Classification-based strategies to simplify complex traditional Chinese medicine (TCM) researches through liquid chromatography-mass spectrometry in the last decade (2011-2020): theory, technical route and difficulty

Yue-Hua Chen , Jing-Hua Bi , Ming Xie , Hui Zhang , Zi-Qi Shi , Hua Guo , Hai-Bo Yin , Jia-Nuo Zhang , Gui-Zhong Xin , Hui-Peng Song

PII: S0021-9673(21)00431-3
DOI: https://doi.org/10.1016/j.chroma.2021.462307
Reference: CHROMA 462307

To appear in: *Journal of Chromatography A*

Received date: 13 February 2021
Revised date: 27 May 2021
Accepted date: 29 May 2021

Please cite this article as: Yue-Hua Chen , Jing-Hua Bi , Ming Xie , Hui Zhang , Zi-Qi Shi , Hua Guo , Hai-Bo Yin , Jia-Nuo Zhang , Gui-Zhong Xin , Hui-Peng Song , Classification-based strategies to simplify complex traditional Chinese medicine (TCM) researches through liquid chromatography-mass spectrometry in the last decade (2011-2020): theory, technical route and difficulty, *Journal of Chromatography A* (2021), doi: https://doi.org/10.1016/j.chroma.2021.462307

Classification-based strategies to simplify complex traditional Chinese medicine (TCM) researches through liquid chromatography-mass spectrometry in the last decade (2011-2020): theory, technical route and difficulty

Yue-Hua Chen[a,1], Jing-Hua Bi[c,1], Ming Xie[a], Hui Zhang[a], Zi-Qi Shi[d], Hua Guo[a], Hai-Bo Yin, Jia-Nuo Zhang, Gui-Zhong Xin[b,*], Hui-Peng Song[a,*]

[a] Key Laboratory for Identification and Quality Evaluation of Traditional Chinese Medicine of Liaoning Province, Liaoning University of Traditional Chinese Medicine, Dalian 116600, China

[b] State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China

[c] Shanxi Medical University, Taiyuan 030001, China

[d] Jiangsu Province Academy of Traditional Chinese Medicine, Nanjing 210028, China

[1] These authors contributed equally to this work.

* Corresponding author:

Gui-Zhong Xin

State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China

E-mail: xingz@cpu.edu.cn (G.Z. Xin)

Highlights

1. A universal methodology was proposed to simplify complex TCM researches.
2. Set theory and LC-MS were combined to study TCM for the first time.
3. The proposed theory is a link between traditional medicines and modern

medicines.

4. Three important technical routes of TCM researches by LC-MS were introduced.
5. This review is a summary of TCM research experience in the past 10 years.

Hui-Peng Song

Key Laboratory for Identification and Quality Evaluation of Traditional Chinese Medicine of Liaoning Province, Dalian 116600, China

E-mail: songhuipeng15@163.com (H.P. Song)

Abstract: The difficulty of traditional Chinese medicine (TCM) researches lies in the complexity of components, metabolites, and bioactivities. For a long time, there has been a lack of connections among the three parts, which is not conducive to the systematic elucidation of TCM effectiveness. To overcome this problem, a classification-based methodology for simplifying TCM researches was refined from literature in the past 10 years (2011-2020). The theoretical basis of this methodology is set theory, and its core concept is classification. Its starting point is that "although TCM may contain hundreds of compounds, the vast majority of these compounds are structurally similar". The methodology is composed by research strategies for components, metabolites and bioactivities of TCM, which are the three main parts of the review. Technical route, key steps and difficulty are introduced in each part. Two perspectives are highlighted in this review: set theory is a theoretical basis for all strategies from a conceptual perspective, and liquid chromatography-mass spectrometry (LC-MS) is a common tool for all strategies from a technical perspective. The significance of these strategies is to simplify complex TCM researches, integrate isolated TCM researches, and build a bridge between traditional medicines and modern medicines. Potential research hotspots in the future, such as discovery of bioactive ingredients from TCM metabolites, are also discussed. The classification-based methodology is a summary of research experience in the past 10 years. We believe it will definitely provide support and reference for the following TCM researches.

Keywords: traditional Chinese medicine; liquid chromatography; mass spectrometry; research strategy; set theory

Table of contents:

1. Introduction

The research on traditional Chinese medicine (TCM) has gone through another 10 years (2011-2020). In ancient times, people found out that some plants had medicinal value [i,ii]. For instance, ginseng has long been found to be beneficial to the overall function of human body, which was recorded in the earliest known Chinese pharmaceutical work of "Shen Nong's Materia Medica" [iii]. Many modern drugs such as artemisinin and ephedrine have undergone the development process from traditional medicines to modern medicines. It proves that discovery of active components is an important way to clarify the effectiveness of TCM [iv,v]. However, TCM are still too complicated because they usually contain hundreds of compounds. Compared with modern drugs which usually have a specific target, the biological targets of TCM are a complex network [vi,vii]. Nowadays, traditional medicines still play an important role in disease treatment around the world. According to the World Health Organization, 80% of people around the world are willing to choose traditional medicines [viii]. Therefore, revealing the effectiveness of traditional medicines is a significant and challenging task.

Liquid chromatography (LC) and mass spectrometry (MS) have been the most important tools for TCM research in the past decade [ix,x]. LC is used to separate complex chemical components in TCM, while MS is applied for structural identification. The development of liquid chromatography has gone from thin-layer liquid chromatography to ultra-high performance liquid chromatography. The enhanced separation capability contributes to the exposure of trace components in TCM [xi]. Mass spectrometry develops rapidly in two dimensions, namely, from low resolution to high resolution and from low sensitivity to high sensitivity [xii,xiii]. In the past decade, the application of LC-MS to TCM mainly focused on components, metabolites and bioactive components, which are the core contents of TCM research. Herein, metabolites refer to metabolized substances produced from the components of TCM in the human or animal body. One of the difficulties of TCM research is the complexities, including (1) the complexity of components: there may be hundreds or even thousands of components in TCM [xiv]; (2) the complexity of metabolic

processes: the components in TCM may undergo a series of reactions in the body such as oxidation, reduction and hydrolysis [xv]; (3) the complexity of bioactive components: compared with the "one drug for one target" mode of modern medicines, the bioactive components in TCM act through multiple targets [xvi]. Because each of the above sections contains a large amount of work [xvii,xviii], they rarely appear in a single research paper at the same time. In other words, the three parts are not closely related. The lack of integration makes TCM research more complex, which is not conducive to systematically elucidating the effectiveness of TCM. It is necessary to propose a methodology to simplify and integrate TCM research from a global perspective.

By reviewing literature in the past decade, the author found that set theory which comes from the field of mathematics can be used to simplify and integrate TCM research [xix]. Although TCM may contain thousands or hundreds of compounds, most of them are structurally similar [xx,xxi]. In set theory (**Fig. 1**), the compounds with similar structures can be regarded as a set, and each compound is an element in this set. For example, more than 120 saponins have been found from ginseng, and they can be divided into three groups according to the aglycone structure, namely, protopanaxadiol type, protopanaxtriol type, and oleanolic acid type [xxii,xxiii]. From the perspective of set theory, saponins in ginseng is a set, and each saponin is an element in the set. The saponin set can be further divided into protopanaxadiol-type subset, proginsentriol-type subset and oleanolic acid-type subset. From this perspective, the composition of ginseng is composed of three subsets rather than hundreds of compounds. In this way, the complexity of TCM composition will be greatly simplified. When the components of TCM are absorbed by the body, more metabolites will be produced. Actually, many metabolic processes such as methylation and hydroxylation do not affect the chemical skeleton of compounds. In this case, the number of metabolite sets remains unchanged [xxiv,xxv]. In addition, based on the general rule that structure determines activity, compounds in the same set have similar activities due to their similar structures [xxvi]. This will facilitate the study on biological activity of compounds in TCM. Therefore, from the perspective of set

theory, the main contents of TCM research through LC-MS can be considered as component sets [xxvii], metabolite sets [xxviii] and bioactive component sets [xxix].

The purpose of this paper is to review the research strategies on component sets, metabolite sets and bioactive component sets of TCM in the past 10 years. "How to simplify and integrate methods for TCM research" is the key question of this review. To answer this question, the concept and idea of sets were used to bridge component sets, metabolite sets and bioactive component sets of TCM. The advantages of this approach are as follows. (1) The core idea of set theory is classification, which has been an important strategy for TCM research in the past decade. (2) The concepts in set theory such as subset and union can be used to accurately describe the research strategies of TCM. (3) From the Venn diagram of multiple sets, it is easy to find blank fields of current research. There are two characteristics in this review: on the one hand, set theory is used to integrate all strategies from a conceptual point of view; on the other hand, LC-MS is a common tool for all strategies from a technical point of view. To the best of our knowledge, set theory and LC-MS were integrated to simplify TCM researches for the first time. The limitations of research strategies in the past ten years and the potential research hotspots in the future were also discussed in this review.

2. Strategies for component analysis of TCM

2.1 Technical route: from set to subset and from training set to testing set

The key to the strategy for qualitative analysis of TCM components is classification, which is the core idea of set theory. There are two technical routes (**Fig. 2**): the first is from set to subset, which refers to the classification of compounds according to structural similarity; the second is from training set to testing set, which refers to inferring the structure of unknown compounds in TCM based on the structural information of known chemical reference substances. The two technical routes are integrated rather than isolated. The general method is as follows. Compounds in a TCM are classified into different subsets according to structural similarity. Some available compounds are selected from each subset as the training set, and their characteristic cleavage rules in mass spectra are summarized. Based on the

information obtained from the training set, unknown compounds with similar structures in TCM are speculated. Discovery of common MS rules is the key of the technical route. On the one hand, common MS rules are the basis for dividing a set into subsets. Compounds belonging to the same subset have common MS fragmentation pathways because they have the same skeleton. On the other hand, common MS rules are the link between the training set and the testing set. The unknown compounds in the testing set were analyzed according to the skeleton information provided by MS [xxx]. Therefore, diagnostic ion filtration (DIF) and neutral fragment filtration (NFF) are commonly used for the analysis of TCM components [xxxi]. Besides, mass defect filtration (MDF) is also an alternative method for rapid identification of similar components. The mass defect refers to the difference between the exact mass and the nominal mass of a compound. In general, structurally similar compounds possess similar mass defects [xxxii]. Some TCM and TCM prescriptions are used as examples to describe the technical routes in detail (**Table 1**).

Flavonoids Thirteen flavonoids were used as the training set by Cheng *et al.* to explore the cleavage rules of components in *Spatholobus suberectus*, which is a traditional herbal medicine for the treatment of blood stasis. The results show that diagnostic ions from retro-Diels-Alder reactions and neutral fragments of C_2H_2O (42 Da), CO (28 Da) and CO_2 (44 Da) are helpful to component analysis. Finally, 36 flavonoids, including 13 target compounds and 23 non-target compounds were characterized [xxxiii]. *Scutellaria baicalensis* is an antimicrobial and antiviral TCM. UHPLC-orbitrap-MS was used to analyze chemical reference substances and *Scutellaria baicalensis*, which were used as training set and testing set respectively. In the training set, 34 chemical reference substances were divided into 6 subsets according to structural similarity, including 13 flavone *O*-glycosides, 8 free flavones, 7 flavone *C*-glycosides, 2 phenylethanoid glycosides, 2 flavanones, and 2 flavanone *O*-glycoside. Through the filtration of key ions and fragments, 132 compounds were characterized from Scutellaria baicalensis, 59 of which were discovered from it for the first time [xxxiv]. Zhang *et al.* obtained diagnostic ions from the training set of 11

hydroxylated polymethoxyflavonoid glycosides (OH-PMFGs). By filtering diagnostic ions, a total of 54 OH-PMFGs derived from 3 subsets of flavone glycosides, flavanone glycosides and chalcone glycosides were identified from *Murraya Paniculata L.* [xxxv]. In another work, this team combined MDF with DIF to characterize 81 polymethoxylated flavonoids from the leaves of *Citrus reticulata Blanco* [xxxvi].

Alkaloids *Uncaria rhynchophylla* is widely used to treat hypertension and giddiness. Fourteen alkaloids in the training set were divided into 4 subsets, including 4 tetracyclic monoterpene oxindole alkaloids, 2 tetracyclic monoterpene indole N-oxides, 6 tetracyclic monoterpene indole alkaloids, and 2 pentacyclic monoterpene indole alkaloids. Based on MS data of the training set, m/z 160.08, m/z 184.10, m/z 144.08 and a neutral fragment 162.05 Da were applied to identify above 4 types of alkaloids in *Uncaria rhynchophylla*, respectively. Finally, the structures of 92 alkaloids were inferred, of which 56 were found for the first time [xxxvii]. Chen *et al*. proposed a stepwise diagnostic fragment ion and neutral loss-dependent algorithm to analyze isoquinolines in *Stephania tetrandra*. Fourteen chemical standards were used as the training set to explore MS fragmentation pathways, and finally 393 isoquinoline alkaloids distributed in more than 20 subsets were characterized [xxxviii]. Shi *et al.* established a self-feedback network to analyze β-carboline alkaloids in *Picrasma quassioides*. Actually, this network is a combination of MDF, DIF, and NFF. MDF was used for ion classification to find potential β-carboline alkaloids, and DIF/NFF was applied for further structural identification. Finally, 89 alkaloids were identified in *Picrasma quassioides*, of which 24 were potential new compounds [xxxix]. The three filtration methods are widely used in alkaloid-containing herbs such as *Stephania hainanensis* [xl], *Semen Strychni* [xli], *Corydalis species* [xlii], and *Plumula nelumbinis* [xliii].

Saponins Qi e*t al.* used 13 ginsenosides as the training set to obtain the common MS rules of triterpenoid saponins in ginseng. By filtering diagnostic ions, the saponins can be divided into three main subsets according to the aglycone structure, namely, protopanaxadiol-type subset, protopanaxatriol-type subset and ocotillol-type

subset. The sugar moieties of saponins can be inferred by NFF. Based on this strategy, a total of 70 saponins were characterized from American ginseng. This strategy can be used to identify aglycones, sugar moieties and their linkages of saponins [xliv]. Similarly, Ma *et al.* used 11 compounds as the training set to explore the common MS rules of triterpenoid saponins in *Akebiae Fructus*. Through DIF and NFF, 85 triterpenoid saponins that distributed in 9 subsets were finally identified [xlv]. Besides, MDF has also been proved to be an effective method for saponin analysis. By setting the mass range and mass defect range, 62 components from *Ophiopogon japonicus* were assigned to ophiopogonins by Xie *et al.*, which is a kind of steroidal saponins. Finally, the chemical structures of 50 ophiopogonins were characterized, and the accuracy of MDF in finding structural analogues was over 80% [xlvi].

TCM prescriptions A TCM prescription consists of multiple herbs, and it has more kinds and larger quantities of compounds. Shuxiong tablet is a modern prescription consisting of *Notoginseng Radix et Rhizoma*, *Carthami Flos*, and *Chuanxiong Rhizoma*. It is commonly used to treat coronary heart disease and angina pectoris. As many as 72 compounds in the training set were divided into 5 subsets including 39 ginsenosides, 17 flavonoid *O*-glucosides, 7 phenolic acids, 4 phthalide derivatives and 5 quinochalcone *C*-glycosides. Considering the difference in MS response of compounds, both the positive-ion mode and the negative-ion mode were used to obtain MS data. The negative-ion mode was used to analyze ginsenosides, flavonoid *O*-glucosides, quinochalcone *C*-glycosides and phenolic acids; the positive-ion mode was applied to analyze phthalide derivatives. Taking ginsenosides and flavonoid *O*-glucosides as examples, their characteristic diagnostic ions and neutral losses are as follows. Since the glycosyl moieties of the two types of compounds are similar, they have some common neutral losses such as 162.05 (Glc) and 146.06 (Rha). Their aglycone ions are different: the characteristic ions of ginsenosides are m/z 459.38 (protopanaxadiol), 475.38 (protopanaxatriol), 455.35 (oleanolic acid) and 457.37 (5,6-didehydro-protopanaxadiol); the characteristic ions of flavonoid *O*-glucosides are m/z 317.03 (6-OH quercetin), 301.04 (6-OH kaempferol or quercetin), 287.06 (4',5,6,7-OH flavanone) and 285.04 (kaempferol).

Based on the fragmentation rules obtained from the training set, a total of 250 components were characterized from Shuxiong tablet [xlvii].

Due to the complexity of TCM prescriptions, the coordination of MS data, retention time, and isotope intensity distribution data is necessary to eliminate interference from irrelevant ions. Er-xian decoction, a TCM prescription for menstrual disorder treatment, is composed of *Curculiginis Rhizoma*, *Epimedii Herba*, *Angelicae Sinensis Radix*, *Morindae Officinalis Radix*, *Phellodendri Chinensis Cortex*, and *Anemarrhenae Rhizoma*. *Wang et al.* aims to study prenylflavonoids, phenylpropanoids, and alkaloids in Er-xian Decoction, and a strategy was proposed to remove interfering ions [xlviii]. Two methods, namely HRMS/MS^n-based mass spectral tree similarity filter and HRMS/RT-based discriminant analysis, were separately used to find potential ions of the three types of compounds in Er-xian Decoction. By intersecting the ions obtained by the two methods, 553 ions were picked out. By further subtracting the background ions, the number of candidate ions was reduced from 553 to 240. Due to the reference of template molecules in the training set, 45 compounds were quickly identified from these ions. Isotope intensity distribution data was finally used to confirm the structures of compounds. This strategy is helpful to eliminate interfering ions and accelerate the identification of target compounds.

2.2 Key steps in component analysis

Two key factors that influence component analysis of TCM are the data acquisition capability of MS and the separation capability of LC. Ideal chromatographic separation will lay a good foundation for subsequent MS data acquisition by reducing mutual interference of ions. A variety of data acquisition modes of mass spectrometry can provide abundant raw data for the component analysis. The two aspects are introduced in detail as follows.

2.2.1 Improve the data acquisition capability of MS

A direct way to improve the component analysis of TCM is to use a variety of data acquisition modes of mass spectrometry. Many mass spectrometers such as

quadrupole time-of-flight (Q-TOF), triple quadrupole linear ion trap (Qtrap), linear trap quadropole Orbitrap (LTQ-Orbitrap), ion trap time-of-flight (IT-TOF), triple quadrupole (QQQ) and Fourier transform ion cyclotron resonance (FTICR) are widely used to collect MS data. In order to obtain more abundant data, two points need to be noted.

First, it is necessary to make full use of the different data acquisition modes of a mass spectrometer. Q-TOF is the most commonly used mass spectrometer for component analysis of TCM. There are two main data acquisition modes. Data-independent acquisition can quickly obtain integrated MS data of TCM, but some molecular ions and fragment ions with the same retention time may interfere with each other. Data-dependent acquisition can be used to selectively detect an ion without interference, but its acquisition channels are usually limited. The combination of the above two acquisition modes is an optimized method for data acquisition. Xu *et al.* used this method to analyze the chemical components in licorice. Based on 15 flavonoids in the training set, 51 components were finally characterized from the licorice extract, of which 3 were reported for the first time [xlix]. In addition to Q-TOF, other mass spectrometers also have similar modes of data acquisition. For example, Qtrap-MS has non-selective EMS-IDA-EPI scan and selective Prec-IDA-EPI scan. The former can scan non-targeted compounds, and the latter can scan targeted compounds with the same substructure. The two methods were jointly used by Li *et al.* for component analysis of Danhong injection. Six compounds were used as the training set, and 90 compounds were inferred from Danhong injection [l].

Second, it is a good choice to use multiple types of mass spectrometers for component analysis of TCM. For example, there is a good complementarity between Q-TOF-MS and Qtrap-MS. The former has high resolution but low sensitivity, while the latter has high sensitivity but low resolution. The two mass spectrometers were combined to analyze chemical components in Baoyuan Decoction: the MS^E mode of Q-TOF-MS was used to analyze principal components, and the data-dependent acquisition mode of Qtrap-MS was applied to analyze trace components. Forty-nine compounds including 32 saponins, 14 flavonoids, and 3 diterpenes were used as the

training set, and finally 236 components were inferred from Baoyuan Decoction [li]. A similar method was also used to analyze components in raw and processed *Alismatis Rhizoma*. In this example, 7 triterpenoids were used as the training set. Finally, 80 and 87 compounds were identified from raw and processed *Alismatis Rhizoma*, respectively [lii].

2.2.2 Improve the separation capability of LC

Multidimensional chromatography is an effective method to improve the separation of components in TCM. In this way, the number of compounds identified in TCM will be significantly increased. For instance, Yang *et al.* used MCI gel×silica gel orthogonal column chromatography to obtain fractions from TCM extracts, and then each fraction was separated on a C18 column by HPLC. Ion-trap mass spectrometer and Q-TOF mass spectrometer were used to collect data. Based on 18 saponins in the training set, 334, 262, and 309 saponins were inferred from *Panax ginseng*, *Panax quinquefolium*, and *Panax notoginseng*, respectively [liii]. In another study, Pan *et al.* adopted offline two-dimensional chromatography to separate alkaloids in five herbs with the same Chinese name of Gou-Teng. A positively charged C18 column and a traditional C18 column were used for the first and the second dimensional separations, respectively. LTQ-Orbitrap mass spectrometer was applied for data acquisition. Based on the 44 alkaloids in the training set, 1227 indole alkaloids were finally characterized [liv]. The above researches show that the improvement of chromatographic separation capability is helpful to increase the number of compounds identified in TCM.

2.3 Difficulty in component analysis: distinguishing isomers

Isomers are widely found in TCM. For many mass spectrometers, molecular ions and fragment ions are insufficient to distinguish isomers due to their high structural similarity. Some methods such as quantitative structure-retention relationship, ion mobility spectrometry, and nuclear magnetic resonance could be applied to assist the identification of isomers.

Quantitative structure-retention relationship (QSRR) Because isomers have the same chemical composition, their MS data are very similar. Thus, it is difficult to

distinguish isomers only by MS data. For some isomers, their retention times on HPLC column are different. In this case, the quantitative structure-retention relationship (QSRR) can be used to distinguish isomers. The principle is to establish a model based on the physical-chemical parameters and retention times of known compounds, which can predict retention times of unknown compounds. In this way, isomers with different retention times can be matched with candidate compounds. For example, Wu *et al.* tried to analyze organic acid isomers in Mai-Luo-Ning injection and *Lonicerae Flos* injection. A total of 30 organic acids were divided into a training set (24 compounds) and a validation set (6 compounds) for establishing a QSRR model. The training set contains 4 groups of isomers, and the validation set contains 1 group of isomers. Multiple linear regression (MLR) model and back propagation neural network (BNN) model were used to simulate the relationship between chemical structures and retention times. The results show that the predicted elution order of the five groups of isomers based on the models is consistent with the experimental result, indicating the reliability of the models in predicting retention times of organic acid isomers. Finally, 45 organic acids containing many isomers were identified from Mai-Luo-Ning injection and *Lonicerae Flos* injection [lv]. A similar study was successfully applied to distinguish isomers in Chuanxiong [lvi].

Ion mobility spectrometry In fact, the retention times of most isomers on HPLC column are the same. For such isomers, the method of quantitative structure-retention relationship (QSRR) is invalid. To solve this problem, HPLC tandem ion mobility mass spectrometry was proposed as an effective method for isomer discrimination. Besides a first-dimensional separation (retention time) by HPLC, and ion mobility mass spectrometry can provide a second-dimensional separation (drift time) based on the charged state, shape and size of isomers in the gas phase. Therefore, although the retention times of some isomers are the same, their drift times are different. For example, there are 8 lanostane-type triterpene acids with the same molecular formula of $C_{30}H_{45}O_4$ in *Poria cocos*. Among them, one compound was identified by comparing with the chemical reference substance. Based on the relationship between drift times and structures, the other 7 isomers were divided into

stereo isomers and positional isomers. By searching SciFinder Scholar and related documents, they can be matched with known compounds. In short, HPLC tandem ion mobility spectrometry can provide information on 4 dimensions, including retention times, drift times, accurate molecular weights and characteristic fragment ions, which helps to identify the structure of isomers from multiple perspectives [lvii].

NMR Nuclear magnetic resonance Generally speaking, the information based on LC-MS is not sufficient to support the absolute identification of isomers. Nuclear magnetic resonance (NMR) is an important supplementary method for isomer identification. Unlike the good compatibility between LC and MS, the application of NMR is sometimes limited by its sensitivity. At present, NMR analysis can be divided into on-line mode and off-line mode, which are illustrated by the following two examples respectively. α-boswellic acids and β-boswellic acids are a pair of isomers in frankincense. Their structural difference lies in the position of a methyl group, and MS data is insufficient for their structural identification. As a supplement, on-line HPLC-NMR was used to distinguish them, and the results showed that their methyl signals were significantly different [lviii]. Compared with online NMR, offline NMR requires more pretreatment steps. The main components in Carthamus tinctorius are flavonoid glycosides, of which the aglycone moieties are difficult to identify due to the aglycone isomers with different positions of hydroxyl groups. To satisfy the detection sensitivity of NMR, a series of pretreatment steps including acid hydrolysis, HPLC separation and SPE enrichment were applied to obtain sufficient amount of flavonoid aglycones. By combining off-line NMR and MS, 103 flavonoid glycosides were finally identified from Carthamus tinctorius [lix].

3. Strategies for metabolite analysis of TCM

3.1 Technical route: from element to set and from prototype set to metabolite set

When a single prototype compound is absorbed into the body, it will undergo a series of biological processes to produce metabolites [lx]. In the concept of set theory, the relationship between prototypes and metabolites is analogous to mapping. The complexity of identifying metabolites of TCM lies in three aspects. (1) A TCM extract

is a complex multi-component set; (2) The generated metabolites are also a complex multi-component set; (3) The generation of metabolites from prototypes is a complex process [lxi]. To address this problem, two technical routes were proposed for the identification of metabolites of TCM (**Fig. 3A**). The first is from element to set. Based on metabolic regularity of a representative compound (element), metabolites and metabolic pathways of TCM (set) could be inferred. The second is from prototype set to metabolite set. Based on clear metabolic pathways such as hydroxylation and methylation, metabolites can be derived from prototypes. Representative works for metabolite analysis of TCM based on LC-MS are shown in **Table 2**.

From element to set Animal experiment and software simulation are commonly used to explore the metabolism of representative compounds. They are introduced through the following two examples respectively.

In the first example, the main components in licorice are flavonoids and saponins. Among them, flavonoids can be divided into various chemical subsets such as flavanones, chalcones, isoflavones, and flavones. The compounds in each subset can be further divided into aglycone form and glycoside form. The diversity of chemical structures increases the complexity of metabolite identification of licorice. According to the route of "from element to set", representative compounds from each chemical subset were selected for *in vivo* metabolism research. For example, 7,4-dihydroxyflavone, isoangustone A, and ononin were used as representative compounds of flavone, isoflavones, and isoflavone glycosides, respectively. The results showed that all the representative compounds could be absorbed into the blood. After obtaining their metabolic rules, the licorice extract was administered orally to rats. Finally, a total of 12 prototype components and 78 metabolites were detected in the blood and urine of rats [lxii].

In the second example, the main components of *Glehniae Radix* are coumarins, which contains sub-types such as simple coumarins and linear-type furocoumarins. According to structural similarity, the coumarins in *Glehniae Radix* were divided into 5 chemical subsets. A representative compound was selected from each chemical subset to explore metabolic regularity. Unlike the first example of licorice, the

metabolism of representative compounds was studied through metabolic simulation software rather than animal experiments. In this example, MetabolitePilot software was used to derive phase I metabolites and phase II metabolites of representative compounds. The information obtained from these representative compounds was used for metabolite analysis of *Glehniae Radix*. Finally, 111 prototypes and metabolites were identified from the urine, plasma and bile of rats [lxiii]. From this example, it can be seen that some reliable software can greatly reduce the workload of metabolite analysis. However, for some less-studied compounds, animal experiments are still indispensable.

From prototype set to metabolite set In theory, a metabolite set was generated from a prototype set through different metabolic reactions. Therefore, metabolic reactions are the key to metabolite analysis. A target-group-change strategy that can reveal metabolic reactions in the body was proposed. The principle of this strategy is as follows: the process from a prototype to a metabolite is often caused by fixed metabolic reactions such as glucuronidation and hydroxylation, and the target group change caused by some reaction will result in a fixed m/z difference. By comparing the m/z values of a prototype and a metabolite, the metabolic reaction can be inferred. For example, the m/z difference of 14.0157 corresponds to methylation (CH_2), the m/z difference of 176.0321 corresponds to glucuronide conjugation ($C_6H_8O_6$), and the m/z difference of 15.9949 corresponds to hydroxylation (O). This strategy was used to a TCM prescription of Fufang-Xialian-Capsule. The results showed that 33 metabolites were derived from 11 prototype components through methylation, demethylation, glucuronide conjugation and hydroxylation [lxiv]. In another similar study, 89 prototypes and 82 metabolites of *Sophora flavescens* in rats were identified by the target-group-change strategy [lxv]. Based on this strategy, chem-metab explorer software was developed to match prototypes and metabolites in Mai-Luo-Ning injection. Finally, 162 metabolites in the urine of rats were connected with 87 prototypes, and the metabolic network of Mai-Luo-Ning injection was revealed [lxvi].

3.2 Key steps in metabolite analysis

3.2.1 Discrimination of exogenous substances from endogenous substances

When a TCM extract is absorbed into the body, exogenous substances are mixed with endogenous substances. Both exogenous substances and endogenous substances are complex, and thus it is challenging to distinguish them. The strategy to distinguish exogenous substances from endogenous substances is shown in **Fig. 3B**. Three samples including TCM-treated blood (exogenous substances + endogenous substances), TCM (exogenous substances), and blank blood (endogenous substances) are regarded as three sets. The logical relationship for finding prototypes and metabolites is as follows: (1) the intersection of TCM-treated blood and blank blood represents prototypes and metabolites (intersection I). (2) the intersection of TCM-treated blood and TCM represents prototypes (intersection II); (3) the intersection of intersection I and intersection II represents metabolites. It could be seen that finding intersections is a key step. In this process, multivariate statistical analysis is an effective tool. Principal component analysis (PCA), partial least squared discriminant analysis (PLS-DA) and orthogonal partial least squared discriminant analysis (OPLS-DA) are the most commonly used data-processing methods to distinguish exogenous substances from endogenous substances. These methods have been used to recognize prototypes and metabolites of Sini decoction [lxvii], Zhi-Zi-Hou-Po decoction [lxviii], Da-Huang-Xiao-Shi decoction [lxix], Shuang-Huang-Lian Injection [lxx], and Qingkailing injection [lxxi]. It is worth noting that some professional software plays an important role. For example, Luo *et al*. embedded MarkerView software into Microsoft Office to find out exogenous substances from biosamples [lxxii].

3.2.2 Cooperation of various databases

Sufficient database support is necessary for the identification of prototypes and metabolites. Among them, the identification of prototypes needs to refer to compound databases of TCM, and the identification of metabolites needs to refer to known metabolite databases. Databases are divided into two types, including self-built databases and public/commercial databases. For example, Liang *et al*. tried to study prototypes and metabolites of *Paeoniae Radix Rubra* decoction by combining

self-built databases and public databases. A digital library containing parent compound database, known metabolite database, and characteristic neutral loss database was established through ChemFinder Ultra software. Some online databases such as Chemical Abstracts Service database, METLIN and Chemspider were also used. With the support of above databases, 15 prototypes and 90 metabolites of *Paeoniae Radix Rubra* decoction were identified in plasma and urine of rats [lxxiii].

It is a time-consuming and labor-intensive task to manually retrieve databases. Some information-processing software such as UNIFI has been developed to solve this problem. The UNIFI software can be used to predict metabolites from prototypes or prototypes from metabolites. In addition to the online database of UNIFI, self-built databases can be imported into UNIF to assist the analysis of prototypes and metabolites. To some extent, the UNIF software is a search engine that links various databases. It can greatly reduce time and effort consumed by manual retrieval. The software has been widely used to analyze prototypes and metabolites of TCM. For example, based on the UNIFI software, 51 prototypes and 119 metabolites of Dan Zhi Tablet were identified from rat plasma [lxxiv]; 18 prototypes and 35 metabolites of Run-zao-zhi-yang capsule were identified from rat plasma, urine and bile [lxxv]; and 22 prototypes and 19 metabolites of Qi-Lin pills were identified from rat plasma [lxxvi]. The development of intelligent data-processing methods for metabolism study has become one of the research hotspots.

3.3 Difficulty in metabolite analysis: absolute identification of metabolites

It is worth noting that metabolite structures identified by LC-MS are not absolutely accurate. The determination of metabolite structures requires a variety of data provided by MS, H-NMR, C-NMR, etc. However, due to the low content of metabolites in the body, it is difficult to obtain sufficient metabolites for structural identification. Some researchers tried to purify metabolites of *Saussurea laniceps* from rat urine. The extract of *Saussurea laniceps* was orally given to three rats at a dose of 2.0 g/kg/d for 10 days, and a total of 420 mL of rat urine was collected. Four metabolites of umbelliferone glucuronide, scopoletin glucuronide, scopoletin glucuronide, and scopoletin glucuronide were purified by high-speed countercurrent

chromatography. The mass of each metabolite is more than 200 mg, and the purity is more than 90%. Their structures were determined by MS, H-NMR, and C-NMR [lxxvii]. In this work, only 4 high-content metabolites were purified, and a lot of low-content metabolites are difficult to purify.

Compared with the direct separation of low-content metabolites from animal urine or plasma, it is easy to purify metabolites through *in vitro* microsomal metabolism. In this way, Li *et al.* obtained 7 metabolites of atractylenolide I, which is a main component of *Atractylodis macrocephalae rhizoma*. The quality of the purified metabolites is in the milligram level, and their structures were identified by NMR and MS [lxxviii]. Similarly, Tang *et al.* discovered that duciformine and lankongensisine A were two main metabolites of acetylduciformine in rat liver microsomes, indicating that intramolecular esterification is an important way of acetylduciformine metabolism [lxxix]. Wu *et al.* explored the glucuronidation reaction of five anthraquinones from rhubarb in microsomes and found that β-OH was an important site of glucuronidation reaction [lxxx]. Ideally, the combination of *in vitro* microsomal metabolism and *in vivo* metabolism is an effective method for metabolite identification. For example, Liu *et al.* studied the metabolites of limonin in rat liver microsomes, bile and urine by LC-Q-TOF-MS. The structures of two metabolites were confirmed by chemical synthesis and NMR analysis, which show that hydrolysis and reduction are important ways of limonin metabolism [lxxxi].

4. Strategies for discovery of bioactive components from TCM

4.1 Technical route: from bioactive component set to element-element interaction

TCM have always played a role in curing diseases in the form of bioactive component sets [lxxxii]. In the process of discovering bioactive compounds from TCM, researchers are pursuing monomeric compounds with better activity [lxxxiii,lxxxiv], and the concept of set is completely ignored. The key to the therapeutic effect of TCM lies in the diversity of chemical structures and the synergistic effect between compounds [lxxxv,lxxxvi]. If each component is studied independently, the overall activity of TCM is still unclear. In other words, the

biological activity of any monomeric component cannot represent the overall activity of TCM. The bioactive component set is the key connection between TCM and its efficacy.

A bioactive component set has the following characteristics. (1) Qualitative characteristic: each compound in the bioactive component set exists in the TCM extract. (2) Quantitative characteristic: the content of each compound in the bioactive component set is the same as that in the TCM extract. (3) Bioactive characteristic: the bioactive component set plays a dominant role in the bioactivity of the TCM extract. The technical route for discovery of bioactive components from TCM is from bioactive component set to element-element interaction (**Fig. 4**). It includes: (1) discovering bioactive compounds from the perspectives of principal components, binding affinity, bioinformatics, and chromatogram-bioctivity correlation; (2) preparing potential bioactive component sets through knockout/knockin and recombination; (3) ranking the elements according to their bioactivity; (4) exploring element-element interactions. Representative works for the discovery of bioactive components from TCM based on LC-MS are shown in **Table 3**.

4.2 Key steps in discovery of bioactive components

4.2.1 Discovery of bioactive components

The most critical issue in the discovery of bioactive components is "how to select compounds from TCM as a member of the bioactive component set". The selection criteria include principal components, binding affinity, bioinformatics, and chromatogram-bioctivity correlation.

Principal components For some TCM, principal components may be bioactive components. Relative contents and absolute contents of components are two potential indicators, which are introduced by the following two examples.

Tanshinone extract is derived from *Salvia miltiorrhizae Radix* for the treatment of cardiovascular diseases. Guo *et al.* tried to find bioactive components from tanshinone extract. Normalized peak area ratio was used as the indicator to select potential bioactive components. By referring to this indicator, 4 principal components of cryptotanshinone, tanshinones IIA, dihydrotanshinone I and tanshinone I were

chosen to form a compound set, which were prepared by recombining the corresponding reference compounds according to their portions and contents in tanshinone extract. Finally, this set showed a similar activity with the original tanshinone extract in a series of in vitro and in vivo models, including hypoxia/reoxygenation-induced injury in H9c2 cardiomyocytes, H_2O_2-induced injury in H9c2 cardiomyocytes, and myocardial ischemia-reperfusion injury in rats. The result suggests that principal components are possible key bioactive components in TCM [lxxxvii].

In another example, Pang *et al.* tried to find a bioactive component set from Yindan Xinnaotong soft capsule, which is a TCM prescription composed of 8 herbs for the treatment of stroke and its sequelae. A total of 124 compounds in this prescription were characterized, of which the absolute contents of 67 components were detected. The absolute content was used as the indicator for selecting bioactive ingredients. Three content limits of > 1, > 0.75 and > 0.5 mg/g were set, and three candidate bioactive component sets of 7, 13, and 16 compounds were generated respectively. Finally, the set of 16 compounds showed similar effects to Yindan Xinnaotong soft capsule in the models of oxygen-glucose derivation/reoxygenation-induced injury in N2A cells and glutamate-induced toxicity in PC12 cells [lxxxviii]. Notably, TCM prescriptions usually contain numerous compounds, which will bring challenges to the discovery of bioactive component sets.

Binding affinity Many bioactive components in TCM take effect by binding to biological targets such as enzymes, receptors and pathological cells. The general flow of this strategy is as follows. A TCM extract and a biological target are incubated to form target-compound complexes. Then, the target-compound complexes are separated from free compounds and dissociated to release bound compounds. The released compounds are candidate members of the bioactive component set.

Pathological cells are used as examples to introduce this strategy. Er-Xian decotion, a TCM prescription composed of 6 herbs, is used to treat aging diseases such as osteoporosis. H_2O_2-treated osteoblast cells were incubated with Er-Xian decotion for 24 hours. Phosphate buffer was used to wash the cells to remove

unbound compounds. Followed by acetonitrile-assisted ultrasonication, compounds were released from the compound-cell complexes and analyzed by HPLC-Q-TOF/MS. As a result, forty compounds in Er-Xian decoction could bind to osteoblast cells. Network pharmacology was further applied to exclude non-specific binding, and 13 components were found to act on targets associated with osteoporosis. Finally, the set of 13 components showed comparable anti-osteoporosis activity to Er-Xian decoction in the models of ovariectomized rats, prednisolone treated-zebrafish and H_2O_2-treated osteoblasts [lxxxix]. In a similar study, three compounds in turmeric were found as ligands to A549 human lung cancer cells, and their combination showed comparable activity to the turmeric extract [xc].

Bioinformatics The development of bioinformatics has made it possible to predict the bioactive components of TCM. Even if some TCM has not been studied, the biological information of its chemical compositions may have been included in various databases. For example, Zhang *et al*. used ChEMBL database to predict the bioactive components in Qishen Yiqi dripping pills, which is a TCM prescription composed of 4 herbs for the treatment of cardiovascular diseases. Based on the biological information related to cardiovascular diseases such as anti-platelet aggregation and anti-atherosclerosis, a total of 24 components were selected as potential bioactive components. Since Qishen Yiqi dripping pills contains many types of compounds, a two-dimensional chromatographic separation system composed of XAmide column and C18 column was used to prepare the combination of 24 components. Finally, the combination showed a similar effect to Qishen Yiqi dripping pills on animal and cell models of heart failure [xci]. In another example, Yang *et al*. proposed Adjusted Efficacy Score, a parameter based on component contents and multiple bioactivities, to find effective compounds in Xuesaitong injection. With this parameter, five saponins were considered as members of the bioactive component set. In the rat model of myocardial infarction, the set showed comparable activity to Xuesaitong injection [xcii]. The above examples demonstrated that bioinformatics can provide guidance for the discovery of bioactive components from TCM. Considering its simplicity and convenience, it will play an increasingly important role in this field.

Chromatogram-bioctivity correlation Chromatogram-bioactivity correlation is a method to discover bioactive compounds from a holistic perspective. The principle is to find chromatographic peaks closely related to bioactivity through chromatogram change. Li *et al*. developed a standard addition method to find anti-inflammatory components from Heishunpian. Four compounds of hypaconitine, deoxyaconitine, chasmanine, and mesaconitine were added to the Heishunpian extract separately. The anti-inflammatory activity of the changed extract was tested on a mouse inflammation model. As a result, the first 3 compounds are positively correlated with anti-inflammatory activity, while the last compound is negatively correlated with it. Interestingly, chasmanine alone did not show strong anti-inflammatory activity, but addition of chasmanine to the Heishunpian extract caused an increase of the activity. Mesaconitine alone has anti-inflammatory activity, but addition of mesaconitine to the Heishunpian extract resulted in a decrease of the activity. This work shows that the activity of compounds may be affected by the complex chemical environment of a TCM extract [xciii]. Therefore, it is necessary to explore bioactive components of TCM from a holistic perspective.

In addition to the standard addition method, multivariate statistical analysis is often used to establish the correlation between chromatogram and bioactivity. Song *et al*. used high-performance liquid chromatography and label-free receptor biosensors to obtain chromatogram and muscarinic acetylcholine receptor activity of 20 batches of *Daturae Flos*, respectively. Pearson correlation coefficient and hierarchical cluster analysis were applied to establish the correlation between chromatogram and bioactivity. As a result, two chromatographic peaks of scopolamine and scopolamine showed an obvious positive correlation with the bioactivity. They are demonstrated to be the main bioactive components of *Daturae Flos* [xciv].

4.2.2 Knockout/knockin and recombination of components

A bioactive component set has qualitative, quantitative and bioactive characteristics. Therefore, a clear chemical composition of TCM is the prerequisite for discovering the bioactive component set. So far, the qualitative and quantitative analysis of TCM components has not been a problem. Based on the known chemical

composition, there are two methods for the preparation of bioactive component sets, namely knockout/knockin and recombination.

From the perspective of set, the components in TCM are a set, the bioactive components are a subset, and each component in the subset is an element. The technical route of knockout/knockin is to prepare a bioactive component group from TCM, that is, from set (TCM) to subset (the bioactive component group). The technical route of recombination is to form a bioactive component group according to proportion and content of each compound, that is, from element (each compound) to subset (the bioactive component group). This is the difference in technical routes between the two methods. In the above examples [87-94], knockout/knockin was used for Er-Xian decotion, turmeric, Qishen Yiqi dripping pills, and Heishunpian; recombination was used for tanshinone extract, Yindan Xinnaotong soft capsule, Xuesaitong injection, and *Daturae Flos*.

4.2.3 Ranking of bioactivity contribution of compounds

Ranking compounds according to their bioactivity is of great significance for understanding the main medicinal substances of TCM. To achieve this goal, chromatographic peaks corresponding to compounds are knocked out one by one from the original TCM extract. It should be emphasized that the bioactivities of compounds are not directly compared. From the perspective of set theory, bioactivity loss caused by the removal of each compound from the set is the criterion for evaluating its contribution to the overall activity. If the bioactivity of the original extract drops significantly after removing a peak, the bioactivity contribution of the peak is great. If the bioactivity of the original extract hardly changes after removing a peak, the bioactivity contribution of the peak is small. For example, the antioxidant components in longan seed were ranked by Chen *et al*. After finding the active fraction, 8 peaks were removed from the fraction one by one, and the remaining parts were assayed by DPPH and ORAC assays. The result showed that some phenols and p-coumaric acid-glycoside were the top compounds in the order of antioxidant activity. They are the key antioxidant components of longan seed [xcv]. Similarly, Qu *et al*. removed 11 components from safflower one by one. Their bioactivity contributions

were evaluated on models of antiplatelet aggregation, antioxidation and antiproliferation. Some flavonoid glycosides were the top components in the bioactivity ranking of safflower [xcvi]. In addition to component removal, component addition is sometimes used to confirm the conclusion. Yan *et al.* removed alkaloids from the extract of *Coptidis Rhizoma*, and their inhibitory activity on *Shigella dysenteriae* was tested. The results showed the antibacterial activity of the extract was significantly weakened after the removal of berberine and coptisine. When berberine and coptisin were added back to the sample, the activity of the extract was significantly enhanced. Through the removal and addition of components, berberine and coptisin were determined to be the main antibacterial components of *Coptidis Rhizoma* [xcvii]. In summary, bioactivity ranking helps to reveal the most important medicinal substances in TCM.

4.2.4 Exploration of the interaction among compounds

There are three main interactions between two compounds, including addition, antagonism, and synergy, which are usually represented by 1+1=2, 1+1<2, and 1+1>2, respectively. A TCM extract may contain hundreds of compounds, and the interaction network among them is very complex. The following examples are used to illustrate the different interactions in TCM.

Synergy The synergistic effect is considered to be one of the advantages of TCM, and the reports on synergy are the most among the three interactions. Chromatogram-bioactivity correlation was used by Shi *et al.* to discover compounds with acetylcholinesterase inhibitory activity in *Lycoridis Radiatae Bulbus*. Multiple linear regression and Back propagation-artificial neural network were used to establish the correlation between components and bioactivity. By a classic method of combination index, ungerimine and galanthamine were confirmed as a synergistic combination in *Lycoridis Radiatae Bulbus* for acetylcholinesterase inhibition [xcviii]. In a similar way, Martviset *et al.* found that atractylodin, β-eudesmol and hinesol in *Atractylodes Lancea* could synergistically inhibit cholangiocarcinoma cells [xcix], and Gong *et al.* found that iridoid glycosides and phenylpropanoid glycosides in *Radix Scrophulariae* had synergistic immunomodulatory and antioxidant effects [c].

Compared with a single herb, a TCM prescription may possess more synergistic effects because it usually contains more compounds. Evodiamine and berberine are two representative components in a TCM prescription of Zuojinwan. Guan *et al.* found that the two components showed synergistic anti-tumor activity against human colorectal cancer Caco-2 cells [ci].

Addition *Long et al.* explored the anti-inflammatory set and element-element interactions in Cardiotonic Pill on the model of LPS-stimulated RAW264.7 macrophages. Taking NO and IL-6 as indicators, a set of 6 components including 4 phenols and 2 diterpene quinones was found to be a dominant activity contributor for the anti-inflammatory effect of Cardiotonic Pill. The combination index shows that the interaction between phenols and diterpene quinones is mainly addition [cii]. A method of post-column on-flow biochemical detection was established by Wu *et al.* to discover peroxidase inhibitors from *Salvia Miltiorrhizae Radix* and evaluate interactions between bioactive components. After chromatographic separation, the components in TCM enter Q-TOF-MS for structural analysis and on-line biochemical reaction system for product detection. If a compound has peroxidase-inhibitory activity, an inverted peak will appear at the corresponding retention time due to the reduction of product. With this on-line method, 9 peroxidase inhibitors were found from the extract of *Salvia Miltiorrhizae Radix*. This method was also used to detect interactions between bioactive compounds. The results show that the interaction between lithospermic acid and tanshinol is addition. Besides, the interaction between salvianolic acid B and tanshinol is also addition. The addition may be related to the common structural units of these compounds [ciii].

Antagonism The reports on antagonism of TCM components are the least, which may be related to the fact that antagonism is considered as a negative result. A strategy for researching multi-component interactions of TCM was proposed by Song *et al*. First, the compounds in TCM were divided into different families according to their chemical structures. Then, inter-family interactions and inner-family interactions were evaluated by combination index. Dan-Qi pair, which consists of *Radix Salvia miltiorrhiza* and *Radix Panax notoginseng*, was used as an example to illustrate this

strategy. The compounds of Dan-Qi pair were divided into three chemical families of phenols, saponins, and tanshinones. Interestingly, tanshinones inhibited the activity of thrombin, while saponins promoted the activity of thrombin. When a tanshinone and a saponin were mixed, the saponin antagonized or even reversed the inhibition of thrombin caused by the tanshinone, indicating that the interaction between the two chemical families were antagonism. The interactions within the tanshinone family were studied by combination index, and the results showed an interaction of addition. Through structural analysis and molecular docking, the reason for the antagonism between saponins and tanshinones was that tanshinones and the hydrophobic part of saponins are similar in shape, and they were in a competitive relationship when binding to the active site of thrombin. When the hydrophobic part of saponins binds to the active site, the hydrophilic part of saponins will repel tanshinones. The addition effect within tanshinones may be due to their similar structures [civ]. It is worth noting that additive, synergistic, and antagonistic effects may exist in the same herb or herbal prescription. *Cheung et al.* studied the anti-atherogenic effect of an herbal prescription composed of *Pueraria lobata* and *Salvia miltiorrhiza*. The interactions of the two herbs were discovered as antagonism, synergy, and addition on inhibiting inflammation, foam cell formation and vascular smooth muscle cell proliferation, respectively. It can be speculated that the interactions of chemical components in the two herbs will be more complicated [cv].

4.3 Difficulty in discovering bioactive components: integration of different bioactive component sets

TCM may have multiple bioactive component sets, which are closely related to the complexity of models for bioactivity evaluation. For example, an *in vitro* enzymatic reaction is simple, while an *in vivo* animal model is complex. Generally, a simple model may correspond to one class of compounds, while a complex model may correspond to multiple classes of compounds. In other words, when the models for bioactivity evaluation are different, the bioactive component sets are also different. For instance, Dan-Qi formula is a TCM prescription for the treatment of coronary heart diseases, of which the main components are composed of phenols, tanshinones

and saponins. The models of DPPH free radical, superoxide anion, and hydroxyl free radical were used to evaluate antioxidant activity. The result showed that 10 phenols in Dan-Qi formula were the bioactive component set for scavenging free radicals, while tanshinones and saponins did not possess this bioactivity [cvi]. However, in complex models of endothelial cell injury, macrophage inflammation, cardiomyocyte injury and rat myocardial infarction, the bioactive component set consists of 10 phenols, 4 tanshinones and 4 saponins. It is worth noting that any one of the three classes of components is indispensable [cvii]. It can be seen that complex models require the participation of more classes of compounds in TCM.

A similar conclusion comes from the comparison of another pair of studies. *Ginkgo Folium* is a traditional medicine used to treat cerebrovascular diseases. In recent years, xanthine oxidase has been considered as a therapeutic target for cerebrovascular diseases. Twenty-six flavonoids in *Ginkgo Folium* were found to be the bioactive component set that inhibits xanthine oxidase [cviii]. In a complex model of middle cerebral artery occlusion (MCAO)-treated mice, the bioactive component set includes flavonoids, lactones and organic acids [cix]. By comparing the above two studies, it can be found that the bioactive component sets of the same TCM on different models may be different. Theoretically, the bioactive component set on a complex model is the union of bioactive component sets on simple models. However, due to the lack of some models, the bioactive components of TCM on different targets are still not fully understood.

5. The significance for TCM researches from the perspective of set theory

5.1 Simplify complex TCM researches

Although TCM contain numerous chemical components, most of them are analogs with similar structures. Classification of chemical components into different sets based on structural similarity is an effective means to simplify TCM researches. The simplicity is reflected in the following aspects. (1) Compounds can be quickly analyzed based on the common MS cleavage rules of analogs; (2) Metabolites can be quickly identified based on the common metabolic rules of analogs; (3) The

bioactivity of compounds can be predicted by structure-activity relationship of analogs. The action mode of TCM is considered to be "multiple components acting on multiple targets". From the perspective of set theory, this mode can be simplified to "several sets acting on several targets". Compared with TCM researches at the compound level, the workload and complexity of researches at the set level will be greatly reduced. In addition, the concepts of element, subset, set, intersection and union in set theory can be used to vividly describe complex issues and strategies in TCM researches. All of the above help to simplify TCM researches.

5.2 Integrate isolated TCM researches

For a long time, there is a lack of systematic strategies for researching components, metabolites and bioactive components of TCM. To some extent, the three parts are not closely related, which is not conducive to revealing the efficacy of TCM. In this review, a methodology based on set theory was proposed for the systematic research of TCM components, metabolites and bioactive components. Technical routes based on set theory were used to connect the three parts: (1) the route for components is from training set to testing set and from set to subset; (2) the route for metabolites is from element to set and from prototype set to metabolite set; (3) the route for bioactive components is from bioactive component set to element-element interaction. In short, set theory not only provides overall guidance and strategic support for TCM research, but also helps explain the effectiveness of TCM from a holistic perspective.

5.3 Build a bridge between traditional medicines and modern medicines

A modern medicine is usually a single compound. In many cases, a modern medicine works through its metabolites, and the metabolites produced by this medicine are a series of analogs [cx,cxi]. For a modern medicine, a single compound enters the body and produces multiple metabolites, and the change is from an element to a set. For a traditional medicine, a set enters the body and produces multiple metabolites, and the change is from a set to a set. In other words, traditional medicines and modern medicines are different in the initial forms, and their final effective forms are the same. Therefore, set theory can build a bridge between them, which helps to

understand the effectiveness of TCM.

6. Conclusions and future perspectives

This article reviews the research strategies of TCM in the past decade from the perspective of set theory. A systematic methodology was proposed to integrate and simplify the complex researches on components, metabolites and bioactive components of TCM. The technical route, key steps and difficulty of each strategy were summarized and introduced. A series of representative examples were provided to fully explain each strategy. We believe that these universal strategies will play an increasingly important role in simplifying complex TCM researches in the next decade.

The possible key issues in future TCM researches are introduced in a form of question and answer as follows. (1) How to discover compounds with multiple bioactivities or selective bioactivity to meet the different drug development needs? From the perspective of set theory, the concepts of intersection and complementary set may provide guidance for discovering the two types of bioactive compounds. (2) How to find bioactive components from metabolites? Development of highly sensitive methods for bioactivity screening or highly efficient methods for metabolite purification may help solve this problem. In conclusion, the discovery of various lead compounds from TCM for modern drug development will continue to be the focus of future research.

Declaration of competing interests

The authors declare that they have no competing interests.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (81803724), Liaoning Natural Science Foundation (2020-BS-167), Liaoning Provincial Education Department project (L202010), Jiangsu Natural Science Foundation (BK20180550), and High-level Talent Research Startup (Recruitment and Employment) Foundation of Liaoning University of

Traditional Chinese Medicine. The authors thank Enos Mais and Sakib Moinuddin for their advice on English.

Credit Author Statement

Hui-Peng Song and **Gui-Zhong Xin:** Conceptualization, Supervision, Writing-Reviewing and Editing, Project administration, Funding acquisition. **Yue-Hua Chen** and **Jing-Hua Bi**: Writing-Original draft preparation. **Ming Xie** and **Hui Zhang**: Writing-Reviewing and Editing. **Zi-Qi Shi** and **Hua Guo**: Visualization. **Hai-Bo Yin** and **Jia-Nuo Zhang**: Validation.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] B. He, C. Lu, M.L. Wang, G. Zheng, G. Chen, M. Jiang, X.J. He, Z.X. Bian, G. Zhang, A.P. Lu, Drug discovery in traditional Chinese medicine: from herbal fufang to combinatory drugs, Science 350 (2015) S74-S76.

[1] Y. Han, H. Sun, A. Zhang, G. Yan, X.J. Wang, Chinmedomics, a new strategy for evaluating the therapeutic efficacy of herbal medicines, Pharmacol. Ther. 216 (2020) 107680.

[1] K. Murata, D. Iida, Y. Ueno, K. Samukawa, T. Ishizaka, T. Kotake, H. Matsuda, Novel polyacetylene derivatives and their inhibitory activities on acetylcholinesterase obtained from Panax ginseng roots, J. Nat. Med. 71 (2017) 114 -122.

[1] Y. Tu, Artemisinin-a gift from traditional Chinese medicine to the world (Nobel lecture), Angew. Chem. Int. Edit. 55 (2016) 10210-10226.

[1] E.L.H. Leung, S.W. Xu, Traditional Chinese medicine in cardiovascular drug discovery, Pharmacol. Res. 160 (2020) 105168.

[1] X. Ma, B. Lv, P. Li, X. Jiang, Q. Zhou, X. Wang, X. Gao, Identification of "multiple components-multiple targets-multiple pathways" associated with Naoxintong capsule in the treatment of heart diseases using UPLC/Q-TOF-MS and network pharmacology, Evid.-Based Compl. Alt. Med. 2016 (2016) 9468087.

[1] W.J. Zhang, Y. Huai, Z.P. Miao, A.R. Qian, Y.H. Wang, Systems pharmacology for

investigation of the mechanisms of action of traditional Chinese medicine in drug discovery, Front. Pharmacol. 10 (2019) 743.

[1] S. Singh, D.B. Singh, S. Singh, R. Shukla, P.W. Ramteke, K. Misra, Exploring medicinal plant legacy for drug discovery in post-genomic era, Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci. 89 (2019) 1141-1151.

[1] A. Stavrianidi, A classification of liquid chromatography mass spectrometry techniques for evaluation of chemical composition and quality control of traditional medicines, J. Chromatogr. A 1609 (2020) 460501.

[1] B.Y. Pang, Y. Zhu, L.Q. Lu, F.B. Gu, H.L. Chen, The applications and features of liquid chromatography-mass spectrometry in the analysis of traditional Chinese medicine, Evid.-Based Compl. Alt. Med. 2016 (2016) 3837270.

[1] S. Ji, S. Wang, H. Xu, Z. Su, D. Tang, X. Qiao, M. Ye, The application of on-line two-dimensional liquid chromatography (2DLC) in the chemical analysis of herbal medicines, J. Pharm. Biomed. Anal. 160 (2018) 301-313.

[1] A. Zhang, H. Sun, X. Wang, Mass spectrometry-driven drug discovery for development of herbal medicine, Mass Spectrometry Reviews 37 (2018) 307-320.

[1] Z. Zhang, T. Bo, Y. Bai, M. Ye, R. An, F. Cheng, H. Liu, Quadrupole time-of-flight mass spectrometry as a powerful tool for demystifying traditional Chinese medicine, Trac-Trend. Anal. Chem. 72 (2015) 169 -180.

[1] H. Lei, Y. Zhang, X. Zu, J. Ye, Y. Liang, T. Cheng, W. Zhang, Comprehensive profiling of the chemical components and potential markers in raw and processed Cistanche tubulosa by combining ultra-high-performance liquid chromatography coupled with tandem mass spectrometry and MS/MS-based molecular networking, Anal. Bioanal. Chem. 413 (2020) 129-139.

[1] T. Xu, S. Li, Y. Sun, Z. Pi, S. Liu, F. Song, Z. Liu, Systematically characterize the absorbed effective substances of Wutou Decoction and their metabolic pathways in rat plasma using UHPLC-Q-TOF-MS combined with a target network pharmacological analysis, J. Pharm. Biomed. Anal. 141 (2017) 95-107.

[1] J. Fang, L. Wang, T. Wu, C. Yang, L. Gao, H. Cai, J. Lin, S. Fang, Y. Chen, W. Tan, Q. Wang, Network pharmacology-based study on the mechanism of action for herbal medicines in Alzheimer treatment, J. Ethnopharmacol. 196 (2017) 281-292.

[1] Y. Song, Q. Song, J. Li, N. Zhang, Y. Zhao, X. Liu, Y. Jiang, P. Tu, An integrated strategy to quantitatively differentiate chemome between Cistanche deserticola and C-tubulosa using high performance liquid chromatography-hybrid triple quadrupole-linear ion trap mass spectrometry, J. Chromatogr. A 1429 (2016) 238-247.

[1] G. Cao, N. Wang, D. He, X. Wang, Y. Tian, N. Wan, W. Yan, H. Ye, H. Hao, Intestinal mucosal metabolites-guided detection of trace-level ginkgo biloba extract metabolome, J. Chromatogr. A 1608 (2019) 460417.

[1] R.T. Cook, The foundations of mathematics in the theory of sets, Brit. J. Philos. Sci. 54 (2003) 347-352.

[1] Z.M. Jiang, L.J. Wang, H.Q. Pang, Y. Guo, P.T. Xiao, C. Chu, L. Guo, E.H. Liu, Rapid profiling of alkaloid analogues in Sinomenii Caulis by an integrated characterization strategy and quantitative analysis, J. Pharm. Biomed. Anal. 174 (2019) 376-385.

[1] D. Li, O.J. Schmitz, Comprehensive two-dimensional liquid chromatography tandem diode array detector (DAD) and accurate mass QTOF-MS for the analysis of flavonoids and iridoid glycosides in Hedyotis diffusa, Anal. Bioanal. Chem. 407 (2015) 231-240.

[1] B.K. Shin, S.W. Kwon, J.H. Park, Chemical diversity of ginseng saponins from Panax ginseng, J. Ginseng Res. 39 (2015) 287-298.

[1] L.W. Qi, C.Z. Wang, C.S. Yuan, Isolation and analysis of ginseng: advances and challenges, Nat. Prod. Rep. 28 (2011) 467-495.

[1] L.F. Hu, Z.H. Yao, Z.F. Qin, L.Y. Liu, X.J. Song, Y. Dai, H. Kiyohara, H. Yamada, X.S. Yao, In vivo metabolic profiles of Bu-Zhong-Yi-Qi-Tang, a famous traditional Chinese medicine prescription, in rats by ultra-high-performance liquid chromatography coupled with quadrupole time-of-flight tandem mass spectrometry, J. Pharm. Biomed. Anal. 171 (2019) 81-98.

[1] T. Wang, X.C. Feng, L.Q. Ding, K. Wang, M. Qiao, L.W. Chai, Y. Li, F. Qiu, Metabolic profiling of icariin in rat feces, urine, bile and plasma after oral administration using ultra-high performance liquid chromatography/quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 168 (2019) 155-162.

[1] L.J. Wang, Z.M. Jiang, P.T. Xiao, J.B. Sun, Z.M. Bi, E.H. Liu, Identification of anti-inflammatory components in Sinomenii Caulis based on spectrum-effect relationship and chemometric methods, J. Pharm. Biomed. Anal. 167 (2019)

38-48.

[1] X. Xu, W. Li, T. Li, K. Zhang, Q. Song, L. Liu, P. Tu, Y. Wang, Y. Song, J. Li, Direct infusion-three-dimensional-mass spectrometry enables rapid chemome comparison among herbal medicines, Anal. Chem. 92 (2020) 7646-7656.

[1] H.S. Fan, S.J. Liu, W.X. Shen, A. Kang, J.N. Tan, L. Li, X. Liu, C.L. Xu, X.F. Xu, Y.Y. Lai, H.B. Cheng, D.D. Sun, Identification of the absorbed components and metabolites of Xiao-Ai-Jie-Du decoction and their distribution in rats using ultra high-performance liquid chromatography/quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 179 (2020) 112984.

[1] M.L. Chen, W.Q. Chang, J.L. Zhou, Y.H. Yin, W.R. Xia, J.Q. Liu, L.F. Liu, G.Z. Xin, Comparison of three officinal species of Callicarpa based on a biochemome profiling strategy with UHPLC-IT-MS and chemometrics analysis, J. Pharm. Biomed. Anal. 145 (2017) 666 -674.

[1] M. Liao, A. Li, C. Chen, H. Ouyang, Y. Zhang, Y. Xu, Y. Feng, H. Jiang, Systematic identification of shikonins and shikonofurans in medicinal Zicao species using ultra-high performance liquid chromatography quadrupole time of flight tandem mass spectrometry combined with a data mining strategy, J. Chromatogr. A 1425 (2015) 158-172.

[1] Z. Xue, C. Lai, L. Kang, A. Kotani, H. Hakamata, Z. Jing, H. Li, W. Wang, B. Yang, Profiling and isomer recognition of phenylethanoid glycosides from Magnolia officinalis based on diagnostic/holistic fragment ions analysis coupled with chemometrics, J. Chromatogr. A 1611 (2020) 460583.

[1] X.Y. Zhu, J.Z. Liu, Z.H. Dong, F. Feng, W.Y. Liu, Identification and screening of cardiac glycosides in Streptocaulon griffithii using an integrated data mining strategy based on high resolution mass spectrometry, Chin. J. Nat. Medicines 16 (2018) 546-560.

[1] X.L. Cheng, J.Y. Wan, P. Li, L.W. Qi, Ultrasonic/microwave assisted extraction and diagnostic ion filtering strategy by liquid chromatography-quadrupole time-of-flight mass spectrometry for rapid characterization of flavonoids in Spatholobus suberectus, J. Chromatogr. A 1218 (2011) 5774-5786.

[1] X. Qiao, R. Li, W. Song, W.J. Miao, J. Liu, H.B. Chen, D.A. Guo, M. Ye, A targeted strategy to analyze untargeted mass spectral data: rapid chemical profiling of Scutellaria baicalensis using ultra-high performance liquid

chromatography coupled with hybrid quadrupole orbitrap mass spectrometry and key ion filtering, J. Chromatogr. A 1441 (2016) 83-95.

[1] J.Y. Zhang, J.Q. Lu, F. Wang, G.S. Cao, P.F. Tu, Y.J. Qiao, A strategy of EIC-MS coupled with diagnostic product ions analysis for efficient discovery of new hydroxylated polymethoxyflavonoid glycosides from the leaves of Murraya paniculata L. using HPLC-DAD-MS/MS, Anal. Methods 5 (2013) 2880-2891.

[1] J.Y. Zhang, F. Wang, H. Zhang, J.Q. Lu, Y.J. Qiao, Rapid identification of polymethoxylated flavonoids in traditional Chinese medicines with a practical strategy of stepwise mass defect filtering coupled to diagnostic product ions analysis based on a hybrid LTQ-Orbitrap mass spectrometer, Phytochem. Analy. 25 (2014) 405-414.

[1] H. Pan, W. Yang, Y. Zhang, M. Yang, R. Feng, W. Wu, D. Guo, An integrated strategy for the systematic characterization and discovery of new indole alkaloids from Uncaria rhynchophylla by UHPLC/DAD/LTQ-Orbitrap-MS, Anal. Bioanal. Chem. 407 (2015) 6057-6070.

[1] J. Chen, Q. Zhao, D. Si, A. Nie, Y. Wang, Z. Deng, Y. Wen, F. Chen, L. Zhang, B. Dong, J. Yang, Comprehensive profiling of Stephania tetrandra (Fangji) by stepwise DFI and NL-dependent structure annotation algorithm-based UHPLC-Q-TOF-MS and direct authentication by LMJ-HRMS, J. Pharm. Biomed. Anal. 185 (2020) 113225.

[1] Y.Y. Shi, R.Q. Wang, X.Y. Zhu, D.G. Xu, W.Y. Liu, F. Feng, A self-feedback network based on liquid chromatography-quadrupole-time of flight mass spectrometry for system identification of beta-carboline alkaloids in Picrasma quassioides, Sci. Rep. 7 (2017) 13841.

[1] J.Y. He, Y. Liu, Y. Kang, P. Yang, Y.Q. Wang, J.X. Guo, J.M. Huang, Identification of alkaloids in Stephania hainanensis by liquid chromatography coupled with quadrupole time-of-flight mass spectrometry, Phytochem. Analy. 27 (2016) 206-216.

[1] J.X. Tian, Y. Tian, L. Xu, J. Zhang, T. Lu, Z.J. Zhang, Characterisation and identification of dihydroindole-type alkaloids from processed Semen Strychni by high-performance liquid chromatography coupled with electrospray ionisation ion trap time-of-flight mass spectrometry, Phytochem. Analy. 25 (2014) 36-44.

[1] E.K. Jeong, S.Y. Lee, S.M. Yu, N.H. Park, H.S. Lee, Y.H. Yim, G.S. Hwang, C. Cheong, J.H. Jung, J. Hong, Identification of structurally diverse alkaloids in

Corydalis species by liquid chromatography/electrospray ionization tandem mass spectrometry, Rapid Commun. Mass Sp. 26 (2012) 1661-1674.

[1] Y.P. Jiang, R. Liu, M. Liu, L.Z. Yi, S. Liu, An integrated strategy to rapidly characterize non-targeted benzylisoquinoline alkaloids from Plumula nelumbinis ethanol extract using UHPLC/Q-orbitrap HRMS, Int. J. Mass Spectrom. 432 (2018) 26-35.

[1] L.W. Qi, H.Y. Wang, H. Zhang, C.Z. Wang, P. Li, C.S. Yuan, Diagnostic ion filtering to characterize ginseng saponins by rapid liquid chromatography with time-of-flight mass spectrometry, J. Chromatogr. A 1230 (2012) 93-99.

[1] B.L. Ma, S.L. Yang, J.M. Li, H. Ouyang, M.Z. He, Y.L. Feng, T. Tan, A four-step filtering strategy based on ultra-high-performance liquid chromatography coupled to quadrupole-time-of-flight tandem mass spectrometry for comprehensive profiling the major chemical constituents of Akebiae Fructus, Rapid Commun. Mass Sp. 33 (2019) 1464-1474.

[1] T. Xie, Y. Liang, H.P. Hao, J.A, L. Xie, P. Gong, C. Dai, L.S. Liu, A. Kang, X. Zheng, G.J. Wang, Rapid identification of ophiopogonins and ophiopogonones in Ophiopogon japonicus extract with a practical technique of mass defect filtering based on high resolution mass spectrometry, J. Chromatogr. A 1227 (2012) 234-244.

[1] C. Yao, W. Yang, W. Si, H. Pan, S. Qiu, J. Wu, X. Shi, R. Feng, W. Wu, D. Guo, A strategy for establishment of practical identification methods for Chinese patent medicine from systematic multi-component characterization to selective ion monitoring of chemical markers: Shuxiong tablet as a case study, Rsc Adv. 6 (2016) 65055-65066.

[1] C. Wang, J. Zhang, C. Wu, Z. Wang, A multiple-dimension liquid chromatography coupled with mass spectrometry data strategy for the rapid discovery and identification of unknown compounds from a Chinese herbal formula (Er-xian decoction), J. Chromatogr. A 1518 (2017) 59-69.

[1] T. Xu, M. Yang, Y. Li, X. Chen, Q. Wang, W. Deng, X. Pang, K. Yu, B. Jiang, S. Guan, D.A. Guo, An integrated exact mass spectrometric strategy for comprehensive and rapid characterization of phenolic compounds in licorice, Rapid Commun. Mass Sp. 27 (2013) 2297-2309.

[1] C. Li, J. Yang, X. Tong, C. Zhao, Y. He, H. Wan, Precursor ion scan enhanced rapid identification of the chemical constituents of Danhong injection by liquid

chromatography-tandem mass spectrometry: an integrated strategy, J. Chromatogr. A 1602 (2019) 378-385.

[1] X. Ma, X. Guo, Y. Song, L. Qiao, W. Wang, M. Zhao, P. Tu, Y. Jiang, An integrated strategy for global qualitative and quantitative profiling of traditional Chinese medicine formulas: Baoyuan decoction as a case, Sci. Rep. 6 (2016) 38379.

[1] S. Li, S. Jin, C. Song, S. Jia, Y. Zhang, Y. Feng, Z. Du, H. Jiang, The strategy for establishment of the multiple reaction monitoring based characteristic chemical profile of triterpenes in Alismatis rhizoma using two combined tandem mass spectrometers, J. Chromatogr. A 1524 (2017) 121-134.

[1] W.Z. Yang, M. Ye, X. Qiao, C.F. Liu, W.J. Miao, T. Bo, H.Y. Tao, D.A. Guo, A strategy for efficient discovery of new natural compounds by integrating orthogonal column chromatography and liquid chromatography/mass spectrometry analysis: Its application in Panax ginseng, Panax quinquefolium and Panax notoginseng to characterize 437 potential new ginsenosides, Anal. Chim. Acta 739 (2012) 56-66.

[1] H. Pan, C. Yao, W. Yang, S. Yao, Y. Huang, Y. Zhang, W. Wu, D. Guo, An enhanced strategy integrating offline two-dimensional separation and step-wise precursor ion list-based raster-mass defect filter: Characterization of indole alkaloids in five botanical origins of Uncariae Ramulus Cum Unicis as an exemplary application, J. Chromatogr. A 1563 (2018) 124-134.

[1] L. Wu, P. Gong, Y. Wu, K. Liao, H. Shen, Q. Qi, H. Liu, G. Wang, H. Hao, An integral strategy toward the rapid identification of analogous nontarget compounds from complex mixtures, J. Chromatogr. A 1303 (2013) 39-47.

[1] Q. Zhang, M. Huo, Y. Zhang, Y. Qiao, X. Gao, A strategy to improve the identification reliability of the chemical constituents by high-resolution mass spectrometry-based isomer structure prediction combined with a quantitative structure retention relationship analysis: phthalide compounds in Chuanxiong as a test case, J. Chromatogr. A 1552 (2018) 17-28.

[1] G.F. Feng, Y. Zheng, Y. Sun, S. Liu, Z.F. Pi, F.R. Song, Z.Q. Liu, A targeted strategy for analyzing untargeted mass spectral data to identify lanostane-type triterpene acids in Poria cocos by integrating a scientific information system and liquid chromatography-tandem mass spectrometry combined with ion mobility

spectrometry, Anal. Chim. Acta 1033 (2018) 87-99.

[1] Y. Liu, Z. Liu, C. Lu, J. Li, Z. Ning, Z. Song, C. Wang, Z. Du, X. Lu, S. Zhao, A. Lu, Comprehensive identification of active triterpenoid metabolites in frankincense using a coupling strategy, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 963 (2014) 90-98.

[1] C.L. Yao, W.Z. Yang, W. Si, Y. Shen, N.X. Zhang, H.L. Chen, H.Q. Pan, M. Yang, W.Y. Wu, D.A. Guo, An enhanced targeted identification strategy for the selective identification of flavonoid O-glycosides from Carthamus tinctorius by integrating offline two-dimensional liquid chromatography/linear ion-trap-Orbitrap mass spectrometry, high-resolution diagnostic product ions/neutral loss filtering and liquid chromatography-solid phase extraction-nuclear magnetic resonance, J. Chromatogr. A 1491 (2017) 87-97.

[1] X. Wan, J. Liu, X. Yang, Q. Zhang, Y. Zhang, Q. Li, K. Bi, Development of a systematic strategy for the global identification and classification of the chemical constituents and metabolites of Kai-Xin-San based on liquid chromatography with quadrupole time-of-flight mass spectrometry combined with multiple data-processing approaches, J. Sep. Sci. 41 (2018) 2672-2680.

[1] G. Feng, Z. Liu, S. Liu, J. Xing, F. Song, Z. Pi, A target integration strategy for analyzing multidimensional chemical and metabolic substance groups of Ding-Zhi-Xiao-Wan prescription by using ultra-high performance liquid chromatography tandem mass spectrometry, J. Chromatogr. A 1608 (2019) 460412.

[1] C. Xiang, X. Qiao, Q. Wang, R. Li, W. Miao, D. Guo, M. Ye, From single compounds to herbal extract: a strategy to systematically characterize the metabolites of licorice in rats, Drug Metab. Dispos. 39 (2011) 1597-1608.

[1] S. Wang, P. Qi, N. Zhou, M. Zhao, W. Ding, S. Li, M. Liu, Q. Wang, S. Jin, A pre-classification strategy based on UPLC-Triple-TOF/MS for metabolic screening and identification of Radix glehniae in rats, Anal. Bioanal. Chem. 408 (2016) 7423-7436.

[1] S. Li, S. Liu, Z. Liu, Z. Pi, F. Song, Y. Jin, A target-group-change strategy based on the UPLC-Q-TOF-MSE method for the metabolites identification of Fufang-Xialian-Capsule in rat's plasma, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 1085 (2018) 42-53.

[1] X. Dong, X. Li, N. Li, H. Zhao, A. GuLa, X. Zhang, P. Zhang, B. Bao, A

target-group-change couple with mass defect filtering strategy to identify the metabolites of "Dogel ebs" in rats plasma, urine and bile, J. Sep. Sci. 42 (2019) 3382-3389.

[1] P. Gong, N. Cui, L. Wu, Y. Liang, K. Hao, X. Xu, W. Tang, G. Wang, H. Hao, Chemicalome and metabolome matching approach to elucidating biological metabolic networks of complex mixtures, Anal. Chem. 84 (2012) 2995-3002.

[1] G. Tan, M. Liu, X. Dong, S. Wu, L. Fan, Y. Qiao, Y. Chai, H. Wu, A strategy for rapid analysis of xenobiotic metabolome of Sini decoction in vivo using ultra-performance liquid chromatography-electrospray ionization quadrupole-time-of-flight mass spectrometry combined with pattern recognition approach, J. Pharm. Biomed. Anal. 96 (2014) 187-196.

[1] K. Luo, F. Feng, Identification of absorbed components and metabolites of Zhi-Zi-Hou-Po decoction in rat plasma after oral administration by an untargeted metabolomics-driven strategy based on LC-MS, Anal. Bioanal. Chem. 408 (2016) 5723-5735.

[1] D. Wang, Q. Shi, F. Feng, Rapid discovery and identification of the prototypes and their metabolites of Da-Huang-Xiao-Shi decoction in rat plasma by an integrative strategy based on liquid chromatography coupled with mass spectrometry, J. Sep. Sci. 40 (2017) 2722-2731.

[1] M. Guo, B. Zhao, H. Liu, L. Zhang, L. Peng, L. Qin, Z. Zhang, J. Li, C. Cai, X. Gao, A metabolomic strategy to screen the prototype components and metabolites of Shuang-Huang-Lian injection in human serum by ultra performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry, J. Anal. Methods Chem. 2014 (2014) 241505.

[1] M.X. Guo, L. Zhang, H.Y. Liu, L.L. Qin, Z.X. Zhang, X. Bai, X.Y. Gao, A metabolomic strategy to screen the prototype components and metabolites of Qingkailing injection in rat urine by high-performance liquid chromatography with tandem mass spectrometry, J. Sep. Sci. 37 (2014) 2844-2850.

[1] Y. Luo, C.J.S. Lai, J. Zhang, Y. Feng, Q. Wen, T. Tan, Comprehensive metabolic profile of phenolic acids and flavonoids in Glechomae Herba using ultra-high-performance liquid chromatography coupled to quadrupole-time-of-flight tandem mass spectrometry with diagnostic ion filtering strategy, J. Pharm. Biomed. Anal. 164 (2019) 615-629.

[1] J. Liang, F. Xu, Y.Z. Zhang, S. Huang, X.Y. Zang, X. Zhao, L. Zhang, M.Y. Shang,

D.H. Yang, X. Wang, S.Q. Cai, The profiling and identification of the absorbed constituents and metabolites of Paeoniae Radix Rubra decoction in rat plasma and urine by the HPLC-DAD-ESI-IT-TOF-MS^n technique: a novel strategy for the systematic screening and identification of absorbed constituents and metabolites from traditional Chinese medicines, J. Pharm. Biomed. Anal. 83 (2013) 108-121.

[1] N. Mi, T. Cheng, H. Li, P. Yang, X. Mu, X. Wang, X. Zu, X. Qi, X. Guo, J. Ye, W. Zhang, Metabolite profiling of traditional Chinese medicine formula Dan Zhi Tablet: An integrated strategy based on UPLC-QTOF/MS combined with multivariate statistical analysis, J. Pharm. Biomed. Anal. 164 (2019) 70-85.

[1] S. Lin, X. Yue, D. Ouyang, Q. Li, P. Yang, The profiling and identification of chemical components, prototypes and metabolites of Run-zao-zhi-yang capsule in rat plasma, urine and bile by an UPLC-Q-TOF/MS^E-based high-throughput strategy, Biomed. Chromatogr. 32 (2018) e4261.

[1] X.Y. Tang, J.X. Zeng, Z.Q. Dai, M.H. Chen, M.N. Ye, Z.H. Yao, Y. Dai, X.S. Yao, Identification and characterization of chemical constituents in Qi-Lin pills and their metabolites in rat bio-samples after oral administration using ultra-high performance liquid chromatography with quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 188 (2020) 113402.

[1] T. Yi, L. Zhu, G.Y. Zhu, Y.N. Tang, J. Xu, J.Y. Fan, Z.Z. Zhao, H.B. Chen, HSCCC-based strategy for preparative separation of in vivo metabolites after administration of an herbal medicine: Saussurea laniceps, a case study, Sci. Rep. 6 (2016) 33036.

[1] Y. Li, X.W. Yang, Five new eudesmane-type sesquiterpenoid lactones biotransformed from atractylenolide I by rat hepatic microsomes, Fitoterapia 85 (2013) 95-100.

[1] J. Tang, Z.T. Wang, T. Akao, M. Hattori, Further Evidence on the intramolecular lactonization in rat liver microsomal metabolism of 12-O-acetylated retronecine-type pyrrolizidine alkaloids, Nat. Prod. Commun. 8 (2013) 1545-1546.

[1] W.J. Wu, N. Hu, Q.W. Zhang, Y.P. Li, P. Li, R. Yan, Y.T. Wang, In vitro glucuronidation of five rhubarb anthraquinones by intestinal and liver microsomes from humans and rats, Chem.-Biol. Interact. 219 (2014) 18-27.

[1] S.J. Liu, G.L. Dai, L.N. Sun, B.T. Sun, D. Chen, L. Zhu, Y. Wang, L. Zhang, P.D.

Chen, D. Zhou, W.Z. Ju, Biotransformation and metabolic profile of limonin in rat liver microsomes, bile, and urine by high-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry, J. Agr. Food Chem. 66 (2018) 10388-10393.

[1] H. Yuan, Q. Ma, L. Ye, G. Piao, The traditional medicine and modern medicine from natural products, Molecules 21 (2016) 559.

[1] H.P. Song, H. Zhang, Y. Fu, H.Y. Mo, M. Zhang, J. Chen, P. Li, Screening for selective inhibitors of xanthine oxidase from Flos Chrysanthemum using ultrafiltration LC-MS combined with enzyme channel blocking, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 961 (2014) 56-61.

[1] H.P. Song, J. Chen, J.Y. Hong, H.P. Hao, L.W. Qi, J. Lu, Y. Fu, B. Wu, H. Yang, P. Li, A strategy for screening of high-quality enzyme inhibitors from herbal medicines based on ultrafiltration LC-MS and in silico molecular docking, Chem. Commun. 51 (2015) 1494-1497.

[1] J.F. Zhang, W.C. Zhong, Y.C. Li, Y.Q. Song, G.Y. Xia, G.H. Tian, G.B. Ge, S. Lin, Bioactivity-guided discovery of human carboxylesterase inhibitors from the roots of Paeonia lactiflora, J. Nat. Prod. 83 (2020) 2940-2949.

[1] Y. Ma, W. Wang, J. Yang, S. Zhang, Z. Li, F. Li, S. Huang, L. Lei, K. Wang, A. Wen, Y. Ding, A network pharmacology technique to investigate the synergistic mechanisms of Salvia miltiorrhiza and Radix puerariae in treatment of cardio-cerebral vascular diseases, Evid.-Based Compl. Alt. Med. 2020 (2020) 6937186.

[1] L. Guo, L. Duan, L.L. Dou, L.L. Liu, H. Yang, E.H. Liu, P. Li, Quality standardization of herbal medicines using effective compounds combination as labeled constituents, J. Pharm. Biomed. Anal. 129 (2016) 320-331.

[1] H.Q. Pang, H.M. An, H. Yang, S.Q. Wu, J.L. Fan, L. Mi, H. Wang, P. Li, W. Gao, Comprehensive chemical profiling of Yindan Xinnaotong soft capsule and its neuroprotective activity evaluation in vitro, J. Chromatogr. A 1601 (2019) 288-299.

[1] N. Wang, P. Xu, X. Wang, W. Ya, Z. Yu, R. Wu, X. Huang, Y. Si, D. Shou, Integrated pathological cell fishing and network pharmacology approach to investigate main active components of Er-Xian decotion for treating osteoporosis, J. Ethnopharmacol. 241 (2019) 111977.

[1] J.L. Zhou, J.Y. Zheng, X.Q. Cheng, G.Z. Xin, S.L. Wang, T. Xie, Chemical

markers' knockout coupled with UHPLC-HRMS-based metabolomics reveals anti-cancer integration effects of the curcuminoids of turmeric (Curcuma longa L.) on lung cancer cell line, J. Pharm. Biomed. Anal. 175 (2019) 112738.

[1] Y. Zhang, J. Yu, W. Zhang, Y. Wang, Y. He, S. Zhou, G. Fan, H. Yang, Y. Zhu, P. Li, An integrated evidence-based targeting strategy for determining combinatorial bioactive ingredients of a compound herbal medicine Qishen Yiqi dripping pills, J. Ethnopharmacol. 219 (2018) 288-298.

[1] Z. Yang, Q. Shao, Z. Ge, N. Ai, X. Zhao, X. Fan, A bioactive chemical markers based strategy for quality assessment of botanical drugs: Xuesaitong injection as a case study, Sci. Rep. 7 (2017) 2410.

[1] Y. Li, Y. Wang, B. Yang, Y. Wang, Z. Hou, A. Li, Y. Xu, L. Ju, H. Wu, Y. Zhang, A practical and novel "standard addition" strategy to screen pharmacodynamic components in traditional Chinese medicine using Heishunpian as an example, Rsc Adv. 5 (2015) 22209-22216.

[1] H.P. Song, H. Wang, X. Zhao, L. He, H. Zhong, S.Q. Wu, P. Li, H. Yang, Label-free pharmacological profiling based on dynamic mass redistribution for characterization and authentication of hazardous natural products, J. Hazard. Mater. 333 (2017) 265-274.

[1] J. Chen, Z.Z. Ge, W. Zhu, Z. Xu, C.M. Li, Screening of key antioxidant compounds of longan (Dimocarpus longan Lour.) seed extract by combining online fishing/knockout, activity evaluation, fourier transform ion cyclotron resonance mass spectrometry, and high-performance liquid chromatography electrospray ionization mass spectrometry methods, J. Agr. Food Chem. 62 (2014) 9744-9750.

[1] C. Qu, L.Y. Wang, H. Lin, E.X. Shang, Y.P. Tang, S.J. Yue, Y. Jin, W.W. Tao, S.P. Li, Y.Q. Hua, P. Liu, S.L. Su, H. Zhou, D.W. Qian, J.A. Duan, Hierarchical identification of bioactive components in a medicinal herb by preparative high-performance liquid chromatography and selective knock-out strategy, J. Pharm. Biomed. Anal. 135 (2017) 206-216.

[1] D. Yan, J. Li, Y. Xiong, C. Zhang, J. Luo, Y. Han, R. Wang, C. Jin, H. Qian, J. Li, L. Qiu, C. Peng, Y. Lin, X. Song, X. Xiao, Promotion of quality standard of herbal medicine by constituent removing and adding, Sci. Rep. 4 (2014) 3668.

[1] Z.Q. Shi, D.F. Song, R.Q. Li, H. Yang, L.W. Qi, G.Z. Xin, D.Q. Wang, H.P. Song,

J. Chen, H. Hao, P. Li, Identification of effective combinatorial markers for quality standardization of herbal medicines, J. Chromatogr. A 1345 (2014) 78-85.

[1] P. Martviset, W. Chaijaroenkul, P. Muhamad, K. Na-Bangchang, Bioactive constituents isolated from Atractylodes lancea (Thunb.) DC. rhizome exhibit synergistic effect against cholangiocarcinoma cell, J. Exp. Pharmacol. 10 (2018) 59-64.

[1] P.Y. Gong, Y.W. He, J. Qi, C.Z. Chai, B.Y. Yu, Synergistic nourishing 'Yin' effect of iridoid and phenylpropanoid glycosides from Radix Scrophulariae in vivo and in vitro, J. Ethnopharmacol. 246 (2020) 112209.

[1] X. Guan, X. Zheng, V. Chi Teng, J. Zhao, J. Xiao, Y. Wang, Z. Zhong, Combined effects of berberine and evodiamine on colorectal cancer cells and cardiomyocytes in vitro, Eur. J. Pharmacol. 875 (2020) 173031.

[1] F. Long, H. Yang, Y. Xu, H. Hao, P. Li, A strategy for the identification of combinatorial bioactive compounds contributing to the holistic effect of herbal medicines, Sci. Rep. 5 (2015) 12361.

[1] S.Q. Wu, H.P. Song, B. Li, R.Z. Liu, H. Yang, L. He, P. Li, A fast and accurate method for the identification of peroxidase inhibitors from Radix Salvia Miltiorrhizae by on-flow biochemical assay coupled with LC/Q-TOF-MS: comparison with ultrafiltration-based affinity selection, Anal. Bioanal. Chem. 410 (2018) 4311-4322.

[1] H.P. Song, S.Q. Wu, H. Hao, J. Chen, J. Lu, X. Xu, P. Li, H. Yang, A chemical family-based strategy for uncovering hidden bioactive molecules and multicomponent interactions in herbal medicines, Sci. Rep. 6 (2016) 23840.

[1] D.W.S. Cheung, C.M. Koon, C.F. Ng, P.C. Leung, K.P. Fung, S.K.S. Poon, C.B.S. Lau, The roots of Salvia miltiorrhiza (Danshen) and Pueraria lobata (Gegen) inhibit atherogenic events: A study of the combination effects of the 2-herb formula, J. Ethnopharmacol. 143 (2012) 859-866.

[1] H.P. Song, S.Q. Wu, L.W. Qi, F. Long, L.F. Jiang, K. Liu, H. Zeng, Z.M. Xu, P. Li, H. Yang, A strategy for screening active lead compounds and functional compound combinations from herbal medicines based on pharmacophore filtering and knockout/knockin chromatography, J. Chromatogr. A 1456 (2016) 176-186.

[1] P. Liu, H. Yang, F. Long, H.P. Hao, X. Xu, Y. Liu, X.W. Shi, D.D. Zhang, H.C. Zheng, Q.Y. Wen, W.W. Li, H. Ji, X.J. Jiang, B.L. Zhang, L.W. Qi, P. Li,

Bioactive equivalence of combinatorial components identified in screening of an herbal medicine, Pharm. Res. 31 (2014) 1788-1800.

[1] H.P. Song, H. Wang, J.X. Liang, C. Qian, S.Q. Wu, W.J. Xu, B. Wu, X.G. Liu, P. Li, H. Yang, Integration of multiple analytical and computational tools for the discovery of high-potency enzyme inhibitors from herbal medicines, ChemMedChem 11 (2016) 2588-2597.

[1] X.G. Liu, C.Y. Cheng, J.X. Wang, H. Luo, L.F. Tu, L. Lin, B. Wu, H.Y. Wang, K. Liu, P. Li, H. Yang, A metabolic exposure-oriented network regulation strategy for the identification of effective combination in the extract of Ginkgo biloba L, J. Pharm. Biomed. Anal. 149 (2018) 151-159.

[1] A.C. Servais, B. Janicot, A. Takam, J. Crommen, M. Fillet, Liquid chromatography separation of the chiral prodrug eslicarbazepine acetate and its main metabolites in polar organic mode. Application to their analysis after in vitro metabolism, J. Chromatogr. A 1467 (2016) 306-311.

[1] Y. Sun, X. Lu, Y. Gai, C. Sha, G. Leng, X. Yang, W. Liu, LC-MS/MS method for the determination of the prodrug aripiprazole lauroxil and its three metabolites in plasma and its application to in vitro biotransformation and animal pharmacokinetic studies, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 1081 (2018) 71-79.

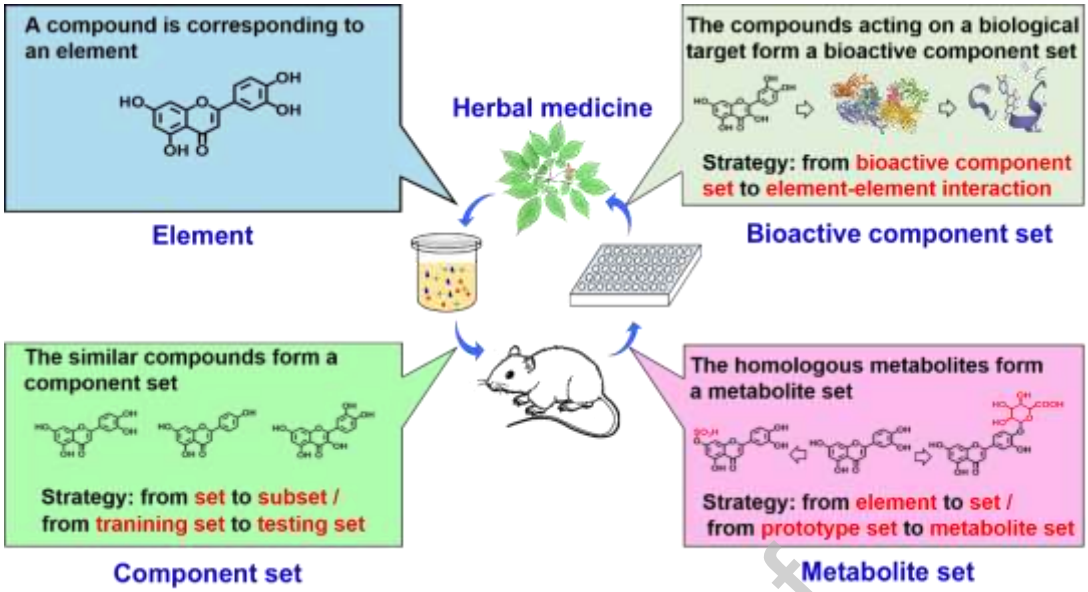

Fig. 1

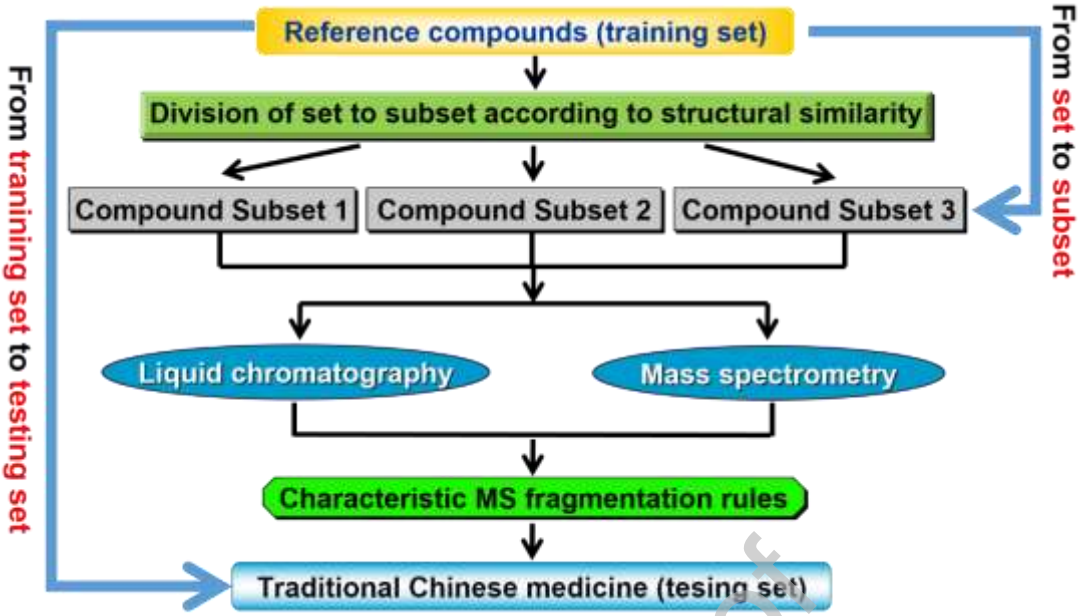

Fig. 2

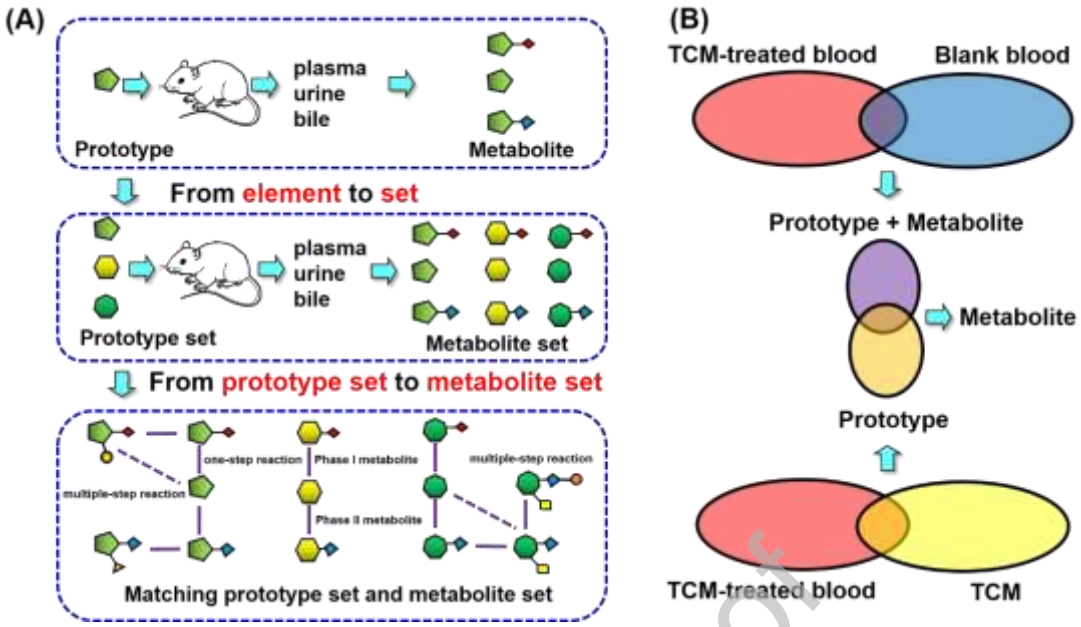

Fig. 3

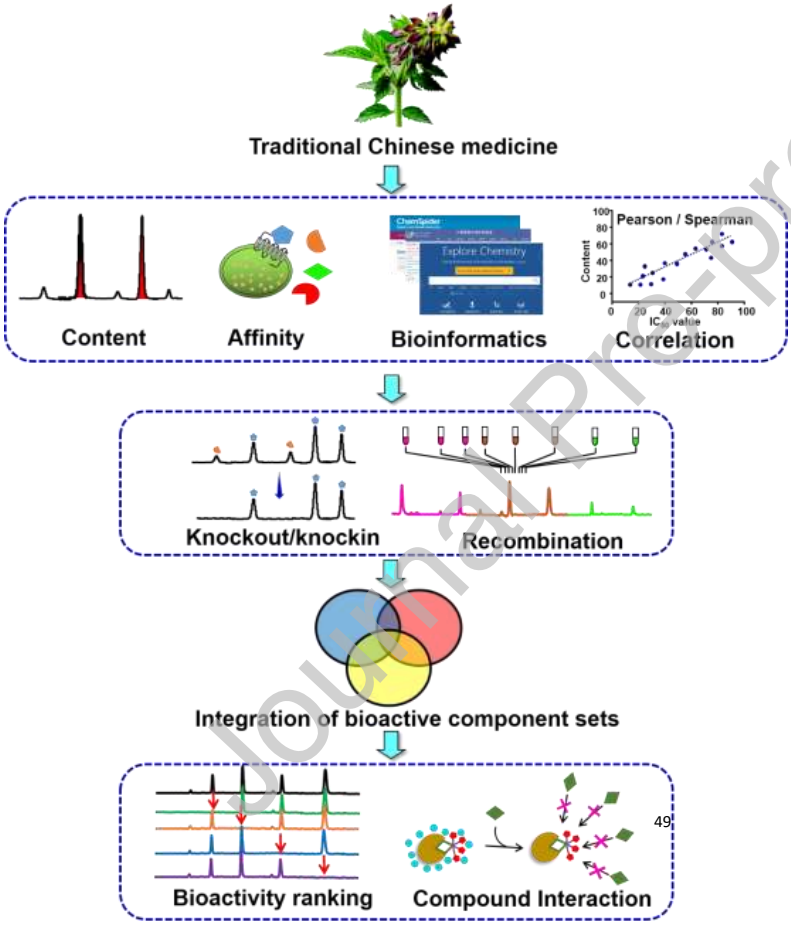

Fig. 4

Figure captions

Fig. 1 Schematic diagram of research strategies for traditional Chinese medicine based on set theory.

Fig. 2 Technical route for component analysis of traditional Chinese medicine.

Fig. 3 (A) Technical route for metabolite analysis of traditional Chinese medicine. (B) Venn diagram for identification of prototypes and metabolites of traditional Chinese medicine.

Fig. 4 Technical route for discovery of bioactive components from traditional Chinese medicine.

Table 1 Representative works for component analysis of TCM based on LC-MS

No.	TCM	LC	MS	Training set	Testing set	The type of set	Reference
1	*Lithospermum erythrorhizon*, *Arnebia euchroma*, *Arnebia guttata*	Welch Ultimate UHPLC C18 column (2.1 mm × 100 mm, 1.7 μm)	Triple TOF MS (AB SCIEX 5600)	17	58	shikonins, shikonofurans	[30]
2	*Spatholobus suberectus*	Agilent Extent C18 column (4.6 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6520)	13	36	flavonoids	[33]
3	*Scutellaria baicalensis*	Agilent Eclipse Plus C18 column (2.1 mm × 150 mm, 1.8 μm)	Q-Orbitrap MS (Thermo Fisher)	34	132	flavonoid, phenylethanoid glycoside	[34]
4	*Uncaria rhynchophylla*	Waters Acquity UPLC BEH C18 column (2.1 mm × 100 mm,1.7 μm)	LTQ-Orbitrap MS (Thermo Fisher)	14	92	alkaloids	[37]
5	*Stephania tetrandra*	Waters XBridge (R) C18 column (2.1 mm × 150 mm, 3.5 μm)	Triple TOF MS (AB SCIEX 6600)	14	393	isoquinoline alkaloids	[38]
6	*Stephania hainanensis*	XBH phenyl column (2.1 mm × 150 mm, 5 μm)	Triple TOF MS (AB SCIEX 5600+)	7	37	alkaloids	[40]
7	Processed *Semen Strychni*	Agilent Zorbax Extend C18 column (4.6 mm × 150 mm, 5 μm)	IT-TOF MS (Shimadzu)	2	24	dihydroindole-type alkaloids	[41]
8	*Corydalis yanhusuo*, *Corydalis ternata*, *Corydalis decumbens*	Agilent Poroshell 120 EC-C18 column (2.1 mm × 150 mm, 2.7 μm)	QQQ MS (AB SCIEX 3200) FTICR MS (IonSpec Corporation)	9	33	alkaloids	[42]
9	Raw/steamed American ginseng roots/berries	Agilent Zorbax Extend-C18 column (4.6 mm × 50 mm, 1.8 μm)	Q-TOF MS (Agilent 6530)	13	70	saponins	[44]
10	*Akebiae Fructus*	Welch Ultimate XB−C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters SYNAPT G2-Si)	11	85	saponins	[45]

11	Shuxiong tablet	Phenomenex Kinetex XB-C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters Xevo G2-S)	72	250	saponins, quinochalcone C-glycosides, flavonoid O-glycoside *et al.*	[47]
12	Er-xian decoction	Thermo BDS HYPERSIL C18 column (2.1 mm × 150 mm, 3 μm)	LTQ-FT MS (Thermo Fisher)	35	71	prenylflavonoids, phenylpropanoids, alkaloids	[48]
13	Licorice	Waters Acquity HSS T3 column (2.1 mm × 100 mm, 1.8 μm)	Q-TOF MS (Waters Xevo G2)	15	51	flavonoids	[49]
14	Danhong injection	Agilent Poroshell 120 EC-C18 column (2.1 mm ×150 mm, 2.7 μm)	Qtrap MS (AB SCIEX 4500)	6	90	phenolic acids, flavonoids	[50]
15	Baoyuan Decoction	Waters CORTECS UPLC C18 column (2.1 mm × 100 mm, 1.6 μm)	Q-TOF MS (Waters Xevo G2); Qtrap MS (AB SCIEX 4500)	49	236	saponins, flavonoids *et al.*	[51]
16	Raw *Alismatis Rhizoma*, Processed *Alismatis Rhizoma*	Welch Ultimate UHPLC C18 column (2.1 mm × 100 mm, 1.8 μm)	Triple TOF MS (AB SCIEX 5600); Qtrap MS (AB SCIEX 4000)	7	80, 87	triterpenes	[52]
17	*Panax ginseng*, *Panax quinquefolium*, *Panax notoginseng*	MCI gel (75-150 μm) × silica gel (200-300 mesh); YMC-Pack ODS-A column (4.6 mm × 250 mm, 5 μm)	IT MS (Thermo Fisher); Q-TOF MS (Agilent 6510)	18	334, 262, 309	ginsenosides	[53]
18	*Uncaria rhynchophylla*, *Uncaria hirsuta*, *Uncaria macrophylla*, *Uncaria sinensis*, *Uncaria sessilifructus*	Acchrom XCharge C18 column (4.6 mm × 250 mm, 5 μm); Phenomenex Kinetex® EVO C18 column (2.1 mm × 150 mm, 2.6 μm)	LTQ-Orbitrap MS (Thermo Fisher)	44	1227	indole alkaloids	[54]
19	Mai-Luo-Ning injection, *Lonicerae Flos* injection, Shen-Mai injection	Agilent Zorbax SB-C18 column (2.1 mm × 150 mm, 3.5 μm)	IT-TOF MS (Shimadzu)	30, 17	45, 46	organic acids, ginsenosides	[55]

20	*Ligusticum chuanxiong*	Waters Acquity BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters SYNAPT G2-Si)	23	52	phthalides	[56]
21	*Poria cocos*	Waters Acquity BEH C18 column (2.1 mm × 50 mm, 1.7 μm)	Q-TOF MS (Waters SYNAPT G2-Si)	8	121	lanostane-type triterpene acids	[57]
22	Frankincense	Agilent Zorbax SB-C18 column (21.2 mm × 250 mm, 7 μm) Agilent XDB-C18 column (4.6 mm × 150 mm, 5 μm)	Q-TOF MS (Agilent 6520)	11	75	triterpenoid	[58]
23	*Carthamus tinctorius*	Offline 2D-LC: Acchrom XAmide column (4.6 mm × 150 mm, 5 μm); Waters BEH Shield RP-18 column (2.1 mm × 100 mm, 1.7 μm)	LTQ-Orbitrap MS (Thermo Fisher)	14	103	flavonoid O-glycosides	[59]

Table 2 Representative works for metabolite analysis of TCM based on LC-MS

No.	TCM	LC	MS	Biological sample	Prototype set	Metabolite set	The type of set	Reference
1	Kai-Xin-San	Phenomenex Kinetex XB C18 column (4.6 mm × 100 mm, 2.6 μm)	Triple TOF MS (AB SCIEX 5600)	plasma	47	22	ginsenosides, polygala saponins, oligosaccharide esters *et al.*	[60]
2	Ding-Zhi-Xiao-Wan prescription	Thermo Hypersil GOLD C18 column (2.1 mm × 100 mm, 1.9 μm)	Q-TOF MS (Waters SYNAPT G2-Si)	plasma, liver, brain, feces, urine	150	51	saponins, xanthones, oligoesters *et al.*	[61]
3	Licorice	Agilent ZORBAX Extent-C18 column (4.6 mm × 250 mm, 5 μm)	IT MS (Thermo Fisher); Q-TOF MS (Agilent 6510); QQQ MS (Thermo Fisher)	plasma, urine	12	78	flavonoids, saponins	[62]
4	*Glehniae Radix*	Phenomenex Kinelex C18 100A column (2.1 mm × 100 mm, 2.6 μm)	Triple TOF MS (AB SCIEX 5600+)	urine, plasma, bile	34, 11, 11	57, 15, 11	coumarins	[63]
5	Fufang-Xialian-Capsule	Waters Acquity UPLC BEH C18 Column (2.1 mm × 50 mm, 1.7 μm)	Q-TOF MS (Waters SYNAPT G2)	plasma	11	33	flavones, alkaloids, ginsenosides	[64]
6	*Sophora flavescens*	Agilent Zorbax SB-Aqcolumn (4.6 mm × 150 mm, 5 μm)	Q-Orbitrap MS (Thermo Fisher)	plasma, urine, bile	89	82	alkaloids, flavonoids	[65]
7	Mai-Luo-Ning injection	Phenomenex Synergi C18 Hydro-RP 80A column (4.6 mm × 250 mm, 4 μm)	IT-TOF MS (Shimadzu)	urine	87	162	organic acids, glycosides, flavones *et al.*	[66]

8	Sini decoction	Waters Acquity UPLC BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Agilent 6530)	urine	53	49	alkaloids, flavonoids, saponins *et al.*	[67]
9	Zhi-Zi-Hou-Po decoction	Lichrospher C18 column (4.6 mm × 250 mm, 5 μm)	TOF MS (Agilent 6224)	plasma	35	26	iridoid glycosides, flavonoids, lignans *et al.*	[68]
10	Da-Huang-Xiao-Shi decoction	Lichrospher C18 column (4.6 mm × 250 mm, 5 μm)	TOF MS (Agilent 6224) QQQ MS (Thermo Finnigan)	plasma	32	30	iridoid glycosides, monoterpenoids, alkaloids *et al.*	[69]
11	Shuang-Huang-Lian injection	Waters Acquit UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 μm)	Q-TOF MS (Waters Evoe G2)	serum	23	12	flavonoids, phenylethanoid glycosides, iridoids *et al.*	[70]
12	Qingkailing injection	Diamonsil C18 column (4.6 mm × 250 mm, 5 μm)	LTQ-Orbitrap MS (Thermo Fisher)	urine	18	19	organic acids, flavonoids, iridoids	[71]
13	*Glechomae Herba*	Waters Acquity UPLC® BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Triple TOF MS (AB SCIEX 5600+)	urine, bile, plasma, feces	13	78	phenolic acids, flavonoids	[72]
14	*Paeoniae Radix Rubra*	Phenomenex Gemini C18 column (4.6 mm × 250 mm, 5 μm)	IT-TOF MS (Shimadzu)	plasma, urine	15	90	monoterpene glycosides, phenolic compounds	[73]
15	Dan Zhi Tablet	Waters Acquity UPLC HSS T3 column (2.1 mm × 150 mm, 1.8 μm)	QTOF MS (Waters SYNAPT G2-Si)	plasma, cerebrospinal fluid	51, 17	119, 14	flavonoids, tanshinones, phthalides *et al.*	[74]

16	Run-zao-zhi-yang capsule	Waters Acquity UPLC BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters Xevo G2-XS)	plasma, urine, bile	18	35	alkaloids, flavonoids, stilbene glycosides *et al.*	[75]
17	Qi-Lin Pills	Waters Acquity UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 μm)	Q-TOF MS (Waters SYNAPT^{TM} G2)	plasma, urine, feces	22, 47, 27	19, 97, 23	flavonoids, monoterpenes, triterpenoids *et al.*	[76]

Table 3 Representative works for the discovery of bioactive components from TCM based on LC-MS

No.	TCM	LC	MS	Bioactivity	Bioactive com
1	Tanshinones extract	Agilent Zorbax Extend C18 column (4.6 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6530)	cardioprotective activity	tanshinones IIA, cryptotanshinone I, dihydrotanshin
2	Yindan Xinnaotong soft capsule	Agilent Zorbax Eclipse Plus C18 column (4.6 mm × 150 mm, 1.8 μm)	Q-TOF MS (Agilent 6530)	neuroprotective activity	borneol, citric ac scutellarin *et al.*
3	Er-Xian decotion	Agilent ZORBAX Eclipse XDB-C18 column (4.6 mm × 250 mm, 5 μm), Agilent ZORBAX Eclipse XDB-C18 column (9.4 mm × 250 mm, 5 μm)	Q-TOF MS (Waters Synapt)	anti-oxidative activity, anti-osteoporotic activity	obaculactone, ber ferulic acid *et al.*
4	Turmeric	Kromasil C18 column (4.6 mm × 250 mm, 5 μm)	LTQ-Orbitrap MS (Thermo Fisher)	anti-cancer activity	curcumin, demethoxycurcu bisdemethoxycur
5	Qishen Yiqi dripping pills	Acchrom XAmide column (30 mm × 150 mm, 10 μm), Agilent SB-C18 column (9.6 mm × 250 mm, 7 μm), Waters BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters Acquity Synapt)	cardioprotective activity	danshensu, proto aldehyde, salvian *et al.*
6	Xuesaitong Injection	Agilent Zorbax SB-C18 column (4.6 mm × 250 mm, 5 μm)	IT MS (Thermo Finnigan)	cardioprotective activity	notoginsenoside ginsenoside Rg1, Re *et al.*
7	Heishunpian	Waters Acquity UPLC BEH C18 column (2.1 mm × 100 mm, 1.7 μm)	Q-TOF MS (Waters Xevo G2)	anti-inflammatory activity	hypaconitine, deoxyaconitine, chasmanine
8	*Daturae Flos*	Agilent C18 column (4.6 mm × 250 mm, 5 μm)	TOF MS (Agilent)	M2 receptor antagonistic activity	scopolamine, hyc
9	Longan seed	Agilent ZORBAX TC-C18 column (4.6 mm × 250 mm, 5 μm), Agilent ZORBAX SB-C18 column (9.4 mm × 250 mm, 5 μm)	FT-ICR MS (Bruker), IT MS (Agilent 6300)	antioxidant activity	ellagic acid and i p-coumaric acid-
10	Safflower	XBridge™ Prep C18 OBD™ column (30 mm × 150 mm, 5 μm)	Q-TOF MS (Waters SYNAPT™)	antiplatelet aggregation activity, anticoagulation activity, antioxidant activity, antiproliferative activity	anhydrosafflor ye hydroxysafflor ye kaempferol 3-O- *et al.*
11	*Coptidis Rhizoma*	Kromasil™ C18 column	Q-TOF MS	antibacterial activity	berberine, coptisi

		(4.6 mm × 250 mm, 5 μm)	(Waters Xevo G2)		
12	*Lycoridis Radiatae Bulbus*	Agilent ZorBax Extend-C18 column (4.6 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6530)	acetylcholinesterase inhibitory activity	ungerimine, galar
13	*Salvia Miltiorrhizae Radix*	Agilent Zorbax Extend-C18 column (4.6 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6530)	peroxidase inhibitory activity	tanshinol, protoca aldehyde, lithosp *al.*
14	Dan-Qi pair	Agilent Zorbax Extend-C18 column (4.6 mm × 250 mm, 5 μm), Agilent Zorbax SB-C18 column (9.4 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6520)	thrombin inhibitory activity	tanshinones IIA, cryptotanshinone tanshinone I, dihydrotanshinon
15	Dan-Qi formula	Agilent Zorbax Extend-C18 column (4.6 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6520)	free radical scavenging activity	salvianolic acid A lithospermic acid tanshinol *et al.*
16	Cardiotonic Pill	Agilent Zorbax SB-C18 column (4.6 mm ×250 mm, 5 μm), Agilent Zorbax SB-C18 column (9.4 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6530)	cardioprotective activity	tanshinol, protoca aldehyde, caffeic
17	*Ginkgo Folium*	Agilent Zorbax Eclipse-AAA reversed-phase column (4.6 mm × 150 mm, 3.5 μm), Agilent Zorbax SB-C18 column (9.4 mm × 250 mm, 5 μm)	Q-TOF MS (Agilent 6520)	xanthine oxidase inhibitory activity	quercetin 3-O-rut isorhamnetin 3-C apigenin *et al.*
18	*Ginkgo Folium*	Agilent Zorbax Extent C18 column (2.1 mm × 150 mm, 5 μm), Agilent Zorbax SB-C18 column (9.4 mm × 250 mm, 5 μm)	QQQ MS (Shimadzu 8050)	neuroprotective activity	quercetin, chloro ginkgolide B *et a*

[i] B. He, C. Lu, M.L. Wang, G. Zheng, G. Chen, M. Jiang, X.J. He, Z.X. Bian, G. Zhang, A.P. Lu, Drug discovery in traditional Chinese medicine: from herbal fufang to combinatory drugs, Science 350 (2015) S74-S76.

[ii] Y. Han, H. Sun, A. Zhang, G. Yan, X.J. Wang, Chinmedomics, a new strategy for evaluating the therapeutic efficacy of herbal medicines, Pharmacol. Ther. 216 (2020) 107680.

[iii] K. Murata, D. Iida, Y. Ueno, K. Samukawa, T. Ishizaka, T. Kotake, H. Matsuda, Novel polyacetylene derivatives and their inhibitory activities on

acetylcholinesterase obtained from Panax ginseng roots, J. Nat. Med. 71 (2017) 114 -122.

[iv] Y. Tu, Artemisinin-a gift from traditional Chinese medicine to the world (Nobel lecture), Angew. Chem. Int. Edit. 55 (2016) 10210-10226.

[v] E.L.H. Leung, S.W. Xu, Traditional Chinese medicine in cardiovascular drug discovery, Pharmacol. Res. 160 (2020) 105168.

[vi] X. Ma, B. Lv, P. Li, X. Jiang, Q. Zhou, X. Wang, X. Gao, Identification of "multiple components-multiple targets-multiple pathways" associated with Naoxintong capsule in the treatment of heart diseases using UPLC/Q-TOF-MS and network pharmacology, Evid.-Based Compl. Alt. Med. 2016 (2016) 9468087.

[vii] W.J. Zhang, Y. Huai, Z.P. Miao, A.R. Qian, Y.H. Wang, Systems pharmacology for investigation of the mechanisms of action of traditional Chinese medicine in drug discovery, Front. Pharmacol. 10 (2019) 743.

[viii] S. Singh, D.B. Singh, S. Singh, R. Shukla, P.W. Ramteke, K. Misra, Exploring medicinal plant legacy for drug discovery in post-genomic era, Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci. 89 (2019) 1141-1151.

[ix] A. Stavrianidi, A classification of liquid chromatography mass spectrometry techniques for evaluation of chemical composition and quality control of traditional medicines, J. Chromatogr. A 1609 (2020) 460501.

[x] B.Y. Pang, Y. Zhu, L.Q. Lu, F.B. Gu, H.L. Chen, The applications and features of liquid chromatography-mass spectrometry in the analysis of traditional Chinese medicine, Evid.-Based Compl. Alt. Med. 2016 (2016) 3837270.

[xi] S. Ji, S. Wang, H. Xu, Z. Su, D. Tang, X. Qiao, M. Ye, The application of on-line two-dimensional liquid chromatography (2DLC) in the chemical analysis of herbal medicines, J. Pharm. Biomed. Anal. 160 (2018) 301-313.

[xii] A. Zhang, H. Sun, X. Wang, Mass spectrometry-driven drug discovery for development of herbal medicine, Mass Spectrometry Reviews 37 (2018) 307-320.

[xiii] Z. Zhang, T. Bo, Y. Bai, M. Ye, R. An, F. Cheng, H. Liu, Quadrupole time-of-flight mass spectrometry as a powerful tool for demystifying traditional

Chinese medicine, Trac-Trend. Anal. Chem. 72 (2015) 169 -180.

[xiv] H. Lei, Y. Zhang, X. Zu, J. Ye, Y. Liang, T. Cheng, W. Zhang, Comprehensive profiling of the chemical components and potential markers in raw and processed Cistanche tubulosa by combining ultra-high-performance liquid chromatography coupled with tandem mass spectrometry and MS/MS-based molecular networking, Anal. Bioanal. Chem. 413 (2020) 129-139.

[xv] T. Xu, S. Li, Y. Sun, Z. Pi, S. Liu, F. Song, Z. Liu, Systematically characterize the absorbed effective substances of Wutou Decoction and their metabolic pathways in rat plasma using UHPLC-Q-TOF-MS combined with a target network pharmacological analysis, J. Pharm. Biomed. Anal. 141 (2017) 95-107.

[xvi] J. Fang, L. Wang, T. Wu, C. Yang, L. Gao, H. Cai, J. Lin, S. Fang, Y. Chen, W. Tan, Q. Wang, Network pharmacology-based study on the mechanism of action for herbal medicines in Alzheimer treatment, J. Ethnopharmacol. 196 (2017) 281-292.

[xvii] Y. Song, Q. Song, J. Li, N. Zhang, Y. Zhao, X. Liu, Y. Jiang, P. Tu, An integrated strategy to quantitatively differentiate chemome between Cistanche deserticola and C-tubulosa using high performance liquid chromatography-hybrid triple quadrupole-linear ion trap mass spectrometry, J. Chromatogr. A 1429 (2016) 238-247.

[xviii] G. Cao, N. Wang, D. He, X. Wang, Y. Tian, N. Wan, W. Yan, H. Ye, H. Hao, Intestinal mucosal metabolites-guided detection of trace-level ginkgo biloba extract metabolome, J. Chromatogr. A 1608 (2019) 460417.

[xix] R.T. Cook, The foundations of mathematics in the theory of sets, Brit. J. Philos. Sci. 54 (2003) 347-352.

[xx] Z.M. Jiang, L.J. Wang, H.Q. Pang, Y. Guo, P.T. Xiao, C. Chu, L. Guo, E.H. Liu, Rapid profiling of alkaloid analogues in Sinomenii Caulis by an integrated characterization strategy and quantitative analysis, J. Pharm. Biomed. Anal. 174 (2019) 376-385.

[xxi] D. Li, O.J. Schmitz, Comprehensive two-dimensional liquid chromatography tandem diode array detector (DAD) and accurate mass QTOF-MS for the analysis of flavonoids and iridoid glycosides in Hedyotis diffusa, Anal. Bioanal. Chem. 407 (2015) 231-240.

[xxii] B.K. Shin, S.W. Kwon, J.H. Park, Chemical diversity of ginseng saponins from Panax ginseng, J. Ginseng Res. 39 (2015) 287-298.

[xxiii] L.W. Qi, C.Z. Wang, C.S. Yuan, Isolation and analysis of ginseng: advances and challenges, Nat. Prod. Rep. 28 (2011) 467-495.

[xxiv] L.F. Hu, Z.H. Yao, Z.F. Qin, L.Y. Liu, X.J. Song, Y. Dai, H. Kiyohara, H. Yamada, X.S. Yao, In vivo metabolic profiles of Bu-Zhong-Yi-Qi-Tang, a famous traditional Chinese medicine prescription, in rats by ultra-high-performance liquid chromatography coupled with quadrupole time-of-flight tandem mass spectrometry, J. Pharm. Biomed. Anal. 171 (2019) 81-98.

[xxv] T. Wang, X.C. Feng, L.Q. Ding, K. Wang, M. Qiao, L.W. Chai, Y. Li, F. Qiu, Metabolic profiling of icariin in rat feces, urine, bile and plasma after oral administration using ultra-high performance liquid chromatography/quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 168 (2019) 155-162.

[xxvi] L.J. Wang, Z.M. Jiang, P.T. Xiao, J.B. Sun, Z.M. Bi, E.H. Liu, Identification of anti-inflammatory components in Sinomenii Caulis based on spectrum-effect relationship and chemometric methods, J. Pharm. Biomed. Anal. 167 (2019) 38-48.

[xxvii] X. Xu, W. Li, T. Li, K. Zhang, Q. Song, L. Liu, P. Tu, Y. Wang, Y. Song, J. Li, Direct infusion-three-dimensional-mass spectrometry enables rapid chemome comparison among herbal medicines, Anal. Chem. 92 (2020) 7646-7656.

[xxviii] H.S. Fan, S.J. Liu, W.X. Shen, A. Kang, J.N. Tan, L. Li, X. Liu, C.L. Xu, X.F. Xu, Y.Y. Lai, H.B. Cheng, D.D. Sun, Identification of the absorbed components and metabolites of Xiao-Ai-Jie-Du decoction and their distribution in rats using ultra high-performance liquid chromatography/quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 179 (2020) 112984.

[xxix] M.L. Chen, W.Q. Chang, J.L. Zhou, Y.H. Yin, W.R. Xia, J.Q. Liu, L.F. Liu, G.Z. Xin, Comparison of three officinal species of Callicarpa based on a biochemome profiling strategy with UHPLC-IT-MS and chemometrics analysis, J. Pharm. Biomed. Anal. 145 (2017) 666 -674.

[xxx] M. Liao, A. Li, C. Chen, H. Ouyang, Y. Zhang, Y. Xu, Y. Feng, H. Jiang, Systematic identification of shikonins and shikonofurans in medicinal Zicao

species using ultra-high performance liquid chromatography quadrupole time of flight tandem mass spectrometry combined with a data mining strategy, J. Chromatogr. A 1425 (2015) 158-172.

[xxxi] Z. Xue, C. Lai, L. Kang, A. Kotani, H. Hakamata, Z. Jing, H. Li, W. Wang, B. Yang, Profiling and isomer recognition of phenylethanoid glycosides from Magnolia officinalis based on diagnostic/holistic fragment ions analysis coupled with chemometrics, J. Chromatogr. A 1611 (2020) 460583.

[xxxii] X.Y. Zhu, J.Z. Liu, Z.H. Dong, F. Feng, W.Y. Liu, Identification and screening of cardiac glycosides in Streptocaulon griffithii using an integrated data mining strategy based on high resolution mass spectrometry, Chin. J. Nat. Medicines 16 (2018) 546-560.

[xxxiii] X.L. Cheng, J.Y. Wan, P. Li, L.W. Qi, Ultrasonic/microwave assisted extraction and diagnostic ion filtering strategy by liquid chromatography-quadrupole time-of-flight mass spectrometry for rapid characterization of flavonoids in Spatholobus suberectus, J. Chromatogr. A 1218 (2011) 5774-5786.

[xxxiv] X. Qiao, R. Li, W. Song, W.J. Miao, J. Liu, H.B. Chen, D.A. Guo, M. Ye, A targeted strategy to analyze untargeted mass spectral data: rapid chemical profiling of Scutellaria baicalensis using ultra-high performance liquid chromatography coupled with hybrid quadrupole orbitrap mass spectrometry and key ion filtering, J. Chromatogr. A 1441 (2016) 83-95.

[xxxv] J.Y. Zhang, J.Q. Lu, F. Wang, G.S. Cao, P.F. Tu, Y.J. Qiao, A strategy of EIC-MS coupled with diagnostic product ions analysis for efficient discovery of new hydroxylated polymethoxyflavonoid glycosides from the leaves of Murraya paniculata L. using HPLC-DAD-MS/MS, Anal. Methods 5 (2013) 2880-2891.

[xxxvi] J.Y. Zhang, F. Wang, H. Zhang, J.Q. Lu, Y.J. Qiao, Rapid identification of polymethoxylated flavonoids in traditional Chinese medicines with a practical strategy of stepwise mass defect filtering coupled to diagnostic product ions analysis based on a hybrid LTQ-Orbitrap mass spectrometer, Phytochem. Analy. 25 (2014) 405-414.

[xxxvii] H. Pan, W. Yang, Y. Zhang, M. Yang, R. Feng, W. Wu, D. Guo, An integrated strategy for the systematic characterization and discovery of new indole alkaloids

from Uncaria rhynchophylla by UHPLC/DAD/LTQ-Orbitrap-MS, Anal. Bioanal. Chem. 407 (2015) 6057-6070.

[xxxviii] J. Chen, Q. Zhao, D. Si, A. Nie, Y. Wang, Z. Deng, Y. Wen, F. Chen, L. Zhang, B. Dong, J. Yang, Comprehensive profiling of Stephania tetrandra (Fangji) by stepwise DFI and NL-dependent structure annotation algorithm-based UHPLC-Q-TOF-MS and direct authentication by LMJ-HRMS, J. Pharm. Biomed. Anal. 185 (2020) 113225.

[xxxix] Y.Y. Shi, R.Q. Wang, X.Y. Zhu, D.G. Xu, W.Y. Liu, F. Feng, A self-feedback network based on liquid chromatography-quadrupole-time of flight mass spectrometry for system identification of beta-carboline alkaloids in Picrasma quassioides, Sci. Rep. 7 (2017) 13841.

[xl] J.Y. He, Y. Liu, Y. Kang, P. Yang, Y.Q. Wang, J.X. Guo, J.M. Huang, Identification of alkaloids in Stephania hainanensis by liquid chromatography coupled with quadrupole time-of-flight mass spectrometry, Phytochem. Analy. 27 (2016) 206-216.

[xli] J.X. Tian, Y. Tian, L. Xu, J. Zhang, T. Lu, Z.J. Zhang, Characterisation and identification of dihydroindole-type alkaloids from processed Semen Strychni by high-performance liquid chromatography coupled with electrospray ionisation ion trap time-of-flight mass spectrometry, Phytochem. Analy. 25 (2014) 36-44.

[xlii] E.K. Jeong, S.Y. Lee, S.M. Yu, N.H. Park, H.S. Lee, Y.H. Yim, G.S. Hwang, C. Cheong, J.H. Jung, J. Hong, Identification of structurally diverse alkaloids in Corydalis species by liquid chromatography/electrospray ionization tandem mass spectrometry, Rapid Commun. Mass Sp. 26 (2012) 1661-1674.

[xliii] Y.P. Jiang, R. Liu, M. Liu, L.Z. Yi, S. Liu, An integrated strategy to rapidly characterize non-targeted benzylisoquinoline alkaloids from Plumula nelumbinis ethanol extract using UHPLC/Q-orbitrap HRMS, Int. J. Mass Spectrom. 432 (2018) 26-35.

[xliv] L.W. Qi, H.Y. Wang, H. Zhang, C.Z. Wang, P. Li, C.S. Yuan, Diagnostic ion filtering to characterize ginseng saponins by rapid liquid chromatography with time-of-flight mass spectrometry, J. Chromatogr. A 1230 (2012) 93-99.

[xlv] B.L. Ma, S.L. Yang, J.M. Li, H. Ouyang, M.Z. He, Y.L. Feng, T. Tan, A four-step filtering strategy based on ultra-high-performance liquid chromatography

coupled to quadrupole-time-of-flight tandem mass spectrometry for comprehensive profiling the major chemical constituents of Akebiae Fructus, Rapid Commun. Mass Sp. 33 (2019) 1464-1474.

[xlvi] T. Xie, Y. Liang, H.P. Hao, J.A, L. Xie, P. Gong, C. Dai, L.S. Liu, A. Kang, X. Zheng, G.J. Wang, Rapid identification of ophiopogonins and ophiopogonones in Ophiopogon japonicus extract with a practical technique of mass defect filtering based on high resolution mass spectrometry, J. Chromatogr. A 1227 (2012) 234-244.

[xlvii] C. Yao, W. Yang, W. Si, H. Pan, S. Qiu, J. Wu, X. Shi, R. Feng, W. Wu, D. Guo, A strategy for establishment of practical identification methods for Chinese patent medicine from systematic multi-component characterization to selective ion monitoring of chemical markers: Shuxiong tablet as a case study, Rsc Adv. 6 (2016) 65055-65066.

[xlviii] C. Wang, J. Zhang, C. Wu, Z. Wang, A multiple-dimension liquid chromatography coupled with mass spectrometry data strategy for the rapid discovery and identification of unknown compounds from a Chinese herbal formula (Er-xian decoction), J. Chromatogr. A 1518 (2017) 59-69.

[xlix] T. Xu, M. Yang, Y. Li, X. Chen, Q. Wang, W. Deng, X. Pang, K. Yu, B. Jiang, S. Guan, D.A. Guo, An integrated exact mass spectrometric strategy for comprehensive and rapid characterization of phenolic compounds in licorice, Rapid Commun. Mass Sp. 27 (2013) 2297-2309.

[l] C. Li, J. Yang, X. Tong, C. Zhao, Y. He, H. Wan, Precursor ion scan enhanced rapid identification of the chemical constituents of Danhong injection by liquid chromatography-tandem mass spectrometry: an integrated strategy, J. Chromatogr. A 1602 (2019) 378-385.

[li] X. Ma, X. Guo, Y. Song, L. Qiao, W. Wang, M. Zhao, P. Tu, Y. Jiang, An integrated strategy for global qualitative and quantitative profiling of traditional Chinese medicine formulas: Baoyuan decoction as a case, Sci. Rep. 6 (2016) 38379.

[lii] S. Li, S. Jin, C. Song, S. Jia, Y. Zhang, Y. Feng, Z. Du, H. Jiang, The strategy for establishment of the multiple reaction monitoring based characteristic chemical profile of triterpenes in Alismatis rhizoma using two combined tandem mass

spectrometers, J. Chromatogr. A 1524 (2017) 121-134.

[liii] W.Z. Yang, M. Ye, X. Qiao, C.F. Liu, W.J. Miao, T. Bo, H.Y. Tao, D.A. Guo, A strategy for efficient discovery of new natural compounds by integrating orthogonal column chromatography and liquid chromatography/mass spectrometry analysis: Its application in Panax ginseng, Panax quinquefolium and Panax notoginseng to characterize 437 potential new ginsenosides, Anal. Chim. Acta 739 (2012) 56-66.

[liv] H. Pan, C. Yao, W. Yang, S. Yao, Y. Huang, Y. Zhang, W. Wu, D. Guo, An enhanced strategy integrating offline two-dimensional separation and step-wise precursor ion list-based raster-mass defect filter: Characterization of indole alkaloids in five botanical origins of Uncariae Ramulus Cum Unicis as an exemplary application, J. Chromatogr. A 1563 (2018) 124-134.

[lv] L. Wu, P. Gong, Y. Wu, K. Liao, H. Shen, Q. Qi, H. Liu, G. Wang, H. Hao, An integral strategy toward the rapid identification of analogous nontarget compounds from complex mixtures, J. Chromatogr. A 1303 (2013) 39-47.

[lvi] Q. Zhang, M. Huo, Y. Zhang, Y. Qiao, X. Gao, A strategy to improve the identification reliability of the chemical constituents by high-resolution mass spectrometry-based isomer structure prediction combined with a quantitative structure retention relationship analysis: phthalide compounds in Chuanxiong as a test case, J. Chromatogr. A 1552 (2018) 17-28.

[lvii] G.F. Feng, Y. Zheng, Y. Sun, S. Liu, Z.F. Pi, F.R. Song, Z.Q. Liu, A targeted strategy for analyzing untargeted mass spectral data to identify lanostane-type triterpene acids in Poria cocos by integrating a scientific information system and liquid chromatography-tandem mass spectrometry combined with ion mobility spectrometry, Anal. Chim. Acta 1033 (2018) 87-99.

[lviii] Y. Liu, Z. Liu, C. Lu, J. Li, Z. Ning, Z. Song, C. Wang, Z. Du, X. Lu, S. Zhao, A. Lu, Comprehensive identification of active triterpenoid metabolites in frankincense using a coupling strategy, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 963 (2014) 90-98.

[lix] C.L. Yao, W.Z. Yang, W. Si, Y. Shen, N.X. Zhang, H.L. Chen, H.Q. Pan, M. Yang, W.Y. Wu, D.A. Guo, An enhanced targeted identification strategy for the selective identification of flavonoid O-glycosides from Carthamus tinctorius by

integrating offline two-dimensional liquid chromatography/linear ion-trap-Orbitrap mass spectrometry, high-resolution diagnostic product ions/neutral loss filtering and liquid chromatography-solid phase extraction-nuclear magnetic resonance, J. Chromatogr. A 1491 (2017) 87-97.

[lx] X. Wan, J. Liu, X. Yang, Q. Zhang, Y. Zhang, Q. Li, K. Bi, Development of a systematic strategy for the global identification and classification of the chemical constituents and metabolites of Kai-Xin-San based on liquid chromatography with quadrupole time-of-flight mass spectrometry combined with multiple data-processing approaches, J. Sep. Sci. 41 (2018) 2672-2680.

[lxi] G. Feng, Z. Liu, S. Liu, J. Xing, F. Song, Z. Pi, A target integration strategy for analyzing multidimensional chemical and metabolic substance groups of Ding-Zhi-Xiao-Wan prescription by using ultra-high performance liquid chromatography tandem mass spectrometry, J. Chromatogr. A 1608 (2019) 460412.

[lxii] C. Xiang, X. Qiao, Q. Wang, R. Li, W. Miao, D. Guo, M. Ye, From single compounds to herbal extract: a strategy to systematically characterize the metabolites of licorice in rats, Drug Metab. Dispos. 39 (2011) 1597-1608.

[lxiii] S. Wang, P. Qi, N. Zhou, M. Zhao, W. Ding, S. Li, M. Liu, Q. Wang, S. Jin, A pre-classification strategy based on UPLC-Triple-TOF/MS for metabolic screening and identification of Radix glehniae in rats, Anal. Bioanal. Chem. 408 (2016) 7423-7436.

[lxiv] S. Li, S. Liu, Z. Liu, Z. Pi, F. Song, Y. Jin, A target-group-change strategy based on the UPLC-Q-TOF-MSE method for the metabolites identification of Fufang-Xialian-Capsule in rat's plasma, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 1085 (2018) 42-53.

[lxv] X. Dong, X. Li, N. Li, H. Zhao, A. GuLa, X. Zhang, P. Zhang, B. Bao, A target-group-change couple with mass defect filtering strategy to identify the metabolites of "Dogel ebs" in rats plasma, urine and bile, J. Sep. Sci. 42 (2019) 3382-3389.

[lxvi] P. Gong, N. Cui, L. Wu, Y. Liang, K. Hao, X. Xu, W. Tang, G. Wang, H. Hao, Chemicalome and metabolome matching approach to elucidating biological metabolic networks of complex mixtures, Anal. Chem. 84 (2012) 2995-3002.

[lxvii] G. Tan, M. Liu, X. Dong, S. Wu, L. Fan, Y. Qiao, Y. Chai, H. Wu, A strategy for rapid analysis of xenobiotic metabolome of Sini decoction in vivo using ultra-performance liquid chromatography-electrospray ionization quadrupole-time-of-flight mass spectrometry combined with pattern recognition approach, J. Pharm. Biomed. Anal. 96 (2014) 187-196.

[lxviii] K. Luo, F. Feng, Identification of absorbed components and metabolites of Zhi-Zi-Hou-Po decoction in rat plasma after oral administration by an untargeted metabolomics-driven strategy based on LC-MS, Anal. Bioanal. Chem. 408 (2016) 5723-5735.

[lxix] D. Wang, Q. Shi, F. Feng, Rapid discovery and identification of the prototypes and their metabolites of Da-Huang-Xiao-Shi decoction in rat plasma by an integrative strategy based on liquid chromatography coupled with mass spectrometry, J. Sep. Sci. 40 (2017) 2722-2731.

[lxx] M. Guo, B. Zhao, H. Liu, L. Zhang, L. Peng, L. Qin, Z. Zhang, J. Li, C. Cai, X. Gao, A metabolomic strategy to screen the prototype components and metabolites of Shuang-Huang-Lian injection in human serum by ultra performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry, J. Anal. Methods Chem. 2014 (2014) 241505.

[lxxi] M.X. Guo, L. Zhang, H.Y. Liu, L.L. Qin, Z.X. Zhang, X. Bai, X.Y. Gao, A metabolomic strategy to screen the prototype components and metabolites of Qingkailing injection in rat urine by high-performance liquid chromatography with tandem mass spectrometry, J. Sep. Sci. 37 (2014) 2844-2850.

[lxxii] Y. Luo, C.J.S. Lai, J. Zhang, Y. Feng, Q. Wen, T. Tan, Comprehensive metabolic profile of phenolic acids and flavonoids in Glechomae Herba using ultra-high-performance liquid chromatography coupled to quadrupole-time-of-flight tandem mass spectrometry with diagnostic ion filtering strategy, J. Pharm. Biomed. Anal. 164 (2019) 615-629.

[lxxiii] J. Liang, F. Xu, Y.Z. Zhang, S. Huang, X.Y. Zang, X. Zhao, L. Zhang, M.Y. Shang, D.H. Yang, X. Wang, S.Q. Cai, The profiling and identification of the absorbed constituents and metabolites of Paeoniae Radix Rubra decoction in rat plasma and urine by the HPLC-DAD-ESI-IT-TOF-MS^n technique: a novel strategy for the systematic screening and identification of absorbed constituents

and metabolites from traditional Chinese medicines, J. Pharm. Biomed. Anal. 83 (2013) 108-121.

[lxxiv] N. Mi, T. Cheng, H. Li, P. Yang, X. Mu, X. Wang, X. Zu, X. Qi, X. Guo, J. Ye, W. Zhang, Metabolite profiling of traditional Chinese medicine formula Dan Zhi Tablet: An integrated strategy based on UPLC-QTOF/MS combined with multivariate statistical analysis, J. Pharm. Biomed. Anal. 164 (2019) 70-85.

[lxxv] S. Lin, X. Yue, D. Ouyang, Q. Li, P. Yang, The profiling and identification of chemical components, prototypes and metabolites of Run-zao-zhi-yang capsule in rat plasma, urine and bile by an UPLC-Q-TOF/MS^E-based high-throughput strategy, Biomed. Chromatogr. 32 (2018) e4261.

[lxxvi] X.Y. Tang, J.X. Zeng, Z.Q. Dai, M.H. Chen, M.N. Ye, Z.H. Yao, Y. Dai, X.S. Yao, Identification and characterization of chemical constituents in Qi-Lin pills and their metabolites in rat bio-samples after oral administration using ultra-high performance liquid chromatography with quadrupole time-of-flight mass spectrometry, J. Pharm. Biomed. Anal. 188 (2020) 113402.

[lxxvii] T. Yi, L. Zhu, G.Y. Zhu, Y.N. Tang, J. Xu, J.Y. Fan, Z.Z. Zhao, H.B. Chen, HSCCC-based strategy for preparative separation of in vivo metabolites after administration of an herbal medicine: Saussurea laniceps, a case study, Sci. Rep. 6 (2016) 33036.

[lxxviii] Y. Li, X.W. Yang, Five new eudesmane-type sesquiterpenoid lactones biotransformed from atractylenolide I by rat hepatic microsomes, Fitoterapia 85 (2013) 95-100.

[lxxix] J. Tang, Z.T. Wang, T. Akao, M. Hattori, Further Evidence on the intramolecular lactonization in rat liver microsomal metabolism of 12-O-acetylated retronecine-type pyrrolizidine alkaloids, Nat. Prod. Commun. 8 (2013) 1545-1546.

[lxxx] W.J. Wu, N. Hu, Q.W. Zhang, Y.P. Li, P. Li, R. Yan, Y.T. Wang, In vitro glucuronidation of five rhubarb anthraquinones by intestinal and liver microsomes from humans and rats, Chem.-Biol. Interact. 219 (2014) 18-27.

[lxxxi] S.J. Liu, G.L. Dai, L.N. Sun, B.T. Sun, D. Chen, L. Zhu, Y. Wang, L. Zhang, P.D. Chen, D. Zhou, W.Z. Ju, Biotransformation and metabolic profile of limonin in rat liver microsomes, bile, and urine by high-performance liquid

chromatography coupled with quadrupole time-of-flight mass spectrometry, J. Agr. Food Chem. 66 (2018) 10388-10393.

[lxxxii] H. Yuan, Q. Ma, L. Ye, G. Piao, The traditional medicine and modern medicine from natural products, Molecules 21 (2016) 559.

[lxxxiii] H.P. Song, H. Zhang, Y. Fu, H.Y. Mo, M. Zhang, J. Chen, P. Li, Screening for selective inhibitors of xanthine oxidase from Flos Chrysanthemum using ultrafiltration LC-MS combined with enzyme channel blocking, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 961 (2014) 56-61.

[lxxxiv] H.P. Song, J. Chen, J.Y. Hong, H.P. Hao, L.W. Qi, J. Lu, Y. Fu, B. Wu, H. Yang, P. Li, A strategy for screening of high-quality enzyme inhibitors from herbal medicines based on ultrafiltration LC-MS and in silico molecular docking, Chem. Commun. 51 (2015) 1494-1497.

[lxxxv] J.F. Zhang, W.C. Zhong, Y.C. Li, Y.Q. Song, G.Y. Xia, G.H. Tian, G.B. Ge, S. Lin, Bioactivity-guided discovery of human carboxylesterase inhibitors from the roots of Paeonia lactiflora, J. Nat. Prod. 83 (2020) 2940-2949.

[lxxxvi] Y. Ma, W. Wang, J. Yang, S. Zhang, Z. Li, F. Li, S. Huang, L. Lei, K. Wang, A. Wen, Y. Ding, A network pharmacology technique to investigate the synergistic mechanisms of Salvia miltiorrhiza and Radix puerariae in treatment of cardio-cerebral vascular diseases, Evid.-Based Compl. Alt. Med. 2020 (2020) 6937186.

[lxxxvii] L. Guo, L. Duan, L.L. Dou, L.L. Liu, H. Yang, E.H. Liu, P. Li, Quality standardization of herbal medicines using effective compounds combination as labeled constituents, J. Pharm. Biomed. Anal. 129 (2016) 320-331.

[lxxxviii] H.Q. Pang, H.M. An, H. Yang, S.Q. Wu, J.L. Fan, L. Mi, H. Wang, P. Li, W. Gao, Comprehensive chemical profiling of Yindan Xinnaotong soft capsule and its neuroprotective activity evaluation in vitro, J. Chromatogr. A 1601 (2019) 288-299.

[lxxxix] N. Wang, P. Xu, X. Wang, W. Ya, Z. Yu, R. Wu, X. Huang, Y. Si, D. Shou, Integrated pathological cell fishing and network pharmacology approach to investigate main active components of Er-Xian decotion for treating osteoporosis, J. Ethnopharmacol. 241 (2019) 111977.

[xc] J.L. Zhou, J.Y. Zheng, X.Q. Cheng, G.Z. Xin, S.L. Wang, T. Xie, Chemical

markers' knockout coupled with UHPLC-HRMS-based metabolomics reveals anti-cancer integration effects of the curcuminoids of turmeric (Curcuma longa L.) on lung cancer cell line, J. Pharm. Biomed. Anal. 175 (2019) 112738.

[xci] Y. Zhang, J. Yu, W. Zhang, Y. Wang, Y. He, S. Zhou, G. Fan, H. Yang, Y. Zhu, P. Li, An integrated evidence-based targeting strategy for determining combinatorial bioactive ingredients of a compound herbal medicine Qishen Yiqi dripping pills, J. Ethnopharmacol. 219 (2018) 288-298.

[xcii] Z. Yang, Q. Shao, Z. Ge, N. Ai, X. Zhao, X. Fan, A bioactive chemical markers based strategy for quality assessment of botanical drugs: Xuesaitong injection as a case study, Sci. Rep. 7 (2017) 2410.

[xciii] Y. Li, Y. Wang, B. Yang, Y. Wang, Z. Hou, A. Li, Y. Xu, L. Ju, H. Wu, Y. Zhang, A practical and novel "standard addition" strategy to screen pharmacodynamic components in traditional Chinese medicine using Heishunpian as an example, Rsc Adv. 5 (2015) 22209-22216.

[xciv] H.P. Song, H. Wang, X. Zhao, L. He, H. Zhong, S.Q. Wu, P. Li, H. Yang, Label-free pharmacological profiling based on dynamic mass redistribution for characterization and authentication of hazardous natural products, J. Hazard. Mater. 333 (2017) 265-274.

[xcv] J. Chen, Z.Z. Ge, W. Zhu, Z. Xu, C.M. Li, Screening of key antioxidant compounds of longan (Dimocarpus longan Lour.) seed extract by combining online fishing/knockout, activity evaluation, fourier transform ion cyclotron resonance mass spectrometry, and high-performance liquid chromatography electrospray ionization mass spectrometry methods, J. Agr. Food Chem. 62 (2014) 9744-9750.

[xcvi] C. Qu, L.Y. Wang, H. Lin, E.X. Shang, Y.P. Tang, S.J. Yue, Y. Jin, W.W. Tao, S.P. Li, Y.Q. Hua, P. Liu, S.L. Su, H. Zhou, D.W. Qian, J.A. Duan, Hierarchical identification of bioactive components in a medicinal herb by preparative high-performance liquid chromatography and selective knock-out strategy, J. Pharm. Biomed. Anal. 135 (2017) 206-216.

[xcvii] D. Yan, J. Li, Y. Xiong, C. Zhang, J. Luo, Y. Han, R. Wang, C. Jin, H. Qian, J. Li, L. Qiu, C. Peng, Y. Lin, X. Song, X. Xiao, Promotion of quality standard of herbal medicine by constituent removing and adding, Sci. Rep. 4 (2014) 3668.

[xcviii] Z.Q. Shi, D.F. Song, R.Q. Li, H. Yang, L.W. Qi, G.Z. Xin, D.Q. Wang, H.P. Song, J. Chen, H. Hao, P. Li, Identification of effective combinatorial markers for quality standardization of herbal medicines, J. Chromatogr. A 1345 (2014) 78-85.

[xcix] P. Martviset, W. Chaijaroenkul, P. Muhamad, K. Na-Bangchang, Bioactive constituents isolated from Atractylodes lancea (Thunb.) DC. rhizome exhibit synergistic effect against cholangiocarcinoma cell, J. Exp. Pharmacol. 10 (2018) 59-64.

[c] P.Y. Gong, Y.W. He, J. Qi, C.Z. Chai, B.Y. Yu, Synergistic nourishing 'Yin' effect of iridoid and phenylpropanoid glycosides from Radix Scrophulariae in vivo and in vitro, J. Ethnopharmacol. 246 (2020) 112209.

[ci] X. Guan, X. Zheng, V. Chi Teng, J. Zhao, J. Xiao, Y. Wang, Z. Zhong, Combined effects of berberine and evodiamine on colorectal cancer cells and cardiomyocytes in vitro, Eur. J. Pharmacol. 875 (2020) 173031.

[cii] F. Long, H. Yang, Y. Xu, H. Hao, P. Li, A strategy for the identification of combinatorial bioactive compounds contributing to the holistic effect of herbal medicines, Sci. Rep. 5 (2015) 12361.

[ciii] S.Q. Wu, H.P. Song, B. Li, R.Z. Liu, H. Yang, L. He, P. Li, A fast and accurate method for the identification of peroxidase inhibitors from Radix Salvia Miltiorrhizae by on-flow biochemical assay coupled with LC/Q-TOF-MS: comparison with ultrafiltration-based affinity selection, Anal. Bioanal. Chem. 410 (2018) 4311-4322.

[civ] H.P. Song, S.Q. Wu, H. Hao, J. Chen, J. Lu, X. Xu, P. Li, H. Yang, A chemical family-based strategy for uncovering hidden bioactive molecules and multicomponent interactions in herbal medicines, Sci. Rep. 6 (2016) 23840.

[cv] D.W.S. Cheung, C.M. Koon, C.F. Ng, P.C. Leung, K.P. Fung, S.K.S. Poon, C.B.S. Lau, The roots of Salvia miltiorrhiza (Danshen) and Pueraria lobata (Gegen) inhibit atherogenic events: A study of the combination effects of the 2-herb formula, J. Ethnopharmacol. 143 (2012) 859-866.

[cvi] H.P. Song, S.Q. Wu, L.W. Qi, F. Long, L.F. Jiang, K. Liu, H. Zeng, Z.M. Xu, P. Li, H. Yang, A strategy for screening active lead compounds and functional compound combinations from herbal medicines based on pharmacophore

filtering and knockout/knockin chromatography, J. Chromatogr. A 1456 (2016) 176-186.

[cvii] P. Liu, H. Yang, F. Long, H.P. Hao, X. Xu, Y. Liu, X.W. Shi, D.D. Zhang, H.C. Zheng, Q.Y. Wen, W.W. Li, H. Ji, X.J. Jiang, B.L. Zhang, L.W. Qi, P. Li, Bioactive equivalence of combinatorial components identified in screening of an herbal medicine, Pharm. Res. 31 (2014) 1788-1800.

[cviii] H.P. Song, H. Wang, J.X. Liang, C. Qian, S.Q. Wu, W.J. Xu, B. Wu, X.G. Liu, P. Li, H. Yang, Integration of multiple analytical and computational tools for the discovery of high-potency enzyme inhibitors from herbal medicines, ChemMedChem 11 (2016) 2588-2597.

[cix] X.G. Liu, C.Y. Cheng, J.X. Wang, H. Luo, L.F. Tu, L. Lin, B. Wu, H.Y. Wang, K. Liu, P. Li, H. Yang, A metabolic exposure-oriented network regulation strategy for the identification of effective combination in the extract of Ginkgo biloba L, J. Pharm. Biomed. Anal. 149 (2018) 151-159.

[cx] A.C. Servais, B. Janicot, A. Takam, J. Crommen, M. Fillet, Liquid chromatography separation of the chiral prodrug eslicarbazepine acetate and its main metabolites in polar organic mode. Application to their analysis after in vitro metabolism, J. Chromatogr. A 1467 (2016) 306-311.

[cxi] Y. Sun, X. Lu, Y. Gai, C. Sha, G. Leng, X. Yang, W. Liu, LC-MS/MS method for the determination of the prodrug aripiprazole lauroxil and its three metabolites in plasma and its application to in vitro biotransformation and animal pharmacokinetic studies, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 1081 (2018) 71-79.

Figure captions

Fig. 1 Schematic diagram of research strategies for traditional Chinese medicine based on set theory.

Fig. 2 Technical route for component analysis of traditional Chinese medicine.

Fig. 3 (A) Technical route for metabolite analysis of traditional Chinese medicine. (B) Venn diagram for identification of prototypes and metabolites of traditional Chinese medicine.

Fig. 4 Technical route for discovery of bioactive components from traditional Chinese medicine.