bene, M. C. Skin repair properties of d-Limonene and perillyl alcohol in murine models. *Antiinflamm Antiallergy Agents Med Chem* **13**, 29–35 (2014).
30. Ball, E. M., Gibson, D. S., Bell, A. L. & Rooney, M. R. Plasma IL-6 levels correlate with clinical and ultrasound measures of arthritis in patients with systemic lupus erythematosus. *Lupus* **23**, 46–56 (2014).
31. Wong, M. M. & Fish, E. N. Chemokines: attractive mediators of the immune response. *Semin Immunol* **15**, 5–14 (2003).
32. Piccoli, G. B. *et al.* NEMO syndrome (incontinentia pigmenti) and systemic lupus erythematosus: a new disease association. *Lupus* **21**, 675–681 (2012).
33. Kawai, T. & Akira, S. TLR signaling. *Semin Immunol* **19**, 24–32 (2007).
34. Ma, K., Li, J., Fang, Y. & Lu, L. Roles of B Cell-Intrinsic TLR Signals in Systemic Lupus Erythematosus. *Int J Mol Sci* **16**, 13084–13105 (2015).
35. Oxer, D. S. *et al.* PPARgamma expression is increased in systemic lupus erythematosus patients and represses CD40/CD40L signaling pathway. *Lupus* **20**, 575–587 (2011).
36. Hu, J. B. *et al.* Metabolomics Study of the Jieduquyuziyin Recipe for treatment of Systemic Lupus Erythematosus in Mice. *Chin J Tradit Chin Med Pharm* **29**, 60–65 (2014).
37. Widner, B. *et al.* Degradation of tryptophan in patients with systemic lupus erythematosus. *Adv Exp Med Biol* **467**, 571–577 (1999).
38. Ouyang, X., Dai, Y., Wen, J. L. & Wang, L. X. (1) H NMR-based metabolomic study of metabolic profiling for systemic lupus erythematosus. *Lupus* **20**, 1411–1420 (2011).
39. Suryaprabha, P., Das, U. N., Ramesh, G., Kumar, K. V. & Kumar, G. S. Reactive oxygen species, lipid peroxides and essential fatty acids in patients with rheumatoid arthritis and systemic lupus erythematosus. *Prostaglandins Leukot Essent Fatty Acids* **43**, 251–255 (1991).
40. Pullmann, R. J. *et al.* Apolipoprotein E polymorphism in patients with neuropsychiatric SLE. *Clin Rheumatol* **23**, 97–101 (2004).
41. Gonzalez, T. P. *et al.* ABCB1 C1236T, G2677T/A and C3435T polymorphisms in systemic lupus erythematosus patients. *Braz J Med Biol Res* **41**, 769–772 (2008).
42. Koga, T. *et al.* CaMK4-dependent activation of AKT/mTOR and CREM-alpha underlies autoimmunity-associated Th17 imbalance. *J Clin Invest* **124**, 2234–2245 (2014).
43. Vinh quoc Luong, K. & Nguyen, L. T. The beneficial role of vitamin D in systemic lupus erythematosus (SLE). *Clin Rheumatol* **31**, 1423–1435 (2012).
44. Li, J. *et al.* Activation of JAK-STAT1 signal transduction pathway in lesional skin and monocytes from patients with systemic lupus erythematosus. *J Cent Sout Univ (Med Sci)* **36**, 109–115 (2011).
45. Campbell, I. K. *et al.* Autoimmune regulator controls T cell help for pathogenetic autoantibody production in collagen-induced arthritis. *Arthritis Rheum* **60**, 1683–1693 (2009).
46. van Schaarenburg, R. A. *et al.* Marked variability in clinical presentation and outcome of patients with C1q immunodeficiency. *J Autoimmun* **62**, 39–44 (2015).
47. Bock, M., Heijnen, I. & Trendelenburg, M. Anti-C1q antibodies as a follow-up marker in SLE patients. *PLoS One* **10**, doi: 10.1371/journal.pone.0123572 (2015).
48. Verona, E. V. *et al.* Expression, purification and characterization of BG(E)RII: a novel pan-TGFbeta inhibitor. *Protein Eng Des Sel* **21**, 463–473 (2008).
49. Kucuk, C. *et al.* Global promoter methylation analysis reveals novel candidate tumor suppressor genes in natural killer cell lymphoma. *Clin Cancer Res* **21**, 1699–1711 (2015).
50. Ma, H. *et al.* Dynamic changes in the numbers of different subsets of peripheral blood NK cells in patients with systemic lupus erythematosus following classic therapy. *Clin Rheumatol* **33**, 1603–1610 (2014).
51. Manolova, I., Gerenova, J. & Ivanova, M. Serum levels of transforming growth factor-beta1 (TGF-beta1) in patients with systemic lupus erythematosus and Hashimoto's thyroiditis. *Eur Cytokine Netw* **24**, 69–74 (2013).
52. Florey, O. J., Johns, M., Esho, O. O., Mason, J. C. & Haskard, D. O. Antiendothelial cell antibodies mediate enhanced leukocyte adhesion to cytokine-activated endothelial cells through a novel mechanism requiring cooperation between Fc{gamma}RIIa and CXCR1/2. *Blood* **109**, 3881–3889 (2007).
53. Schiffer, L., Worthmann, K., Haller, H. & Schiffer, M. CXCL13 as a new biomarker of Systemic Lupus Erythematosus (SLE) and Lupus Nephritis (LN) – from bench to bedside? *Clin Exp Immunol* **179**, 85–89 (2014).
54. Besliu, A. N. *et al.* PI3K/Akt signaling in peripheral T lymphocytes from systemic lupus erythematosus patients. *Roum Arch Microbiol Immunol* **68**, 69–79 (2009).
55. Wang, J., Nuite, M. & McAlindon, T. E. Association of estrogen and aromatase gene polymorphisms with systemic lupus erythematosus. *Lupus* **19**, 734–740 (2010).
56. Watkins, A. A., Boneggio, R. G. & Rifkin, I. R. Evaluating the role of nucleic acid antigens in murine models of systemic lupus erythematosus. *Methods Mol Biol* **1169**, 143–158 (2014).
57. Lin, L. *et al.* Microfilament regulatory protein MENA increases activity of RhoA and promotes metastasis of hepatocellular carcinoma. *Exp Cell Res* **327**, 113–122 (2014).
58. Manjarrez-Orduno, N. *et al.* CSK regulatory polymorphism is associated with systemic lupus erythematosus and influences B-cell signaling and activation. *Nat Genet* **44**, 1227–1230 (2012).
59. Fraser, L. D. *et al.* Immunoglobulin light chain allelic inclusion in systemic lupus erythematosus. *Eur J Immunol* **24** JUN doi: 10.1002/eji.201545599 (2015).
60. Hiepe, F. *et al.* Long-lived autoreactive plasma cells drive persistent autoimmune inflammation. *Nat Rev Rheumatol* **7**, 170–178 (2011).
61. Lu, M., Yu, S., Xu, W., Gao, B. & Xiong, S. HMGB1 Promotes Systemic Lupus Erythematosus by Enhancing Macrophage Inflammatory Response. *J Immunol Res* **2015**, 946748 (2015).
62. Solano, F. G., Bellei, E., Cuoghi, A., Caiazza, M. & Bruni, F. Radical improvement of signs and symptoms in systemic lupus erythematosus when treated with hemodiafiltration with endogenous reinfusion dialysis. *Case Rep Nephrol Dial* **5**, 106–112 (2015).
63. Lieberman, L. A. *et al.* Complement receptor of the immunoglobulin superfamily reduces murine lupus nephritis and cutaneous disease. *Clin Immunol* **160**, 286–291 (2015).
64. Becker, K. G., Barnes, K. C., Bright, T. J. & Wang, S. A. The genetic association database. *Nat Genet* **36**, 431–432 (2004).
65. Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**, D199–205 (2014).
66. Huang, D. 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).
67. Keshava Prasad, T. S. *et al.* Human Protein Reference Database--2009 update. *Nucleic Acids Res* **37**, D767–772 (2009).
68. Szklarczyk, D. *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **39**, D561–568 (2011).

69. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
70. Cao, W. J., Li, Y. M. & Chen, C. S. Progress on Enrichment Analysis Approach of Gene Expression Profiles. *Biotechnol Lett* **19**, 4 (2008).

Acknowledgements

This work was supported by the Natural Science Foundation of China (81173249, 81373633 and 81273680), the Traditional Chinese Medicine Scientific Research Fund Project of Zhejiang Province (2012ZB062) and the Important Zhejiang Provincial Science and Technology Specific Projects (2014C03046-1). We thank Rong Cheng, Guanting Zhang and Huan Wang for their help in data collection and programming.

Author Contributions

T.S. and C.W. conceived and designed the experiments; H.L. and Q.L. performed the experiments, analyzed the data and contributed to the writing of the manuscript. F.L. contributed analysis tools.

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: Huang, L. *et al.* A Systems Biology-Based Investigation into the Pharmacological Mechanisms of Sheng-ma-bie-jia-tang Acting on Systemic Lupus Erythematosus by Multi-Level Data Integration. *Sci. Rep.* **5**, 16401; doi: 10.1038/srep16401 (2015).



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/>